Search for First-Generation Leptoquarks in the Jets and Missing Transverse Energy Topology

D. Acosta, J. Konigsberg, A. Moorhead, D. Tsybychev, S.M. Wang

*University of Florida, Gainesville, USA*

**Abstract**

We report on a search for the pair production of first-generation leptoquarks using 76 pb$^{-1}$ of proton-antiproton collision data recorded by the CDF experiment during Run 2 of the Tevatron. The leptoquarks are sought via their decay into a neutrino and quark, which yields missing transverse energy and several high-$E_T$ jets. Several control regions are studied to check the background estimation from Standard Model sources, with good agreement observed in data. In the leptoquark signal region, 42 events are observed with $42 \pm 11$ expected from background. Therefore, no evidence for leptoquark production is observed, and limits are set on the cross section times squared branching ratio. Using the next-to-leading order cross section for leptoquark production, we exclude the mass interval 60 to 107 GeV/$c^2$ at the 95% confidence level for 100% branching ratio into neutrino plus quark.
1 Introduction

Leptoquarks are color-triplet bosons carrying both lepton and baryon quantum numbers that are predicted in many extensions of the Standard Model (e.g., Grand Unification models, Technicolor, and Supersymmetry with R-parity violation). They may be produced in pairs in $p\bar{p}$ collisions with a cross section essentially independent of the Yukawa coupling of the leptoquark to a lepton and quark. The branching ratio to a charged lepton, denoted by $\beta$, is model-dependent. In the general phenomenological framework of Buchmüller, Rückl, and Wyler [1], which respects the symmetries of the Standard Model, this branching ratio is either 0, 1/2, or 1. Usually it is assumed that leptoquarks couple to one generation only to avoid flavor-changing neutral current constraints, which allows one to classify leptoquarks as first-, second-, or third-generation.

Limits on leptoquark production from the Tevatron and from HERA as of 1999 are summarized in [2]. In particular, both CDF [3] and DØ [4] published limits on first-generation leptoquarks in the $eejj$ channel shortly after the HERA experiments announced [5] an anomalous excess of electron plus jet events in 1997. The combined mass limit [6] of 242 GeV/$c^2$ for scalar leptoquarks with $\beta = 1$ from CDF and DØ ruled out a leptoquark interpretation of the HERA anomaly for large $\beta$. More recently, DØ has published [7] an improved limit on the search for first-generation leptoquarks in the $\nu\nujj$ final state, and CDF has published [8] limits on a search for second- and third-generation leptoquarks in the MET plus heavy-flavor jets final state ($\nu\nuQQ$). Preliminary limits from searches by CDF for first-generation leptoquarks decaying in the $evjj$ channel [9] and in the $\nu\nujj$ channel [10] were never published, but are similar to the published DØ limits [7].

Figure 1 shows the DØ combined limit for first-generation scalar leptoquarks in the $\beta$ vs. mass plane. Also shown is the CDF limit from the published $eejj$ channel interpreted for values of $\beta$ other than 1. The aim of this study is to improve the sensitivity near the $\beta = 0$ region by conducting a search for the decay of first-generation leptoquarks in the $\nu\nujj$ channel, that is, in the missing transverse energy (MET) and multi-jets channel.

2 Theory

Figure 2 shows the leading-order (LO) diagrams for the pair production of leptoquarks at the Tevatron. For Yukawa couplings of electromagnetic strength or less, diagram (b) contributes less than 1% to the total cross section; so the cross section is essentially determined by the known QCD couplings. Krämer et al. [11] have calculated the next-to-leading order (NLO) contributions to the cross section for scalar leptoquark production\footnote{The NLO cross section for vector leptoquark production has not been calculated, since the possibility of anomalous couplings complicates the situation.}, which significantly reduces the dependence of the cross section on the renormalization and factorization scales. Programs to compute the cross section from these calculations have been obtained [12] from the authors [11]. The leading order cross section obtained from these programs, and from
Figure 1: Run 1 limits from CDF and DØ on scalar leptoquarks in the $\beta$ vs. mass plane.

the Pythia[15] event generator for comparison\textsuperscript{2}, is shown in Fig. 3 as a function of the leptoquark mass for $\sqrt{s} = 1.96$ TeV and for both the CTEQ4L and CTEQ5L parameterizations of the parton densities. The cross section is 15\% larger at leading order for CTEQ4L versus CTEQ5L for the mass range shown.

Figure 4 shows the next-to-leading order scalar leptoquark cross section for $\sqrt{s} = 1.96$ TeV, again for CTEQ4 and CTEQ5. The cross section for a leptoquark mass of 100 GeV/$c^2$ is 15.8 pb using CTEQ4, and is 8\% smaller using CTEQ5 at this mass; however, for masses above about 150 GeV/$c^2$ the CTEQ5 cross section becomes larger. Figure 5 shows the next-to-leading order scalar leptoquark cross section for 3 choices of the factorization/renormalization scale: $\mu = m_{\text{LQ}}$, $\mu = 2m_{\text{LQ}}$, and $\mu = 0.5m_{\text{LQ}}$. The dependence on $\mu$ is significantly reduced using the NLO calculation, where the cross section varies by $\pm 15\%$ for the range of $\mu$ chosen at $M = 100$ GeV/$c^2$. The ratio of the NLO to LO cross sections is shown in Fig. 6.

Figure 7 shows the ratio of the cross sections for $\sqrt{s} = 1.96$ TeV and $\sqrt{s} = 1.8$ TeV. As can be seen, the increase in the center-of-mass energy increases the cross section by 30\% for a mass of 100 GeV/$c^2$, which offers the possibility to set stronger limits on leptoquark production even for a data sample equivalent in size to Run 1.

\textsuperscript{2}For this study, the number of flavors contributing to $\alpha_s$ was set to 5 in Pythia.
Figure 2: Leading-order Feynman diagrams for leptoquark pair production at the Tevatron.

Figure 3: Leading order cross section for scalar leptoquark production for 2 choices of the parton density parameterizations: CTEQ4L and CTEQ5L. Good agreement is seen between the Pythia event generator and the programs from Plehn and Krämer.
Figure 4: Next-to-leading order cross section for scalar leptoquark production for both the CTEQ4 and CTEQ5 parton densities.

Figure 5: Next-to-leading order cross section for scalar leptoquark production for 3 choices of the factorization/renormalization scale: $\mu = m_{LQ}$, $\mu = 2m_{LQ}$, and $\mu = 0.5m_{LQ}$. CTEQ5 parton densities are chosen.
Figure 6: Ratio of the NLO to LO cross sections for scalar leptoquark production for both CTEQ4 and CTEQ5.

Figure 7: Ratio of the NLO cross section for scalar leptoquark production at $\sqrt{s} = 1.96$ TeV to $\sqrt{s} = 1.8$ TeV.
3 Data Selection

The search for leptoquark pair-production and decay into the $\nu\nujj$ final state centers on selecting events with large MET, so the inclusive $\text{MET45}$ data sample is used for this search. The data collected from Feb 9, 2002 through Jan 9, 2003—runs 138815 through 156369—are used in this analysis. The data from the E stream were processed with offline software version 4.8.4a and written to emet08 dataset.

