CDF/ANAL/JET/PUB/9609
November 19, 2008
Version 1.0

Search for Quark Substructure in the angular distribution of dijets produced in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV.

The CDF Collaboration
URL http://www-cdf.fnal.gov

November 19, 2008

Abstract

CDF dijet data have been analyzed to look for evidence for quark substructure in the angular distributions. The data consist of 11,903,325 events obtained with the single jet $E_T > 100$ GeV trigger, and an integrated luminosity of $1109 pb^{-1}$. No evidence for quark substructure was found, and a limit on the coupling parameter $\Lambda > 2.4$TeV (95% confidence) was obtained. Systematic uncertainties dominate the limit.

1 Introduction

This note describes a comparison between the CDF jet $E_T > 100$ GeV trigger data, with only minimal cuts and corrections, to the expectations of the PYTHIA Monte Carlo simulation [1] together with the CDF detector simulation (CDFSim). PYTHIA allows the incorporation of some models of new physics beyond the Standard Model, including quark substructure [2]. The sensitivity of the data to the presence of quark substructure was determined by comparing PYTHIA plus the CDFSim with the substructure switched on to PYTHIA plus CDFSim without quark substructure ('default PYTHIA'). Limits were then placed on the presence of substructure from the ratio of the data distributions to those of the default PYTHIA.Similar analyses from CDF Run 1 were reported in reference [3], and from D0 Run 1 in reference [4]. D0 has recently posted a Run 2 analysis [5].

PYTHIA bases hard $\bar{p} + p$ scattering on the parton model, in which the cross section $\sigma(ij \to k)$ is written in terms of a parton sub cross section $\hat{\sigma}_{ij\to k}$, and parton distribution functions (PDF's) $f_i(x)$ as

$$\sigma_{ij\to k} = \int dx_1 \int dx_2 f_i^1(x_1) f_j^2(x_2) \hat{\sigma}_{ij\to k}. \tag{1}$$

The PDF's $f^1(x)$ and $f^2(x)$ are supplied from fits to experimental data. The default set used is CTEQ5L [6].

Standard QCD matrix elements for $2 \to 2$ hard scattering at the parton level were evaluated by Combridge [7]. These include all permutations of $\bar{q}$ q, q q, q g, and g g reactions, and are contained in PYTHIA code. The resulting jet-jet angular distribution is remarkably similar to Rutherford's formula.

Quark substructure terms, which can be switched on for inclusion in the calculation in PYTHIA , are derived from Eichten,*et al* [2], with appropriate interference between QCD and quark substructure. The quark substructure effective contact color singlet Lagrangian is:

$$L = \pm(g^2/2\Lambda^2(\bar{q}_L\gamma_\mu q_L)(\bar{q}_L\gamma^\mu q_L)), \tag{2}$$

where Λ is a parameter with dimensions of energy which characterizes the strength of the term. The coupling constant $g^2/4\pi = 1$. This term affects only the u and d quarks, and $1/\Lambda$ is a measure of quark size. Since this Lagrangian is basically isotropic, its presence enhances the QCD cross section near 90° in the diquark (dijet) center of mass. The amplitude is proportional to $\hat{s}/\Lambda^2$,which increases with increasing jet-jet invariant mass. This experiment is sensitive only to the square of the amplitude, or $(\hat{s}/\Lambda^2)^2$.

2 Apparatus

The CDF II detector is described in detail elsewhere [8]. A polar coordinate system (r, θ, ϕ) is used, centered at the detector origin, with the z axis pointing towards the proton beam direction. Pseudorapidity $\eta = -\log(\tan(\theta/2))$. The most important components for the angular distribution measurement are the electromagnetic and hadronic calorimeters. The central calorimeters were part of the original CDF detector, and cover $0 < |\eta| < 1.3$. The electromagnetic calorimeters (CEM) are constructed of sheets of lead, and the central hadronic calorimeters (CHA) of iron plates, interspersed with scintillators,and coupled to wavelength shifters. The scintillation light is detected by photomultiplier tubes. The scintillating tile plug calorimeters, covering $1.3 < |\eta| < 3.0$ are new to CDF II. For most of the calorimeters the segmentation is $\Delta\eta = 0.1$ and $\Delta\phi = 15°$, but $\Delta\phi = 7.5°$ in the plug. Energy resolutions are $\sigma(\mathrm{E_T})/\mathrm{E_T} = 13.5\%/\sqrt{\mathrm{E_T}}$ (GeV) for the CEM, and $50\%/\sqrt{E_T}$ for the CHA.

