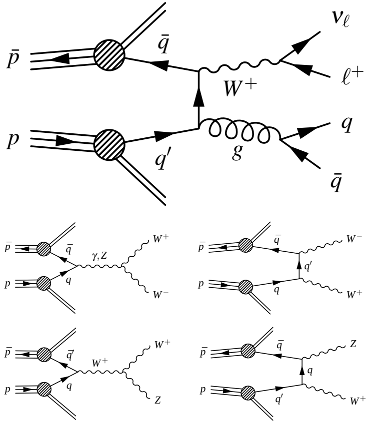
# *Invariant Mass Distribution of Jet Pairs Produced in Association with a $W$ boson in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*

Viviana Cavaliere
University of Illinois at Urbana-Champaign

for the CDF collaboration

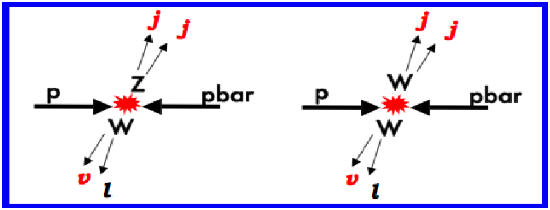
20th May 2011

# Theoretical interest of the final state




- Measurements of associate production of a W boson and jets are an important 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

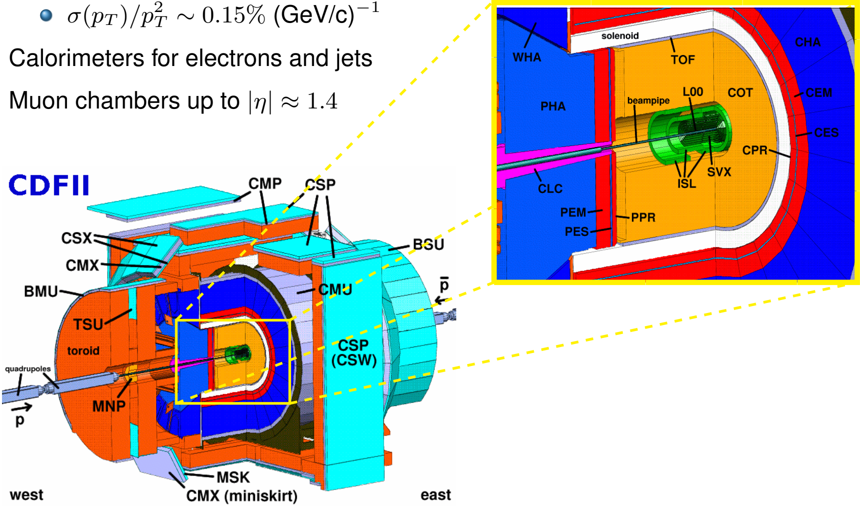
# *Diboson with jets at CDF*




- Diboson represent a Standard Model reference: a starting point for searches in $W$+jets events.
- D0 had strong evidence of $WW + WZ \to \ell\nu$ +jets, PRL 202, 161801 (2009).
- $WW, WZ, ZZ$ observed in $\not\!\!E_T$ + jets mode at CDF, PRL 103, 091803 (2009).
- $WW, WZ$ observed in $\ell\nu$ +jets final state. Two analysis:
  1. Using matrix elements technique
  2. Looking for the $W \to jj$ peak in the dijet invariant mass
- Phys. Rev. Lett. 104, 101801 (2010)
- Results have been updated with 4.3 $fb^{-1}$ ( Analysis page )

# The CDF II Detector in a Nutshell

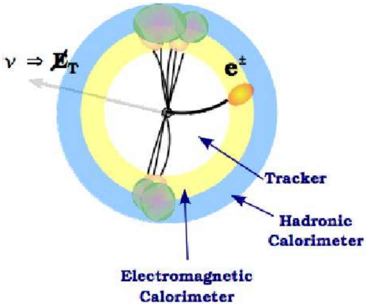
- Silicon detectors (L00+SVX+ISL) and central drift chamber (COT) in $1.4$ T magnetic field
  - I.P. resolution $\sim 40\ \mu$m
  - $\sigma(p_T)/p_T^2 \sim 0.15\%$ (GeV/c)$^{-1}$
- Calorimeters for electrons and jets
- Muon chambers up to $|\eta| \approx 1.4$



$$WW + WZ \to l\nu + jj$$

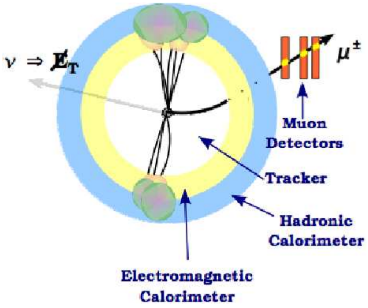
# *High $p_T$ Electron and Muon selection*