We reprocessed the data with offline software 4.9.1 and created stntuples version dev.239. A number of corrections were applied to the data:

- tower to tower gain corrections in CEM
- global and time dependent energy scale in CEM
- global energy scale 4% in CHA

All the objects that rely on calorimetry were dropped and remade with the corrected energy scales. Also, we dropped hit information from layer 0 of the silicon detector and refit all tracks and remade all vertices.

Missing transverse energy and jets were reconstructed using the highest $\sum |P_T|$ vertex from $\text{ZVertexCollection}$ with quality 12 or higher [14] (i.e. at least two tracks with COT hits are required). We use jets reconstructed with JetClu algorithm of cone size 0.4.

Also $\text{JET20}$, $\text{JET50}$ and $\text{JET100}$ secondary data samples created by QCD group[13] were used for QCD multijet background normalization and trigger efficiency studies.

3.1 Trigger Path

The data were collected using inclusive missing $E_T$ trigger, $\text{MET45}$. The online MET is calculated by summing over trigger towers with energies above 1 GeV. The Level-1 trigger requires online MET to be greater than 25 GeV. The Level-2 trigger automatically accepts events that passed Level-1. At Level-3, the MET is recalculated using full calorimeter information and using an event vertex of $z = 0$. The individual energy threshold is 100 MeV. The Level-3 MET threshold is 45 GeV.

3.2 Good Run List

For the data used in this analysis we require that the detector components that are relevant to the leptoquark search in MET+2 jets signature to be in proper working condition. A good run list was made by requiring that several online and offline data quality bits to be set. These bits are:

- $\text{RUNCONTROL\_STATUS}$
- $\text{SHIFTCREW\_STATUS}$
- $\text{OFFLINE\_STATUS}$

7
• GOODRUN\_STATUS :
  - CLC\_STATUS
  - L1T\_STATUS
  - L2T\_STATUS
  - L3T\_STATUS
  - CAL\_STATUS
  - COT\_STATUS
  - CMU\_STATUS
  - SMX\_STATUS

• Detector offline bits :
  - CAL\_OFFLINE
  - COT\_OFFLINE
  - CMU\_OFFLINE
  - CMP\_OFFLINE
  - CMX\_OFFLINE (only for run \( \geq \) 150145, the status of CMX is not well understood before run 150145 ???)

These bits are applied on runs taken between Feb 9, 2002 and Jan 12, 2003 (runs 138819–156487). Although run 154654 is marked good in the CAL\_OFFLINE status bit, there is a CEM PMT (left) in wedge 23 East which was very noisy in this particular run. We excluded this run from our good run list. The total integrated luminosity of the good runs used for this analysis is 75.9 pb\(^{-1}\) (after scaling up by factor 1.019 [16]).

3.3 Data Clean Up

Several sources of non-collision backgrounds and detector inefficiencies can spoil the energy measurement in the calorimeter and cause events to appear to have large missing \( E_T \). If the MET is large enough, the event can end up in our sample. Special care needs to be taken to discard such events in searches for new physics in the tails of the missing-\( E_T \) distribution. Some of these sources are:

• Beam halo muons that bremsstrahlung in the calorimeter in coincidence with a real \( p\bar{p} \) collision.

• Cosmic ray showers in the calorimeter.

• \( p\bar{p} \) collisions produced far from the nominal collision point. In this case some of the energy could escape detection.
• QCD multi-jet events where one of the jets falls into a crack in the calorimeter or an non-instrumented region.

• Energy response mis-calibration in the calorimeter.

• Two or more $p\bar{p}$ collision in the same bunch crossing.

A secondary dataset was derived after applying clean-up cuts similar to those of Run I [17]:

- At least one vertex from ZVertexCollection of quality 12 or higher with $|Z_{\text{vertex}}| < 60$ cm.
- At least one central jet, $E_T > 10$ GeV, $\eta_d < 0.9$
- Event electromagnetic fraction (EEMF) greater than 0.1, where EEMF is defined as

$$EEMF = \frac{\sum_{j=1}^{N_{\text{jet}}} E_T^j \cdot EMF_j}{\sum_{j=1}^{N_{\text{jet}}} E_T^j},$$

and $EMF_j$ is the electromagnetic fraction of the jet. Only jets with $E_T > 10$ GeV are considered.

- Event charge fraction (ECHF) greater than 0.1, where ECHF is defined as

$$ECHF = \frac{\sum_{j=1}^{N_{\text{jet}}} \left( \frac{\sum_{i=1}^{\text{tracks}} P_{T,i}}{E_{T,j}} \right)_j}{N_{\text{jet}}}.$$  

Tracks containing COT hits, with $P_T > 0.5$ GeV and within 10 cm from event vertex are associated to the jet by requiring that the track is within an $\eta-\phi$ cone of radius 0.4 from the jet centroid. Jets with $E_T > 10$ and $\eta_d < 0.9$ are considered.

- Total calorimeter energy less than 2 TeV.

Table 1 shows the sample reduction with each cut.

<table>
<thead>
<tr>
<th>Cut</th>
<th># of Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>emet08 sample</td>
<td>1,164,771</td>
</tr>
<tr>
<td>Good Run</td>
<td>758830</td>
</tr>
<tr>
<td>Vertex requirement</td>
<td>193298</td>
</tr>
<tr>
<td>Central jet</td>
<td>159410</td>
</tr>
<tr>
<td>EEMF&gt;0.1 and ECHF&gt;0.1</td>
<td>66342</td>
</tr>
<tr>
<td>$E_T &gt; 45$ GeV</td>
<td>39389</td>
</tr>
</tbody>
</table>
3.4 Calorimeter Mis-calibration and Jet Fiduciality Requirement

It has been observed that the relative energy scale for the CHA part of the calorimeter is not uniform in $\phi$ in the same $\eta$ ring [18]. Figure 8 shows the leading-jet $\phi$ distribution for di-jet events with large missing $E_T$ and jets back-to-back in $\phi$, which is expected to be dominated by QCD multi-jet background. All other selection cuts used in the leptoquark analysis and described later in section 6.1 are applied, except for the topological cut on the angle between the missing $E_T$ direction and the leading jet, and the minimum angle between the missing $E_T$ direction and any jet. Also, the requirement on the minimum number of the tracks in the jet was removed. Two big peaks are observed in the $\phi$ distribution at $\phi = 1.7$ rad and $\phi = 4.5$ rad, which are traced to an energy mis-calibration in the calorimeter. The peaks translate to the non-uniform missing $E_T$ $\phi$ distribution shown on Fig. 9. Figure 10 shows the calorimeter occupancy in the $\eta$-$\phi$ space for the direction of the leading jet. Some “hot” regions are readily seen on this plot.

To reduce the effect of the calorimeter mis-calibration on our background normalization, we remove events where the leading jet contains as a seed one of the towers listed in Tab. 2. This requirement removes about 10% of the events. The same fiduciality requirement is applied to the Monte Carlo samples that are used for the background estimation and for the leptoquark signal acceptance calculation.