The three level jet100 trigger was formed from Level1(single tower $\mathrm{E_T} >$ 10 GeV) $\rightarrow$ Level2(cluster $\mathrm{E_T} >$ 90 GeV) $\rightarrow$ Level3(jet $\mathrm{E_T} >$ 100 GeV). The transverse energies were calculated from calorimeter signals assuming the event origin coincided with the origin of the CDF coordinates. The cross section for the jet100 trigger was $\sigma(\bar{\mathrm{p}} + \mathrm{p} \rightarrow$ single jet $\mathrm{E_T} >$ 100 GeV)=10 nb. Triggers were accepted throughout the $|\eta|$ range of the CDF detector. The jets were reconstructed using a cone algorithm of radius 0.7 in (η, ϕ) space.

3 The Data Sample

The analysis code first applied energy corrections to the jet $\mathrm{E_T}$'s [9]. The first correction smoothed out the detector response as a function of $|\eta|$. A correction for extra energy in the jets from multiple $\bar{\mathrm{p}}\,\mathrm{p}$ interactions per bunch crossing was then applied. Next an overall scale factor of about 1.16 was applied to the calorimeter energies to correct back to the hadrons generated by fragmentation of the hard scattered parton. Two other corrections were a subtraction of the measured energy from the underlying event, spectator remnants of the hard scattering, and 'out-of-cone' corrections obtained from PYTHIA Monte Carlo. These last two effects were about 2% for 100 GeV $\mathrm{E_T}$ jets.

The cuts applied to the data after jet energy corrections were $\not{E}_T^{\rm sig} < 5\sqrt{\rm GeV}$, and $|{\rm Zvertex}| < 60$ cm. The parameter $\not{E}_T^{\rm sig} = \not{E}_T/\sqrt{\rm E_T}$, and is a measure of the statistical significance of the missing $\rm E_T$. The cut value was chosen to eliminate events where the $\not{E}_T$ was larger than expectations from jet energy measurement fluctuations [10]. After the cuts the jet-jet invariant mass was calculated from the formula

$$\rm M^2 = 2 \times E_{T1} \times E_{T2}(\cosh(\eta_1 - \eta_2) - \cos(\phi_1 - \phi_2)), \qquad (3)$$

where indices (1,2) refer to the highest $\rm E_T$ jet and the next highest $\rm E_T$ jet in the event respectively. No cut was made on the $\rm E_T$ of a third jet. About 20% of the events had a third jet with $\rm E_T > 15GeV$.

Another variable used in the analysis was $\chi = \cot^2(\theta^*/2)$, where θ^* is the polar angle in the jet-jet center of mass. χ was calculated from the formula

$$\chi = \exp(|\eta_1 - \eta_2|). \qquad (4)$$

χ removes the coulomb pole in Rutherford scattering. Thus if $\rm dN/d\cos(\theta) = 1/\sin^4(\theta/2)$, then $\rm dN/d\chi$ = flat. The distribution is folded back at $\chi = 1$, or symmetric about $\theta^* = 90°$.

