7.3 Benchmark Channel: new physics from dijets

Editor(s): R. Harris
Contributor(s): N. Akchurin, S. Esen, K. Gumus, R. Harris

Inclusive dijet production ($pp \rightarrow 2\,\text{jets} + X$) is the dominant LHC process. Simple to observe, and rich in potential signals of new physics, dijets are expected to be one of the earliest CMS measurements.

7.3.1 Dijet Resonances and Contact Interactions

Dijet resonances and contact interactions are the two major signals of new physics with dijets. Dijet resonances are direct and compelling observations of a new physical object at a mass M, requiring an incoming parton-parton collision energy equal to the mass. Contact interactions are indirect observations of an energy scale of new physics, Λ, which can be significantly larger than the available collision energy. Resonances have the benefit of being clear and irrefutable signals of new physics, within the accessible energies, while contact interactions allow one to probe larger energy scales. For any new phenomena, we expect contact interactions to show up in the data first. For a massive resonance that is kinematically accessible at LHC, the peak will be observed after contact interactions affect the lower mass data.

7.3.2 Dijet Mass Analysis

We use samples produced during Data Challenge 2004, generated using Pythia dijet processes mixed with pileup of minimum bias interactions for an assumed luminosity of 2×10^{33} cm^{-2} s^{-1}, simulated with OSCAR 2_4_5, and later reconstructed with ORCA 8_7_1. Jets are reconstructed as localized energy depositions in the CMS calorimeters arranged in a projective tower geometry. 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. Both the jet energy and momentum are corrected back to the particles in the jet cone originating from the hard interaction excluding pileup [3]. We define the dijet system as the two jets with the highest p_{T} in an event (leading jets) and define the dijet mass $m = \sqrt{(E_1 + E_2)^2 - (\vec{P}_1 + \vec{P}_2)^2}$. We select events in which the leading jets each have $|\eta| < 1$. This cut enhances our sensitivity to new physics, produced at low $|\eta|$, compared to the predominantly t-channel processes from the QCD background. In all plots that are a function of dijet mass, we plot in bins of width equal to the measured mass resolution.

7.3.3 Single Jet Triggers (MOVE ME TO TRIGGER CHAPTER)

In this section we propose the single jet trigger paths. We have been driven by the needs of the inclusive jet and dijet analysis. To measure jet spectra down to low jet E_T and dijet mass requires multiple triggers, of roughly equal total rate, and with appropriately chosen E_T thresholds and prescales. In figure 7.1 we show estimates of the L1 and HLT single jet trigger rates vs. corrected jet E_T. In table 7.4 we show the single jet trigger paths from L1 to HLT including thresholds, prescales and estimates of the rates. We find that the maximum

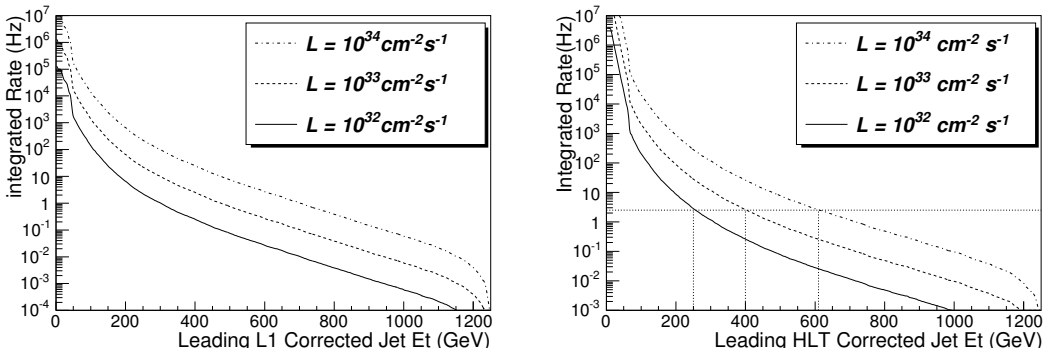

Figure 7.1: The integrated trigger rates at L1 (left) and HLT (right) above the E_T thresholds for the highest E_T jet is plotted versus the E_T threshold for three luminosity scenarios: $\mathcal{L} = 10^{32}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ (solid) , and $\mathcal{L} = 10^{33}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ (dashed), and $\mathcal{L} = 10^{34}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ (dot-dashed). HLT thresholds that give 2.5 Hz are shown by vertical dotted lines.

allowed HLT rate is the constraining factor for trigering on jets. For luminosity scenario 1, 2 and 4 the highest E_T threshold at HLT was chosen to give a rate of roughly 2.5 Hz, as illustrated in figure 7.1, so that four triggers would saturate an allowed jet rate of roughly 10 Hz at HLT. The highest E_T threshold in each scenario is not prescaled. Lower thresholds are prescaled and are chosen at roughly half the E_T of the next highest threshold. This allows reasonable statistics in the overlap between the two samples, necessary for measuring trigger efficiencies and producing a continuous jet spectrum. Note that the total L1 jet rate required is only around 0.1 KHz, a small fraction of the L1 total bandwidth. Since we are limited by HLT, not L1, for each trigger path the L1 thresholds are chosen low enough to have a L1 trigger efficiency of more than 95% at the corresponding HLT threshold in the path, as shown in Figure 7.2. This strategy utilizes ten times more bandwidth at L1 than at HLT to insure that all of the resulting HLT sample has high enough trigger efficiency to be useful for analysis.

