

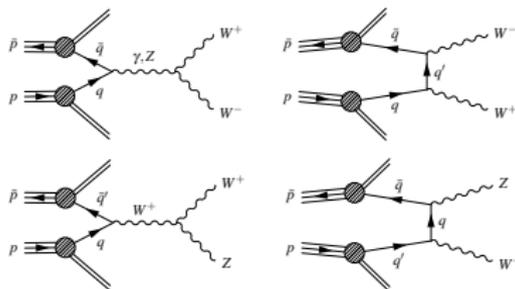
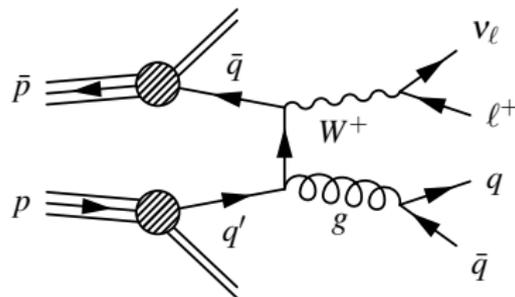
*Invariant Mass Distribution of Jet Pairs Produced in
Association
with a W boson in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*

Viviana Cavaliere for the CDF Collaboration

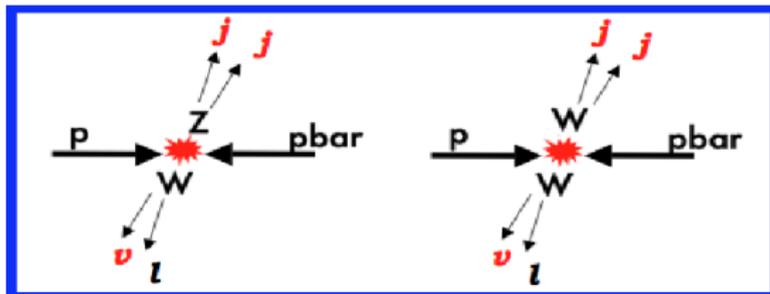
University of Illinois at Urbana-Champaign

04/06/2011





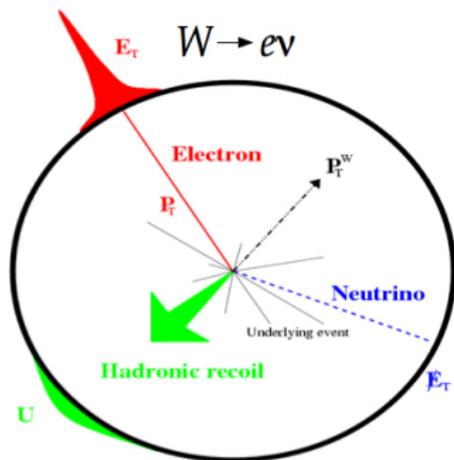
- Measurements of associate production of a W boson and jets are a fundamental test of the Standard Model
- $\ell\nu$ +jets signature shared by several important processes
 - 1 $WW + WZ$, single top, WH
- Essential starting point for physics beyond the SM
- Important to understand the tools used in high energy physics



- Diboson represent a Standard Model reference that can be used as a starting point for searches.
- WW, WZ, ZZ observed in $\cancel{E}_T + \text{jets}$ mode at CDF in 2009 PRL 103, 091803 (2009)
- WW, WZ observed in $\ell\nu + \text{jets}$ final state. Two analysis:
 - ① Using matrix elements technique
 - ② Looking for the $W \rightarrow jj$ peak in the dijet invariant mass
- Phys. Rev. Lett. 104, 101801 (2010)
- Results have been updated with 4.3 fb^{-1}

- Trigger on the charged lepton
- Require presence of neutrino: undetected neutrino manifests as an imbalance in transverse momentum: “missing” transverse energy

$$\vec{E}_T = - \sum_i \vec{E}_T^i$$



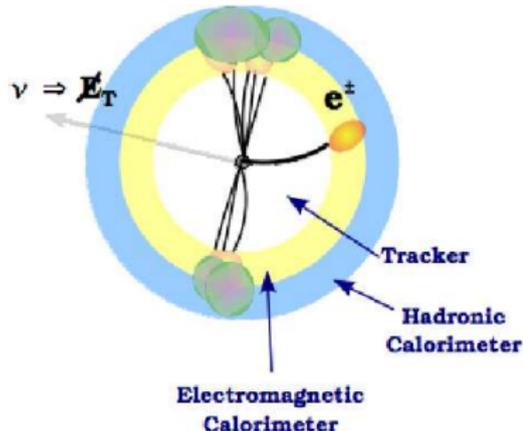
- Get a sample of inclusive W
- Look for at least two jets and form the invariant mass
- Fit the expected W/Z and background dijet mass shape to data to extract the diboson fraction



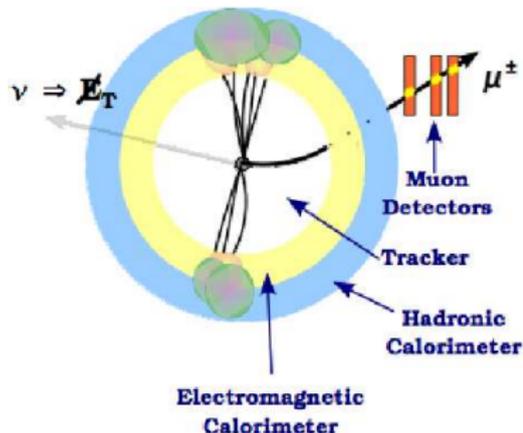
- The analysis uses 4.3 fb^{-1} of data collected by CDF.



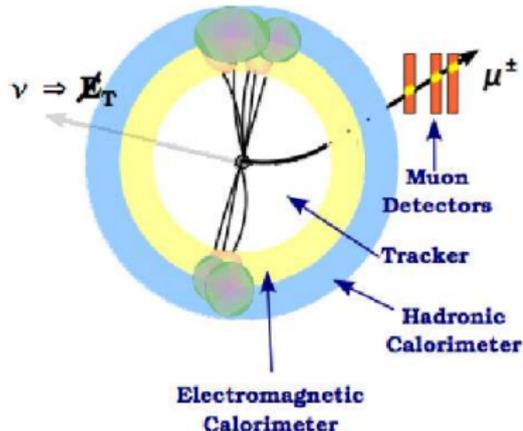
- The analysis uses 4.3 fb^{-1} of data collected by CDF.
- Electrons $\rightarrow E_T > 20 \text{ GeV}/c^2$ (GeV/c) and $|\eta| < 1.0$
 - 1 Require calorimeter showers consistent with electromagnetic interactions
 - 2 Require that 90% of energy is deposited in the EM calorimeter



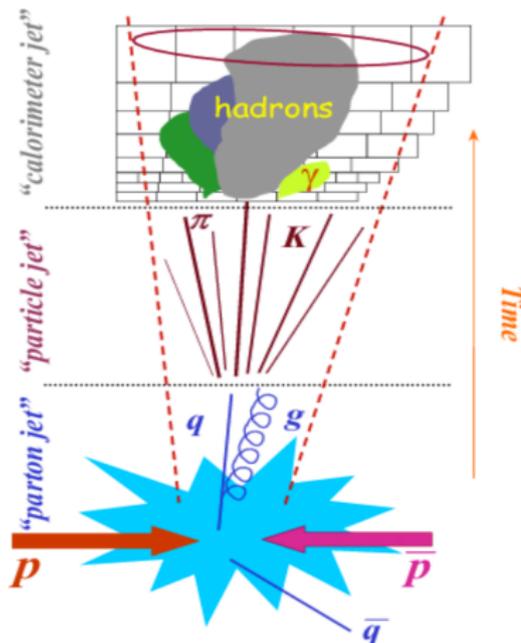
- The analysis uses 4.3 fb^{-1} of data collected by CDF.
- Electrons $\rightarrow E_T > 20 \text{ GeV}/c^2$ (GeV/c) and $|\eta| < 1.0$
 - 1 Require calorimeter showers consistent with electromagnetic interactions
 - 2 Require that 90% of energy is deposited in the EM calorimeter
- Muons $\rightarrow p_T > 20 \text{ GeV}/c^2$ (GeV/c) and $|\eta| < 1.0$
 - 1 Require high quality track and good matching between the track and the hit in the muon chambers



- The analysis uses 4.3 fb^{-1} of data collected by CDF.
- Electrons $\rightarrow E_T > 20 \text{ GeV}/c^2$ (GeV/c) and $|\eta| < 1.0$
 - 1 Require calorimeter showers consistent with electromagnetic interactions
 - 2 Require that 90% of energy is deposited in the EM calorimeter
- Muons $\rightarrow p_T > 20 \text{ GeV}/c^2$ (GeV/c) and $|\eta| < 1.0$
 - 1 Require high quality track and good matching between the track and the hit in the muon chambers
- Both are required to be isolated to reject leptons from semi-leptonic decays of heavy flavor hadrons
- We further require $E_T > 25 \text{ GeV}$ and $M_T^W > 30 \text{ GeV}$ to ensure the presence of a real W