The angular distributions were collected in four bins in the dijet invariant mass, as follows: 550-650, 650-750, 750-850, and 850-950 GeV. $\rm E_T$, χ, and the dijet mass, M, are related by the formula:

$$\rm E_T = (M\sqrt{\chi})/(1 + \chi). \qquad (5)$$

4 PYTHIA and CDF Simulation

PYTHIA allows the user to select various PDF's in Eq (1). Since the PDF's are fits to experimental data with attendant uncertainties, there are differences among the available PDF's which can lead to systematic uncertainty. The default PDF's in this analysis were CTEQ5L. The PDF's are functions of two variables, x and $\rm Q^2$. For many subprocesses the definition of $\rm Q^2$, on which the PDF's have a logarithmic dependence, is also ambiguous, leading to another systematic uncertainty. An independent scale which appears in the formula for the running strong coupling constant $\alpha_{\rm s}(\mu^2)$ is chosen such that $\rm Q^2 = \mu^2$.

PYTHIA models the real events, which involve jets of particles rather than partons, by adding initial and final state processes to the basic $2 \to 2$ hard collision. Initial state gluon radiation adds extra partons to the final state, and affects the kinematics of the hard collision. In the final state the partons shower to become physical hadrons, resulting in the PYTHIA version of the event. The simulation of the response of the CDF detector to this event then follows, to give Monte Carlo simulated events which can be compared to the real data. The calorimeter response is modeled using a shower average program called GFLASH, written for the DESY H1 detector [11].

5 Simulation Including Quark Substructure

A total of 8 million Monte Carlo simulated events were generated and run through CDFSim. Half of the events had the default $\mathrm{Q}^2 = \mathrm{p}_\mathrm{T}^2$, and half had $\mathrm{Q}^2 = \hat{\mathrm{s}}$. Three million events had no quark substructure. Of the remaining 5 million events, 4 million were evenly distributed among four values of the substructure parameter Λ = -1.0, -1.4, -2.0, and -3.0 TeV in order to determine the sensitivity of the analysis to this extra term. The minus sign for the interference term was used in each case. The remaining 1 million events used the MRST PDF's [12]. Figures 1, 2, 3, and 4 show χ plots, defined in Eq (4), of the PYTHIA Monte Carlo simulation with CTEQ5L plus CDFSim for the two choices for the ambiguous variable Q^2. The sample sizes for the data and the summed Monte Carlo were approximately the same. The data are multiplied by 0.6 on these plots. It is apparent that the two choices of Q^2 give different angular distributions. Even though the dependence is logarithmic, the fact that $\hat{s}$ is constant in a given mass bin, while p_T^2 changes with angle makes an observable difference.

The combinations of $\mathrm{Q}^2 = \hat{\mathrm{s}}$ and $\mathrm{Q}^2 = \mathrm{p}_\mathrm{T}^2$ which minimized χ^2 were determined in each mass bin by fitting the Monte Carlo distributions to the data. The results are given in Table 1, and the plots are shown in Figures 5, 6, 7, and 8. The fits favor $\mathrm{Q}^2 = \mathrm{p}_\mathrm{T}^2$ by about three to one, except for the highest mass bin, where the statistics are limited.

Figures 9, 10, 11, and 12 show the effect of switching on the quark substructure.The angular distributions peak in the range $1 < \chi < 10$ for finite values of Λ, and the peak fraction increases with increasing jet-jet mass.

A ratio method was used to minimize the dependence of the effect on fluctuations in the cross section in a given jet-jet mass bin. The angular distribution was normalized in each bin by taking the ratio R = (1 < χ <

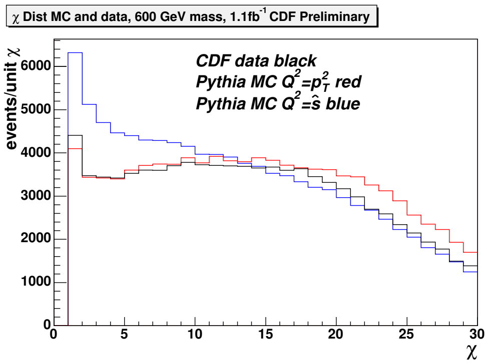

Figure 1: Pythia $Q^2 = p_T^2$ and $Q^2 = \hat{s}$ compared to the data in 600 GeV mass bin