Table 7.4 illustrates a trigger strategy to maintain the continuity of jet analysis as the luminosity increases over a time span of years. The most important feature is that each luminosity scenario maintains the thresholds introduced in the previous scenario, allowing combination of trigger samples over time. For the prescaled thresholds, we may increase the prescales, either in discrete steps or dynamically, to maintain the allowed HLT rate with increasing luminosity. However, to maintain maximum sensitivity to new physics, the highest E_T threshold must never be prescaled. For example, in table 7.2 when the luminosity increases by only a factor of 2 from Scenario 2 to Scenario 3, we double the prescales on the prescaled triggers but don't change either the threshold or the prescale of the highest E_T trigger labelled Ultra. This allows us to maintain stability of the single jet trigger thresholds, and analyses that depend on them, with only modest increases in the total rate for single jets. When the HLT rate in the unprescaled trigger becomes intolerably high, a higher E_T threshold unprescaled trigger is introduced, and the old unprescaled trigger can then be prescaled as necessary.

To commission the calorimeters, or perform a one-time jet study, it may be desirable to have more jets. If we want to write more than roughly 10 Hz of single jets at HLT, we recommend using the thresholds in Table 7.4, but lowering the prescales to obtain more jets at low E_T. This is preferable to moving the threshold for the unprescaled trigger, or any of the triggers,

Path	L1				HLT	
	E_T Cut (GeV)	Unpres. Rate (KHz)	Prescale (N)	Presc. Rate (KHz)	E_T Cut GeV)	Rate (Hz)
Single Jet Triggers in Scenario 1: $\mathcal{L} = 10^{32}$ cm^{-2} s^{-1}						
High	140	0.034	1	0.034	250	2.8
Med	60	0.91	40	0.023	120	2.4
Low	25	3.9×10^1	2,000	0.020	60	2.8
Single Jet Triggers in Scenario 2: $\mathcal{L} = 10^{33}$ cm^{-2} s^{-1}						
Ultra	270	0.016	1	0.016	400	2.6
High	140	0.34	10	0.034	250	2.8
Med	60	9.1	400	0.023	120	2.4
Low	25	3.9×10^2	20,000	0.020	60	2.8
Single Jet Triggers in Scenario 3: $\mathcal{L} = 2 \times 10^{33}$ cm^{-2} s^{-1}						
Ultra	270	0.032	1	0.032	400	5.2
High	140	0.68	20	0.034	250	2.8
Med	60	18	800	0.023	120	2.4
Low	25	7.8×10^2	40,000	0.020	60	2.8
Single Jet Triggers in Scenario 4: $\mathcal{L} = 10^{34}$ cm^{-2} s^{-1}						
Super	450	0.013	1	0.013	600	2.8
Ultra	270	0.16	10	0.016	400	2.6
High	140	3.4	100	0.034	250	2.8
Med	60	9.1×10^1	4,000	0.023	120	2.4
Low	25	3.9×10^3	200,000	0.020	60	2.8

Table 7.4: Single jet trigger table showing path names, trigger thresholds in corrected E_T, prescales, and estimated rates at L1 and HLT for four different luminosity scenarios.

and ending up with a special trigger that is only applicable for a given running period and difficult to combine with other samples.

7.3.4 Rates and Efficiencies from Jet Triggers

We use simulated data from the single jet triggers discussed in section 7.3.3. From the three trigger tables for luminosities of $L = 10^{32}$, 10^{33}, 10^{34} cm^{-2}s^{-1} we expect initial samples of size at least 100 pb^{-1}, 1 fb^{-1}, and 10 fb^{-1} respectively. This is from 10^6 seconds of collisions, equivalent to one month of continuous operation at 40% efficiency. In Fig. 7.3 we show the rate expected from these triggers as a function of dijet mass. By construction there are comparable events in each trigger, and a high statistics overlap between triggers for a given table. We see that the highest mass dijet is expected to be 5, 6 and 7 TeV for samples of size 100 pb^{-1}, 1 fb^{-1}, and 10 fb^{-1} respectively. In Fig. 7.4 we show the trigger efficiency vs. dijet mass, measured for each trigger using the neighboring trigger with a lower p_{T} threshold, and explicitly show the mass cuts that are fully efficient. In Fig. 7.5 we show the data we will use to measure the cross section. We use each trigger where it is fully efficient and stop using the trigger where the next trigger is fully efficient. Fig. 7.5 shows there are adequate numbers of fully efficient events for analysis.