![Figure 8](image_url)

Figure 8: Leading jet $\phi$ distribution for events with large missing $E_T$ and jets back-to-back in $\phi$. 

\[ \text{Norm: 1 Jet: Phi} \]
Figure 9: Missing $E_T$ $\phi$ distribution for events with large missing $E_T$ and jets back-to-back in $\phi$.

Figure 10: Leading jet occupancy in the central hadronic calorimeter in $\eta$-$\phi$ space ($x$ and $y$ axes, respectively.) Numbers show the total number of events in the given bin.
Table 2: Towers marked as bad for the jet fiducial requirement. Offline classification is used [19].

<table>
<thead>
<tr>
<th>$\eta$ Index</th>
<th>$\phi$ Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>16 - 35</td>
<td>17</td>
</tr>
<tr>
<td>25</td>
<td>6</td>
</tr>
<tr>
<td>25</td>
<td>16</td>
</tr>
<tr>
<td>28</td>
<td>6</td>
</tr>
<tr>
<td>29 - 33</td>
<td>0</td>
</tr>
</tbody>
</table>

4 Trigger Efficiency

In order to be able to predict number of background events from Standard Model processes, we need to understand the missing-$E_T$ trigger efficiency. We use the JET50 sample to measure the combined Level-1 and Level-3 $E_T$ trigger efficiency. Since the JET50 trigger path also relies on the calorimeter, it could be biased with respect to $E_T$ trigger path. To cross-check this efficiency, we also use the high-$P_T$ muon sample defined in [20].

We re-process the samples using the highest $\sum |P_T|$ vertex to calculate MET and the transverse energies of the jets. In addition, we require at least one vertex from ZVertexCollection of quality 12 or higher with $|Z_{vtx}| < 60$ cm. Figure 11 shows the offline MET from the JET50 sample before and after the requirement that the L3 MET.45 trigger was set.

The combined Level-1 and Level-3 missing $E_T$ trigger efficiency is defined as:

$$\epsilon(\text{MET45}) = \frac{\# \text{of events passing MET45 and JET50}}{\# \text{total events passing JET50}}$$

Figures 12 and 13 show the trigger efficiency curves from JET50 and high-$P_T$ muon samples respectively. The curve is fit with function:

$$\epsilon(E_T) = \frac{\epsilon_o}{1 + exp(-\frac{E_T - A_o}{\Delta})}.$$  

The fit parameters for both samples are summarized in Table 3. The difference between the two parameterizations is used as a systematic uncertainty on our Standard Model background prediction.

Table 3: Parametrization of the missing $E_T$ trigger efficiency.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\epsilon_o$</th>
<th>$A_o$</th>
<th>$\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>JET50</td>
<td>0.989 ± 0.003</td>
<td>45.61 ± 0.11</td>
<td>3.37 ± 0.04</td>
</tr>
<tr>
<td>Muon Sample</td>
<td>0.979 ± 0.009</td>
<td>45.32 ± 0.15</td>
<td>1.92 ± 0.08</td>
</tr>
</tbody>
</table>
Figure 11: Missing $E_T$ distribution in JET50 sample. Empty histogram—all events, filled histogram—after MET45 trigger requirement.

Figure 12: The MET45 trigger efficiency as a function of offline $E_T$ from the JET50 sample.
Figure 13: The MET45 trigger efficiency as a function of offline $E_T$ from the high $p_T$ muon sample.

5 Monte Carlo Generation

The dominant backgrounds to the leptoquark search are expected to include QCD jet production, top quark pair production, and $W$ and $Z$ boson production with one or more jets. The Pythia event generator[15], version 6.203, was used for the QCD background (discussed below) and the leptoquark signal. Pythia was tuned to R. Field’s “tune A” [21] for the underlying event. The ALPGEN generator [22] was used for the $W/Z$ boson production plus jets. The Herwig MC program[23] was used for parton showering of these electroweak processes, and the cross section for the electroweak processes was later corrected to next-to-leading order using the MCFM program [24]. The specific background channels simulated are listed in Table 4.

The first-generation scalar leptoquark signal was generated in 9 mass bins (60, 75, 90, 100, 110, 125, 150, 175, and 200 GeV/$c^2$) with 30K events each. In addition, mass bins 100 and 150 were generated without final-state QCD radiation and with R. field tune B, for systematic studies.

Several QCD di-jet samples of 500K events were generated by the MC Production group with the hard scattering $p_T^{\text{min}}$ cut at 5, 10, 40, 60, and 90 GeV/$c$, and 3 million events with the $p_T^{\text{min}}$ cut at 18 GeV/$c$. In addition, a large sample of 2 million events with $60 < p_T < 90$ GeV/$c$ was generated to increase the statistics in the region where small number of events passing the analysis cuts.

All Monte Carlo samples used CTEQ5L for the parton densities and were reconstructed
Table 4: List of MC background samples used, number of events generated, leading-order cross section, and equivalent integrated luminosity.

<table>
<thead>
<tr>
<th>Background Channel</th>
<th>Number of events</th>
<th>LO Cross Section (pb)</th>
<th>Luminosity (pb⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W \rightarrow e\nu + 2$ jets</td>
<td>187.5K</td>
<td>258.7</td>
<td>538.33</td>
</tr>
<tr>
<td>$W \rightarrow \mu\nu + 2$ jets</td>
<td>260K</td>
<td>258.7</td>
<td>746.48</td>
</tr>
<tr>
<td>$W \rightarrow \tau\nu + 2$ jets</td>
<td>127.5K</td>
<td>258.7</td>
<td>366.06</td>
</tr>
<tr>
<td>$Z \rightarrow \mu\mu + 2$ jets</td>
<td>161K</td>
<td>23.3</td>
<td>4597</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau + 2$ jets</td>
<td>167.5K</td>
<td>23.3</td>
<td>4785</td>
</tr>
<tr>
<td>$Z \rightarrow \nu\nu + 2$ jets</td>
<td>170K</td>
<td>139.8</td>
<td>809</td>
</tr>
<tr>
<td>$t\bar{t}(m = 175 \text{ GeV})$ inclusive</td>
<td>38K</td>
<td>7</td>
<td>5428</td>
</tr>
</tbody>
</table>

using version 4.9.1 of the cdfsoft2 software.

6 Cut Optimization

The search for leptoquark pair-production and decay into the $\nu\nu jj$ final state has been optimized using the Monte Carlo samples described earlier. The analysis centers on selecting events with large MET and 2 or 3 high $E_T$ jets in the central region of the detector. Parameters that can be tuned include the MET threshold, the minimum $E_T$ of the jets, the $\eta$ reach of the jets, the angle between the two highest $E_T$ jets, and the angles between the two highest $E_T$ jets and the MET direction.

The distributions of these parameters, without applying selection cuts, are shown in Figure 14, 15, 16, 17, 18, 19, and 20. All these distributions are normalized to the same area.

The MET distribution from the pair production of first-generation leptoquarks ($m_{LQ} = 100$ GeV/$c^2$) is compared to the MET distributions from the dominant Standard Model background processes in Figure 14. The MET from leptoquark decays, which is mainly due to the transverse energy carried by the escaping neutrinos from the decay $LQ \rightarrow q\nu$, is larger than that from $W/Z + 2$ jets processes but comparable to that from $t\bar{t}$ production.