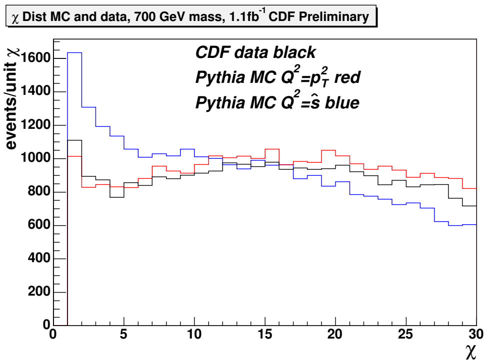

Figure 2: Pythia $Q^2 = p_T^2$ and $Q^2 = \hat{s}$ compared to the data in 700 GeV mass bin

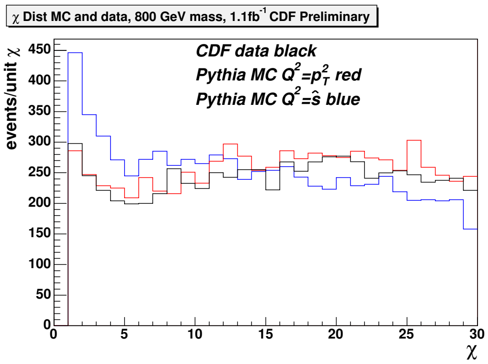

Figure 3: Pythia $Q^2 = p_T^2$ and $Q^2 = \hat{s}$ compared to the data in 800 GeV mass bin

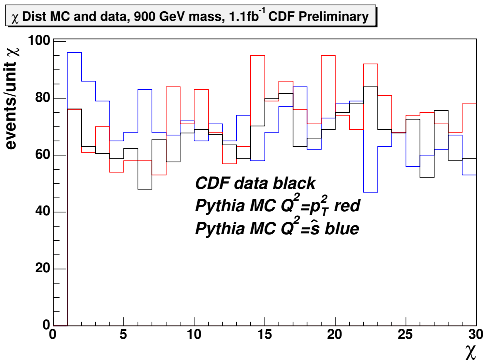

Figure 4: Pythia $Q^2 = p_T^2$ and $Q^2 = \hat{s}$ compared to the data in 900 GeV mass bin

Table 1: Fits of the Monte Carlo simulations to the data in 28 1 unit χ bins: $\mathrm{data(k)} = a_1 \times p_T^2(k) + a_2 \times \hat{s}(k)$

mass bin	events (data)	a_1	a_2	χ^2
550-650 GeV	165926	1.13 ± 0.03	0.43 ± 0.03	104.
650-750 GeV	54404	1.31 ± 0.056	0.28 ± 0.054	15.
750-850 GeV	16112	1.31 ± 0.11	0.25 ± 0.11	23
850-950 GeV	4581	1.02 ± 0.18	0.53 ± 0.18	19.

Table 2: Least Squares Fits to the χ ratios

Λ (TeV)	slope	error	χ^2	int
1.0	5.0	0.03	7.4	.88
1.4	1.86	0.08	1.0	.9
2.0	0.59	0.06	3.2	.96
data	-0.16	0.08	.05	1.0

$10)/(15 < \chi < 25)$. This ratio was calculated for each Monte Carlo sample. Then with no quark substructure as the denominator, $R(\Lambda)/R(\infty)$ was plotted *versus* $(\mathrm{mass})^4$, shown in Figure 13.The straight line fits are shown in Table 2. Figure 14 shows these slopes from Figure 13 plotted *versus* $(1/\Lambda)^4$, and fitted to a quadratic function. The formula for this curve is:

$$\mathrm{s} = 8.24 \times \mathrm{x} - 3.25 \times \mathrm{x}^2 \tag{6}$$

and

$$x = 1.265 \pm (1.60 - 0.308 \times s)^{1/2}. \tag{7}$$

Here s is the slope in units of TeV^{-4}, and x represents Λ^{-4}.This is the formula for converting a measured slope into a value of Λ.

The data are treated in the same way. The ratio R normalized to the no quark substructure Monte Carlo is plotted as a function of $(\mathrm{mass})^4$ in Figure 15. Conclusions from this plot await a discussion of the systematics.