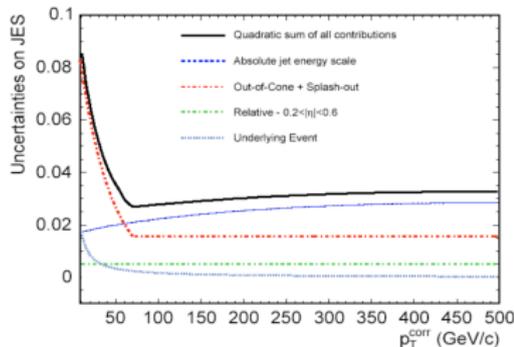


- Jets are selected using the standard CDF JETCLU algorithm
- Cluster energy in cones of $\Delta R < 0.4$
- Calorimeter signature must be inconsistent with electron signatures
- Select exactly two jets with $E_T > 20 \text{ GeV}$ (corrected for detector inefficiencies) and $|\eta| < 2.4$
 - Jet Energy scale known at 3% level
 - Independent check of the scale: W from top \rightarrow [pubic webpage](#)
- Require $p_{T,jj} > 40 \text{ GeV}/c$ \rightarrow in order to improve data/MC modeling



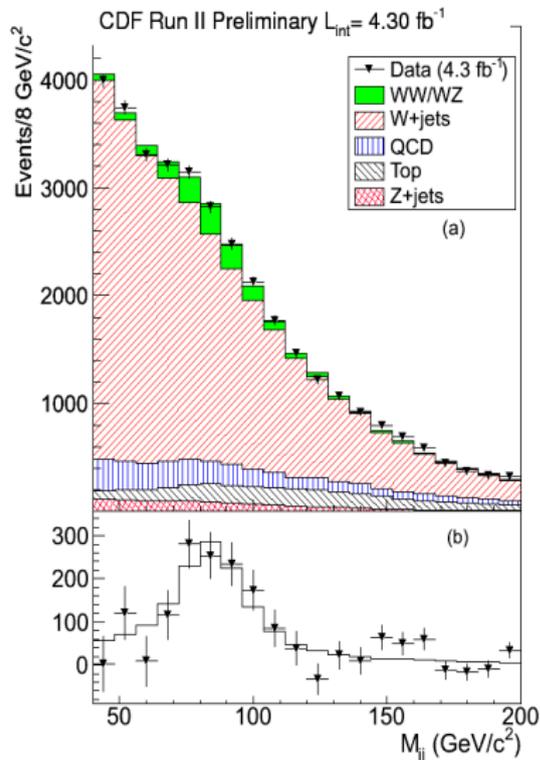


- Jets are selected using the standard CDF JETCLU algorithm
- Cluster energy in cones of $\Delta R < 0.4$
- Calorimeter signature must be inconsistent with electron signatures
- Select exactly two jets with $E_T > 20$ GeV (corrected for detector inefficiencies) and $|\eta| < 2.4$
 - Jet Energy scale known at 3% level
 - Independent check of the scale: W from top \rightarrow [pubic webpage](#)
- Require $p_{T,jj} > 40$ GeV/c \rightarrow in order to improve data/MC modeling





- $W \rightarrow l\nu + \text{jets}$ ($l = e, \mu, \tau$):
 - same signature as signal with a much higher cross section (2066 pb)
 - $\sim 80\%$ of the sample
- $Z \rightarrow ll + \text{jets}$ ($l = e, \mu, \tau$):
 - where one of the two leptons escapes detection and produces \cancel{E}_T
 - cross section 187 pb
- $t\bar{t} + \text{single top}$:
 - final state similar to signal with at least one real W and two jets.
 - $\sigma(t\bar{t}) = 7.5$ pb and $\sigma(\text{single top}) = 2.9$ pb (assuming a mass of 172.5 GeV/c²)
- QCD Multijet:
 - events without a real high p_T lepton
 - e.g a three-jet event in which one jet passes all lepton cuts and, simultaneously, the energies are so badly measured that a large \cancel{E}_T is reported.
 - probability for a jet to mimic a lepton is very small, but QCD processes have very large cross sections
 - estimated from data using orthogonal selection



- Use a fit to dijet mass to extrapolate the WW/WZ contribution
- We estimate $1582 \pm 275(\text{stat.}) \pm 107(\text{syst.})$ events for a significance of 5.2σ .

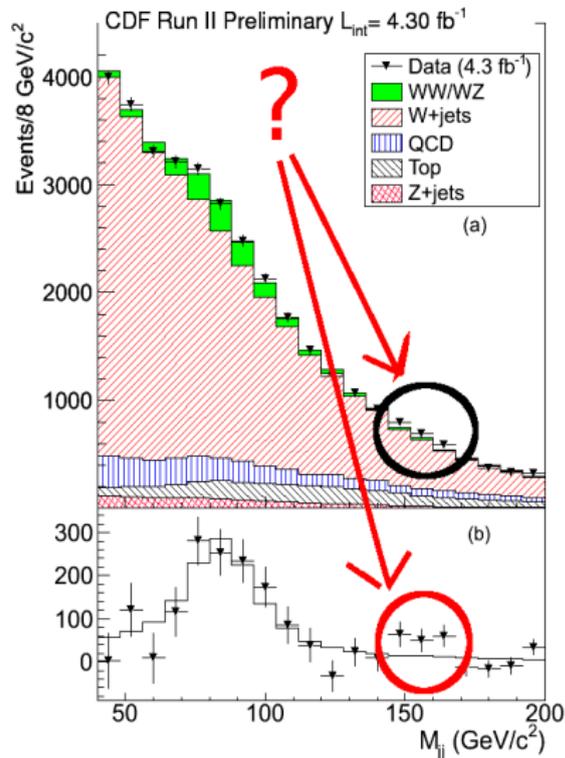
- The resulting cross section is

$$\sigma(WW/WZ) = 18.1 \pm 3.3(\text{stat.}) \pm 2.5(\text{syst.}) \text{ pb}$$

that is in agreement with SM expectation ($15.9 \pm 0.9 \text{ pb}$).



- CDF established Diboson production with jets in the final state
- We can safely state that we are able to observe dijet resonances
- Moreover in the $WW/WZ \rightarrow l\nu + \text{jets}$ analysis we learned:
 - 1 $W + \text{jets}$ normalization compatible with expectations scaled to NLO
 - 2 Jet Energy Scale well under control \rightarrow multiplicative factor to correct diboson position is found to be compatible with 1
 - 3 Our mass fit is sensitive to $t\bar{t}$ + single top with an uncertainty of 10% \rightarrow cross section compatible with expectation



- Observed an interesting discrepancy in the M_{jj}
- Needed work in order to understand the nature and properties of the excess
 - 1 an artifact of background subtraction
 - 2 misunderstanding of one of the backgrounds
 - 3 real physics?



- Using exactly the same kinematical cuts as the diboson analysis but:
- We require both jets to have $E_T > 30$ GeV
 - ① Energetic jets are measured with better accuracy.
 - ② Modeling in this region is expected to be more accurate
 - ③ A possible heavier particle would be characterized by more energetic jets
- Sample modeling using same processes with different relative contribution
- All cuts chosen “a priori”

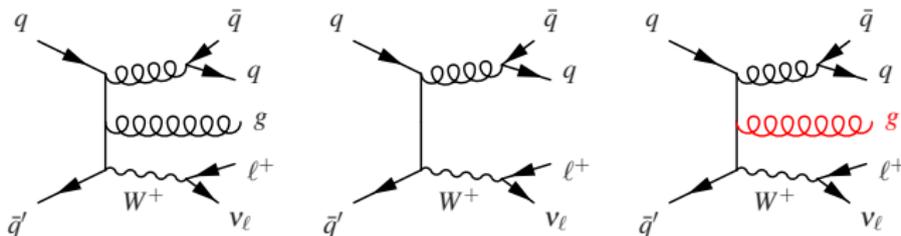


Process	Model	σ (pb)
WW/WZ inclusive	PYTHIA	15.9 ± 0.9
$Z \rightarrow e, \mu, \tau$ +jets	ALPGEN+ PYTHIA	787 ± 85
$t\bar{t}$	PYTHIA	7.5 ± 0.83
single top	MADEVENT + PYTHIA	2.86 ± 0.36
W+jets	ALPGEN+ PYTHIA	from data
QCD multijet	from data	from data

- Correct MC for:
 - Trigger Efficiencies: Data must pass the trigger to be selected \rightarrow apply these efficiencies to the MC
 - Lepton Energy Scale, Energy Resolution, and Identification.
 - Luminosity Profile: not the same as for the data \rightarrow reweight as a function of number of vertices