The $E_T$ distributions of the first two leading jets are shown in Figure 15. The first-leading jet ($J1$) and second-leading jet ($J2$) $E_T$ distributions from the decays of the $LQ$ pair, which respectively peak at $\sim 40$ GeV and $\sim 25$ GeV, are generally higher than the corresponding jets from the $W/Z + 2$ jets processes. The $E_T$ of the first and second leading jets from the QCD di-jet process are large due to the fact that these Monte Carlo events were generated with a cut on the minimum transverse energy at the parton level of 50 GeV.

Figure 16 shows the pseudo-rapidity distributions of the first- and second-leading jets. The first-leading jet from the decays of the $LQ$ pair is more central than the first-leading jet from QCD di-jet and $W/Z + 2$ jets processes. For $t\bar{t}$ production, both the first- and second-leading jets are more central than the other processes shown in the plots.
The distribution of the azimuthal separation between the direction of the jets and MET, and between the first- and second-leading jets are shown in Figure 17, 18, and 19. In Figure 17 one can clearly observe that the two leading jets in QCD di-jet events are back-to-back in the $x$-$y$ plane most of the time. Since the MET in QCD di-jet events is mainly due to jet energy mis-measurement, this means that usually the MET direction is parallel to the direction of the jet that is most mis-measured, as shown in Figure 18.

The other parameter which is used in the optimization studies is the number of tracks in a central jet that has the least number of tracks among all the central jets ($|\eta_{jet}| < 1$). The distribution of this parameter is shown in Figure 20. This parameter is useful in suppressing background contributions from $W(\tau\nu)/Z(\tau\tau) + 2$ jets, where $\tau$ decays hadronically, and from $W(e\nu)/Z(ee) + 2$ jets, where the electron is mis-identified as a jet. In these cases, the jet will typically has fewer tracks ($\sim 1 - 3$).

6.1 Steps in Deriving Cuts

In the previous section, we described the parameters that we would cut on to enhance the leptoquark signal and suppress the Standard Model backgrounds. The selection of the cut values were determined in several steps.

First we applied some base level cuts:

- MET $> 55$ GeV (offline value, corrected for the $z$ position of the primary vertex)
- Lepton veto
- 2 or 3 jets (cone size 0.4, $E_T > 7$ GeV, and the electromagnetic fraction of the jet between 0.1 and 0.9)

The MET cut value is selected at 55 GeV, safely above the turn-on region of the MET45 trigger efficiency curve shown in Figs. 12 and 13. The lepton veto is used for suppressing the contributions from $W/Z + n$ jets. The cuts for selecting electron and muon candidates are:

- Electron Candidate:
  - Central Electron:
    * $E_T > 10$ GeV
    * $E/P < 2$
    * Had/EM $< 0.1$
    * LShr $< 0.2$
    * $|\Delta X| < 3$ cm
    * $|\Delta Z| < 3$ cm
    * $\chi^2 < 10$
  - Plug Electron:
    * $E_T > 10$ GeV
Figure 14: The MET distribution from the pair production of first generation leptoquarks ($m_{LQ} = 100$ GeV/$c^2$) compared to the MET distributions from the dominant Standard Model background processes.
Figure 15: (Top) Comparing the $E_T$ distributions of the first-leading jet (J1) from leptoquark pair production and from dominant Standard Model processes. (Bottom) Comparing the $E_T$ distributions of the second-leading jet (J2) from leptoquark pair production and from dominant Standard Model processes.
Figure 16: (Top) Comparing the pseudo-rapidity distributions of the first-leading jet (J1) from leptoquark pair production and from dominant Standard Model processes. (Bottom) Comparing the pseudo-rapidity distributions of the second-leading jet (J2) from leptoquark pair production and from dominant Standard Model processes.
Figure 17: Comparing the distribution of the azimuthal angular separation between the first- and second-leading jets from leptoquark pair production and from dominant Standard Model processes.

Figure 18: Comparing the distribution of the minimum azimuthal angular separation between the directions of the jets and the MET from leptoquark pair production and from dominant Standard Model processes.
Figure 19: Comparing the distribution of the azimuthal angular separation between the directions of the first leading jet and the MET from leptoquark pair production and from dominant Standard Model processes.

Figure 20: Comparing the distribution of the minimum number of tracks in the central jets from leptoquark pair production and from dominant Standard Model processes.
* Had/EM < 0.1  
* $\chi^2(3 \times 3) < 10$

- Muon Candidate:
  - CMU, CMP, CMUCMP, CMX muon:
    * $P_T > 10$ GeV  
    * EM < 2 GeV  
    * Had < 6 GeV  
    * EM+Had > 0.1 GeV  
    * $|D0| < 0.5$ cm (corrected for beamline offset)  
    * $|Z0| < 60$ cm  
    * if CMU or CMUCMP  
      * $|cmu_{\Delta X}| < 3.0$ cm  
    * if CMP  
      * $|cmp_{\Delta X}| < 6.0$ cm  
    * if CMX  
      * $|cmx_{\Delta X}| < 8.0$ cm

  - CMIO muon:  
    * $P_T > 10$ GeV  
    * EM < 2 GeV  
    * Had < 6 GeV  
    * EM+Had > 0.1 GeV  
    * $|D0| < 0.5$ cm (corrected for beamline offset)  
    * $|Z0| < 60$ cm  
    * Isolation $E_T(\text{cone } 0.4) < 5$ GeV

After applying the base level cuts we looked at the azimuthal angle between the first-leading jet and the MET direction, and the azimuthal angle between the first- and second-leading jets, as shown in Figure 21 and 22. From the $S/\sqrt{S+B}$ distributions we derived the following cuts (OPT1):

- $100^\circ < \Delta \phi(J1, \text{MET}) < 165^\circ$
- $\Delta \phi(J1, J2) < 165^\circ$

In the second step we applied the base level cuts and the OPT1 cuts, and looked at the transverse energy and the pseudo-rapidity of J1 and J2. These distributions are shown in Figure 23, 24, 25, 26, and 27. From the $S/\sqrt{S+B}$ distributions we derived the next set of cuts (OPT2):

- $E_T(J1) > 40$ GeV, $E_T(J2) > 25$ GeV
• $|\eta_{J_1,J_2}| < 1$, $|\eta_{J_3}| < 2.5$

Even though the cut of $|\eta_{J_1,J_2}| < 1$ does not completely maximize $S/\sqrt{S+B}$, we want to minimize the jet energy uncertainty due to the energy scale uncertainty in the Wall and Plug calorimeter. Thus, we restrict that the two highest $E_T$ jets should be in the central region.