6 Systematics

Systematic effects affecting the predictions of the Monte Carlo Simulation are:

- Sensitivity to the PDF's. What PDF's contribute to the quark substructure effect?
- Sensitivity to the Q^2 used in the PDF call. Here the mixture of $\mathrm{Q}^2 = \mathrm{p}_\mathrm{T}^2$ and $\mathrm{Q}^2 = \hat{\mathrm{s}}$ was varied by 1 σ from the best fit..

And those affecting the treatment of the data:

- Sensitivity to the jet energy corrections [9].

6.1 The PDF's

The effect of quark substructure for high E_T jets in $p\bar{p}$ collisions depends on proton valence $\times$ valence quarks in the range $.1 < x < 1$. Proton valence quark PDF's are reasonably well known.

To test the dependence of the quark substructure limit on the choice of PDF's, the 10^6 PYTHIA events generated with MRST were compared to the default CTEQ5L sample.. The angular distributions were identical within statistics. We conclude that the PDF's do not present an important systematic uncertainty.

6.2 Choice of Q^2

To test the sensitivity of the result to the choice of Q^2, an input to the PDF call which is ambiguous, the p_T^2 fraction was varied up and down by 1 σ. The two parameters a_1 and a_2 in Table 2 were strongly negatively correlated, to hold the total sum constant. Thus when a_1 was increased, a_2 was decreased, and vice versa. The χ^2 comparisons to the data in Table 2 each increased by 1 unit as a result of the variation. Each of these three Monte Carlo's can be used to calculate the ratio $R = (1 \leq \chi \leq 10)/(15 \leq \chi \leq 25)$ in each mass bin to compare with experimental data and place a limit on quark substructure. Since the Monte Carlo's are the same, the statistical uncertainties in the three ratios are not independent.

6.3 The Jet Energy Scale

There is a systematic uncertainty assigned to each stage of the jet energy corrections [9]. The entire data sample was reanalyzed with the jet energy corrections changed by $\pm 1\,\sigma$, about a 3% effect at the higher jet E_T.

6.4 Treatment of the Systematic Uncertainties

To include systematic effects in the calculation of the limit on Λ, the quark substructure parameter, the Monte Carlo ratios in each mass bin were averaged, and the uncertainty was calculated from the statistical uncertainties plus the deviations from the average value. The formulas, which assume that the uncertainties are Gaussian distributed, are as follows:

$$< R_j >= \Sigma_{i=1,3}\, R_{i,j}/3, \qquad (8)$$

Table 3: Final ratios as a function of jet-jet mass

mass bin	R_{data}	R_{MC}	R_{data}/R_{MC}
600 GeV	0.974 ± 0.073	0.977 ± 0.010	0.996 ± 0.075
700 GeV	0.815 ± 0.028	0.824 ± 0.017	0.988 ± 0.039
800 GeV	0.765 ± 0.029	0.804 ± 0.032	0.952 ± 0.052
900 GeV	0.721 ± 0.045	0.779 ± 0.049	0.926 ± 0.082

where $< R_j >$ is the average ratio for the jth mass bin, and the sum i runs over the three different data or Monte Carlo samples. Then for the uncertainty

$$\Delta < R_j >= (< R_j >^2 \times \Sigma_{i=1,3}(\Delta R_{i,j}/R_{i,j})^2/3) + \Sigma_{i=1,3}(< R_j > - R_{i,j})^2/2)^{1/2}. \tag{9}$$

The first term is the statistical uncertainty for the ratio. The second term, the systematic uncertainty, dominates the overall uncertainty. The factor of two in the denominator is (n-1) [13]. These same formulas were also applied to the jet100 data sample analyzed with three different sets of jet energy corrections. This averaging technique yielded one number in each mass bin for R_{data} and $R_{MCnoqsub}$, with uncertainties in each calculated as above. Table 3 shows the final ratios, which are plotted in Figure 15.