- Electrons $\rightarrow E_T > 20$ 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



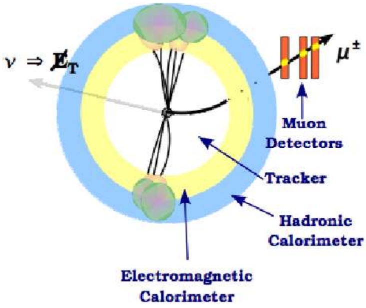
# High $p_T$ Electron and Muon selection

- Electrons $\rightarrow E_T > 20$ 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 $\rightarrow p_T > 20$ 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



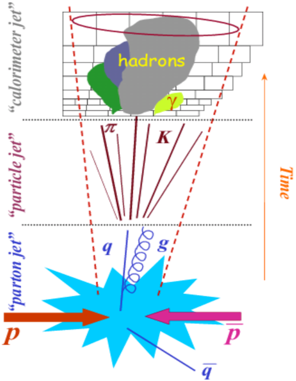
# High $p_T$ Electron and Muon selection

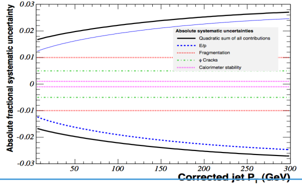
- Electrons $\rightarrow E_T > 20$ GeV/c$^2$ (GeV/c) and $|\eta| < 1.0$
  1. Require calorimeter showers consistent with electromagnetic interactions
  2. Require that 90% of energy is deposited in the EM calorimeter
- Muons $\rightarrow p_T > 20$ GeV/c$^2$ (GeV/c) and $|\eta| < 1.0$
  1. Require high 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 $\not{E}_T > 25$ GeV and $M_T^W > 30$ GeV to ensure the presence of a real $W$



# Jet Definition

- 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 at least two jets with $E_T > 20$ GeV and $|\eta| < 2.4$
- Absolute Jet Energy scale known at 3% level NIM A 375, 354 (2006) (CDF webpage)
- Independent checks:
    - Dijet events.
    - $\gamma$ + jet.
    - $Z$ + jet.
    - $W$ from top (CDF webpage)
    - $Z \to b\bar{b}$ peak. NIM A 596, 354 (2008)
- Require $p_{T,jj}$ >40 GeV/c

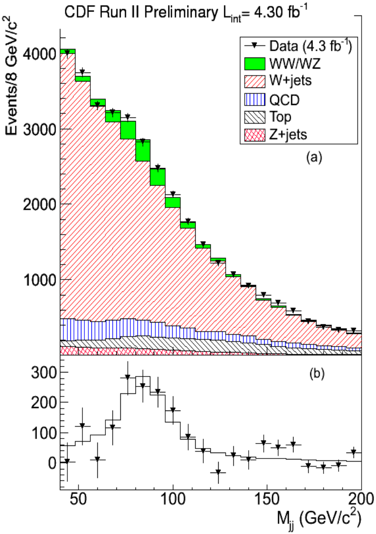


# Sample Composition

- $W \to \ell\nu$ + jets ($l = e, \mu, \tau$):
  - same signature as signal with a much higher cross section (2066 pb)
  - $\sim$ 80% of the sample
- $Z \to ll$ + jets ($l = e, \mu, \tau$):
  - where one of the two leptons escapes detection and produces $\not{E}_T$
  - cross section 187 pb
- $t\bar{t}$ + single top:
  - final state similar to signal with at least one real $W$ and two jets.
  - $\sigma(t\bar{t})$ = 7.5 pb and $\sigma$(single top) = 2.9 pb (assuming a mass of 172.5 GeV/c$^2$)
- QCD Multijet:
  - events without a primary $p_T$ lepton from W decay.
  - 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 $\not{E}_T$ 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.

# Cross sections measurement




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

- The resulting cross section is

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

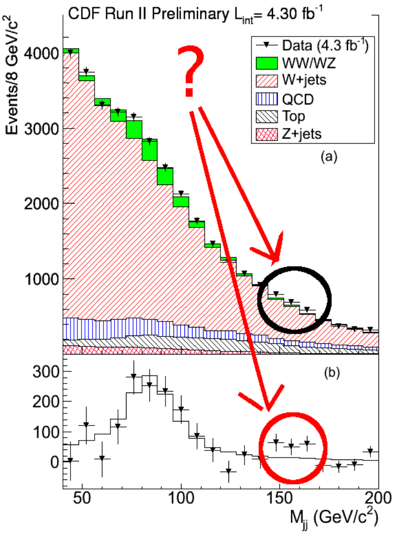
  that is in agreement with SM expectation ($15.9 \pm 0.9$ pb).

# *What did we learn?*

- 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 $WW/WZ \rightarrow \ell\nu$+jets analysis we learned:

1. $W + jets$ 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.

