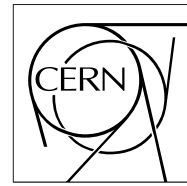**The Compact Muon Solenoid Experiment**

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

May 17, 2006

CMS Sensitivity to Quark Contact Interactions using Dijets

Selda Esen

Cukurova University, Adana, Turkey
Visitor at Fermilab, Batavia, IL, USA

Robert M. Harris

Fermilab, Batavia, IL, USA

Abstract

We estimate CMS sensitivity to quark contact interactions in the dijet final state. The canonical model of a contact interaction among left-handed composite quarks changes the dijet angular distribution at high dijet mass. The dijet ratio variable introduced at the Tevatron is used as a simple measure of the angular distribution as a function of dijet mass. The contact interaction signal and QCD background are estimated for the dijet ratio as a function of dijet mass from 0.3 to 6.5 TeV. Statistical uncertainties are estimated for integrated luminosities of $100 \ \mathrm{pb}^{-1}$, $1 \ \mathrm{fb}^{-1}$, and $10 \ \mathrm{fb}^{-1}$ and a realistic trigger table including multiple thresholds and prescales for the single jet triggers. Systematic uncertainties on the dijet ratio are estimated and are found to be small. The χ^2 between the background and the signal is estimated, including systematics, and is used to find CMS sensitivity to the contact interaction scale Λ^+. For an integrated luminosity of $100 \ \mathrm{pb}^{-1}$, $1 \ \mathrm{fb}^{-1}$, and $10 \ \mathrm{fb}^{-1}$, CMS can expect to exclude at 95% CL a Λ^+ value of 6.2, 10.4, and 14.8 TeV or discover at 5σ significance a Λ^+ value of 4.7, 7.8 and 12.0 TeV, respectively.

1 Introduction

1.1 Physics at a New Scale

New physics at a scale Λ above the mass of the final state is effectively modeled as a contact interaction. This is schematically pictured in Figure 1 for the case of quarks in the initial and final state, where the final state mass is the dijet mass. When Λ is out of direct experimental reach the propagator in the intermediate state effectively shrinks to a point, yielding a contact interaction. This is true regardless of the source of the new physics. The most common example of physics behind contact interactions is quark compositeness, where the quarks are each made of smaller particles called preons, and the quarks can interact by exchanging a preon bound state of mass roughly equal to Λ. This is a specific model of a more generic new interaction where some unknown new particle of mass roughly equal to Λ is exchanged among quarks. Contact interactions are a generic signal of new physics when the energy scale of the new physics is just out of reach. It is therefore a powerful first signal of new physics, which we should expect to observe before the new particle could be directly produced and detected. Given that the LHC energy scale is 7 times that of the Tevatron, as the luminosity of the machine gradually increases there is a significant chance of approaching and perhaps even crossing the energy scale of some new physics. We must prepare to search for contact interactions. They may indicate that a direct observation of new physics could be expected soon.

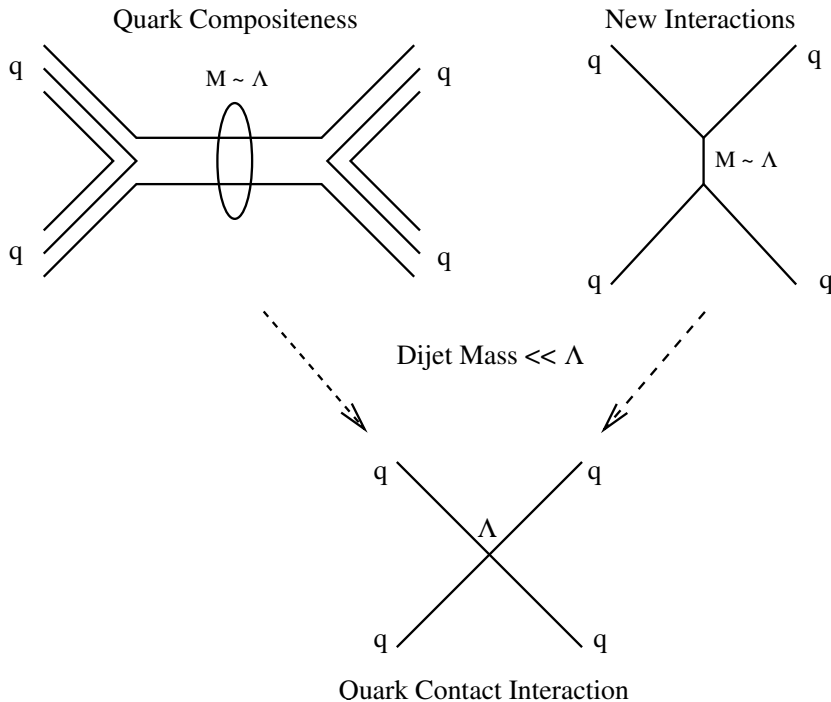In the case of quark compositeness, a simple analogy to the Rutherford experiment comes

Figure 1: Schematic picture of the origin of quark contact interactions from either quark compositeness or any model of new interactions among quarks. In either case an object of mass roughly equal to Λ is exchanged in the intermediate state, and the final state mass is significantly beneath Λ.

immediately to mind. When Rutherford bombarded a thin gold foil with α particles he expected that most of the α particles would emerge in the forward direction. He expected only a gradual change in the electromagnetic scattering as the α particle passed through what he thought were uniformly distributed charges within the atom. When many of the particles actually emerged at much larger angles to the incident beam, he discovered the nucleus, responsible for the hardest scatters of the α particles when they were in direct contact. In hadron colliders, when we scatter quarks we expect that most of the collisions will be of a glancing nature, t-channel exchange of gluons among pointlike quarks, and be peaked heavily in the forward direction. If quarks are pointlike we don't expect to find any distance scale, or any incident quark energy, where the quarks scatter more frequently and at larger angles to the beam. Quark compositeness would manifest itself as something hard within the quark, producing more hard scatters than expected, and more jets perpendicular to the beam. This admittedly simplistic analogy leads one to expect both more jets at high p_T and with a different angular distribution.

1.2 Contact Interaction Searches in Mass or p_T

Contact interactions will always produce a rise in rate relative to QCD at high dijet mass or high inclusive jet p_T. For example, in Figure 2 we show the effect on the dijet mass distribution of the quark contact interaction we will consider in this analysis [1, 2], with three different values of the scale Λ compared to lowest order QCD. The effects in jet p_T are similar. However, observation of contact interactions in the mass distribution alone requires precise understanding of the QCD rate as a function of mass. This is difficult because there are large systematic uncertainties in both the measurement and the QCD calculation of the cross section as a function of dijet mass. Jet energy measurement uncertainties are multiplied by the steeply falling QCD spectrum to give large uncertainties in the cross section. A jet energy scale uncertainty of 5% produces an upper uncertainty in the cross section varying from +30% at a dijet mass of 0.3 TeV to +80% at a dijet mass of 6.5 TeV [3]. This progressive increase of rate with mass is very similar to the effect of a contact interaction, shown in Figure 2, so a jet energy mis-measurement could produce a fake contact interaction signal in the cross section. There are also uncertainties in the parton distributions used to calculate the QCD background. Using parton distribution uncertainties from the CTEQ collaboration [14], we have estimated the uncertainty in the cross section varies from 5% at a dijet mass of 0.3 TeV to 32% at a dijet mass of 6.5 TeV. Although smaller than jet energy measurement uncertainties, this progressive increase of rate with mass is again similar to the effect of a contact interaction, and could produce a fake signal. Indeed, measurements in both the jet p_T and dijet mass distributions at the Tevatron showed an excess of rate at high jet p_T and dijet mass above QCD expectations [6, 7], and these effects were eventually explained by larger than expected parton distributions [4]. After those measurements, the dijet mass and jet p_T distributions at the Tevatron were no longer used to search for or constrain contact interactions.

1.3 Contact Interaction Searches in Angular Distribution

Contact interactions are often more isotropic than the QCD background, since QCD is dominated by t-channel scattering and produces jets predominantly in the forward direction. Isotropic means a flat distribution in $\cos\theta^*$, while the t-channel scattering angular distribution is roughly proportional to $1/(1 - \cos\theta^*)^2$, where θ^* is the angle between the final state jets and the proton beam line in the center of momentum frame. The exact angular distribution of a contact interaction depends on what model is chosen [5]. The model we will search
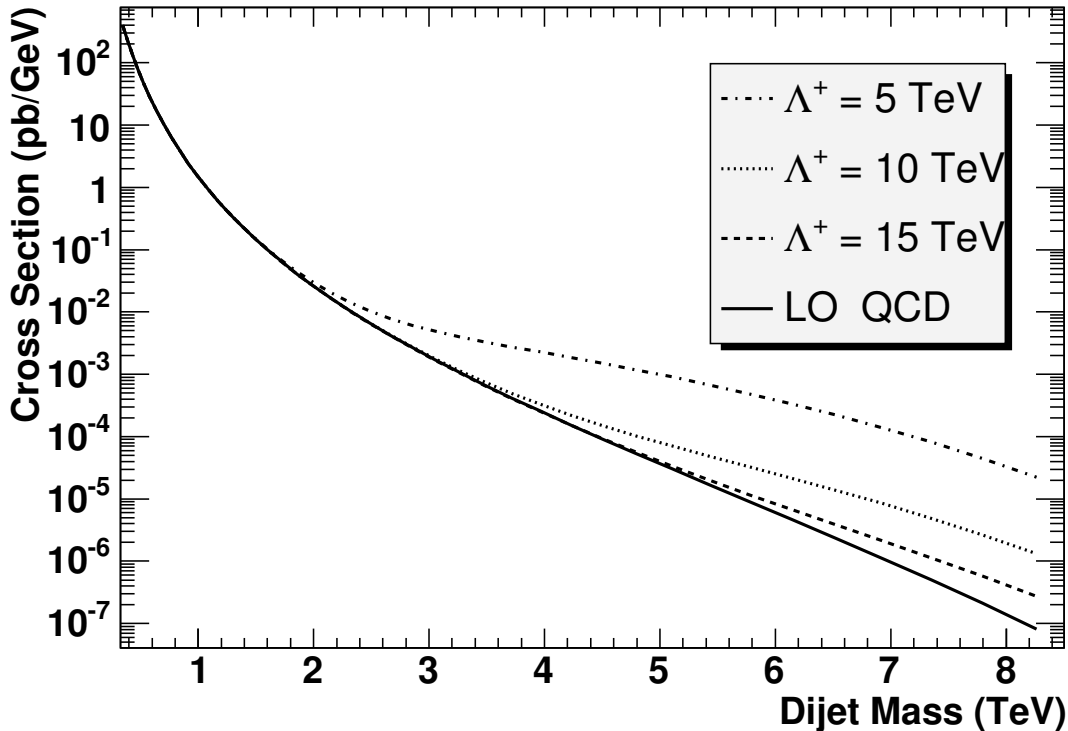
3

Figure 2: A lowest order calculation of the dijet mass distribution from QCD (solid) and from QCD plus a contact interaction with scale Λ of 5 TeV (dot-dash), 10 TeV (dotted), and 15 TeV (dashed).

for here, discussed in detail in Appendix A, has an angular distribution that is pretty close to isotropic, and produces significant changes in the jet angular distributions.