7 Setting the Limit

The fitted slope of the data is shown in Figure 15 is $-.16 \pm 0.08$, which, being negative, is unphysical. The method of Feldman and Cousins is used to interpret this result [14]. In this method, one generates 'pseudoexperiments' with physically allowed slope values, varying the data points within the uncertainties. A two dimensional plot of physically meaningful slopes *versus* experimental results can be used to generate the confidence level contours shown in Figure 16, giving slope $< 0.2495\%$, and $\Lambda > 2.4$ TeV (95% confidence) from Eq 7.The expected limit, obtained if the measured slope is zero, is $\Lambda > 2.2$ TeV (95% confidence).

8 Conclusions

We have performed a search in the CDF Run II jet100 data, looking for evidence of quark substructure in the dijet angular distribution. No evidence for substructure has been seen. The limit $\Lambda > 2.4$ TeV (95% confidence) is dominated by systematic uncertainties.

References

[1] T. Sjostrand *et al,*Computer Physics Communications **135,** 238 (2001).

[2] E. Eichten, K. Lane, and M. Peskin, Phys Rev Lett **50,**811 (1983).

[3] F. Abe *et al,*CDF Collaboration, Phys Rev Lett **77**, 5336 (1996). .

[4] B. Abbott *et al,* D0 Collaboration, Phys Rev D **62,**031101(R) (2000).

[5] D0 Collaboration at www-d0.fnal.gov/Run2Physics/WWW/results/prelim/QCD/Q11.

[6] L.Lai *et al* Eur Phys J. **C12,** 375 (2000).

[7] B. Combridge, J. Kripfganz, and J. Ranft, Phys Lett **70B,** 234 (1977).

[8] A. Abulencia et al.(CDF Collaboration), J.Phys. G **34,**2457 (2007).

[9] A. Bhatti, et al, Nucl Instr Meth A **566,**375 (2006).

[10] The $\not{E}_{\mathrm{T}} = -\Sigma_{\mathrm{i}}(\mathrm{E}_{\mathrm{T,i}} \times \mathbf{n_i})$, where $\mathbf{n_i}$ is a unit vector in the transverse plane pointing to the ith tower.

[11] G. Grindhammer, M. Rudowicz, and S. Peters, Nucl. Instr. Meth.**A290,** 469 (1990).

[12] Martin, Roberts, Stirling, and Thorne, Eur Phys Journal **C4,** 463(1998).

[13] S. Eidelman et al, Particle Data Group, Phys Lett B **592,**1 (2004).See Eqs 32.4 and 32.5 in Section 32.1.1, p 279.

[14] G. Feldman and R. Cousins, Phys Rev D **57** 3873 (1998).

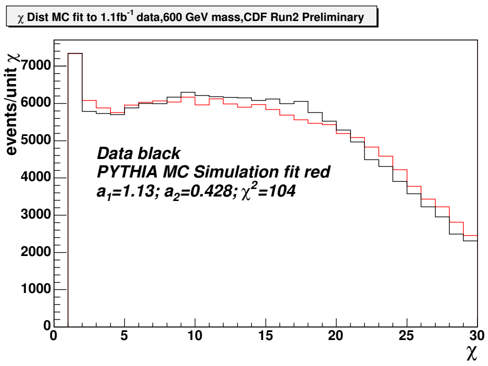

Figure 5: Pythia $Q^2 = p_T^2$ and $Q^2 = \hat{s}$ fit to the data in 600 GeV mass bin

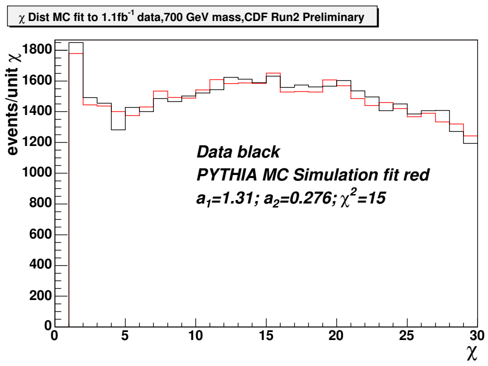

Figure 6: Pythia $Q^2 = p_T^2$ and $Q^2 = \hat{s}$ fit to the data in 700 GeV mass bin