In the third step we applied the base level, OPT1, and OPT2 cuts. The parameters which we then looked at are shown in Figure 28, 29, and 30. The final set of cuts we derived (OPT3) are:

• $80^\circ < \Delta \phi (J_1, J_2) < 165^\circ$

• $30^\circ < \min \Delta \phi (J, \text{MET}) < 135^\circ$ (J: any of the 2 or 3 selected jets)

• Minimum number of tracks in jet $\geq 4$ (jets in $|\eta| < 1$)

The complete list of cuts is therefore:

• MET $> 55$ GeV (offline value, corrected for $Z$ position of the primary vertex)

• Lepton veto

• 2 or 3 jets (cone size 0.4, $E_T > 7$ GeV, electromagnetic fraction of the jet to be between 0.1 and 0.9)
  
  – $E_T(J_1) > 40$ GeV, $E_T(J_2) > 25$ GeV, $E_T(J_3) > 7$ GeV
  
  – $|\eta_{J_1,J_2}| < 1$, $|\eta_{J_3}| < 2.5$

• No other jet with $E_T > 7$ GeV

• $100^\circ < \Delta \phi (J_1, \text{MET}) < 165^\circ$

• $80^\circ < \Delta \phi (J_1, J_2) < 165^\circ$

• $30^\circ < \min \Delta \phi (J, \text{MET}) < 135^\circ$ (J: any of the 2 or 3 selected jets)

• Minimum number of tracks in jet $\geq 4$ (jets in $|\eta| < 1$)

### 6.2 Study of Projected Limits as Function of Cuts and Uncertainties

The derivation of cuts based on obtaining the highest $S/\sqrt{S+B}$ value does not necessarily lead to the best search limits when systematic errors on the signal acceptance and Standard Model background are included in the limit-setting procedure. In this section, we discuss how the limit changes with a variation in some of the cuts, when finite uncertainties in the signal acceptance and background estimation are included.

Figure 31 shows the expected upper limit curves (@ 95% C.L.) for three different sets of selection cuts. We assume that the number of observed events is equal to the number
of expected Standard Model background events. For the “Et(40,25) MET(55)” curve, the lower limit cut on the transverse energy of the first (second) leading jet is 40 (25) GeV, and the missing transverse energy cut is MET > 55 GeV. These are the cut values used in the nominal cuts described in the earlier section 6.1. The cut values used for the other two limit curves have either lower jet transverse energy cut and higher missing transverse energy cut, or higher jet transverse energy cut. For these three curves, we assume that the uncertainty in the signal acceptance is 30%, and the uncertainty in estimating the expected Standard Model contribution is 30%. The trigger efficiency of the MET45 dataset had not been considered in this study. In Figure 31, one can observe that the nominal cut can set a higher leptoquark mass limit compare to the other two sets of cuts.

In Figure 32 we used the nominal cuts as described in section 6.1 for all shown expected limit curves (@ 95% C.L.). We varied the signal acceptance and the expected background uncertainties. The leptoquark mass limit decreases as the uncertainties increase; however, one may observe that the leptoquark mass limit is less sensitive to the uncertainty in the signal acceptance than to the uncertainty of the expected Standard Model background contribution.
Figure 21: (Top Left) The distribution of the azimuthal angle between the first-leading jet and MET direction from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.

Figure 22: (Top Left) The distribution of the azimuthal angle between the first- and second-leading jets from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.
Figure 23: (Top Left) The distribution of the transverse energy of the first-leading jet from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.

Figure 24: (Top Left) The distribution of the transverse energy of the second-leading jet from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.
Figure 25: (Top Left) The distribution of the pseudo-rapidity of the first-leading jet from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.

Figure 26: (Top Left) The distribution of the pseudo-rapidity of the second-leading jet from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.
Figure 27: (Top Left) The distribution of the pseudo-rapidity of the third-leading jet from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.
Figure 28: (Top Left) The distribution of the azimuthal angle between the first- and second-leading jets from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.

Figure 29: (Top Left) The distribution of the minimum azimuthal angle between the selected jets and MET direction from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.
Figure 30: (Top Left) The distribution of the minimum number of tracks in the selected central jets in the from leptoquark pair production, and from Standard Model processes. (Bottom) These are the $S/\sqrt{S+B}$ distributions of the parameter shown in the top left plot.
Figure 31: The expected upper limit curves (© 95% C.L.) for three different sets of selection cuts. For the “Et(40,25) MET(55)” curve, the lower limit cut on the transverse energy of the first (second) leading jet is 40 (25) GeV, and the missing transverse energy cut is MET > 55 GeV. The uncertainties in the signal acceptance and the estimation of the expected Standard Model contribution are both assumed to be 30%.
Figure 32: The expected upper limit curves (@ 95% C.L.) for various uncertainties in the signal acceptance and the estimation of the expected Standard Model contribution. For the “(30%, 40%)” curve, the signal acceptance uncertainty is 30%, and the uncertainty in the estimation of the Standard Model contribution is 40%. Nominal selection cuts were used in this study.
7 Control Regions and Background Normalization

In addition to the signal “box” described in the previous section, several control regions are also defined to normalize the expected QCD jet contribution as well as to check the normalizations of the other backgrounds. These control regions are selected by relaxing or inverting some of the cuts used to enhance the leptoquark signal.

7.1 QCD Di-jet Background Normalization

The Monte Carlo estimation of the QCD di-jet background has been normalized to the JET20 and JET50 data. We select a region where the QCD di-jet background is dominant: low MET and with the two highest-$E_T$ jets back-to-back in the transverse plane.

Since the JET20 and JET50 trigger paths were prescaled at Level-1 and Level-2, we use the JET100 dataset to absolutely normalize the luminosity of the jet samples. Jet samples, especially JET100, also can suffer from non-collision backgrounds such as cosmic rays or beam halo. To reduce such contributions, we select events that satisfy the following criteria:

- at least one reconstructed vertex with $|z| < 60$ cm
- the leading jet is central: $|\eta_d| < 1$
- the electromagnetic fraction of the jet energy is greater than 0.1 for leading jet.
- 2 or more COT tracks with $P_T > 0.5$ GeV/c associated to the leading jet.

Figures 33, 34, 35 and 36 show the distributions of the electromagnetic fraction, number of tracks associated to the jet, pseudo-rapidity and transverse energy for the leading jet, before and after all cuts are applied (but with the vertex cut applied in any case).

To measure the relative prescale of the JET20 sample, we plot the ratio of events that pass the JET50 trigger as function of the leading jet transverse energy, and fit it to a constant. The JET50 prescale is measured using JET100 sample.

Figures 37 and 38 show the prescale normalization. The effective prescale factors for the JET20 and JET50 data samples are 15.67 and 19.64, respectively. The integrated luminosity of the JET100 sample is 69.11 pb$^{-1}$.

The QCD MC jet samples are normalized to each other by weighting events based on the hard scattering $\hat{p}_T$ using following expression:

$$ \hat{p}_T^{\text{cut}} \geq \hat{p}_T^{\text{min,j}} \rightarrow w = \left[ \frac{\hat{p}_T^{\text{min,i}}}{\sum_{i=1} L_i} \right]^{-1}, $$

where $\hat{p}_T^{\text{min,i}}$ - minimum $\hat{p}_T$ at which sample was generated, $L_i$ - luminosity of the $i$ th sample, $\hat{p}_T^{	ext{cut}}$ - reconstructed $\hat{p}_T$ of the event. Figures 39 and 40 show the distribution of the reconstructed $\hat{p}_T$ of the generated samples before and after weighting was applied.