Dijet angular distributions benefit from much smaller systematic uncertainties than dijet mass or jet p_{T} distributions. This is because the angle of the jet is well measured by finely segmented calorimeters, and the relative rates at different angles only depend on the relative response of one calorimeter with respect to another, which can be determined with data. The angular variable $\cos\theta^*$ is not as well matched to the experiment as the angular variable η. Since the calorimeter boundaries are in η we can reduce the systematics, or at least make them easier to understand, by considering angular distributions in terms of η. Also, we expect the effect of compositeness to emerge at high dijet mass, so we must measure the angular distributions as a function of dijet mass. Rather than measure an entire distribution at each mass value, we will define a single number that quantifies the shape of the angular distribution at that mass value. Our analysis uses the dijet ratio, discussed in section 2, to measure the angular distribution as a function of dijet mass.

2   Dijet Ratio: $N(|\eta| < 0.5)/N(0.5 < |\eta| < 1.0)$

The ratio of the number of dijets in which both jets have $|\eta| < 0.5$ to the number of dijets in which both jets have $0.5 < |\eta| < 1.0$ was first introduced by D0 to search for contact interactions as a function of dijet mass [13] at the Tevatron. It is the simplest measure of the

4

most sensitive part of the angular distribution, providing a single number we can measure as a function of dijet mass. The dijet ratio is also convenient for us, since it uses a subset of the data we considered in our previous analysis of the jet triggers and dijet mass distribution in the region $|\eta| < 1$ [3]. Within the region $|\eta| < 1$, the data not used by the dijet ratio analysis are those dijets for which one of the two leading jets has $|\eta| < 0.5$ and the other leading jet has $0.5 < |\eta| < 1.0$. These events are used to inter-calibrate the two pseudorapidity regions, as discussed in section 3.1.2.

2.1 Dijet Ratio from Lowest Order Calculation

In Figure 3 we show our lowest order calculation of the dijet ratio from QCD compared with QCD plus a left-handed contact interaction among quarks [1, 2] at three different values of the contact interaction scale. As discussed in Appendix A, for this calculation we used the same code as a previous search [16] with modern parton distributions [14].
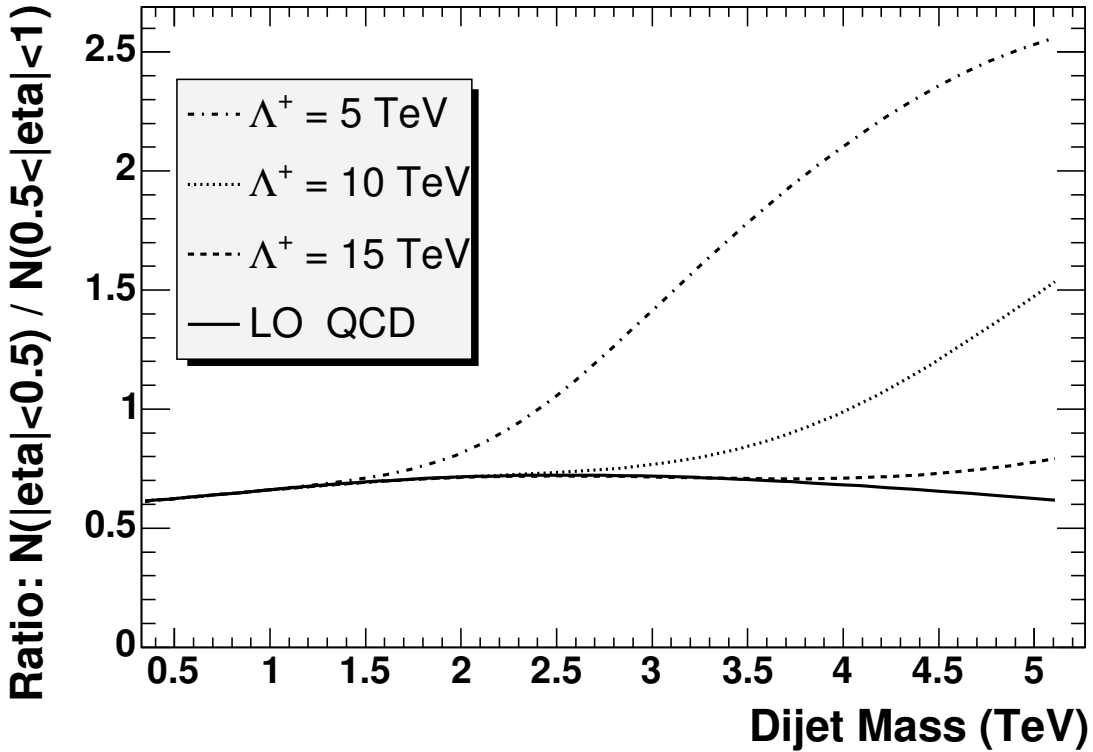

Figure 3: A lowest order calculation of the dijet ratio from QCD (solid curve) is compared with QCD plus a quark contact interaction at a scale Λ^+ of 15 TeV (dashed), 10 TeV (dotted) and 5 TeV (dot-dashed).

Figure 3 shows that lowest order QCD gives a fairly flat dijet ratio around 0.6. This indicates that there are 40% less dijets in the center of the barrel, $|\eta| < 0.5$ than there are nearby at $0.5 < |\eta| < 1$. This is because of the peaking of the dijet angular distribution in the forward direction due to t-channel scattering.

In contrast, pure contact interactions have a dijet ratio that is around 2.6, indicating there are almost three times as many dijets in the center of the barrel, $|\eta| < 0.5$, than there are nearby

at $0.5 < |\eta| < 1$. They are much more centrally produced because they originate from harder interactions than conventional QCD. For the combination of QCD plus a contact interaction, shown in Figure 3, as the dijet mass increases the fraction of the calculation that comes from the contact interaction increases as well, and the dijet ratio rises gradually at high dijet mass to a maximum value of around 2.6. The affect of this model of contact interactions on the dijet ratio is adequately estimated by this lowest order calculation.

2.2 CMS Simulation of Dijet Ratio

2.2.1 Jet Reconstruction

Jets are reconstructed as localized energy depositions in the CMS calorimeters arranged in a projective tower geometry (EcalPlusHcalTowers). The jet energy E is defined as the scalar sum of the calorimeter tower energies inside a cone of radius $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.5$, centered on the jet direction. The jet momentum $\vec{p}$ is the corresponding vector sum of energies, with the vector pointing in the tower direction. The jet transverse energy is $E_T = E\sin\theta$, and the jet transverse momentum is $p_T = p\sin\theta$, where θ is the angle between the jet momentum and the proton beam. Both the jet energy and momentum are corrected back to the particles in the jet cone originating from the hard interaction excluding pileup, as discussed in the next section. We define the dijet system as the two jets with the highest p_T in an event (leading jets) and define the dijet mass $m = \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2}$. As in the dijet mass analysis [3], in all plots that are a function of dijet mass, we plot in bins of width equal to the estimated mass resolution [17].

2.2.2 CMS Simulation

We use a sample of QCD jet events generated with Pythia, passed through the full CMS detector simulation, and reconstructed with the ORCA [8] reconstruction package. A total of 210,000 events were used, from 21 samples each consisting of 10,000 events sub-samples in contiguous intervals in generator level of p_T spanning from 0 to 4000 GeV: 0-15, 15-20, 20-30, 30-50, 50-80, 80-120, 120-170, 170-230, 230-300, 300-380, 380-470, 470-600, 600-800, 800-1000, 1000-1400, 1400-1800, 1800-2200, 2200-2600, 2600-3000, 3000-3500 and 3500-4000. EcalPlusHcalTowers were reconstructed with the default CMS algorithm which had a cut at 0.5 GeV on the energy in each HCAL compartment. Offline jets were reconstructed with the default CMS algorithm [9]: iterative cone algorithm, a cone size of $R = 0.5$, no seed threshold, 0.5 GeV E_T tower threshold, and E-scheme method of constructing jet four vectors. All reconstructed jets with $p_T > 10$ GeV were written to a root tree by the RecJetRootTree program [10], along with a single multiplicative correction factor for the offline jet Lorentz vector [3]. The correction is designed to give a Lorentz vector from the particles in a jet cone of radius $R = 0.5$ before pileup. The correction depends on reconstructed jet p_T and η. For jets in the region $|\eta| < 1$ on average a reconstructed p_T of 75 GeV was corrected by 33% to give 100 GeV, a p_T of 430 GeV by 16% to give 500 GeV, and a p_T of 2.8 TeV by 7% to give 3.0 TeV corrected jet p_T. The sample consists of a mixture of QCD events and minimum bias events corresponding to the anticipated number of multiple interactions for a luminosity of 2×10^{33} cm^{-2} s^{-1}. Each sub-sample has a weight corresponding to the generated cross section per event for that sub-sample. The weights vary significantly from sample to sample, ranging from 5.5×10^6 pb for the 0-15 sample to 9.7×10^{-9} pb for the 3500-4000 sample. When making the dijet ratio histogram all events from each sub-sample are used along with their corresponding weight, and all errors are calculated taking into account the weights.