- While simulating $W/Z + N$ -jets, we need to get the inclusive cross section and the relative cross section for exclusive N -jets
- We simulate by pairing AlpGen (LO matrix elements) and Pythia (parton showers)
 - 1 Matrix Elements: Fixed order. Gives an accurate description of the hard process. Needed for N -jet description
 - 2 Parton showers: Needed for a realistic description of the final state in the detector
- Combine them using MLM scheme to avoid double counting.
- Cluster the showered partons into cone jets. Keep events only if each jet is matched to just one parton



ALPGEN produces events from diagrams (a) and (b). PYTHIA's showering routine will sometimes take events from diagram (b) and add a radiated gluon to produce diagram (c)

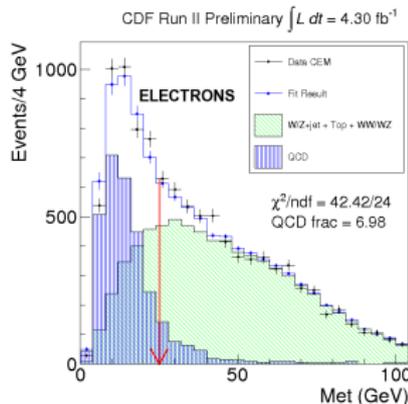
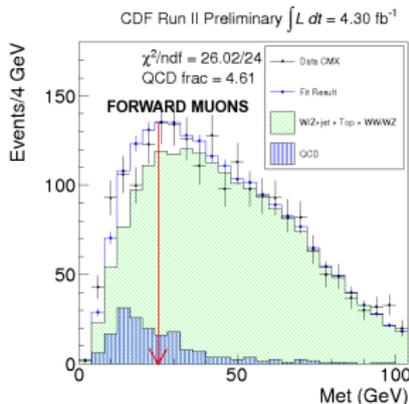
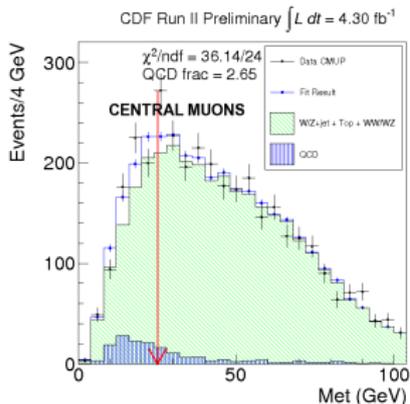


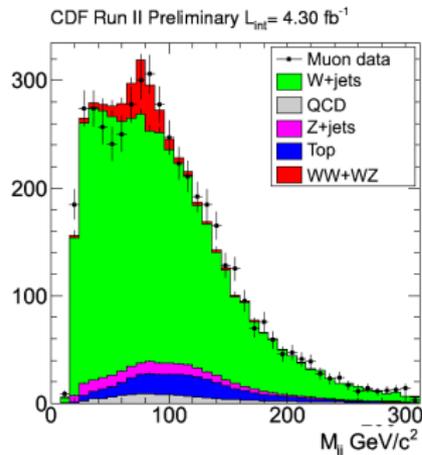
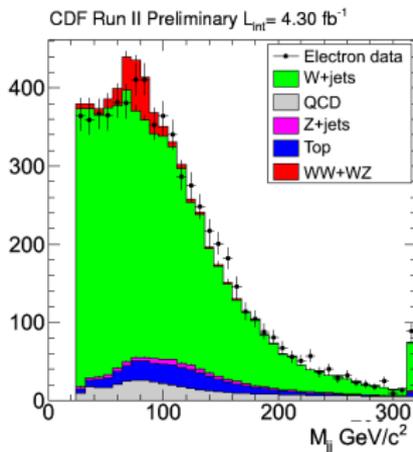
Modeled using data sidebands

- **Non isolated muons:**
 - Using non-isolated events, events which **pass all selection criteria except the requirement of lepton isolation.**
 - Based on the rationale that non-isolated events are typically leptons contained in jets, and jets that contain energetic leptons are mostly non- W events.
- **“AntiElectrons”:**
 - Some non-kinematic cuts for the electron are used to reject fake electrons.
 - Model is constructed with events which **fail at least two of the non-kinematic quality cuts** but pass all the kinematic cuts of the electron.

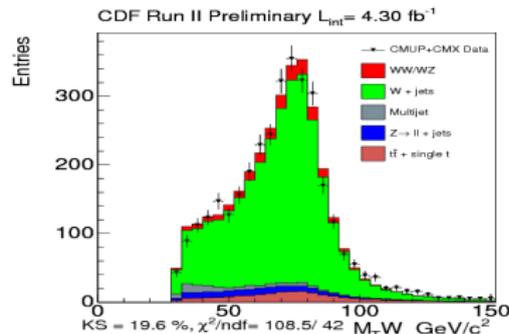
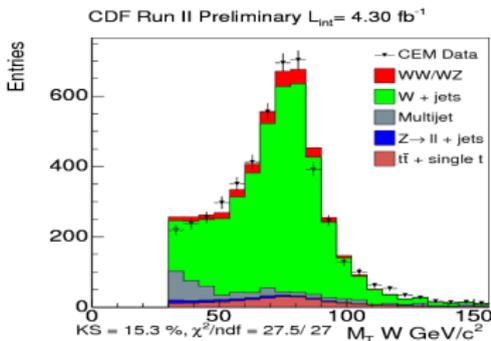
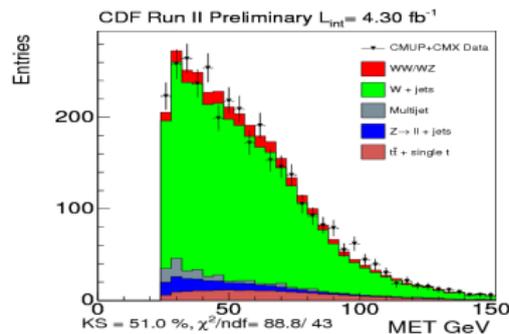
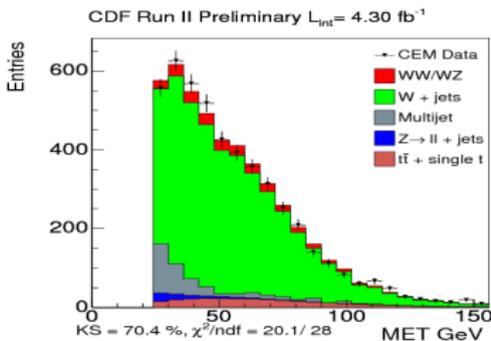


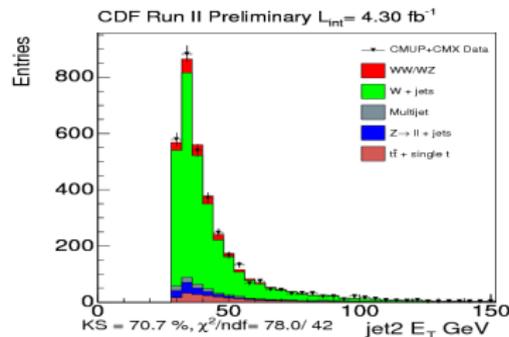
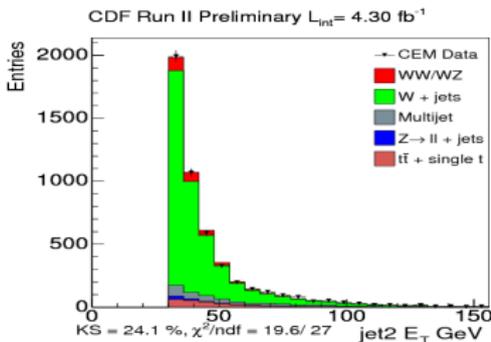
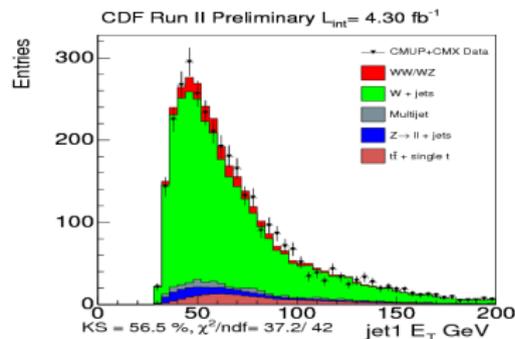
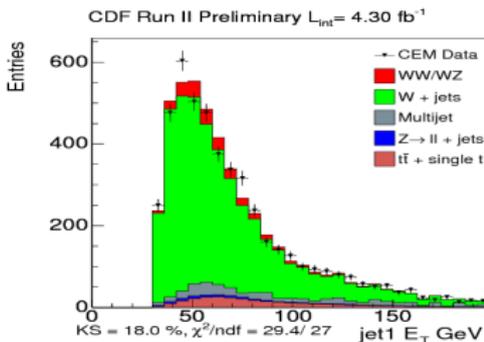
- QCD multijet events are characterized by low \cancel{E}_T , so \cancel{E}_T distribution is completely different from $W + jets$
- Best solution \rightarrow Fit the \cancel{E}_T distribution on data in order to constrain multijet normalization.
- Extract the fraction of QCD and knowing all the others contributions can extract also a preliminary $W + jets$ normalization (left completely free in the final fit)
- Systematic associated with the normalization estimated using different models (25%)