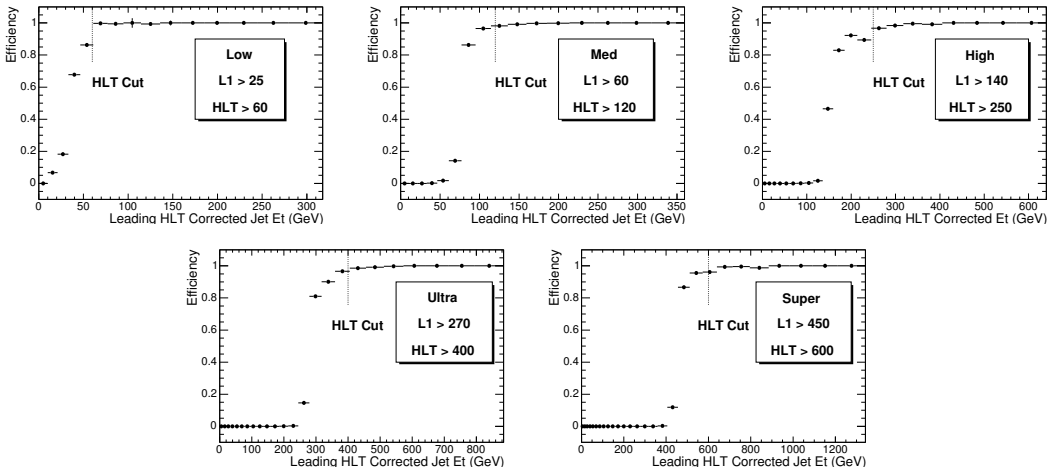

Figure 7.2: The efficiency for passing the L1 jet trigger is shown as a function of HLT corrected jet E_T for each of the trigger paths shown in table 7.4. The L1 thresholds were chosen to give an efficiency of greater than 95% at the corresponding HLT threshold.

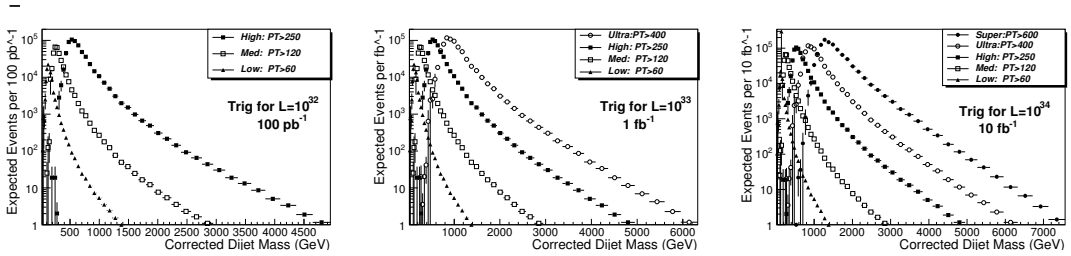

Figure 7.3: Rate of jet trigger as a function of dijet mass. The 3 plots correspond to 3 trigger tables, and each plot shows multiple triggers with various p_{T} thresholds and prescales.

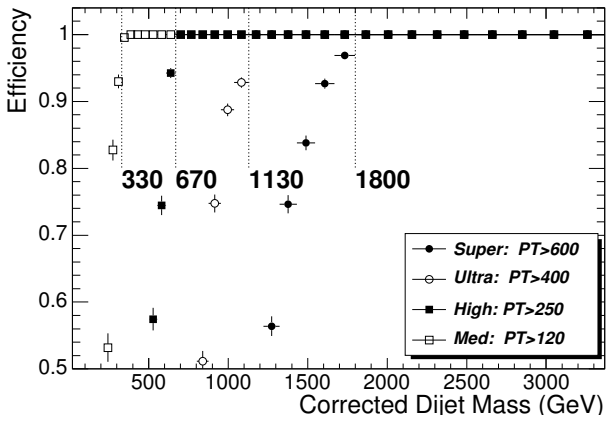

Figure 7.4: Jet trigger efficiency (points) and fully efficient dijet mass cuts (lines).