In Figure 4 we show this full CMS detector simulation of the dijet ratio compared to the value

0.6. The simulated dijet ratio is indistinguishable from a flat ratio of 0.6 within the simulation statistical uncertainty.
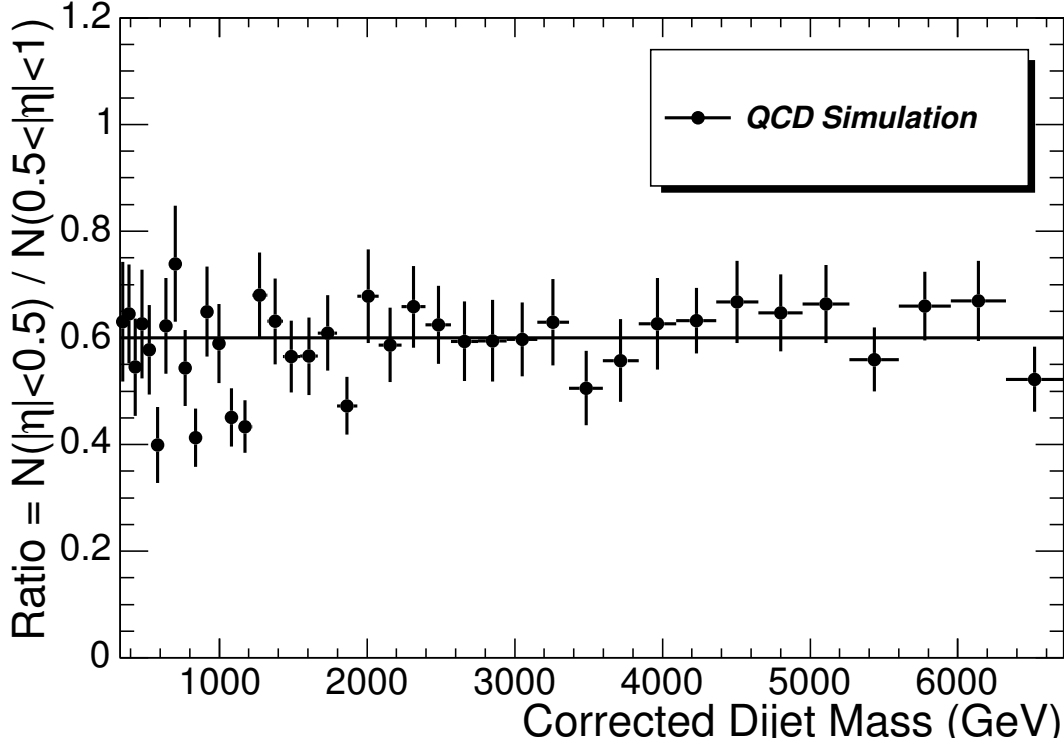

Figure 4: The dijet ratio in the CMS simulation from QCD (points) is compared to the value 0.6 (line). The error bars are the statistical uncertainty on the simulation.

2.3 Dijet Ratio for Sensitivity Estimates

In Figure 5 we show a smooth dijet ratio for QCD, estimated at 0.6 from the fit to the full simulation. The error bars shown in Figure 5 are the statistical uncertainties expected for three different integrated luminosity: $100 \, \mathrm{pb}^{-1}$, $1 \, \mathrm{fb}^{-1}$ and $10 \, \mathrm{fb}^{-1}$. Each integrated luminosity corresponds to a particular luminosity scenario and jet trigger table proposed in reference [3]. In Table 1 we reproduce the trigger table for our three integrated luminosities. Note that the triggers are prescaled at low jet E_T. The trigger prescales affect the analysis statistics, as can be seen by the bunching of the statistical error bars in Figure 5, where the statistical errors bars increase monotonically with mass within each bunch. For integrated luminosities of $100 \, \mathrm{pb}^{-1}$, $1 \, \mathrm{fb}^{-1}$ and $10 \, \mathrm{fb}^{-1}$ Table 1 shows that we employ 2, 3, and 4 triggers for the measurement, and Figure 5 shows a corresponding number of bunches of statistical errors bars.
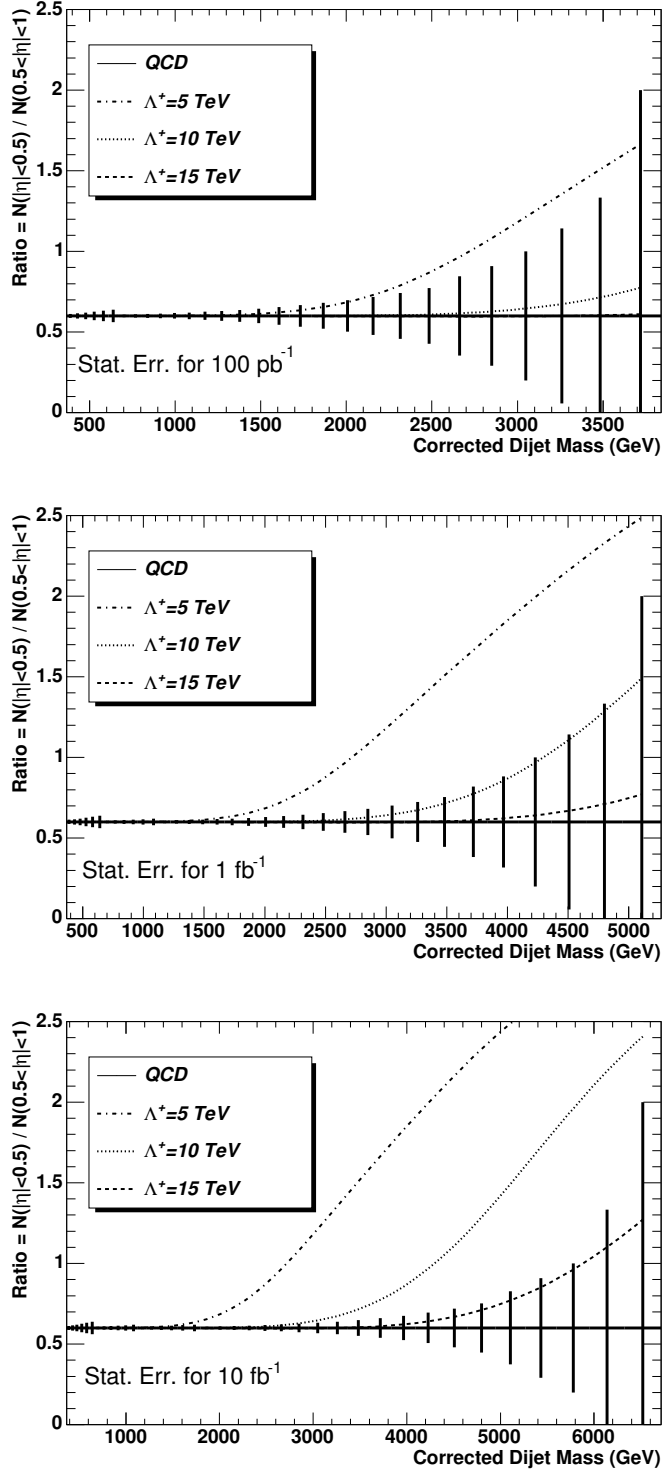
Figure 5: For integrated luminosities of 100 pb^{-1}(top plot), 1 fb^{-1}(middle plot), and 10 fb^{-1}(bottom plot), the expected value and statistical error of the dijet ratio of QCD in the CMS detector (solid) is compared with QCD plus a quark contact interaction at a scale Λ^+ of 15 TeV (dashed), 10 TeV (dotted) and 5 TeV (dot-dashed).

In Table 1 we list the dijet mass range that we analyze from each trigger. The lower value of the mass range is the lowest value of dijet mass for which the trigger is fully efficient, as discussed in reference [3]. The upper value of the mass range is the mass value where the next highest threshold trigger becomes fully efficient. For the unprescaled trigger, where there is no next highest threshold, the upper mass value listed is the upper edge of the last bin we used in the analysis As discussed previously [3] the efficiency of each trigger will be measured using the next trigger in each table with a lower threshold. Since there is no trigger with E_T threshold beneath the path labeled *Low* in any of the tables, we do not know how the trigger efficiency for the *Low* path will be measured, and therefore we do not use the *Low* path to measure the dijet ratio.

Path	L1				HLT		ANA
	E_T Cut (GeV)	Unpres. Rate (KHz)	Prescale (N)	Presc. Rate (KHz)	E_T Cut GeV)	Rate (Hz)	Dijet Mass (TeV)
Triggers for $\mathcal{L} = 10^{32}$ cm^{-2} s^{-1} and integrated luminosity = 100 pb^{-1}							
High	140	0.044	1	0.044	250	2.8	0.67-3.84
Med	60	3.9	40	0.097	120	2.4	0.33-0.67
Low	25	2.9×10^2	2,000	0.146	60	2.8	None
Triggers for $\mathcal{L} = 10^{33}$ cm^{-2} s^{-1} and integrated luminosity = 1 fb^{-1}							
Ultra	270	0.019	1	0.019	400	2.6	1.13-5.27
High	140	0.44	10	0.044	250	2.8	0.67-1.13
Med	60	39	400	0.097	120	2.4	0.33-0.67
Low	25	2.9×10^3	20,000	0.146	60	2.8	None
Triggers for $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1} and integrated luminosity = 10 fb^{-1}							
Super	450	0.014	1	0.014	600	2.8	1.80-6.72
Ultra	270	0.19	10	0.019	400	2.6	1.13-1.80
High	140	4.4	100	0.044	250	2.8	0.67-1.13
Med	60	3.9×10^2	4,000	0.097	120	2.4	0.33-0.67
Low	25	2.9×10^4	200,000	0.146	60	2.8	None

Table 1: The single jet trigger table previously proposed [3],showing path names, trigger thresholds in corrected E_T, prescales, estimated rates at L1 and HLT for three different luminosity scenarios, and here we also list the corresponding range of corrected dijet mass used in this analysis. The trigger in the Low path is not used to measure the dijet ratio.

The statistical uncertainties on the dijet ratio in Figure 5 are smooth estimates for the integrated luminosities shown. The calculation of the statistical uncertainties is discussed in Appendix B. We use Poisson statistics at high dijet mass where few events are expected. For the QCD background, for all three integrated luminosities, the highest mass bin we show in Figure 5 has a mean value of expected events in the numerator of approximately 1.5 events and a mean value of expected events in the denominator of approximately 2.5 events.

In Figure 3 we presented a lowest order calculation of both QCD and a contact interaction among left-handed quarks. The signal in Figure 5 is estimated by scaling the lowest order contact interaction calculation of Figure 3 by the ratio of our full simulation prediction for QCD to the lowest order QCD calculation: signal = contact $\times$ 0.6 / QCD. In this representation a contact interaction is an upward deviation from a flat dijet ratio.