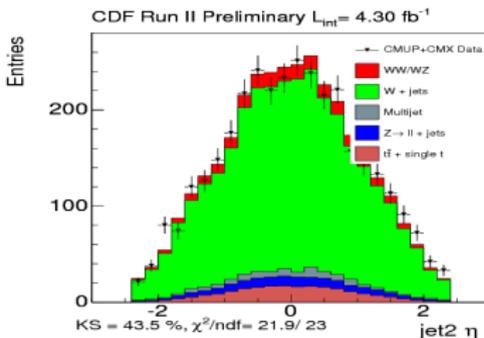
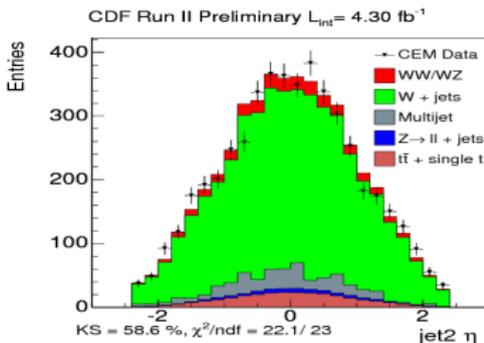
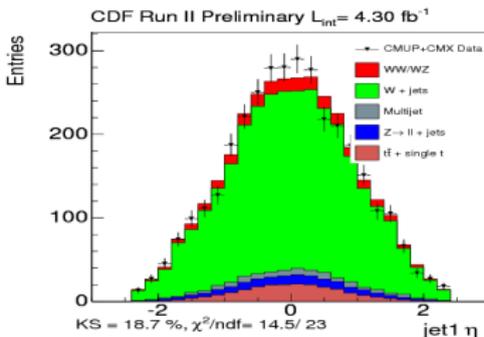
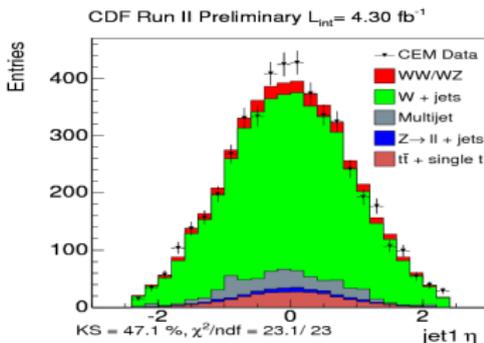


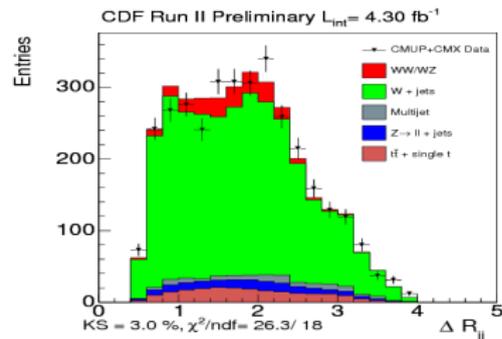
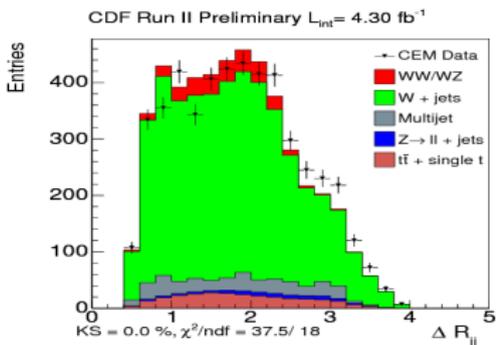
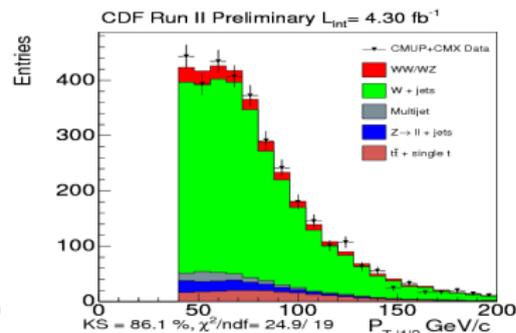
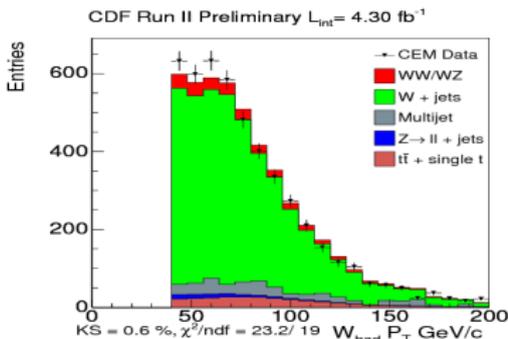


- 1 M_{jj} of data compared to SM predictions
- 2 Want to investigate further \rightarrow Need to check every background carefully. Public Webpage
 - 1 Look at Excess sidebands ($115 > M_{jj} > 175 \text{ GeV}^2$)
 - 2 QCD multijet shape modeling
 - 3 $t\bar{t}$ production
 - 4 $W + jet$ shape

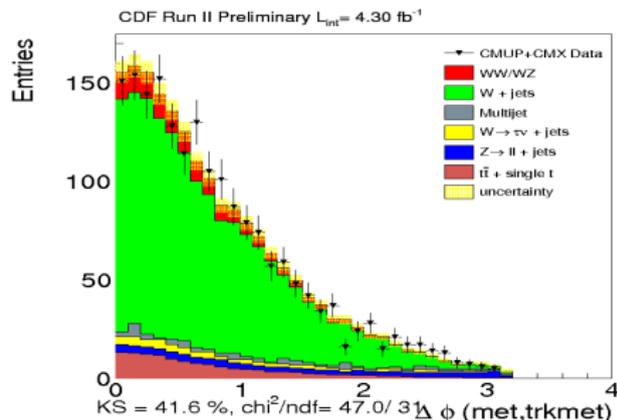
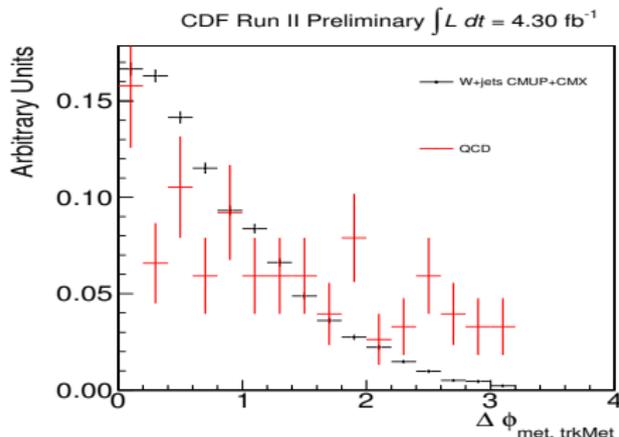
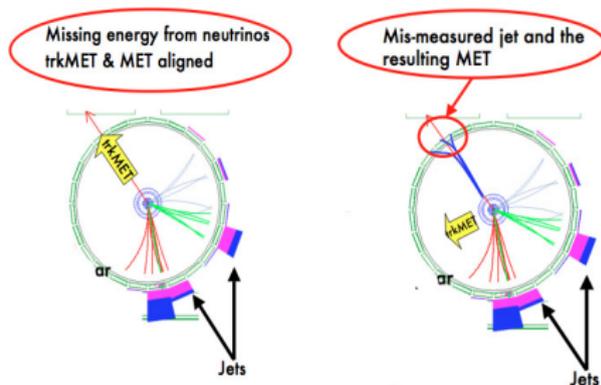