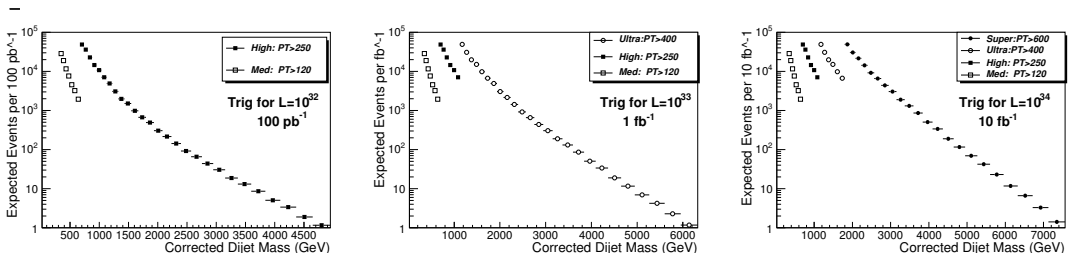

Figure 7.5: Rate of jet trigger for cross section measurement. Same triggers as Fig. 7.3.

7.3.5 Dijet Mass Distribution from QCD

In Fig. 7.6 we combine the triggers to produce a cross section across the full mass spectrum. The prescaled triggers allow us to measure mass down to 300 GeV, or even smaller if we

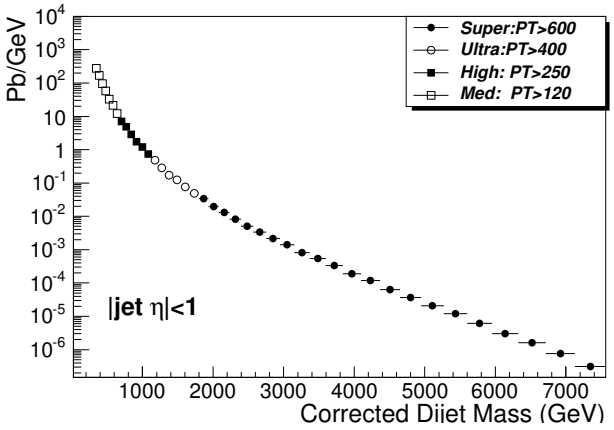

Figure 7.6: Cross section vs. dijet mass and the contributing jet triggers.

can understand the efficiency of the lowest threshold trigger. The mass measured with the prescaled triggers will allow us to connect to dijet masses measured at the Tevatron.

In Fig. 7.7 we show the fractional statistical error on the cross section, the simplest measure of our sensitivity to new physics. Fig. 7.7 shows that our prescaled triggers will allow a

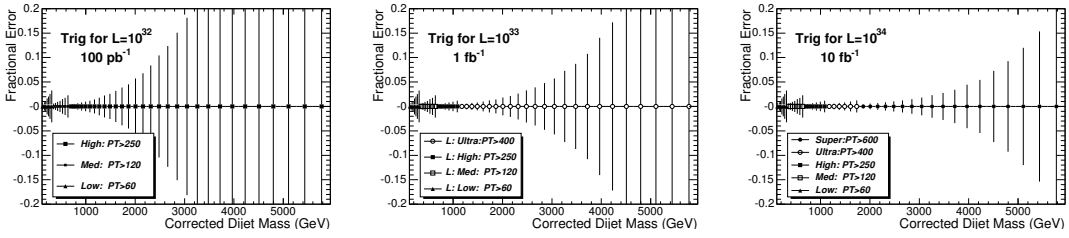

Figure 7.7: Fractional statistical error on the jet cross section for the samples in Fig. 7.7.

measurement of QCD with 1-3% statistical accuracy. The unprescaled triggers will have 1% error at threshold and the first unprescaled sample begins at a mass of 670 GeV, giving us full

sensitivity to new physics in a region that overlaps with previous dijet mass measurements at the Tevatron.

7.3.6 Dijet Resonance Search

We search for processes producing narrow resonances, X, decaying to dijets: $pp \to X \to$ jet + jet (inclusive). Our experimental motivation is that LHC is a parton-parton collider, and resonances made from partons must decay to the same partons giving two jets in the final state. The theoretical motivation is broad, since there are many models that predict narrow dijet resonances. In Figure 7.8 we show the cross section times branching ratio calculated to