Comparing the contact interaction signals in Figure 5 to the QCD background and its statistical uncertainty shows us almost immediately what our level of sensitivity will be for contact interactions. Although we will quantify this more precisely in the remainder of the note, we can essentially read off our sensitivity from Figure 5. Clearly for 100 pb^{-1}it will be difficult to discover or exclude $\Lambda = 10$ TeV, which is too close to QCD, but we will be sensitive to roughly $\Lambda = 5$ TeV because it is beyond the edge of our error bars at high mass. For 1 fb^{-1}our statistical errors are reduced at high mass, and we should be sensitive to roughly $\Lambda = 10$ TeV, since that curve is now at the edge of our error bars. For 10 fb^{-1}the statistical errors are reduced again, and we should roughly be sensitive to $\Lambda = 15$ TeV. This simple sensitivity analysis works because the systematic uncertainties on the dijet ratio are small, as we show in the following section.

3 Systematic Uncertainties

In Figure 6 we present estimates of systematic uncertainties on the dijet ratio. Systematic uncertainties on the cross section versus dijet mass are large, as discussed in reference [3], but most of them simply cancel out in the ratio leaving a very small systematic uncertainty. Figure 6 shows both a plot where the systematics can be compared to the size of the statistical errors for 10 fb^{-1}and a similar plot with a smaller vertical scale where the individual systematics can be seen. Here we discuss these systematics.

3.1 Jet Energy

We divide the determination of the jet energy into the determination of two energy scales, also known as two energy multiplication factors, necessary to correct from measured jet energy to true jet energy. The absolute jet energy scale is the energy multiplication factor needed to correct jets constrained to the region $|\eta| < 1$, where the vast majority of energy is measured in the barrel calorimeter. The absolute jet energy scale is a function of p_{T}. The relative jet energy scale is the energy multiplication factor needed to correct jets as a function of η, relative to the region $|\eta| < 1$, and is also a function p_{T}. The relative jet energy scale is a measurement of the uniformity of jet energy response as a function of η.

3.1.1 Absolute Jet Energy Scale

We have concluded that an overall uncertainty on the jet energy scale in the barrel of $\pm 5\%$ is achievable [3]. This leads to a large error in the dijet mass cross section, because of the steeply falling spectrum. The uncertainty on the cross section in the region $|\eta| < 1$ varies from +30%(-22%) at a dijet mass of 0.3 TeV/c^2 to +80%(-45%) at a dijet mass of 6.5 TeV [3]. This large systematic uncertainty, increasing with dijet mass, is the primary reason we do not use the dijet mass distribution to search for quark contact interactions. For the background dijet ratio the absolute jet energy scale uncertainty has no effect, because the dijet ratio is flat versus dijet mass. The uncertainty affects the numerator and the denominator of the dijet ratio equally, and the uncertainty simply cancels out in the ratio. For the signal the uncertainty of 5% in the jet energy scale gives an uncertainty of 5% in the contact interaction energy scale Λ. The treatment of this systematic is discussed in section 4.2.2.
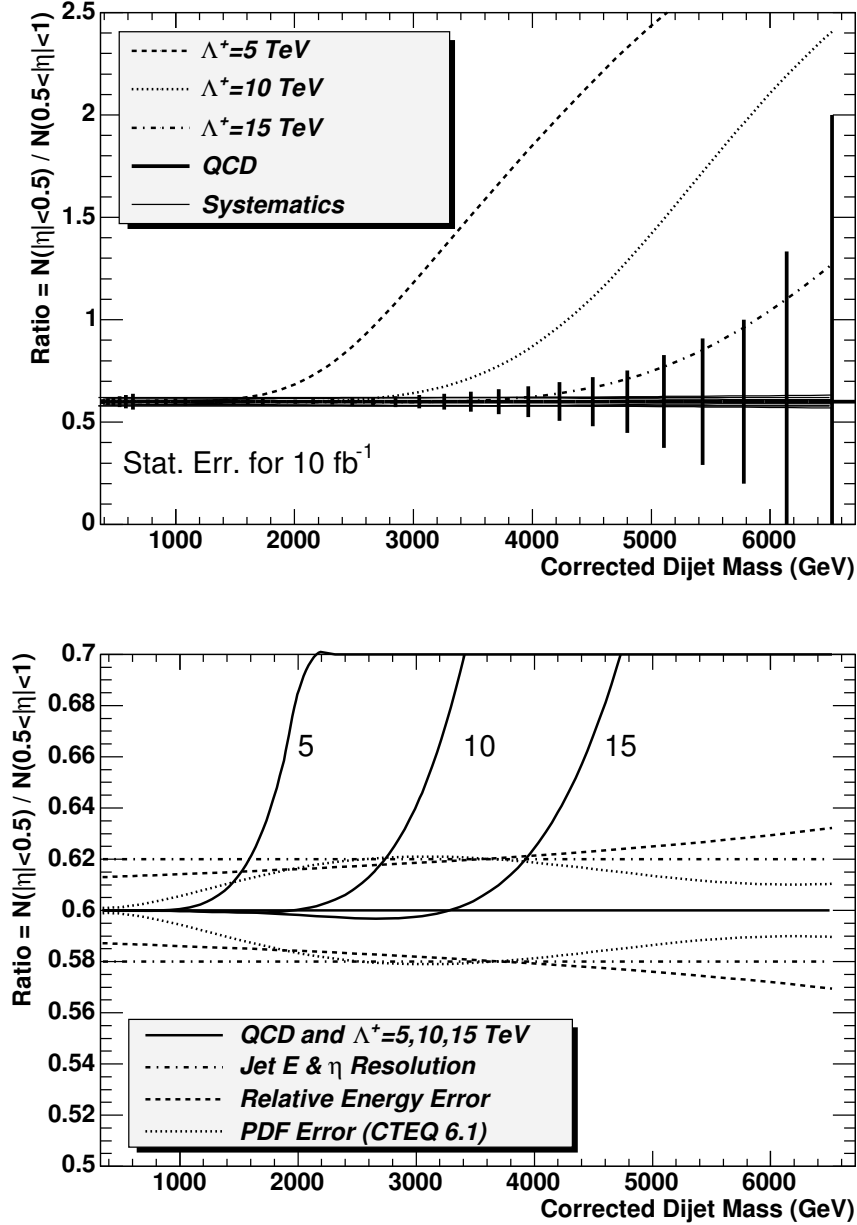
Figure 6: Top Plot) For $10 \, \mathrm{fb}^{-1}$ the expected value and statistical error of the dijet ratio of QCD in the CMS detector (solid) is compared with QCD plus a quark contact interaction at a scale Λ^+ of 15 TeV (dashed), 10 TeV (dotted) and 5 TeV (dot-dashed). The systematic uncertainties (curves) are shown bracketing QCD, but on this scale they are only barely visible. Bottom Plot) Similar to top plot but with a magnified vertical scale. Systematic bounds on the dijet ratio from uncertainties in the relative jet energy scale (dashed curve), parton distributions (dotted curve), and calorimeter energy and eta resolution (dot dash curve), are compared to the expectations of QCD and three contact interaction scales (solid line and curves).

3.1.2 Relative Jet Energy Scale

The relative jet energy scale will be determined using dijet balance [12]. In dijet balance one of the two leading jets is required to be in the region $|\eta| < 1$, the other leading jet can be at any η, and the relative jet energy scale is set by requiring that they have equal p_T on average. We have shown that by using dijet balance an uncertainty of $\pm 0.5\%$ is achievable [12] for the relative jet energy scale as a function of η within the barrel, in 0.1 steps in η. Here we assume that the relative jet energy scale, defined in this analysis as the uniformity in energy scale in the region $0.5 < |\eta| < 1.0$ compared to $|\eta| < 0.5$, can be determined to $\pm 0.5\%$. We have propagated this error to the dijet ratio by measuring the effect of a $\pm 0.5\%$ change in dijet mass for the measurement of $N(0.5 < |\eta| < 1)$ while keeping $N(|\eta| < 0.5)$ unchanged. As shown in Figure 6, the resulting upper uncertainty in the ratio varies from 0.013 (2%) at a mass of 0.3 TeV/c^2 to 0.032 (5%) at a mass of 6.5 TeV.

3.2 Resolution

The effect of calorimeter resolution is the difference between the measurement with jets constructed from MC particles (Gen Jets) and the measurement with jets constructed from calorimeter depositions and corrected (Rec Jets). This difference, often called the smearing due to calorimeter resolution, is taken as a bound on the size of the systematic uncertainty due to resolution. For the cross section in $|\eta| < 1$, the difference between Rec Jets and Gen Jets is small [3]. For the ratio, we show in Figure 7 that there is no change between Gen Jets and corrected Rec Jets within the Monte Carlo statistics. The statistical error on the simulation gives a bound on the systematic of 0.02 (3%) in the ratio, which is shown in Figure 6.
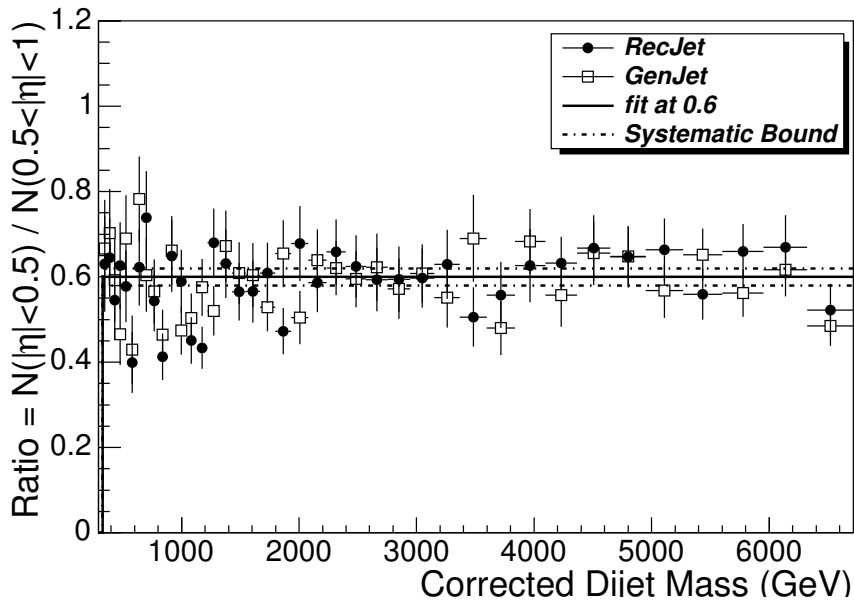

Figure 7: The dijet ratio in the CMS QCD simulation from reconstructed jets (solid circles) is compared to jets clustered from MC particles before interaction in the detector (open boxes). They are both consistent with 0.6 (solid line). The combined statistical error on mean ratio from the two measurements is ± 0.02 and is shown as dot-dash lines bracketing 0.6

3.3 Parton Distributions

Uncertainties in parton distributions produce uncertainties in theoretical predictions of the QCD cross section as a function of dijet mass. When the dijet ratio is measured, we will compare it to the QCD prediction, and this uncertainty in the theory will make the search for new physics more difficult. Therefore, although this uncertainty does not affect our future measurement, it can affect our future search for new physics, and so we consider it here.