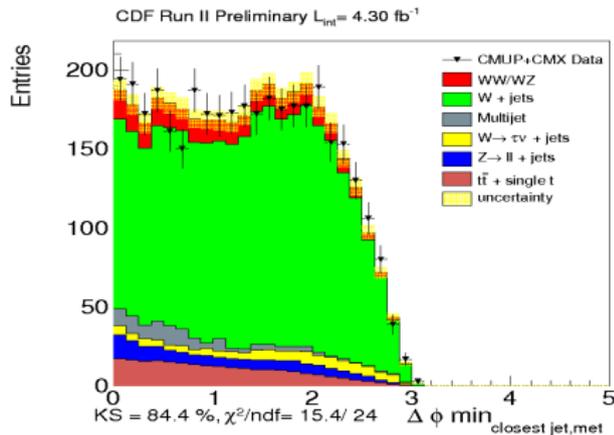
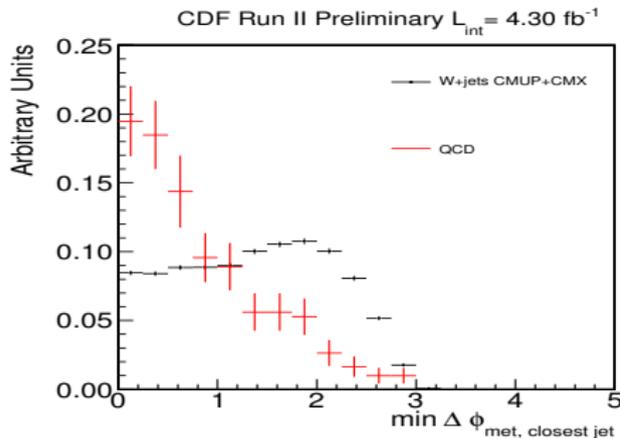


- Check shape and rate using independent ways:
 - 1 $\Delta\phi(E_T, \text{Missing Pt}) \rightarrow$ large for qcd events



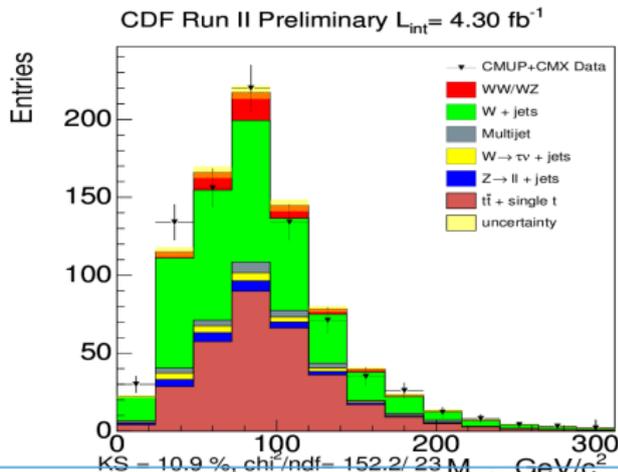
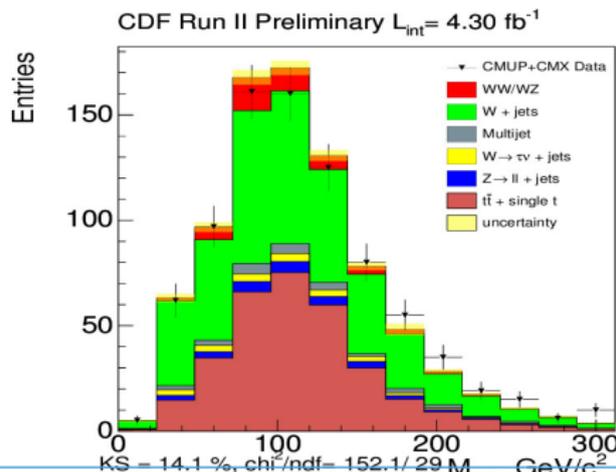


- Check shape and rate using independent ways:
 - $\Delta\phi(\cancel{E}_T, \text{Missing Pt}) \rightarrow$ large for qcd events
 - $\Delta\phi(\cancel{E}_T, \text{closest jet}) \rightarrow$ mis-measured jets tend to be aligned with the \cancel{E}_T



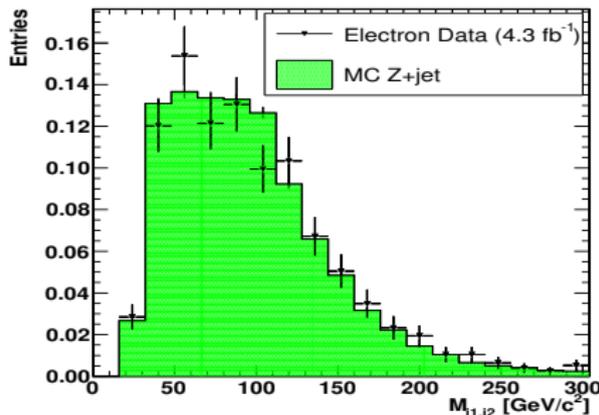
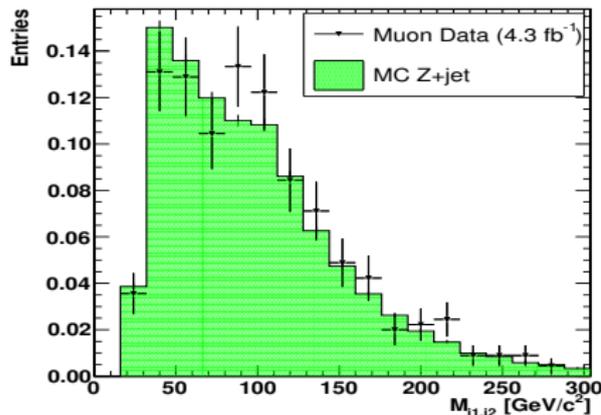
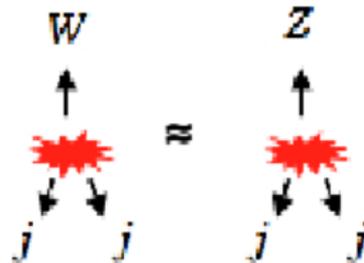


- 1 $t\bar{t}$ production is characterized by a larger number of jets with respect to other processes.
- 2 To test it, we exploit its higher jet multiplicity and look at events with at least 3 jets and events with at least 4 jets.
 - No mismodeling is observed
- 3 Furthermore we consider events with exactly three jets passing selection criteria and we look at the distribution of the invariant mass system of the other combinations of jets, $M_{j1,j3}$ and $M_{j2,j3}$.



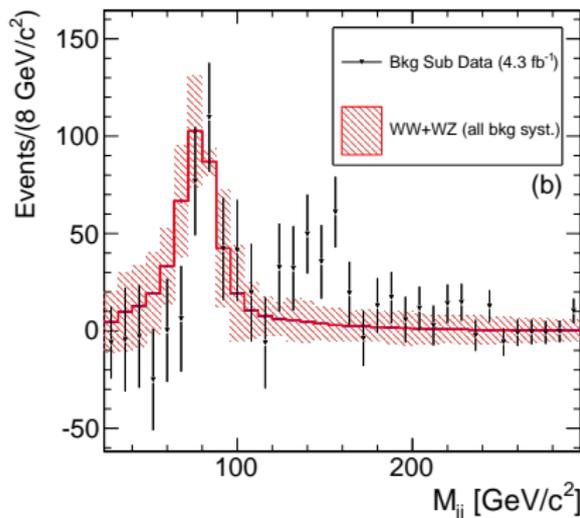
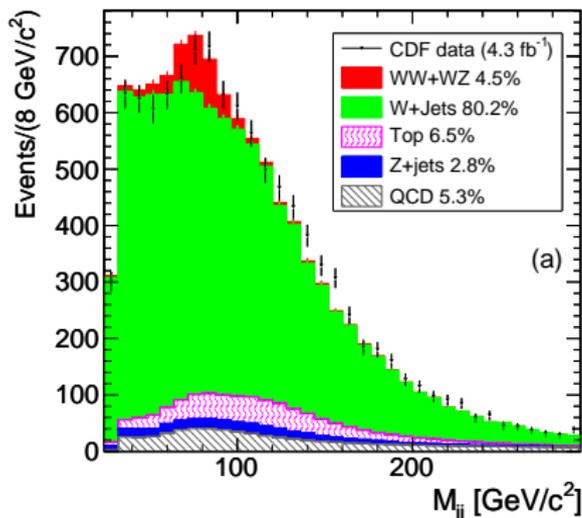


- Use $Z + jets$ data to check W+jets shape
- Require two leptons (one tight and one loose)
- Selection on the hadronic side is the same
- Basic Idea : Similar kinematics
- Due the purity of the sample can be modeled by ALPGEN only
- Problem: 10 times less data





- Combined χ^2 fit to the dijet mass distribution in electron and muon samples.
- 5 templates:
 - 1 $W + \text{jets}$ (unconstrained, normalization determined from the fit)
 - 2 QCD (normalization constrained to its fraction with 25 % error)
 - 3 $Z + \text{jets}$ (normalization constrained to the measured cross section)
 - 4 top & single top (normalization constrained to the theoretical cross section)
 - 5 WW/WZ (normalization constrained to the theoretical cross section)

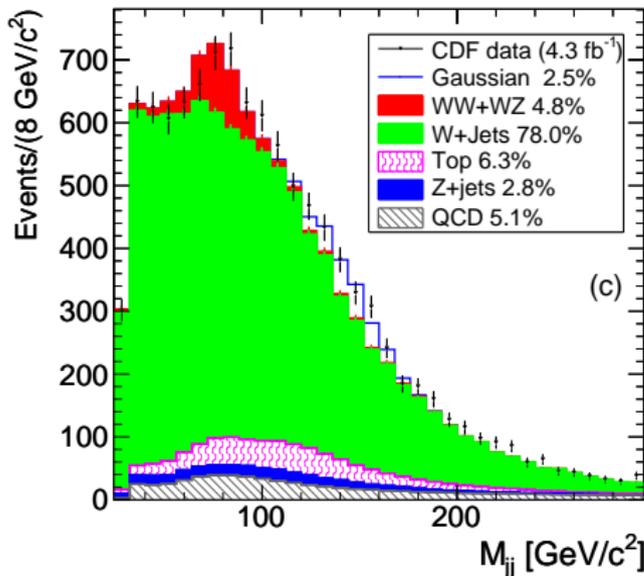




- Chosen to estimate the significance of the excess assuming an additional gaussian component.
- The gaussian assumption is a simplified model, since the exact shape would depend on the specific physics process and the heavy flavor content of the decay products.
- Since the excess looks narrow with respect to the detector resolution, we search for a peak compatible with the detector resolution for a given dijet mass value.