The following control region for the QCD di-jet normalization is defined: $20 < \text{MET} < 55$ GeV, 2 jets with $E_T > 30$ GeV, and the two-highest $E_T$ jets back-to-back in $\phi$ with $\Delta \phi(j_1,j_2) > 165^\circ$. In addition, we require:
Figure 33: Electromagnetic fraction of the jet energy for the leading jet. Empty histogram before cuts, filled histogram after all cuts applied

Figure 34: Number of tracks associated with the jet. Empty histogram before cuts, filled histogram after all cuts applied
Figure 35: Distribution of detector $\eta$ for the leading jet. Empty histogram before cuts, filled histogram after all cuts applied.

Figure 36: Leading jet transverse energy distribution. Empty histogram before cut, filled histogram after all cuts applied.

Figure 37: Ratio of the event rates in JET50 sample over JET20 sample.

Figure 38: Ratio of the event rates in JET100 sample over JET50 sample.
Figure 39: Reconstructed $\hat{p}_T$ distribution of the QCD samples.

- central jets: $|\eta_d| < 1$
- allow a third jet in the event with $E_T > 7$ GeV, $\eta_d < 2.5$
- no additional jets with $E_T > 7$ GeV, anywhere in detector.

For the data, we require that the leading jet satisfy $30 < E_T < 70$ GeV for JET20 sample and $E_T > 70$ GeV for JET50 sample. Figure 41 and 42 show the MET and jet $\Delta \phi$ distribution for the region of QCD normalization.

This normalization can be applied to several other regions to check for consistency. In particular, the back-to-back cut in $\phi$ can be inverted to match the signal search (but with lower MET) by requiring $\Delta \phi < 165^\circ$. In this region, 47117 events are expected by Monte Carlo and 51352 are observed in data.

Similarly, the MET cut can be raised to match closely the signal region, MET> 55 GeV, but keeping the jets back-to-back in $\phi$. In this region, 3051 events are expected by Monte Carlo and 2710 are observed in data.

### 7.2 Large MET Control Regions

Aside from the control regions defined for QCD di-jet background normalization, additional control regions using the MET45 trigger can be defined as well. These regions are used to check the Standard Model background prediction before looking at the signal “box” There
Figure 41: $H_T$ distribution before cut $\Delta \phi$ jets is applied.

Figure 42: $\Delta \phi$ between to leading jets for the low missing $E_T$ region.

are seven control regions which are defined in Tab. 5. To increase the statistics of the regions, we relaxed some of the cuts from their values used for the leptoquark search region:

- $\Delta \phi(J_1, \text{MET}) < 160^\circ$
- $0^\circ < \Delta \phi(J_1, J_2) < 165^\circ$
- $20^\circ < \min \Delta \phi(J, \text{MET})$ (J: any of the 2 or 3 selected jets)
- The minimum number of tracks required per jet is removed

Table 5: Default energy thresholds in the calculation of missing $E_t$.

<table>
<thead>
<tr>
<th>Region definition</th>
<th>$45 &lt; E_T &lt; 55 \text{ GeV}$, $\Delta \phi(j_1, j_2) &gt; 165^\circ$, $N_l = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>$45 &lt; E_T &lt; 55 \text{ GeV}$, $\Delta \phi(j_1, j_2) &lt; 165^\circ$, $N_l = 0$</td>
</tr>
<tr>
<td>2)</td>
<td>$45 &lt; E_T &lt; 55 \text{ GeV}$, $\Delta \phi(j_1, j_2) &gt; 165^\circ$, $N_l &gt; 0$</td>
</tr>
<tr>
<td>3)</td>
<td>$E_T &gt; 55 \text{ GeV}$, $\Delta \phi(j_1, j_2) &gt; 165^\circ$, $N_l = 0$</td>
</tr>
<tr>
<td>4)</td>
<td>$E_T &gt; 55 \text{ GeV}$, $\Delta \phi(j_1, j_2) &gt; 165^\circ$, $N_l &gt; 0$</td>
</tr>
<tr>
<td>5)</td>
<td>$E_T &gt; 55 \text{ GeV}$, $\Delta \phi(j_1, j_2) &lt; 165^\circ$, $N_l &lt; 0$</td>
</tr>
<tr>
<td>6)</td>
<td>$E_T &gt; 55 \text{ GeV}$, $\Delta \phi(j_1, j_2) &lt; 165^\circ$, $N_l &lt; 0$</td>
</tr>
<tr>
<td>7)</td>
<td>$45 &lt; E_T &lt; 55 \text{ GeV}$, $\Delta \phi(j_1, j_2) &lt; 165^\circ$, $N_l &gt; 0$</td>
</tr>
</tbody>
</table>
8 Systematic Uncertainties

8.1 Uncertainties on Background and Control Regions

The systematic uncertainties on the background predictions to the leptoquark signal region, as well as the various control regions, have been determined. The systematic uncertainty on the luminosity is taken to be 6%, and the PDF uncertainty on the cross section for all processes is assumed to be 5%.

The systematic uncertainty on the energy scale of the calorimeter is estimated by mis-calibrating the CHA towers in Monte Carlo events by the amount of mis-calibration that appears to exist in data (see Sec. 3.4) when comparing the CHA MIP signals of muons identified by the CMU/CMP. These incorrect CHA LERs so far have not been corrected offline, and thus affect this analysis. To study the systematic dependence in Monte Carlo events, jet energies were adjusted by the correction factor of the highest $E_T$ tower in each jet. The overall MET also was re-calculated with the tower corrections applied. When the MET and jet $E_T$ thresholds are applied, a 5% change in the event yield is observed for the first control region. Since the calibration corrections obtained from muons are known only for approximately half of the CHA, the overall systematic uncertainty is doubled to 10% and applied to all regions.

The Monte Carlo simulation of the QCD multi-jet background was normalized to the data in a lower MET region and for di-jets that are back-to-back in $\phi$, as explained in Sec. 7.1. The systematic uncertainty on the prediction of the QCD multi-jet background is taken to be the full value of the increase in the the LO prediction of Pythia needed for the normalization, that is 41%.

The uncertainty on the cross section of the vector boson background ($W/Z + 2$ jets) is estimated by varying the renormalization/factorization scale, $\mu$, in the NLO cross section obtained from MCFL [24] by a factor of two larger and smaller than the default scale $\mu = \sqrt{Q^2} \equiv \sqrt{V^2 + p_t^2}$. The uncertainty is 12% for both $W + 2$ jets and $Z + 2$ jets. The statistical error on the contributions from each Monte Carlo sample are also included, and can be as large as 100% for the QCD jet contribution to the leptoquark signal region. But when combined with other background sources, the statistical uncertainty coming from the Monte Carlo samples is 17% for the leptoquark signal region. Overall, the total systematic uncertainty on the background contribution to the leptoquark signal region is 25%.