Previous experiments colliding hadrons with various targets provide information on Parton Distribution Functions (PDFs) that can be evolved to LHC energies. Attempts to extract the PDFs are affected by uncertainties in the experimental measurements and the theory that describes them. The CTEQ collaboration provides a PDF set which includes both the best fit PDF and the errors resulting from the fit between world data and theory [14]. We have used CTEQ 6.1, which has 40 different error PDFs: 20 PDFs at positive error (S_i^+), and 20 PDFs at negative error (S_i^-). They recommend that the error on any observable quantity, X, be calculated using the formula

$$\Delta X = \frac{1}{2} \left(\sum_{i=1}^{20} \left[X(S_i^+) - X(S_i^-) \right]^2 \right)^{1/2},\tag{1}$$

This requires the user to explicitly calculate the observable quantity 40 times in order to calculate the error ΔX. This can be a prohibitively large task for a Monte Carlo and detector simulation. Instead, to simplify the problem ,we have used our own lowest order analytic calculation of the dijet ratio versus dijet mass. This is sufficient to determine uncertainties due to parton distributions for this analysis since our lowest order calculation makes the same use of the parton distributions as does the Pythia Monte Carlo. For each bin of dijet mass we calculated the lowest order dijet ratio in that bin, $R = N(|\eta| < 0.5)/N(0.5 < |\eta| < 1.0)$, using the default CTEQ6.1 set, and then used equation 1 to calculate the uncertainty in the ratio, ΔR. As shown in Figure 6, the resulting uncertainty in the dijet ratio peaks at a value of 0.02 (3%) in the ratio at a mass of around 3.5 Tev, and declines at both lower and higher masses.

3.4 Luminosity, Efficiency and Acceptance

The luminosity uncertainty on the cross section cancels in the dijet ratio. Efficiency is not an issue, since at dijet masses greater than 0.33 TeV/c^2 there is full efficiency for finding a dijet in the region $|\eta| < 1$ with negligible uncertainty. Acceptance is also not an issue, since our acceptance for the two leading jets is defined by our cut on $|\eta|$. Any measurements made with an $|\eta|$ cut must be compared to theoretical predications that also have the same $|\eta|$ cut. There is negligible uncertainty in the relative acceptance of the measured and calculated jet η region.

4 Contact Interaction Sensitivity Estimates

To make quantitative estimates of our sensitivity to contact interactions we employ a χ^2 method. In Figure 5 we visually compared QCD plus a contact interaction to QCD alone and its estimated statistical uncertainties for 100 pb^{-1}, 1 fb^{-1}and 10 fb^{-1}. In Figure 6 we showed the systematic uncertainties on the QCD prediction. In this section we form a χ^2 between QCD plus a contact interaction and QCD alone, and use that χ^2 to estimate the Λ values we expect to be able to exclude at 95% CL and the Λ values we expect to be able to discover

at 5σ. First we use only the expected statistical uncertainties for each sample size. Next we include systematic uncertainties which have very little effect because they are small.

4.1 Sensitivity with Statistical Uncertainties

In Table 2 we show the χ^2 with statistical uncertainties only between QCD plus a contact interaction and QCD alone:

$$\chi^2 = \sum_i \frac{\Delta_i^2}{\sigma_i^2} \tag{2}$$

where for each bin i, Δ_i is the difference between QCD plus a contact interaction and QCD alone, and σ_i is the statistical uncertainty on QCD, as shown in Figure 5. Since all our es-

Luminosity	100 pb^{-1}			1 fb^{-1}			10 fb^{-1}		
Λ^+(TeV)	5	10	15	5	10	15	5	10	15
χ^2 (Stat)	18.3	.090	.0037	316	5.82	.107	3652	133	4.15

Table 2: Chisquared between signal and background. For each luminosity and contact interaction scale considered we list the chisquared between QCD alone and QCD plus a contact interaction, for the case where only statistical uncertainties are included.

timates are smooth, without statistical fluctuations in either the background or the signal, the χ^2 must be exactly zero when the contact interaction scale is very large ($\Lambda \to \infty$) and the signal distribution becomes identical to QCD. This is different than a χ^2 in the presence of actual statistical fluctuations, which is seldom expected to be zero. Our χ^2, like any χ^2, changes by $\Delta\chi^2 = N^2$ units for a difference between signal and background of $N\sigma$, where σ is the RMS of the Gaussian probability of a signal. Since our χ^2 begins at $\chi^2 = 0$ for no signal, the significance in sigma, N, of any signal is simply, $N = \sqrt{\chi^2}$. Our χ^2 is actually a χ^2 difference.

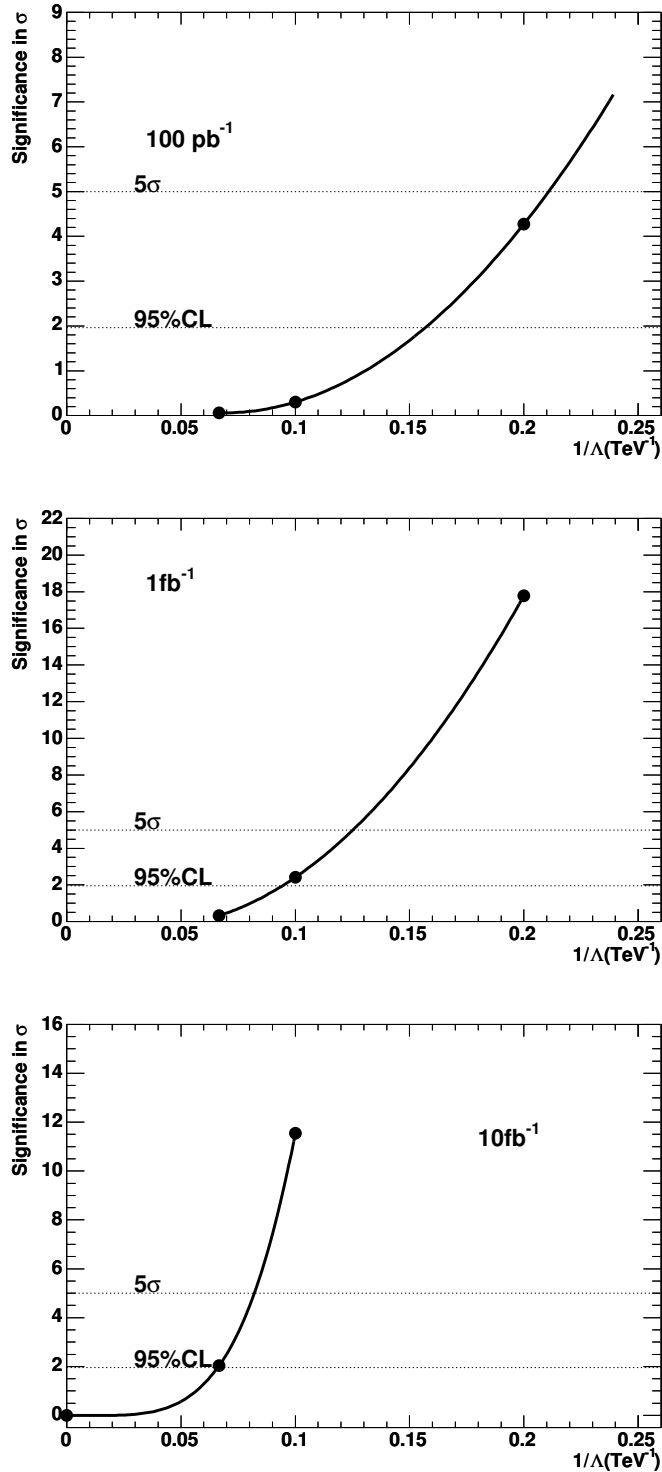Figure 8: The significance with statistical uncertainties only of the difference between QCD alone and QCD plus a quark contact interaction for integrated luminosities of 100 pb^{-1}(top plot), 1 fb^{-1}(middle plot), and 10 fb^{-1}(bottom plot). The significance is plotted vs $1/\Lambda^+$ and fit with a smooth function. Horizontal lines show the 5σ and 95% CL levels.

In Figure 8 we plot the significance versus $1/\Lambda$. As discussed, $1/\Lambda = 0$ corresponds to QCD, no contact interaction signal, and must have a $\chi^2 = 0$ and a significance of 0σ. A 95% CL exclusion corresponds to a significance of 1.96σ, conservatively estimated assuming a two sided Gaussian probability for the signal, and this level is shown by a horizontal dotted line in Figure 8. A 5σ discovery level is also shown by a horizontal dotted line in Figure 8. For integrated luminosities of 100 pb^{-1} and 1 fb^{-1} we compare the significance for the three values of $\Lambda = 5$, 10, and 15 TeV to these levels, and note that the significance of these Λ values is close enough to the desired confidence levels to use them to either interpolate or extrapolate to find the desired confidence level. For 100 pb^{-1} and 1 fb^{-1} a quadratic function is fit to the three points as shown in Figure 8 and is used to find the 95% CL and 5σ values of Λ. For an integrated luminosity of 10 fb^{-1} the 95% CL and 5σ significance levels are both close to the $\Lambda = 15$ TeV value, and we found that including the $\Lambda = 5$ TeV point in the fit biased the interpolation to slightly higher values of Λ. So instead we fit the three points $\Lambda = 10$, 15, and ∞ TeV explicitly using the constraint that $\Lambda = \infty$ must have a significance of 0σ. These three points were well fit with a quartic function, shown in Figure 8, and the quartic was solved for the 95% CL and 5σ values of Λ. In Table 3 we show the resulting Λ values for a 95% CL exclusion or 5σ discovery with statistical uncertainties only.

	95% CL Excluded Scale			5σ Discovered Scale		
	100 pb^{-1}	1 fb^{-1}	10 fb^{-1}	100 pb^{-1}	1 fb^{-1}	10 fb^{-1}
$\Lambda^+(\text{ TeV})$	<6.4	<10.6	<15.1	<4.7	<8.0	<12.2

Table 3: Statistical sensitivity to contact interactions with 100 pb^{-1}, 1 fb^{-1}, and 10 fb^{-1}. We list the largest value of the contact interaction scale we expect to be able to exclude at a confidence level of 95% or greater, and the largest value we expect to be able to discover with a significance of 5σ or greater. Estimates include ONLY statistical uncertainties.