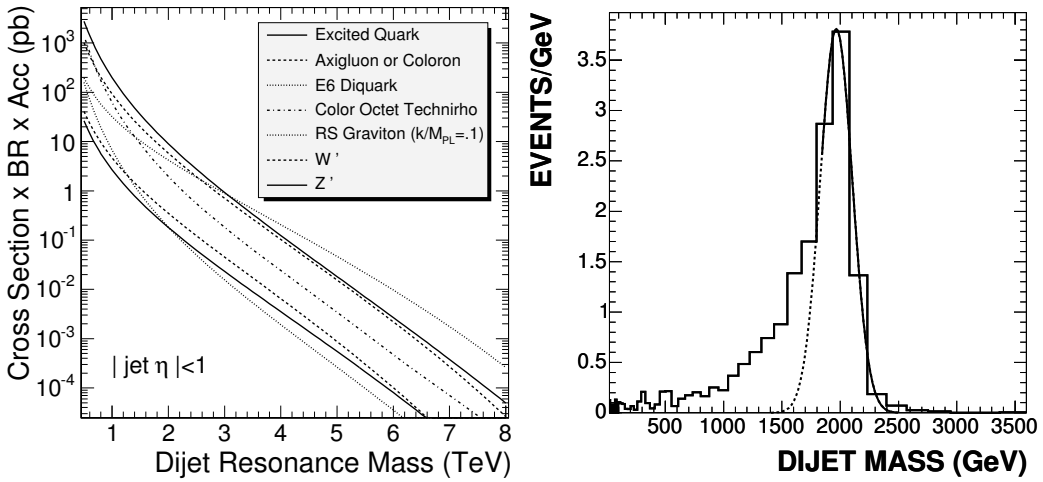

Figure 7.8: Left) The total cross section for dijet resonances from seven different models (see text). Right) The dijet mass distribution from a 2 TeV/c^2 Z' (histogram) is fit with a Gaussian (solid curve) from the peak region to high mass and the Gaussian is extended to lower mass (dashed curve).

lowest order for some models. Here we introduce them in order of descending cross section at low mass. Excited states of composite quarks [4] are strongly produced giving large cross sections ($qg \to q^*$). Axigluons [5] or colorons [6] from an additional color interaction are also strongly produced, but require an antiquark in the initial state ($q\bar{q} \to A$ or C) slightly reducing the cross section compared to excited quarks. Diquarks [7] from superstring inspired E_6 grand unified models are produced with electromagnetic coupling from the valence quarks of the proton ($ud \to D$). The cross section for E_6 diquarks at high mass is the largest of all the models considered, because at high parton momentum the probability of finding a quark in the proton is significantly larger than the probability of finding a gluon or antiquark. Color octet techirhos [8] from topcolor-assisted technicolor are produced for either gluons or quark-antiquark pairs in the initial state through a vector-dominance model of mixing between the gluon and the technirho ($q\bar{q}, gg \to g \to \rho_{T8}$). Randall Sundrum gravitons [9] from a model of large extra dimensions are produced with a significant cross section at masses below 1 TeV/c^2 primarily from gluons in the initial state ($q\bar{q}, gg \to G$). Heavy W bosons [10] inspired by left-right symmetric grand unified models have electroweak couplings and require antiquarks for their production($q_1\bar{q}_2 \to W'$), giving small cross sections. Heavy Z bosons [10] inspired by grand-unified models are widely anticipated by theorists,

but they are electroweakly produced, and require an antiquark in the initial state($q\bar{q} \rightarrow Z'$), so their production cross section is around the lowest of the models considered. Lower limits from CDF [11] and D0 [12] on the mass of these models are between 0.4 and 1.0 TeV/c^2.

7.3.6.1 Narrow Dijet Resonance Shapes

The simulated shape of a narrow dijet resonance in CMS is shown in Figure 7.8. The shape is composed of a Gaussian distribution from jet energy resolution and a long tail to low mass. The measured RMS of the Gaussian component is $\sigma/M = 0.045 + 1.3/\sqrt{M}$. The long tail to low mass comes predominately from final state QCD radiation (extra jets) which reduce the reconstructed mass. All resonances with a natural width significantly less than our resolution should look similar to this in the CMS detector. The model used in Figure 7.8 was a Z' from Pythia.

7.3.6.2 QCD Background

We consider the QCD background from the trigger for a luminosity of $L = 10^{33}$ and a sample size of 1 fb^{-1}. Figure 7.9 compares the Z' signal cross section to the QCD background, and also illustrates the statistical sensitivity for this signal. The vertical error bars on the QCD