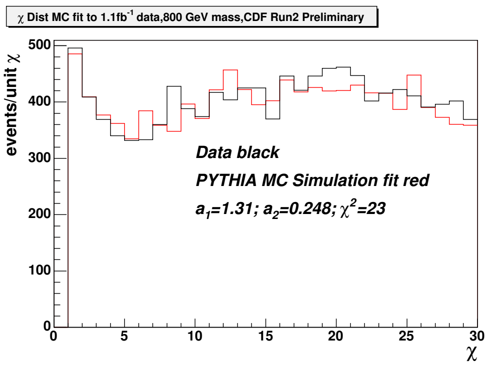

Figure 7: Pythia $Q^2 = p_T^2$ and $Q^2 = \hat{s}$ fit to the data in 800 GeV mass bin

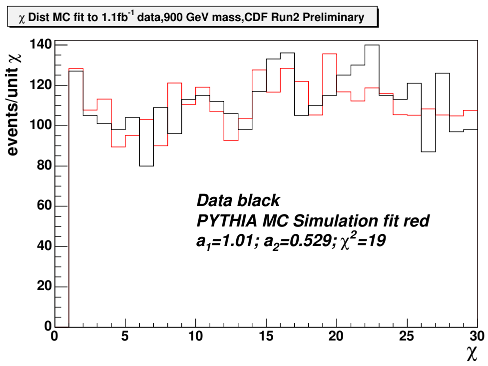

Figure 8: Pythia $Q^2 = p_T^2$ and $Q^2 = \hat{s}$ fit to the data in 900 GeV mass bin

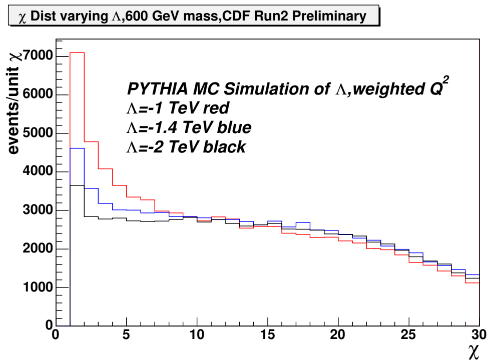

Figure 9: PYTHIA predictions for the sensitivity to the quark substructure parameter $\Lambda = 1$, 1.4, and 2 TeV in the 600 GeV mass bin

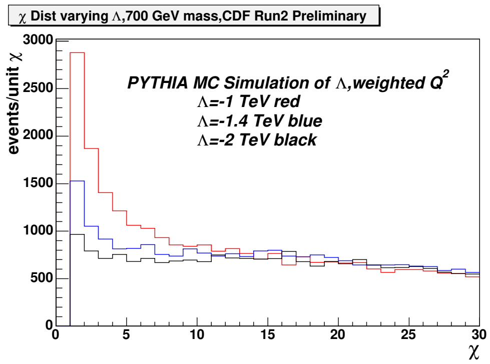

Figure 10: PYTHIA predictions for the sensitivity to the quark substructure parameter $\Lambda = 1$, 1.4, and 2 TeV in the 700 GeV mass bin

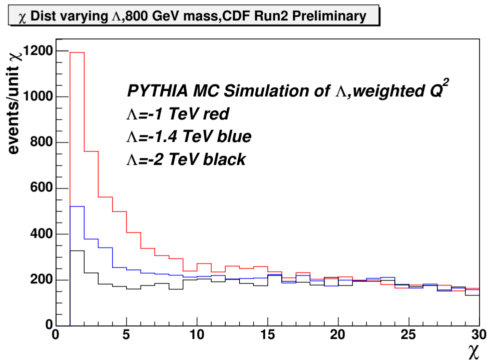

Figure 11: PYTHIA predictions for the sensitivity to the quark substructure parameter $\Lambda = 1$, 1.4, and 2 TeV in the 800 GeV mass bin