4.2 Sensitivity with All Uncertainties

The inclusion of systematic uncertainties in statistical estimators is not a well defined process. There are many different approaches, and they are as fraught with as much systematic bias as the underlying systematic uncertainties themselves. We have tried to pick the simplest possible method, based on one that has been used in the past to search for contact interactions [16]. We note that our systematics are small, and we therefore do not expect them to have much affect on the sensitivity. Indeed, using our method, the systematics have only a small affect, and for our analysis the method is sufficient.

Each individual systematic uncertainty on the dijet ratio is completely correlated as a function of dijet mass. For example, for the systematic uncertainty on the relative jet energy scale, if the ratio increases by a value of 0.013 (2%) at a mass of $0.3 \text{ TeV}/c^2$ it must also increase by an amount of 0.032 (5%) at a mass of 6.5 TeV, and it must also increase by every value in between as shown explicitly by the curve in Figure 6. This is what we mean by a completely correlated systematic uncertainty.

4.2.1 χ^2 Including Systematic Uncertainties on Dijet Ratio

In this section we discuss first the inclusion of a single correlated systematic uncertainty in our χ^2, which is precisely the method used in reference [16]. We then discuss the handling of multiple systematic uncertainties, introducing a simple method of preserving the correla-

tions as a function of mass of each systematic while including all systematics in the χ^2.

The χ^2 in the presence of correlated uncertainties can be written as

$$\chi^2 = \sum_{i,j} \Delta_i (V^{-1})_{ij} \Delta_j \tag{3}$$

Where i and j range over all the bins of dijet mass, and as for equation 2, Δ_i is the difference between QCD plus a contact interaction and QCD alone in bin i, and similarly for Δ_j in bin j. V is the covariance matrix, and equation 3 contains its inverse, V^{-1}. The diagonal elements of the covariance matrix contain the statistical and systematic uncertainty for each bin added in quadrature

$$V_{ii} = \sigma_i^2(stat) + \sigma_i^2(sys) \tag{4}$$

and the off-diagonal elements of the covariance matrix contain the bin-to-bin correlations of the systematic uncertainty

$$V_{ij} = \sigma_i(sys)\sigma_j(sys) \tag{5}$$

where $\sigma_i(sys)$ is the systematic uncertainty on bin i as shown in Figure 6.

When there is no systematic uncertainty equation 3 simply reduces to equation 2. When there is a correlated systematic uncertainty, equation 3 will always give a smaller χ^2 than equation 2. Qualitatively, if the difference Δ has a similar shape as a function of mass as the systematic uncertainty, $\sigma(sys)$, then including the systematic will give a smaller χ^2 than if it has a different shape. In Table 4 we compare the χ^2 with statistical uncertainties only to the χ^2 including a single systematic uncertainty for each of the systematics shown in Figure 6. Note that the effect of any individual systematic on the χ^2 is small for Λ values at the edge

Luminosity	100 pb^{-1}			1 fb^{-1}			10 fb^{-1}		
Λ^+(TeV)	5	10	15	5	10	15	5	10	15
χ^2 (Stat Only)	18.3	.090	.0037	316	5.82	.107	3652	133	4.15
χ^2 (PDF)	17.1	.086	.0024	264	5.65	.082	2353	128	3.86
χ^2 (Jet Energy)	18.1	.088	.0030	303	5.76	.096	3110	131	3.99
χ^2 (Resolution)	18.1	.099	.0031	305	5.77	.097	3181	131	4.01
χ^2 (All)	16.7	.082	.0011	240	5.55	.061	1340	124	3.56

Table 4: Chisquared between signal and background. For each luminosity and contact interaction scale considered we list the χ^2 between QCD alone and QCD plus a contact interaction. The χ^2 where only statistical uncertainties are included (Stat Only), and the χ^2 where a single correlated systematic uncertainty has also been included. The systematics considered are parton distributions (PDF), relative jet energy scale (Jet Energy), and jet energy and η resolution (Resolution). The final χ^2 (All) includes all of the above uncertainties (see text).

of our sensitivity: for 100 pb^{-1}the effect is small for $\Lambda = 5$ TeV, for 1 fb^{-1}the effect is small for $\Lambda = 10$ TeV, and for 10 fb^{-1}the effect is small for $\Lambda = 15$ TeV. For these cases the effect of one systematic is never greater than 7% on the chisquared value, and the significance is proportional to the $\sqrt{\chi^2}$. Still, we learn something by looking at the relative magnitude of the effects. The PDF systematic has largest effect because it varies the most with dijet mass. In particular, Figure 6 shows that the PDF systematic rises versus dijet mass for dijet mass less than 3 TeV, and the $\Lambda = 5$ TeV contact interaction rises in the same mass region, and as a consequence the PDF systematic has the largest effect for a $\Lambda = 5$ TeV contact interaction. For the case of 10 fb^{-1}, where the statistical uncertainties are smallest, when $\Lambda = 5$ TeV the

PDF systematic decreases the χ^2 by over 30%. The systematic uncertainties on the jet energy scale and jet resolution are fairly flat as a function of mass and they have less effect on the χ^2.

We now consider the χ^2 in the presence of multiple systematic uncertainties. Systematic number k, reduces the χ^2 by an amount

$$\Delta\chi_k^2(sys) = \chi^2(stat) - \chi_k^2(sys) \tag{6}$$

where $\chi_k^2(sys)$ includes the systematic, and $\chi^2(stat)$ includes only statistical uncertainties, as listed in Table 4. We estimate the χ^2 including all systematic uncertainties, $\chi^2(All)$, by subtracting from the χ^2 with only statistical uncertainties the sum of the changes in the χ^2 from each of the sources of systematic uncertainty:

$$\chi^2(All) = \chi^2(stat) - \sum_k \Delta\chi_k^2(sys) \tag{7}$$

Each systematic then decreases $\chi^2(All)$, and each systematic has the same effect in the total χ^2 as it would if it were considered in isolation. The value of $\chi^2(All)$ is listed in Table 4.

This is a conservative way to handle the systematics in the calculation of a χ^2. It takes into account the shape as a function of mass of each systematic. The method used in reference [16] only used the shape of a single combined systematic. That method was to first sum all the systematics in quadrature, then consider this combined systematic as a single total systematic, and use that single total systematic in equation 3 to determine the χ^2. We believe our method is an improvement on this, because we do not lose the individual shapes of the systematics by combining them all in quadrature, and we have seen how important the individual shapes of the systematics can be even when the magnitudes of the systematic are comparable. If we had first added the systematics in quadrature we would have flattened out the variation as a function of mass, and this would have led to a smaller change in χ^2.

In Figure 9 we show the sensitivity in σ resulting from $\chi^2(all)$. The determination of the 95% CL and 5σ levels uses the same methodology as discussed previously in section 4.1, and the results are presented in Table 5. The systematic uncertainties on the dijet ratio reduced the CMS sensitivity to a contact interaction between 0.0 and 0.2 TeV/c^2 depending on luminosity and level of significance.
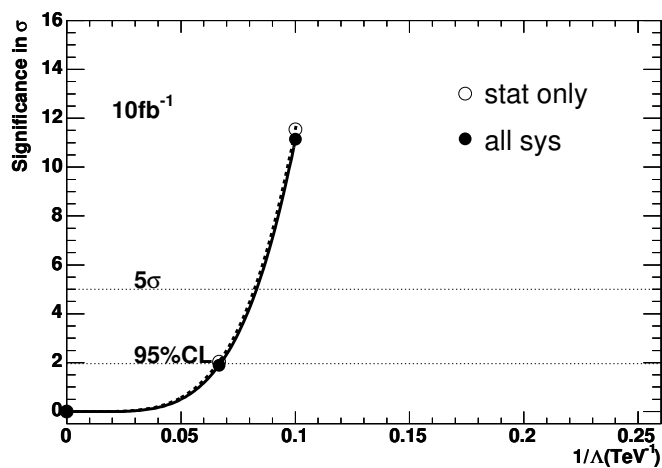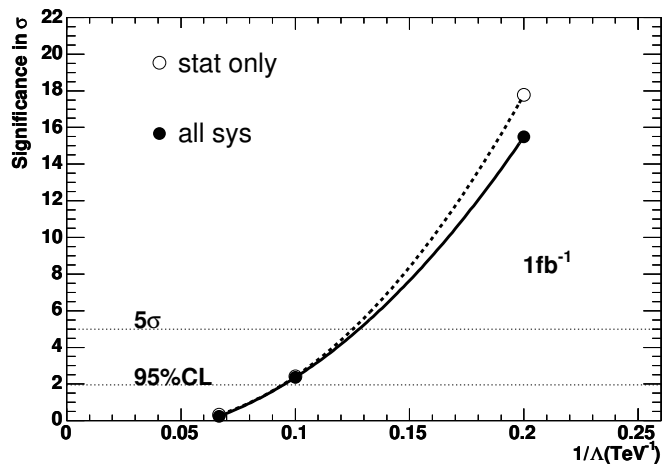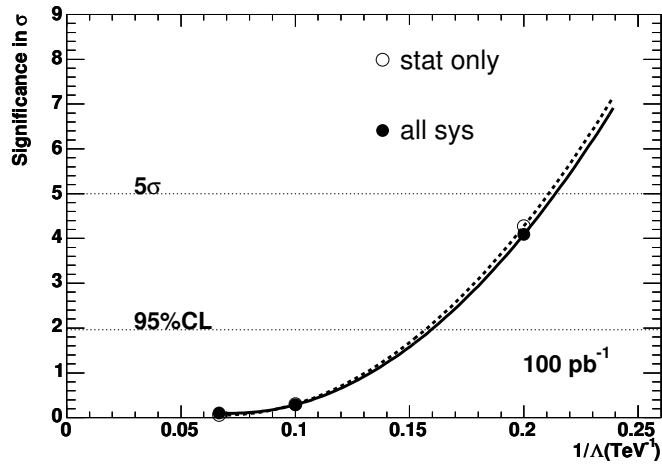
Figure 9: The significance of the difference between QCD alone and QCD plus a quark contact interaction for integrated luminosities of 100 pb^{-1}(top plot), 1 fb^{-1}(middle plot), and 10 fb^{-1}(bottom plot), is shown for both statistical uncertainties only (open points and dashed curve) and for all uncertainties (solid points and curve). The significance is plotted vs $1/\Lambda^{+}$ and fit with a smooth function. Horizontal lines show the 5σ and 95% CL levels.