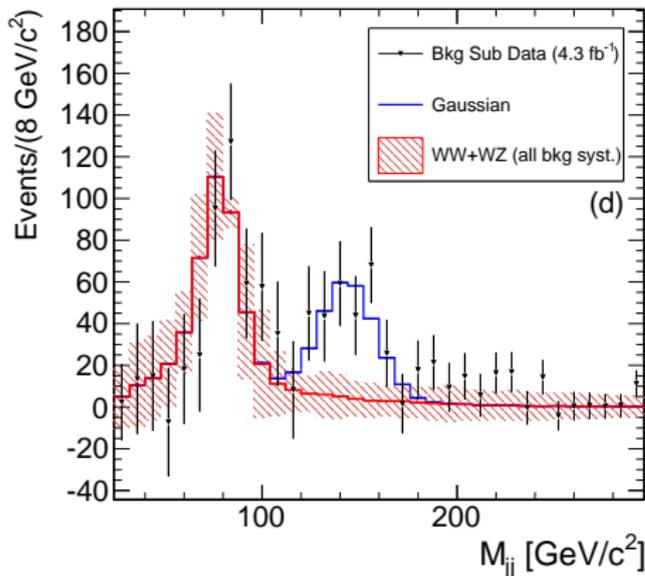
$$\sigma_{gaussian} = \sigma_W \sqrt{\frac{M_{jj}}{M_W}} = 14.3 GeV$$

- Procedure:
 - 1 Fit the data without the gaussian \rightarrow evaluate χ^2
 - 2 Fit the data with the gaussian \rightarrow evaluate χ^2
 - 3 We add 3 degrees of freedom to the fit (mass, separate e/μ yields) so the $\Delta\chi^2$ should have the distribution of a χ^2 with 3 degrees of freedom.
 - 4 Verify the behaviour of the $\Delta\chi^2$ with statistical trials with trial factor.



- Data fitted with SM templates plus a gaussian.
- $\Delta\chi^2$ observed 20.31 that corresponds to a statistical significance of 3.7σ (including trial factor)

	Electrons	Muons
Excess events	156 ± 42	97 ± 38
Excess events / expected diboson	0.60 ± 0.18	0.44 ± 0.18
Mean of the Gaussian component	$144 \pm 5 \text{ GeV}/c^2$	



- Data fitted with SM templates plus a gaussian.
- $\Delta\chi^2$ observed 20.31 that corresponds to a statistical significance of 3.7σ (including trial factor)

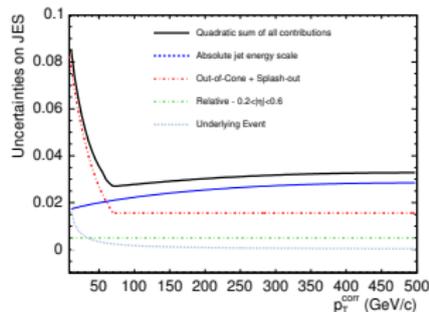
	Electrons	Muons
Excess events	156 ± 42	97 ± 38
Excess events / expected diboson	0.60 ± 0.18	0.44 ± 0.18
Mean of the Gaussian component	$144 \pm 5 \text{ GeV}/c^2$	



- systematics affecting **background shapes**
- Evaluated generating statistical trials with the varied templates and fitting with the standard ones.

Affected Quantity	Source	Uncertainty (%)
Number of Excess Events	QCD shape	± 1.9
	Q^2	± 6.7
	JES	± 6.1
	Total	± 9.3

- Measurement affected by Jet Energy Scale:
 - 1 Apply to all MC modeled processes at the same time
- QCD shape systematic evaluated using different Isolation ranges.





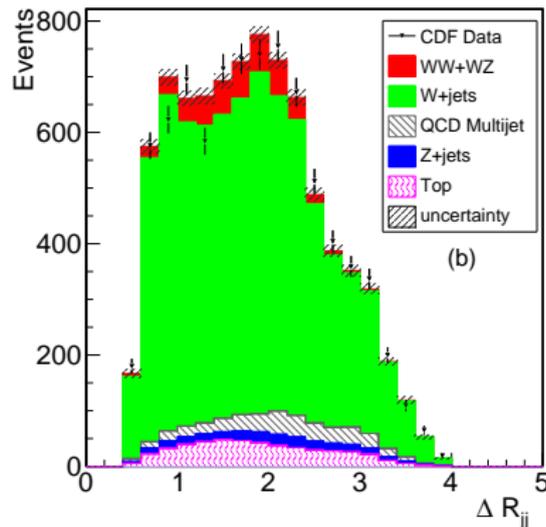
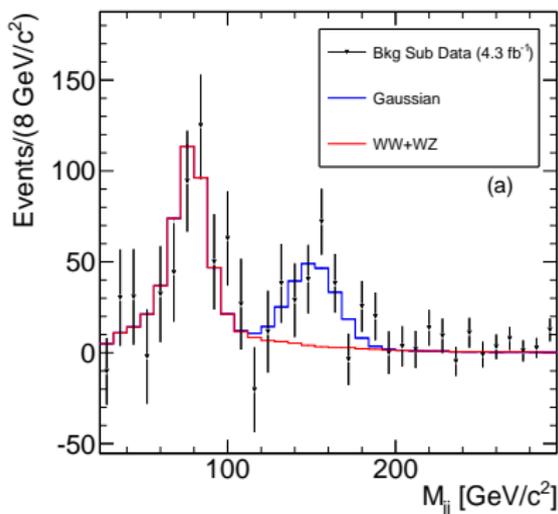
- Alpgen MC depends on some parameters:
 - 1 **Factorization and normalization scale** $Q^2 = M_W^2 + \sum p_{T,j}^2$ which can be varied by a constant factor on an event by event basis
 - 2 **k_T Scale Factor**: Alpgen's scale factor for α_s at each decay vertex.
 - 3 **Parton matching cluster p_T threshold**: the minimum p_T for jet clusters that are used for matching procedure.
 - 4 **Parton matching clustering radius size**: the size of the jet cone used when creating jet clusters for matching procedure \rightarrow Alpgen authors recommend using the generator level cut.
- The only significant effect for this kind of selection is given by the Q^2
- Use standard CDF procedure:
 - Double and halve it to obtain alternative templates
 - This choice is motivated by standard practice based on extensive theoretical work



- To evaluate the significance of the resonance, we apply a procedure called “supremum p-value”
- If N_{syst} is the number of systematics sources, we generate a toy MC sample for each combination of the N_{syst} , i.e. in each sample, some of the systematics are varied.
- For each sample, we evaluate the corresponding p-value using the $\Delta\chi^2$ between the background only and signal hypothesis as test statistics.
- The significance we quote for our final result is the largest among the p-values we obtain.
- To take into account the trial factor, in our toy experiments we scan the mass of the resonance in the search region [120 - 200] GeV/ c^2 using step of 4 GeV/ c^2 and evaluate, at each step, the corresponding χ^2 : for each toy sample, the minimum χ^2 of the scan is used in the $\Delta\chi^2$ evaluation.
- The largest p-value is 7.6×10^{-4} , corresponding to a significance of **3.2 standard deviations**.



- In order to investigate possible mismodeling we consider two control regions, the first defined by events with $M_{jj} < 115$ and $M_{jj} > 175$ GeV/ c^2 and the second defined by events with $p_T < 40$ GeV/ c .
- We use these regions to derive a correction as a function of ΔR_{jj} to reweigh the events in the excess region.
- The reweighings change the statistical significance of the result by plus or minus one sigma.
- However, the ΔR_{jj} distribution is strongly correlated to M_{jj} and the control regions both have significantly different distributions of ΔR_{jj} .
- Reweighting our W +jets sample may be correct ΔR_{jj} mismodeling or may introduce bias in the M_{jj} distribution.
- In addition, the ΔR_{jj} distribution is consistent within the one sigma variation of the systematic uncertainties for events outside the excess mass region.