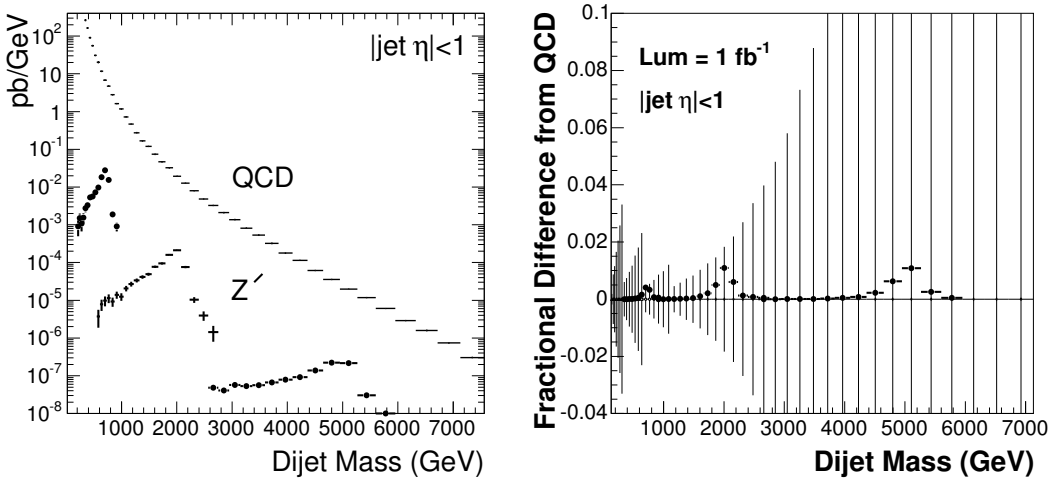

Figure 7.9: Left)The differential cross section as a function of dijet mass for the QCD background and three Z' signals with a mass of 0.7, 2, and 5 TeV/c^2. Right) The fractional difference between data and QCD as a function of dijet mass for 1 fb^{-1} of luminosity. The vertical lines are the statistical error on the QCD points, measured with four triggers with different prescales, and the points are Z' with input masses of 0.7, 2 and 5 TeV/c^2.

background rise as a function of mass in four distinct groups, because the four triggers begin at different masses and have different prescales. With 1 fb^{-1} it will be difficult to observe a Z' due to its small cross section and the limited statistics of the measurement. The differential cross section for the QCD background is well fit by a simple parameterization of the form

$$\frac{d\sigma}{dm} = \frac{p_0(1 - m/\sqrt{s})^{p_1}}{m^{p_2}} \tag{7.1}$$

where m is the dijet mass, $\sqrt{s} = 14000$ GeV/c^2 is the collision energy, and p_0, p_1, p_2 are arbitrary parameters. In the statistical estimators that follow, we use the fit to smooth away background fluctuations in our simulation samples that would distort our likelihoods. In a search with real data, a similar fit could be used to simply model the measured background, as was done by CDF [11], or a full NLO QCD calculation smeared with the jet resolution could be used to model the background, as was done by D0 [12].

7.3.6.3 Resonance Sensitivity Estimates

In Figure 7.10 we show examples of likelihoods for excluding or observing a narrow resonance signal on a QCD background as a function of the signal cross section. In the case where

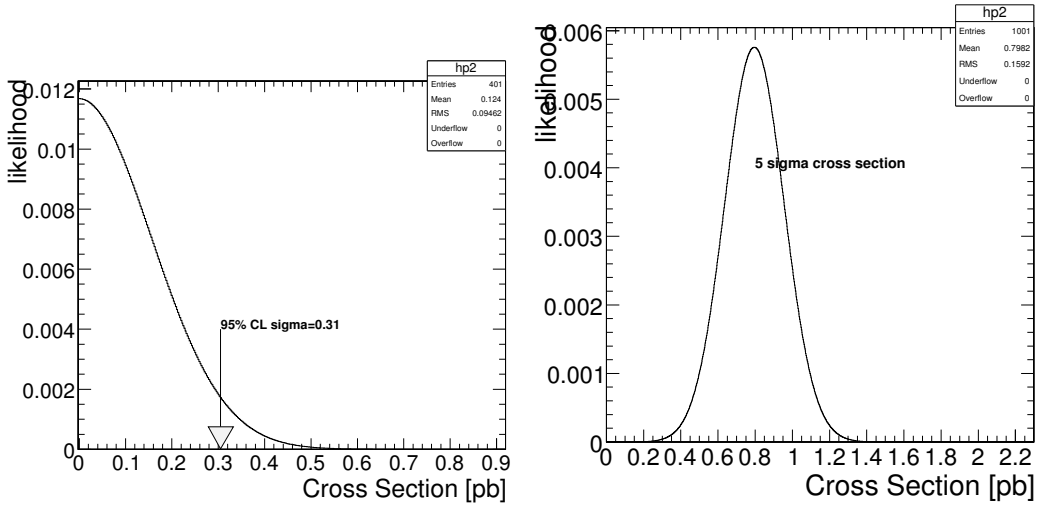

Figure 7.10: Likelihoods for observing a narrow dijet resonance in a 1 fb^{-1} data sample that contains only QCD background (left) and a data sample that also contains a resonance with a significance of 5σ (right).

the observed sample is QCD only, the signal likelihood peaks at zero cross section, and the 95% CL excluded signal cross section is shown. In the case where the observed sample is QCD plus a resonance signal, we have varied the signal size until the Gaussian distributed likelihood is 5 σ above zero.

In Figure 7.11 we compare the cross section for a 95% CL exclusion or a 5σ discovery to the size of the statistical error bars on QCD. From Figure 7.11 we read off the mass limits or discoveries that are possible with 1 fb^{-1} of data, which are tabulated in Table 7.5. We will be sensitive to strongly produced resonances up to many TeV. Further optimizations of this straightforward analysis are possible to attempt to increase our limited sensitivity to electroweakly produced resonances. The models discussed here are merely benchmarks in our search for new physics. The cross section for any other models can be compared with the cross section points in Figure 7.11 to determine CMS sensitivity to those models. We have discussed a generic search for all narrow dijet resonances, because we are preparing for unanticipated new physics at the energy scale of the LHC.