	95% CL Excluded Scale			5σ Discovered Scale		
	100 pb^{-1}	1 fb^{-1}	10 fb^{-1}	100 pb^{-1}	1 fb^{-1}	10 fb^{-1}
Λ^+(TeV)	<6.3	<10.5	<14.9	<4.7	<7.8	<12.0

Table 5: Sensitivity to contact interactions excluding Λ systematic. We list the largest value of the contact interaction scale we expect to be able to exclude at a confidence level of 95% or greater, and the largest value we expect to be able to discover with a significance of 5σ or greater. Estimates include both statistical and systematic uncertainties on the dijet ratio, but do NOT include a 5% uncertainty on the contact interaction scale resulting from the jet energy scale uncertainty.

4.2.2 Sensitivity Including Λ Systematic from Jet Energy Scale

The sensitivity to Λ presented in Table 5 include all systematic uncertainties except an uncertainty of 5% in Λ due to the jet energy scale. This uncertainty does not affect our ability to discover a contact interaction. The Λ value we discovered would just be uncertain by 5% due to our energy scale uncertainty. This uncertainty does, however, affect our ability to exclude a contact interaction. Here we discuss how this uncertainty is included in the determination of the contact interaction scale that we can expect to exclude at 95% CL.

Following reference [13] we employ a Bayesian technique. Motivated by the form of the Lagrangian in Appendix A, a uniform prior is assumed in $\xi = 1/\Lambda^2$ and a Gaussian likelihood function $P \propto e^{-\chi^2/2}$ is used. The likelihood $P(\xi)$ is convoluted with our uncertainty in Λ. The 95% confidence level in Λ is determined by requiring $\int_0^\xi P(\xi')d\xi' = 0.95$. Including the uncertainty in Λ decreased the excluded scale by 0.7%. This change in the excluded value of Λ is smaller than the 5% uncertainty in the scale because the exclusion is dominated by statistical uncertainties.

The sensitivity to a contact interaction including all uncertainties is presented in Table 6.

	95% CL Excluded Scale			5σ Discovered Scale		
	100 pb^{-1}	1 fb^{-1}	10 fb^{-1}	100 pb^{-1}	1 fb^{-1}	10 fb^{-1}
Λ^+(TeV)	<6.2	<10.4	<14.8	<4.7	<7.8	<12.0

Table 6: Final sensitivity to contact interactions with 100 pb^{-1}, 1 fb^{-1}, and 10 fb^{-1}. We list the largest value of the contact interaction scale we expect to be able to exclude at a confidence level of 95% or greater, and the largest value we expect to be able to discover with a significance of 5σ or greater. Estimates include both statistical uncertainties and all systematic uncertainties. We estimate that any value of Λ discovered would have a systematic uncertainty of 5% due to jet energy scale uncertainty.

5 Conclusions

We have used the dijet ratio, a simple measure of the dijet angular distribution, to estimate CMS sensitivity to quark contact interactions. The dijet ratio from a full CMS simulation of the QCD background is flat as a function of dijet mass. The dijet ratio from a contact interaction among left-handed composite quarks rises with increasing dijet mass. With only 100 pb^{-1}of data there is very significant statistical separation of the QCD background and a

quark contact interaction signal that is beyond the reach of the published Tevatron data.

The systematic uncertainties on the dijet ratio are small. The cross section systematics are large as a function of dijet mass [3] but they cancel in the dijet ratio. The impact of systematics on our sensitivity is evaluated and it is also small.

The CMS sensitivity to a contact interaction among left-handed quarks has been presented. For an integrated luminosity of $100 \, \mathrm{pb}^{-1}$, $1 \, \mathrm{fb}^{-1}$, and $10 \, \mathrm{fb}^{-1}$, CMS can expect to exclude at 95% CL a contact interaction scale Λ^+ of 6.2, 10.4, and 14.8 TeV or discover at 5σ significance a scale Λ^+ of 4.7, 7.8 and 12.0 TeV, respectively. This can be compared to the most sensitive search for quark contact interactions at the Tevatron [13], which used the same dijet ratio and $100 \, \mathrm{pb}^{-1}$, and obtained an exclusion of $\Lambda^+ < 2.7$ TeV at 95% CL.

This sensitivity to quark contact interactions implies a sensitivity to quark compositeness. Our Λ sensitivity is equivalent to observing or excluding a quark radius of order 10^{-18} cm. We also note that our contact interaction sensitivity to composite quarks is roughly twice our dijet mass resonance sensitivity to excited states of composite quarks [17]. This is consistent with the general expectation that new physics should be observed first in contact interactions before it is more directly observed in dijet mass resonances.

Quark contact interactions result from new physics coupled to quarks. Contact interactions are generally observed first before any exchanged particles are directly seen. As the LHC opens up a new energy frontier we must prepare to search in those channels which will give us the earliest indications of new physics. This analysis demonstrates that dijets at CMS can provide an early signal of physics beyond the standard model.

A Model and Calculation

This appendix contains details of the contact interaction model and a discussion of the lowest order QCD and contact interaction calculations.

A.1 Contact Interaction Model

We estimate CMS sensitivity to the most commonly used model of contact interactions among composite quarks [1, 2]. The Lagrangian for this contact interaction is formed by the product of left-handed electroweak isoscalar quark currents:

$$\mathcal{L} = \frac{2\pi A}{\Lambda^2)} \sum_{i,j=1}^{6} (\bar{q}_{iL}\gamma^\mu q_{iL})(\bar{q}_{jL}\gamma_\mu q_{jL}) \tag{8}$$

Here $A = \pm 1$ is the sign of the amplitude, and i and j labels the quark flavors and q is the quark field. Early limits in the literature only considered two quark flavors, u and d, but at the Tevatron it has become conventional in dijet searches to consider all quark flavors that result in two final state jets, which is five quark flavors. This Lagrangian is then added to the standard model QCD Lagrangian for jet production and the lowest order subprocess cross sections with quarks in the initial and final state is calculated [2]:

$$\frac{d\hat{\sigma}(q_i q_i \to q_i q_i)}{d\cos\theta^*} = \frac{d\hat{\sigma}(\bar{q}_i\bar{q}_i \to \bar{q}_i\bar{q}_i)}{d\cos\theta^*} = \frac{\pi}{2\hat{s}}\left\{\frac{4}{9}\alpha_S^2\left[\frac{\hat{u}^2+\hat{s}^2}{\hat{t}^2}+\frac{\hat{t}^2+\hat{s}^2}{\hat{u}^2}-\frac{2}{3}\frac{\hat{s}^2}{\hat{t}\hat{u}}\right]+\frac{8}{9}\alpha_S\frac{A}{\Lambda^2}\left[\frac{\hat{s}^2}{\hat{t}}+\frac{\hat{s}^2}{\hat{u}}\right]+\frac{8}{3}\frac{\hat{s}^2}{\Lambda^4}\right\}; \tag{9}$$

$$\frac{d\hat{\sigma}(q_i \bar{q}_i \to q_i \bar{q}_i)}{d\cos\theta^*} = \frac{\pi}{2\hat{s}}\left\{\frac{4}{9}\alpha_S^2\left[\frac{\hat{u}^2+\hat{s}^2}{\hat{t}^2}+\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}-\frac{2}{3}\frac{\hat{u}^2}{\hat{s}\hat{t}}\right]+\frac{8}{9}\alpha_S\frac{A}{\Lambda^2}\left[\frac{\hat{u}^2}{\hat{t}}+\frac{\hat{u}^2}{\hat{s}}\right]+\frac{8}{3}\frac{\hat{u}^2}{\Lambda^4}\right\}; \tag{10}$$

$$\frac{d\hat{\sigma}(q_i \bar{q}_i \to q_j \bar{q}_j)}{d\cos\theta^*} = \frac{\pi}{2\hat{s}}\left\{\frac{4}{9}\alpha_S^2\left[\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}\right]+\frac{\hat{u}^2}{\Lambda^4}\right\}; \tag{11}$$

$$\frac{d\hat{\sigma}(q_i \bar{q}_j \to q_i \bar{q}_j)}{d\cos\theta^*} = \frac{\pi}{2\hat{s}}\left\{\frac{4}{9}\alpha_S^2\left[\frac{\hat{u}^2+\hat{s}^2}{\hat{t}^2}\right]+\frac{\hat{u}^2}{\Lambda^4}\right\}; \tag{12}$$

$$\frac{d\hat{\sigma}(q_i q_j \to q_i q_j)}{d\cos\theta} = \frac{d\hat{\sigma}(\bar{q}_i\bar{q}_j \to \bar{q}_i\bar{q}_j)}{d\cos\theta^*} = \frac{\pi}{2\hat{s}}\left\{\frac{4}{9}\alpha_S^2\left[\frac{\hat{u}^2+\hat{s}^2}{\hat{t}^2}\right]+\frac{\hat{s}^2}{\Lambda^4}\right\}. \tag{13}$$

Where $\sqrt{\hat{s}}$ is the sub-process energy in the center of momentum frame (effectively the dijet mass), $\hat{t} = \hat{s}(1 - \cos\theta^*)/2$ and $\hat{u} = \hat{s}(1 + \cos\theta^*)/2$. In equations 9- 13 the terms proportional to α_S^2 are from QCD, the terms proportional to $1/\Lambda^4$ are from the contact interaction, and the terms proportional to α_S/Λ^2 are from the interference between QCD and the contact interaction. To get a sense for the effective angular distributions, notice that in equation 13 the QCD term is proportional to $1/\hat{t}^2$ (t-channel) which peaks in the forward direction while the pure contact term has no angular dependence (isotropic). Similarly for equations 9, 10 and 13 the QCD terms contain (t-channel) components that peak in the forward direction while the pure contact term is either proportional to $\hat{s}^2$ (isotropic) or proportional to $\hat{u}^2$ which only gradually increases with $\cos\theta^*$. In this analysis we have chosen the sign $A = +1$ for the interference term, so following convention we have designated Λ as Λ^+. Previous searches have shown that this choice for the sign of the interference term maximizes the sensitivity to a contact interaction in angular distribution searches [13, 16], although the difference in the 95% CL limits between choosing $A = +1$ and $A = -1$ has only been roughly 10%.