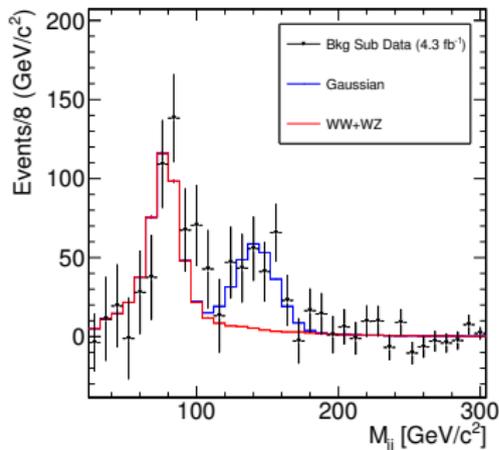
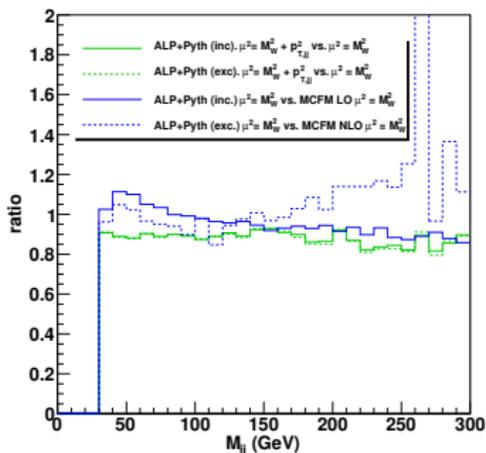


- Fit performed combination of systematic that fits data best: lowest χ^2
- Returns a p-value intermediate between the largest and statistical only
- ΔR_{jj} shown with the same combination of systematic. $\chi^2/ndf=26.6/18$.



- In order to test Next to Leading Order contributions to the $W+2$ partons prediction, we compare (private communication with J.Campbell, E. Eichten, K.Lane, A.Martin) ALPGEN and interfaced to PYTHIA for showering to a sample of $W+2$ partons simulated using the MCFM.
- We extract a correction as a function of M_{jj} that is applied to the ALPGEN + PYTHIA sample used in our background model.
- This procedure returns a statistical significance of 3.4σ .

W+jj : comparing different μ , generators





- We increase the jet E_T threshold in steps of 5 GeV and check the fraction of excess events that are selected as function of the jet E_T .
- The result is compatible with expectation from a Monte Carlo simulation of a W boson plus a particle with a mass of $150 \text{ GeV}/c^2$ and decaying isotropically into two jets.
- In this model, we estimate a cross section times the particle branching ratio into dijets of the order of 4 pb \rightarrow not compatible with Standard Model WH
 $\sigma \times BR(b\bar{b}) = 39 \text{ fb}$



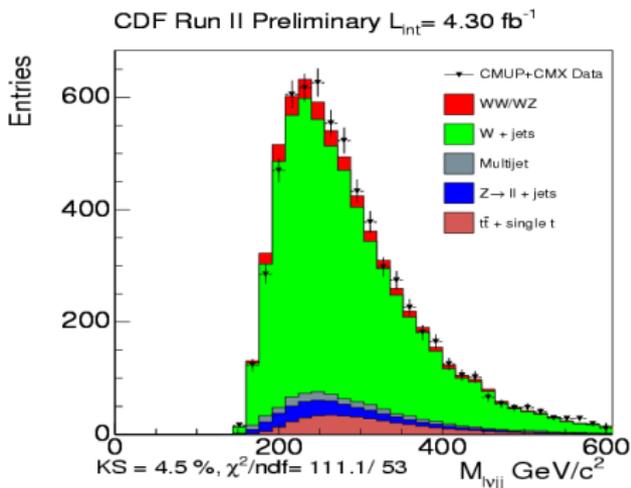
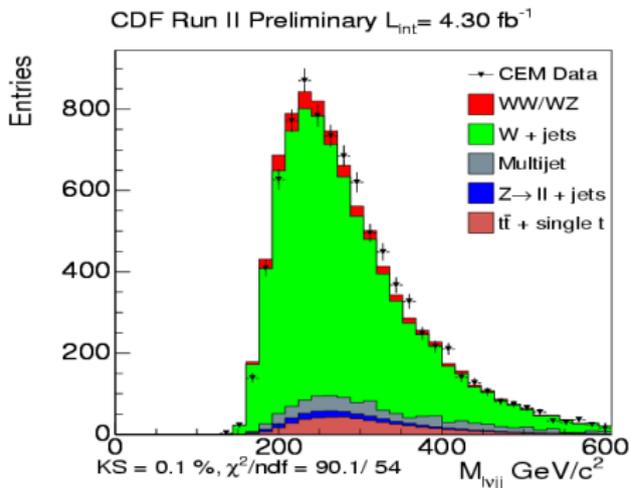
- Tried to look at the flavour composition of this jets.
 - Assuming a branching ratio in $b\bar{b}$ of 100% and the tagging efficiency from WH we expect to be able to see a much stronger signal in the tagged sample
- Compared the **fraction of events with b-jets** in the sidebands ($120 > M_{jj} > 160 \text{ GeV}/c^2$) to that in the excess region

Tag requirement	Excess region	Sideband region
Muons		
1 tag	0.1027 ± 0.0112	0.0813 ± 0.0096
2 tag	0.0078 ± 0.0030	0.0084 ± 0.0030
Electrons		
1 tag	0.0897 ± 0.0088	0.0945 ± 0.0087
2 tag	0.0110 ± 0.0030	0.0095 ± 0.0026

No significant difference is observed



- Finally, to investigate the possibilities of a parent resonance or other quasi-resonant behavior, we consider the $M_{(\text{lepton}, \nu, jj)}$ and the $M_{(\text{lepton}, \nu, jj)} - M_{jj}$ distributions for events with M_{jj} in the range 120-160 GeV/c² and to investigate the Dalitz structure of the excess events, the distribution of $M_{(\text{lepton}, \nu, jj)} - M_{jj}$, in bins of M_{jj} .
- The distributions are compatible in shape with the background-only hypothesis in all cases.

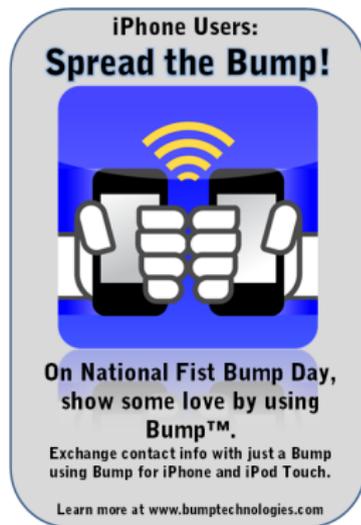




- We study the invariant mass distribution of jet pairs produced in association with a W boson.
- The best fit to the observed dijet mass distribution using known components, and modeling the dominant W +jets background using ALPGEN+PYTHIA Monte Carlo, shows a statistically significant disagreement.
- One possible way to interpret this disagreement is as an excess in the 120-160 GeV/ c^2 mass range.
- If we model the excess as a Gaussian component with a width compatible with the dijet invariant mass resolution we obtain a **p-value of 7.6×10^{-4}** , corresponding to a significance of **3.2 standard deviations**, after accounting for all statistical and systematic uncertainties.

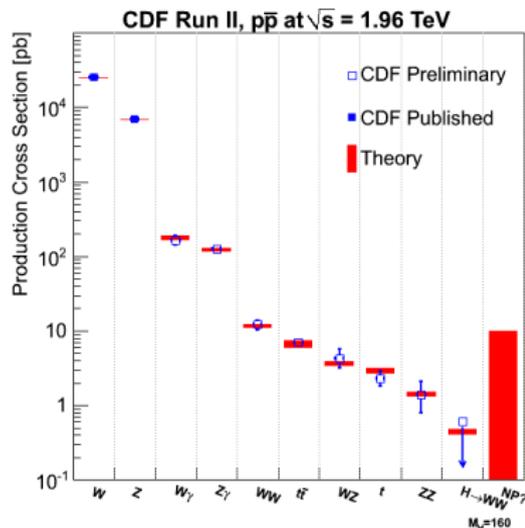


- Even if it is hard





- Even if it is hard
- Need to keep the focus on things we know:
 - 1 CDF now sensitive to very small cross sections and small differences between data and model also in complex final states





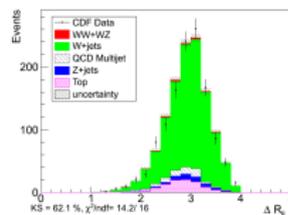
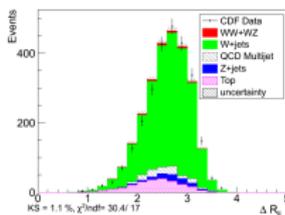
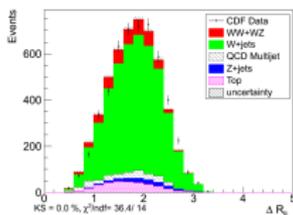
- Even if it is hard
- Need to keep the focus on things we know:
 - 1 CDF now sensitive to very small cross sections and small differences between data and model also in complex final states
- Work to do:
 - 1 Understand if the excess is **truly** new physics or **simply** mismodeling
 - 2 If it is found to be mismodeling need to comprehend why we don't understand this hard kinematical region as we would expect from theory
 - 3 Need to carefully consider both possibilities as it is exciting both ways.