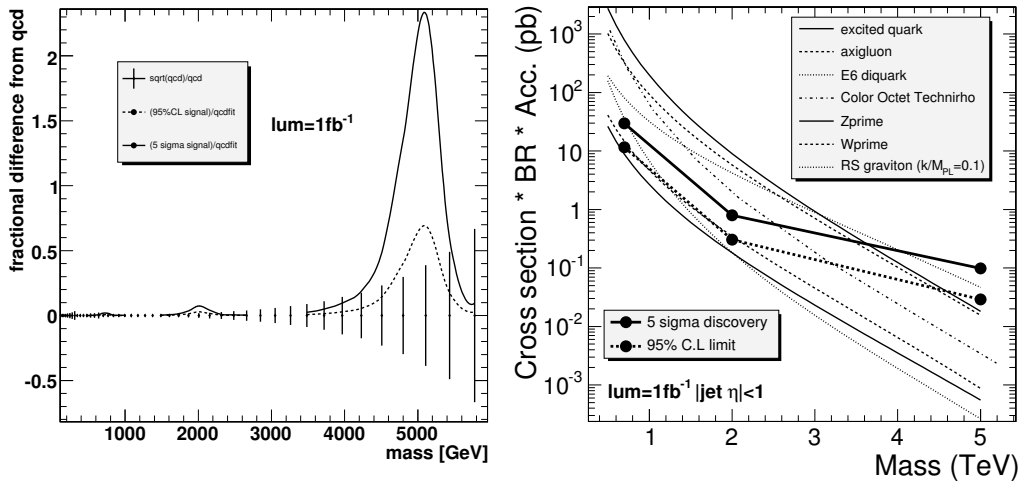

Figure 7.11: The cross section of a resonances signal that could be excluded at 95% CL (dashed) or discovered with 5σ significance (solid) for a sample of size 1 fb^{-1}. Left) As a fraction of the QCD background and compared to the QCD error bars. Right) Compared to the cross section for various resonance models.

Resonance Model	95% CL Excluded Mass (TeV/c^2)	5σ Discovered Mass (TeV/c^2)
Excited Quark	~ 4.5	~ 3.5
Axigluon or Coloron	~ 4.5	~ 3.5
E_6 diquark	~ 5	~ 4
Color Octet Technirho	~ 3	~ 2.4
Randall Sundrum Graviton	~ 1.1	~ 0.7
W′	~ 0.9	N/A
Z′	N/A	N/A

Table 7.5: Sensitivity to dijet resonances from 1 fb^{-1} of data. For each resonance model, we show the expected 95% CL lower limit on the resonances mass, and the highest mass for which a 5σ discovery is possible. All estimates are with statistical uncertainties only.

7.3.7 Contact Interaction Search

7.3.7.1 Compositeness Signal in Mass and Angular Distributions

Contact interactions, for example those that arise from composite quarks [13], produce a rise in rate relative to QCD at high dijet mass or high inclusive jet E_T. However, observation in the mass distribution alone requires precise understanding of QCD rate as a function of dijet mass. This is difficult due to significant jet energy scale uncertainties, which on a sharply falling QCD spectrum produce rate uncertainties roughly one order of magnitude larger. Parton distribution uncertainties may also be significant at high mass, as the LHC will be the first to observe parton distributions at momentum transfers of a TeV and more. Angular distributions benefit from much smaller systematic uncertainties. The contact interaction is often more isotropic than the QCD background, since QCD is dominated by t-channel scattering and produces jets predominantly in the forward direction. Our analysis will use a single number to measure the angular distribution at a given value of dijet mass, so that we can plot it versus dijet mass and see contact interactions as they emerge at high mass.

7.3.7.2 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 [14]. It is the simplest measure of the most sensitive part of the angular distribution as a function of mass. Troubling systematic uncertainties on the absolute scale of jet energy and on parton distributions roughly cancel in this ratio. In Figure 7.12 we show our lowest order calculation of the dijet ratio from QCD compared with a left-handed contact interaction among quarks **??** at three different values of the contact interaction scale. Lowest order QCD gives a fairly flat dijet ratio around $0.6 - 0.7$ while the