8.2 Uncertainties on Signal Acceptance

The systematic uncertainties on the acceptance for the first-generation scalar leptoquark signal are listed in Tab. 6. The uncertainty on the parton density of the proton is estimated from the difference in the NLO cross section between the CTEQ4M and CTEQ5M parameterizations at $m_{Q} = 100$ GeV$/c^2$, where it is found to be 8%. The systematic uncertainty on the initial state radiation model used by Pythia is estimated from the difference in acceptance between the “A” and “B” tunes of R. Field [21]. The systematic uncertainty on final state radiation is estimated by completely switching off final state radiation (since
leptoquarks could hadronize if long-lived), and gives the larger acceptance variation of 10%. The finite statistics of the leptoquark Monte Carlo samples gives a 3% statistical error. The theoretical uncertainties on the renormalization and factorization scales are not included here, since we conservatively choose the NLO cross section for $\mu = 2m_{LQ}$, which is found to reduce the cross section by 15% relative to $\mu = m_{LQ}$.

On the experimental side, the systematic uncertainty on the luminosity is 6%, and the uncertainty on the $z$-vertex cut is 0.5%. The uncertainty on the MET trigger efficiency, 2%, is estimated from the difference in acceptance if the efficiency curve derived from a high-$p_T$ muon sample is used in place of the one derived from a di-jet sample. The uncertainty on the calorimeter energy scale is taken to be the same as for the background study, that is 10%, as explained in the previous subsection. Overall, the total systematic uncertainty assigned to the acceptance is 18%.

Table 6: Systematic uncertainties on the acceptance of the first-generation scalar leptoquark signal.

<table>
<thead>
<tr>
<th>Systematic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF</td>
<td>8%</td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>10%</td>
</tr>
<tr>
<td>MC Statistics</td>
<td>3%</td>
</tr>
<tr>
<td>Luminosity</td>
<td>6%</td>
</tr>
<tr>
<td>Vertex Cut</td>
<td>0.5%</td>
</tr>
<tr>
<td>Trigger Efficiency</td>
<td>2%</td>
</tr>
<tr>
<td>Energy Scale</td>
<td>10%</td>
</tr>
<tr>
<td>Total</td>
<td>18%</td>
</tr>
</tbody>
</table>

9 Results

9.1 Control Regions

The 7 control regions with large MET defined in Tab. 5 were examined before “opening the box” that contains the leptoquark signal. The total background expected from the Monte Carlo simulation, with the statistical and systematic uncertainties listed separately, and the number of events observed in data are listed in Tab. 7 as well as shown in Fig. 43. The cross section for the vector boson production plus 2-jets has been scaled up from the LO cross section of ALPGEN to the NLO cross section of MCFM. As stated earlier, the QCD multi-jet production has been normalized to data in a looser control region. The difference between the data yield and Monte Carlo expectation for region 6 is $-18 \pm 11$, which is 1.5 standard deviations low. Likewise, in region 4, the Poisson probability for a statistical fluctuation from the 1.3 events expected to the 4 observed (or larger) is 4.4%, or about two standard
deviations high; but when convoluted with the systematic uncertainty, the discrepancy is closer to one standard deviation. All other regions show very good agreement.

Table 7: Comparison of the number of expected events from Monte Carlo with the number observed in data for the control regions defined in Tab. 5. The Monte Carlo statistical error is the first uncertainty shown, and the systematic error from other sources is the second.

<table>
<thead>
<tr>
<th>Region</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>QC1</td>
<td>0</td>
<td>14.62</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.69</td>
<td>0</td>
</tr>
<tr>
<td>W → eν + 2 jets</td>
<td>0.30</td>
<td>4.96</td>
<td>0.20</td>
<td>0.30</td>
<td>0.20</td>
<td>22.56</td>
<td>5.46</td>
</tr>
<tr>
<td>W → μν + 2 jets</td>
<td>1.44</td>
<td>8.04</td>
<td>0.62</td>
<td>0.21</td>
<td>7.01</td>
<td>1.24</td>
<td></td>
</tr>
<tr>
<td>W → τν + 2 jets</td>
<td>0.01</td>
<td>0.28</td>
<td>0.01</td>
<td>0.03</td>
<td>3.23</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Z → μμ + 2 jets</td>
<td>0.03</td>
<td>0.36</td>
<td>0</td>
<td>0.03</td>
<td>0.63</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Z → ττ + 2 jets</td>
<td>0.17</td>
<td>5.40</td>
<td>0.25</td>
<td>0</td>
<td>0.17</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>tt</td>
<td>0.03</td>
<td>0.20</td>
<td>0.10</td>
<td>0.08</td>
<td>0.39</td>
<td>4.48</td>
<td>0.83</td>
</tr>
<tr>
<td>Total Events</td>
<td>1.96</td>
<td>35.12</td>
<td>0.88</td>
<td>1.27</td>
<td>2.96</td>
<td>59.27</td>
<td>17.70</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>±0.59±0.37</td>
<td>±10.4±8.33</td>
<td>±0.32±0.15</td>
<td>±0.43±0.22</td>
<td>±0.57±0.49</td>
<td>±3.70±9.87</td>
<td>±1.45±3.21</td>
</tr>
</tbody>
</table>

9.2 Signal Region

Table 8 shows the acceptance of first-generation scalar leptoquarks (with β = 0) as a function of mass after the final selection cuts. Also shown is the number of events expected in 76 pb⁻¹ using the NLO cross section of T. Plehn [12]. The total systematic uncertainty on the leptoquark acceptance is 18%, as described in Sec. 8.2.

The total background to the leptoquark signal region is predicted to be 42.5 ± 10.7, where the breakdown according to each Standard Model source is listed in Tab. 9. Assuming that there is no leptoquark signal, and that the number of observed events in data would equal the expected background, we can derive an expected cross section upper limit using the acceptances in Tab. 8 and the systematic uncertainties just noted for the signal and background. This is shown in Fig. 44, where we have used a Bayesian likelihood method [25] to determine the upper limit on the possible number of signal events at the 95% confidence level. If we conservatively choose the theoretical cross section calculated with μ = 2m₇Q, we expect to exclude the mass interval between about 65 and 110 GeV/c². Also shown in the plot are the expected limits if we were to observe a ±1σ fluctuation in the data, where σ includes both the statistical and systematic errors added in quadrature. For a positive fluctuation of one standard deviation or larger, we may not be able to exclude any mass interval.

The actual number of events observed in data for the leptoquark signal region is 42, in excellent agreement with the number predicted. The following histograms show the distribu-
Background Predictions and Data Around The Signal Region

CDF Run II Preliminary

\[ \int L \, dt = 76 \text{ pb}^{-1} \]

- Data
- SM EWK prediction
- + \( t\bar{t} \) and QCD
- Uncertainty

Control Region

Figure 43: Comparison of the number of events observed in data (solid points) with the number expected from Standard Model sources (histograms) for the control regions defined in Tab. 5. The electroweak contribution is shown separately from the contribution due to QCD multi-jet production and top pair production. The total uncertainty on the predicted background is shown by the hatched region.
Table 8: Acceptance of the first-generation scalar leptoquark signal as a function of mass, as well as the number of events expected in 76 pb$^{-1}$.