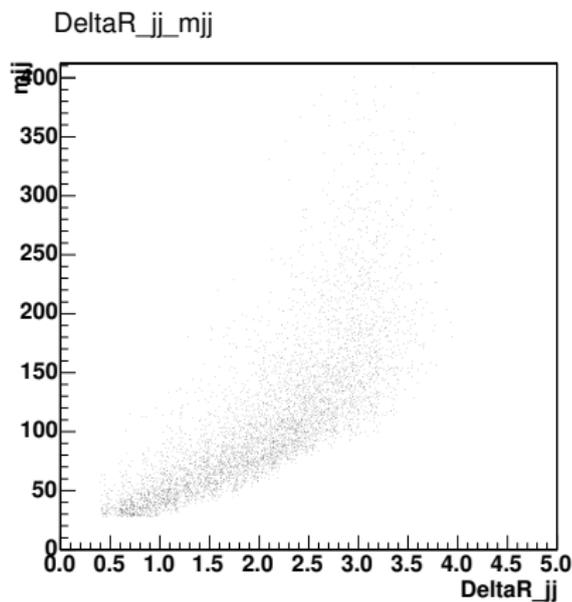
- Would like to thank J.Campbell, E. Eichten, R. K.Ellis, C.Hill, K.Lane, M. Mangano, A. Martin for the useful suggestions and discussions

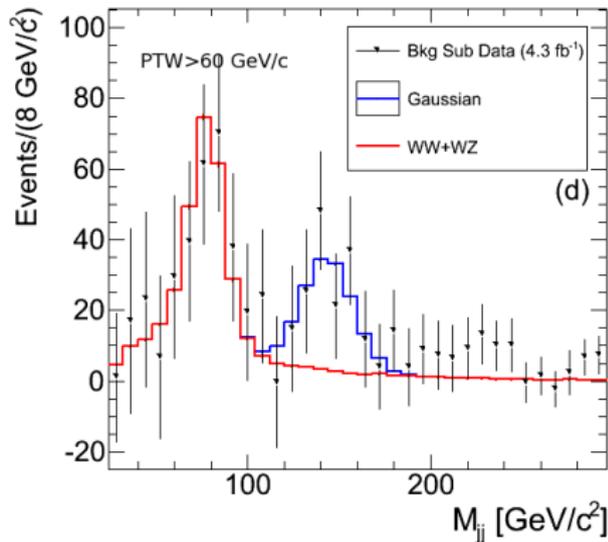
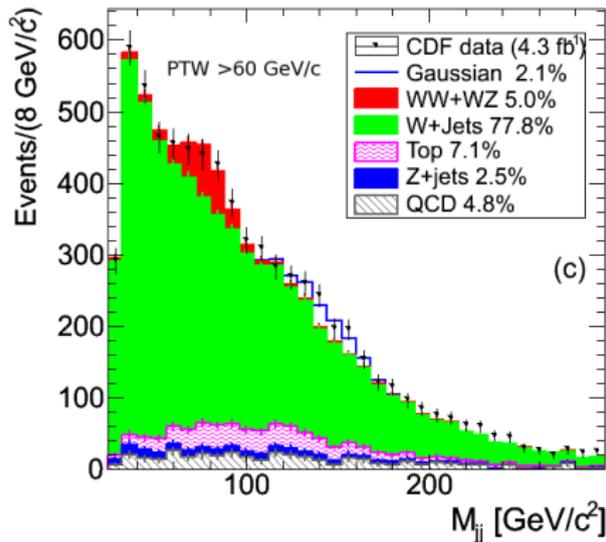
Backup

- If we do reweight to the sidebands we observe that the significance drops to 2.3σ
- Does it really makes sense?
 - 1 The two sidebands are qualitatively different (events are from two different kinematical regions, with different sample composition) $\rightarrow \Delta R_{jj}$ is highly correlated (and the correlation is hard to understand) with the M_{jj} .



- 2 reweighting to the sum of the two sidebands, since the low one has more statistics, is not completely right because we artificially make the m_{jj} look more like the low sideband.

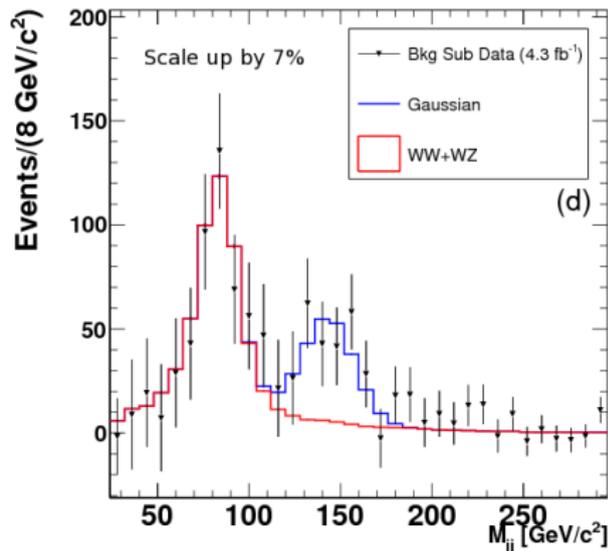
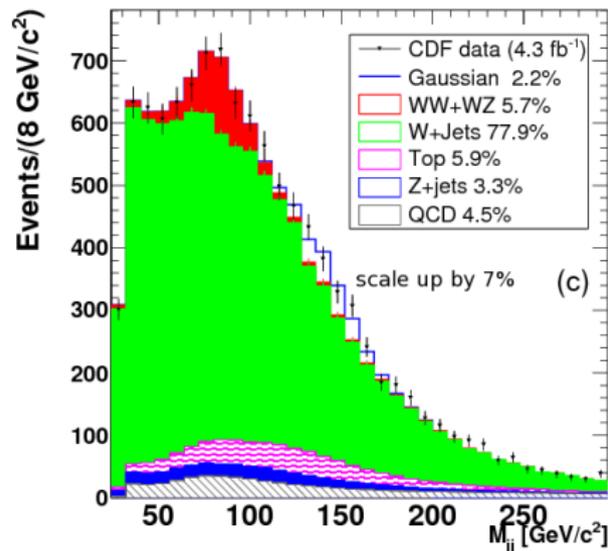




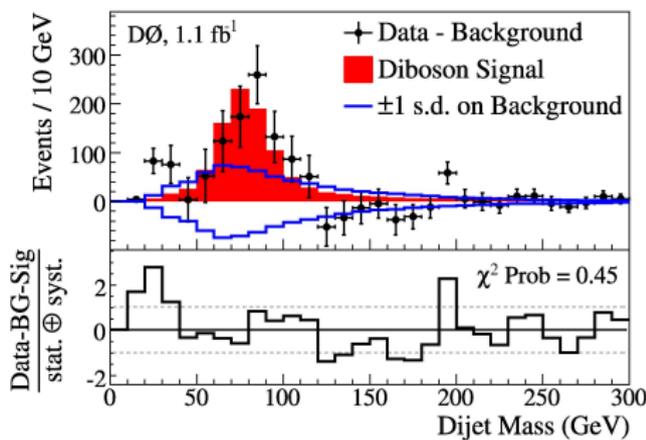
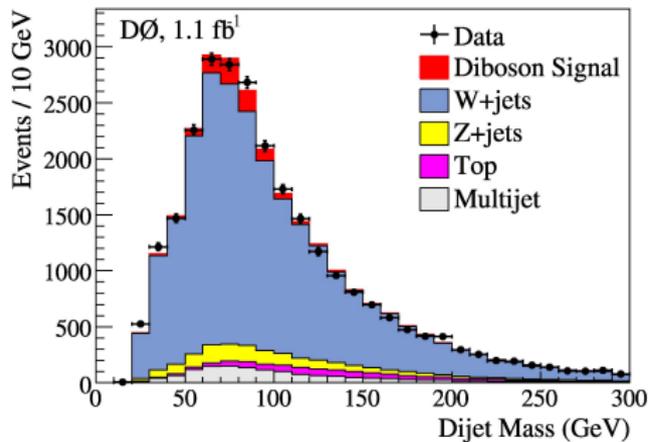
What happens if we change the Jet Energy scale?



Result of the fit scaling JES up by 7%



always above 3σ



paper