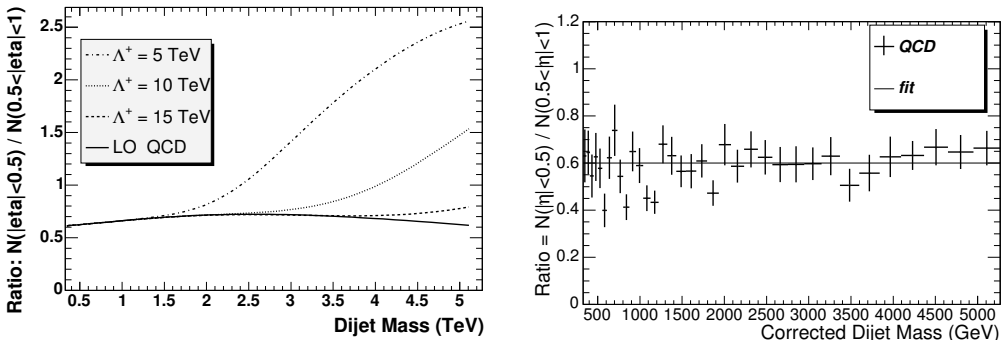

Figure 7.12: Left) A lowest order calculation of the dijet ratio from QCD (solid curve) is compared with QCD plus a quark contact interaction at a scale Λ^+ of 15 TeV (dashed), 10 TeV (dotted) and 5 TeV (dot-dashed). Right) The dijet ratio in the CMS simulation from QCD (points) is compared to the value 0.6 (line).

contact interactions produce an increase in the dijet ratio at high mass. Figure 7.12 also shows that a full CMS detector simulation of the dijet ratio from QCD, using the samples discussed in section 7.3.2, is indistinguishable from a flat ratio of 0.6 within the simulation statistical uncertainty. The agreement between the full simulation and the lowest order parton level

QCD calculation suggest that the dijet ratio is not very sensitive to the effects of QCD radiation, detector response, and jet reconstruction, demonstrating some of the advantages of measuring a ratio.

7.3.7.3 Contact Interaction Sensitivity Estimates

In Figure 7.13 we show the dijet ratio for QCD, estimated at 0.6 from the fit to the full simulation, along with the statistical uncertainties expected with 1 fb^{-1} and our jet triggers. The uncertainties are calculated using Poisson statistics at high dijet mass, where few events

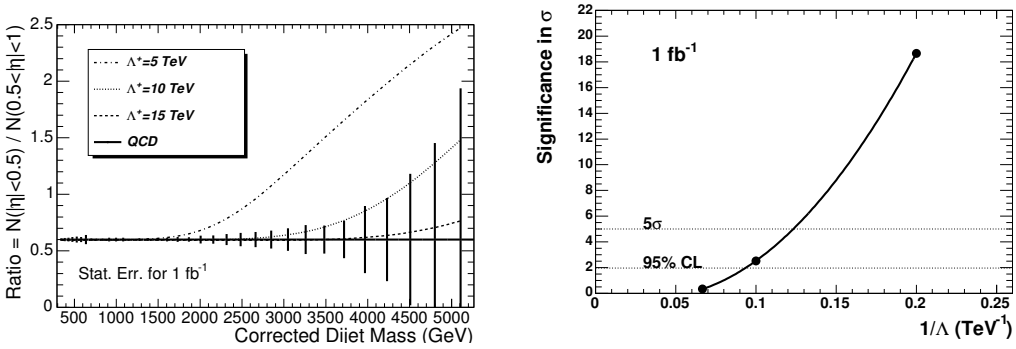

Figure 7.13: Left) The expected value and statistical error of the dijet ratio of QCD in the CMS detector for 1 fb^{-1} (solid) is compared with QCD plus a quark contact interaction at a scale Λ^+ of 15 TeV (dashed), 10 TeV (dotted) and 5 TeV (dot-dashed). Right) The statistical significance of the difference between QCD alone and QCD plus a quark contact interaction (points) is plotted vs $1/\Lambda^+$ and fit with a quadratic function (solid curve). Horizontal lines show the 5σ and 95% CL levels.

are expected and Gaussian statistics is less accurate. The signal in Figure 7.13 is estimated by scaling the lowest order contact interaction calculation by the ratio of our full simulation prediction for QCD to the lowest order QCD calculation: signal = contact $\times$ 0.6 / QCD. The calculated chisquared between QCD and the contact interaction signal is 348 for $\Lambda^+ = 5$ TeV, 6.4 for $\Lambda^+ = 10$ TeV and 0.1 for $\Lambda^+ = 15$ TeV. In Figure 7.13 we show the significance in σ, estimated as $\sqrt{\chi^2}$, compared to a fit as a quadratic function of $1/\Lambda^+$. From this we estimate, that with 1 fb^{-1} of data and including only statistical uncertainties, CMS can either discover quark contact interactions up to a scale of $\Lambda^+ = 8.1$ TeV (5σ significance) or exclude it up to a scale of $\Lambda^+ = 10.7$ TeV (95% CL). This is four times the sensitivity of the D0 search($\Lambda^+ > 2.7$ TeV at 95% CL)[14], more than double our mass sensitivity to excited states of composite quarks in table 7.5, and is equivalent to observing or excluding a quark radius of order 10^{-18} cm.