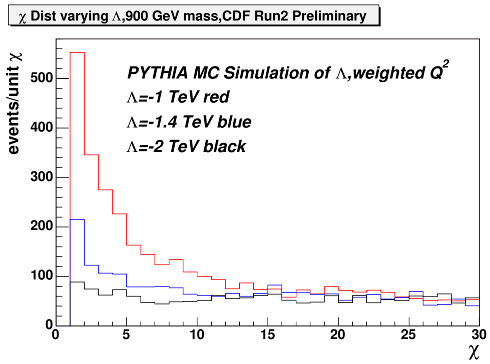

Figure 12: PYTHIA predictions for the sensitivity to the quark substructure parameter $\Lambda = 1$, 1.4, and 2 TeV in the 900 GeV mass bin

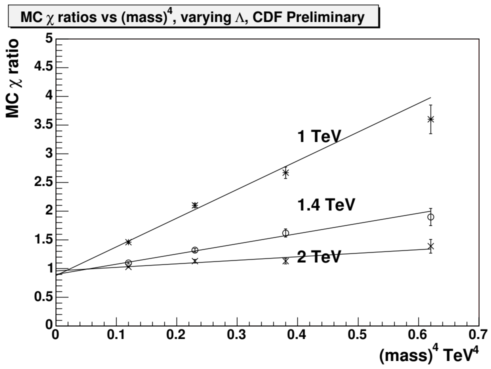

Figure 13: Plots of R $= (1 \leq \chi \leq 10)/(15 \leq \chi \leq 25)$ *versus* (mass)^4 for various values of Λ to the default Monte Carlo.

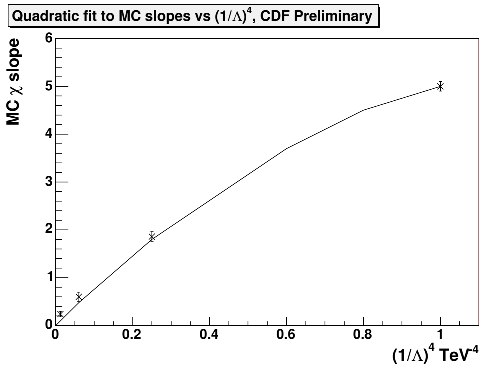

Figure 14: Quadratic fit to the slopes in the previous Figure as a function of $(1/\Lambda)^4$. This curve gives the correspondence between measured slopes and limits on the quark substructure parameter Λ.

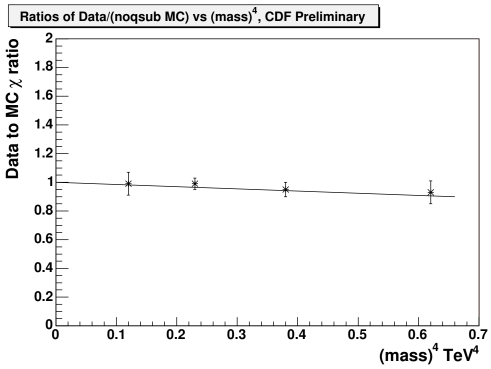

Figure 15: R $= (1 \leq \chi \leq 10)/(15 \leq \chi \leq 25)$ *versus* (mass)^4. The ratio of the averaged data to the averaged Monte Carlo is shown. The fitted slope is -0.16 ± 0.08

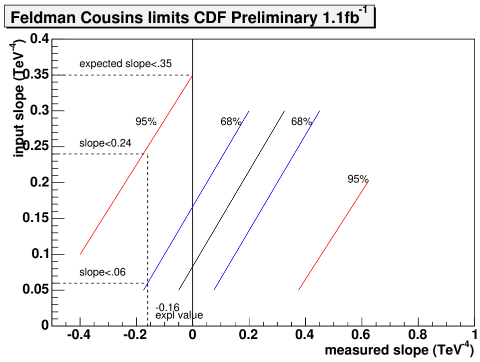

Figure 16: Confidence level limits from the Feldman-Cousins plot