<table>
<thead>
<tr>
<th>LQ Mass (GeV/c$^2$)</th>
<th>Acceptance</th>
<th>Number Expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>0.0020</td>
<td>33.8</td>
</tr>
<tr>
<td>75</td>
<td>0.0085</td>
<td>45.0</td>
</tr>
<tr>
<td>90</td>
<td>0.0231</td>
<td>45.7</td>
</tr>
<tr>
<td>100</td>
<td>0.0334</td>
<td>37.0</td>
</tr>
<tr>
<td>110</td>
<td>0.0464</td>
<td>30.4</td>
</tr>
<tr>
<td>125</td>
<td>0.0662</td>
<td>21.0</td>
</tr>
<tr>
<td>150</td>
<td>0.0909</td>
<td>10.0</td>
</tr>
<tr>
<td>175</td>
<td>0.110</td>
<td>4.74</td>
</tr>
<tr>
<td>200</td>
<td>0.124</td>
<td>2.32</td>
</tr>
</tbody>
</table>

Table 9: The number of expected events from various Standard Model sources in the leptoquark signal region. The Monte Carlo statistical error is the first uncertainty shown, and the systematic error from other sources is the second.

<table>
<thead>
<tr>
<th>Source</th>
<th>Events expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>QCD</td>
<td>7.3</td>
</tr>
<tr>
<td>$W \rightarrow e\nu + 2$ jets</td>
<td>1.7</td>
</tr>
<tr>
<td>$W \rightarrow \mu\nu + 2$ jets</td>
<td>8.3</td>
</tr>
<tr>
<td>$W \rightarrow \tau\nu + 2$ jets</td>
<td>10.3</td>
</tr>
<tr>
<td>$Z \rightarrow \mu\mu + 2$ jets</td>
<td>0.5</td>
</tr>
<tr>
<td>$Z \rightarrow \tau\tau + 2$ jets</td>
<td>0.2</td>
</tr>
<tr>
<td>$Z \rightarrow \nu\nu + 2$ jets</td>
<td>13.4</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>0.7</td>
</tr>
<tr>
<td>Total Events</td>
<td>$42.5 \pm 7.6 \pm 7.5$</td>
</tr>
</tbody>
</table>
Figure 44: The expected cross section upper limit for scalar leptoquark production ($\beta = 0$) assuming that no leptoquark events are observed and that the data will yield exactly the Monte Carlo expectation for the Standard Model background. Also shown are the limits that would be obtained if we observe a $\pm 1\sigma$ fluctuation in the expected background.
tions of several principal kinematic quantities for data (solid points) compared to the Monte Carlo expectations from $W/Z + 2$ jets and top quark pair production (yellow histograms). The contribution from QCD multi-jet production is not shown because of the poor Monte Carlo statistics. Also shown is the expected distribution arising from leptoquark production and decay at a mass of 100 GeV/$c^2$ (red histograms). The distributions shown include MET (Fig. 45), jet multiplicity (Fig. 46), leading-jet $E_T$ (Fig. 47), second-leading jet $E_T$ (Fig. 48), third-leading jet $E_T$ (Fig. 49), the $\phi$ difference between the two highest $E_T$ jets (Fig. 50), the $\phi$ difference between the leading-jet and the MET direction (Fig. 51), and the minimum $\phi$ difference between any jet and the MET direction (Fig. 52).

![Figure 45: The MET distribution in the leptoquark signal region for data (solid points) compared to Standard Model background (yellow histogram). Also shown is the distribution expected for leptoquark production and decay at a mass of 100 GeV/$c^2$ (red histogram).](image)

10 Search Limits

The number of events observed in data for the leptoquark signal region is 42, in good agreement with the $42 \pm 11$ expected from Standard Model backgrounds. Therefore, no evidence for leptoquark production is observed, and an upper limit can be set on the cross section times squared branching ratio. This is shown in Fig. 53, where we have used a Bayesian likelihood method [25] to determine the upper limit on the possible number of signal events at the 95% confidence level using an 18% systematic uncertainty on the signal acceptance and a 25% systematic uncertainty on the background yield. If we conservatively choose the theoretical cross section calculated with $\mu = 2m_{LQ}$, we exclude the mass interval between 60 and 107 GeV/$c^2$ for 100% branching ratio into neutrino plus quark.
Figure 46: The jet multiplicity distribution in the leptoquark signal region for data (solid points) compared to Standard Model background (yellow histogram). Also shown is the distribution expected for leptoquark production and decay at a mass of 100 GeV/c^2 (red histogram).

Figure 47: The leading-jet $E_T$ distribution in the leptoquark signal region for data (solid points) compared to Standard Model background (yellow histogram). Also shown is the distribution expected for leptoquark production and decay at a mass of 100 GeV/c^2 (red histogram).
Figure 48: The second-leading jet $E_T$ distribution in the leptoquark signal region for data (solid points) compared to Standard Model background (yellow histogram). Also shown is the distribution expected for leptoquark production and decay at a mass of 100 GeV/$c^2$ (red histogram).

Figure 49: The third-leading jet $E_T$ distribution in the leptoquark signal region for data (solid points) compared to Standard Model background (yellow histogram). Also shown is the distribution expected for leptoquark production and decay at a mass of 100 GeV/$c^2$ (red histogram).
Figure 50: The $\phi$ difference between the two highest $E_T$ jets in the leptoquark signal region for data (solid points) compared to Standard Model background (yellow histogram). Also shown is the distribution expected for leptoquark production and decay at a mass of 100 GeV/c$^2$ (red histogram).

Figure 51: The $\phi$ difference between the leading-jet and the MET direction in the leptoquark signal region for data (solid points) compared to Standard Model background (yellow histogram). Also shown is the distribution expected for leptoquark production and decay at a mass of 100 GeV/c$^2$ (red histogram).
Figure 52: The minimum $\phi$ difference between any jet and the MET direction in the leptoquark signal region for data (solid points) compared to Standard Model background (yellow histogram). Also shown is the distribution expected for leptoquark production and decay at a mass of 100 GeV/$c^2$ (red histogram).

11 Conclusion

We have performed a search for first-generation leptoquarks in the missing-$E_T$ plus jets channel in the first 76 pb$^{-1}$ of Run II data. We saw no evidence for leptoquarks in the data; therefore, we set an upper limit on the production cross section at 95% CL. We exclude the corresponding mass interval from 60 to 107 GeV/$c^2$ for leptoquark decays into neutrino and quark with 100% branching ratio.

12 Acknowledgments

We would like to thank Pasha Murat and the authors of the Stntuple package for providing us with a powerful analysis tool and for many useful discussions; and Beate Heinemann for her help to sort out all the issues with calorimeter software. We also would like to thank Stephan Lammel for many suggestions during the course of this analysis.

References


Figure 53: The upper limit on the cross section times squared branching ratio for scalar leptoquark production ($\beta = 0$). Also shown is the NLO cross section for $\beta = 0$ for 3 choices of the factorization/renormalization scale: $\mu = m_{LQ}$, $\mu = 2m_{LQ}$, and $\mu = 0.5m_{LQ}$.


J.-F. Arguin, B. Heinemann, A. Yagil, ”The z-Vertex Algorithm in Run II”, CDF Note 6238


[20] E. James et al. “Measurement of $\sigma \cdot B(W \rightarrow \mu \nu)$ and $\sigma \cdot B(Z \rightarrow \mu^+ \mu^-)$ in the Run II High $P_T$ Muon Data Sample”, CDF Note 5925.