22

A.2 Lowest Order Calculation

We have performed a lowest order calculation of both the QCD dijet background and QCD plus a contact interaction. For the sub-process cross section we used the expressions above and also included the purely QCD sub-processes from the remaining parton-parton interactions in proton-proton collisions: $q\bar{q} \to gg$, $gg \to q\bar{q}$, $qg \to qg$, and $gg \to gg$. These do not contribute to a quark contact interaction but they do contribute to the QCD background. The code that performed the numerical integrations was the same we used at the Tevatron in our measurement of the angular distribution and our search for quark contact interactions [16], except we updated it to use modern parton distributions. We used the CTEQ6L parton distribution set [14] for our lowest order calculation. The parton distributions were obtained from the LHAPDF interface [15] which also provides the value of α_S appropriate for the CTEQ 6L parton distributions. For the renormalization scale μ we used the dijet mass $\sqrt{\hat{s}}$. The resulting dijet mass distribution is shown in figure 2 and the dijet ratio versus dijet mass is shown in figure 3. The dijet ratio had been measured at the Tevatron and its ratio from QCD calculated at NLO [13]. As a check of our code we did a lowest order calculation of the dijet ratio for the Tevatron, and obtained a value of 0.7 for all measured masses, consistent with both the Tevatron measurement and the NLO prediction within errors.

B Statistical Uncertainty

B.1 Smooth Estimates

The statistical uncertainty on the QCD dijet ratio requires us to have estimates of the number of events in the numerator and the denominator, or equivalently estimates of the number of events in the denominator and a known value of 0.6 for the ratio. Like our ratio itself, we prefer to use smooth values for the expected statistical uncertainties. We measured the number of events *before* trigger prescales for the denominator, $N_b(0.5 < |\eta| < 1.0)$, using our QCD and detector simulation discussed in section 2.2.2, and then we smoothed it by fitting it to the following parameterization:

$$N_b(0.5 < |\eta| < 1.0) = \frac{p_0(1 - m/\sqrt{s})^{p_1}}{m^{p_2}} \tag{14}$$

where m is the dijet mass in GeV, $\sqrt{s} = 14000$ GeV is the collision energy, and p_0, p_1, p_2 are the fit parameters. We fit the simulation independently for the three different integrated luminosities, using the three different mass ranges of the analysis. The fit parameters found in table 7 give rates which are smoothly varying as a function of dijet mass, measured at the center of each bin, and scale properly with integrated luminosity within the simulation statistics. To obtain the expected number of events in the denominator of the dijet ratio,

	100 pb^{-1}	1 fb^{-1}	10 fb^{-1}
p_0	2.17993e+16	1.1837e+17	1.24253e+18
p_1	6.32830	7.33137	7.26496
p_2	4.21070	4.10893	4.11684

Table 7: Parameters of the fit to the number of events expected before trigger prescales, for the leading dijet in the region $0.5 < |\eta| < 1.0$, as a function of dijet mass, for three different values of integrated luminosity.

$\mu_2 = N(0.5 < |\eta| < 1.0)$, we multiply this smooth parameterization of $N_b(0.5 < |\eta| < 1.0)$ by the trigger prescales given in table 1. The number of events expected in the numerator of the dijet ratio from QCD, $\mu_1 = N(|\eta| < 0.5)$, is then simply 0.6 times μ_2. The majority of the statistical uncertainties shown in figure 5 are for large enough numbers of events to employ Gaussian statistics to estimate the uncertainty on the ratio:

$$\sigma_r = r * \sqrt{\frac{1}{\mu_1} + \frac{1}{\mu_2}} \tag{15}$$

where $r = \mu_1/\mu_2$ is the dijet ratio and σ_r is the Gaussian uncertainty on that ratio.

B.2   Poisson Statistics

The last few bins at high mass in figure 5 are expected to have only a handful of events, and Gaussian statistics are a poor estimator of the statistical uncertainty. To get an improved estimate we use Poisson statistics. The Poisson error on a ratio is not known in closed analytic form. We have written code that does a numerical calculation of the Poisson error on a ratio that works for the case where both the numerator, μ_1, and denominator, μ_2, are expected to have less than 30 events and the ratio,$r = \mu_1/\mu_2$ is in the interval $0.3 < r < 3$, which is sufficient for our analysis.

The Poisson probability of observing n events (an integer) when μ is the mean number expected is

$$P(\mu, n) = e^{-\mu}\mu^n/n! \tag{16}$$

For a fixed value of μ_1 and μ_2 we used equation 16 to determine the probability $P(\mu_1, n_1)$ and $P(\mu_2, n_2)$ of observing n_1 and n_2. We then looped over the values n_1 and n_2 from 0 to 50 observed events, and added the combined probability $P(\mu_1, n_1)P(\mu_2, n_2)$ into a histogram with the observed ratio n_1/n_2 on the horizontal axis and the total probability on the vertical axis. This gave us, for every pair of μ_1 and μ_2, a histogram which was a Poisson probability distribution as a function of observed ratio. For all cases of interest the integrated probability was very close to 1, indicating that 50 was a large enough maximum value for n, and we then forced the integral of the Poisson distribution to be exactly 1. We then found the value of the observed ratio n_1/n_2 on that probability distribution that corresponded to a 1σ upper error on a Gaussian: the observed ratio value where the integrated probability from 0 to n_1/n_2 reached 0.84. This value is defined as being 1σ above the mean expected ratio μ_1/μ_2 using Poisson statistics. As an example, the resulting Poisson error is tabulated in table 8 for the case of the ratio measured with an integrated luminosity of 1 fb^{-1}. There, Poisson statistics begin to be used at a mass of 3.7 TeV, where Poisson statistics only increases the error slightly. However, at a mass of 5.1 TeV Poisson statistics more than doubles the error, which is why we do not use Gaussian statistics at highest dijet mass. For integrated luminosities of 100 pb^{-1}, 1 fb^{-1}, and 10 fb^{-1}we used Poisson statistics for the last 6, 6, and 5 bins respectively, and everywhere else we used Gaussian statistics.

Dijet Mass (GeV)	Expected Events $N(\lvert\eta\rvert < 0.5)$	Expected Events $N(0.5 < \lvert\eta\rvert < 1.0)$	Gaussian Error on Ratio	Poisson Error on Ratio
1175.01	9071.81	15119.7	0.00796827	-
1273.21	6165.81	10276.3	0.00966532	-
1377.71	4197.35	6995.59	0.0117145	-
1488.83	2860.24	4767.07	0.0141909	-
1606.93	1949.78	3249.63	0.0171877	-
1732.35	1328.89	2214.82	0.0208193	-
1865.49	904.926	1508.21	0.0252293	-
2006.73	615.317	1025.53	0.0305958	-
2156.49	417.495	695.825	0.0371437	-
2315.21	282.463	470.772	0.0451576	-
2483.33	190.426	317.376	0.0549982	-
2661.35	127.814	213.024	0.0671308	-
2849.75	85.3445	142.241	0.082153	-
3049.06	56.6386	94.3977	0.100845	-
3259.84	37.3205	62.2008	0.124233	-
3482.65	24.3903	40.6506	0.153675	-
3718.11	15.79	26.3166	0.190994	0.2185
3966.83	10.1126	16.8544	0.23866	0.2825
4229.49	6.3972	10.662	0.300066	0.3995
4506.77	3.99034	6.65057	0.379932	0.5425
4799.4	2.44939	4.08231	0.484934	0.7335
5108.14	1.47614	2.46023	0.624665	1.3995

Table 8: Gaussian and Poisson statistical errors. For an integrated luminosity of 1 fb^{-1}, and the trigger path without prescales, we list the dijet mass at the center of the bin, and for a dijet ratio equal to 0.6 we list the mean expected values of the number of events in the numerator $N(\lvert\eta\rvert < 0.5)$, and the denominator $N(0.5 < \lvert\eta\rvert < 1.0)$, and the statistical uncertainty on the dijet ratio from Gaussian statistics, and the statistical uncertainty on the dijet ratio from Poisson statistics (where we are able to calculate it).

C References

[1] Estia Eichten, Kenneth D. Lane, Michael E. Peskin *New Tests for Quark and Lepton Substructure*, Phys.Rev.Lett.50:811-814 (1983).

[2] Kenneth D. Lane, *Electroweak and Flavor Dynamics at Hadron Colliders*, BUHEP-96-8 (1996), hep-ph/9605257.

[3] Selda Esen and Robert M. Harris, *Jet Triggers and Dijet Mass*, **CMS NOTE-2006/069**.

[4] Daniel Stump, Joey Huston, Jon Pumplin, Wu-Ki Tung, H.L. Lai, Steve Kuhlmann, J.F. Owens, *Inclusive jet production, parton distributions, and the search for new physics*, JHEP 0310:046,2003, hep-ph/0303013.

[5] See R.S. Chivukula, Andrew G. Cohen, E.H. Simmons, *New Strong Interactions at the Tevatron?*, Phys.Lett.B385:209-217 (1996), hep-ph/9606233.

[6] CDF Collaboration (F. Abe et al.) *Inclusive jet cross-section in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV*, Phys.Rev.Lett.77:438-443,1996, hep-ex/9601008.

[7] CDF Collaboration (T. Affolder et al.), *A Measurement of the Differential Dijet Mass Cross-Section in p anti-p Collisions at $\sqrt{s} = 1.8$ TeV*, Phys.Rev.D61:091101 (2000), hep-ex/9912022.

[8] ORCA 8_7_3 at http://cmsdoc.cern.ch/ORCA/.

[9] A. Heister, O. Kodolova, V. Konoplianikov, S. Petrushanko, J. Rohlf, C. Tully, A. Ulyanov, *Measurement of Jets with the CMS Detector at the LHC.*, **CMS NOTE-2006/036**.

[10] The RecJetRootTree code is in the ORCA 8_7_1 package at [8].

[11] CMS Collaboration, D. Acosta et al., *CMS Physics TDR Volume 1*, CERN/LHCC 2006-001 (2006).

[12] PTDR1 [11], Section 11.6.2: *Dijet Balancing* beginning on page 422.

[13] D0 Collaboration (B. Abbott et al.), *The dijet mass spectrum and a search for quark compositeness in anti-p p collisions at $\sqrt{s} = 1.8$ TeV*, Phys.Rev.Lett.82:2457-2462 (1999), hep-ex/9807014.

[14] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP 07 (2002) 012, hep-ph/0201195.

[15] LHAPDF code available online at http://vircol.fnal.gov/.

[16] CDF Collaboration (F. Abe et al.), *Measurement of Dijet Angular Distributions at CDF*, Phys.Rev.Lett.77:5336-5341 (1996), hep-ex/9609011.

[17] K. Gumus, N. Akchurin, S. Esen and R. Harris, *CMS Sensitivity to Dijet Resonances*, **CMS NOTE 2006/070**.