# Cross sections measurements

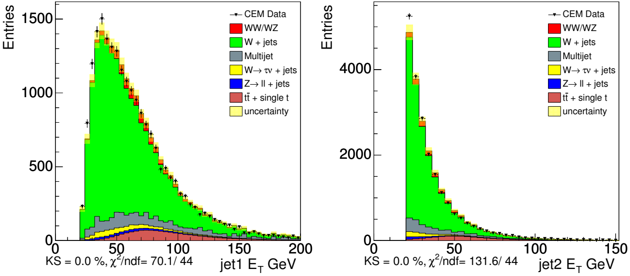



- Observed an interesting discrepancy in the $M_{jj}$
- Needed work in order to understand the nature and properties of the excess
  1. an artifact of background subtraction?
  2. misunderstanding of one of the backgrounds?
  3. real physics?
  4. Is it compatible with a narrow dijet resonance?
  5. Look for model independet answers

# Moving to High Mass

# *Moving to different kinematical region*

- 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
- All cuts chosen “a priori”



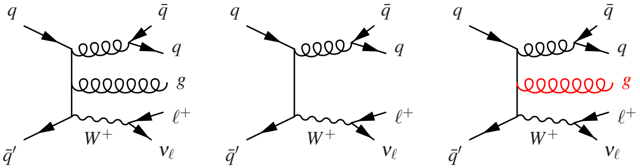
# Modeling of samples

| Process | Model | $\sigma$ (pb) |
|---|---|---|
| WW/WZ inclusive | PYTHIA | $15.9 \pm 0.9$ |
| $Z \to e, \mu, \tau$+jets | ALPGEN+ PYTHIA | $787 \pm 85$ |
| $t\bar{t}$ | PYTHIA | $7.5 \pm 0.83$ |
| single top | MADEVENT + PYTHIA | $2.86 \pm 0.36$ |
| 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

# *Alpgen+Pythia*

- 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 takes events from diagram (b) and add a radiated gluon to produce diagram (c)

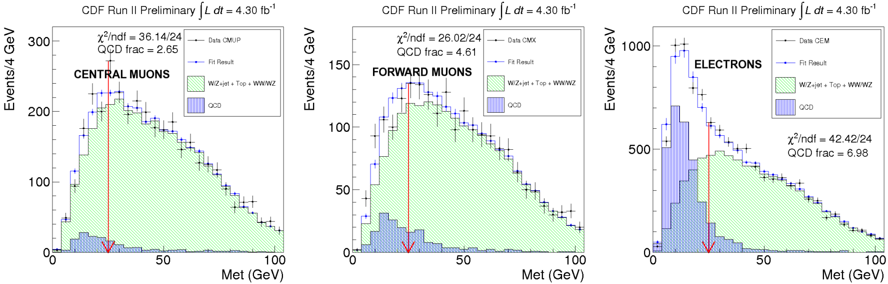
# QCD Multijet Model

Modeled using data sidebands

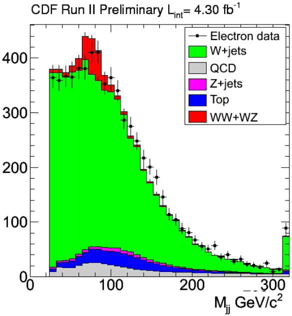
- Non isolated muons:
  - Using non-isolated events, events which pass all selection criteria except the requirement of lepton isolation.
  - based on idea that jets that contain energetic leptons are mostly non-$W$ events.
- "AntiElectrons":
  - Some non-kinematic cuts for the electron are used to reject 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.

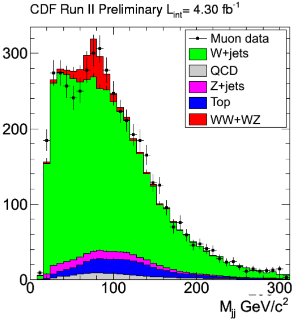
# QCD Multijet estimation

- QCD multijet events are characterized by low $\not\!\!E_T$, so $\not\!\!E_T$ distribution is completely different from $W + jets$
- Best solution $\rightarrow$ Fit the $\not\!\!E_T$ 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%)



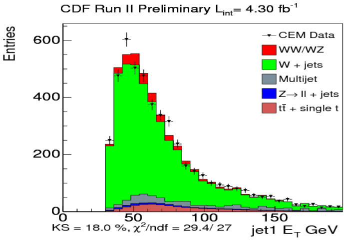
# Putting ingredients together

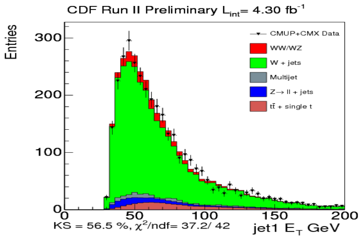


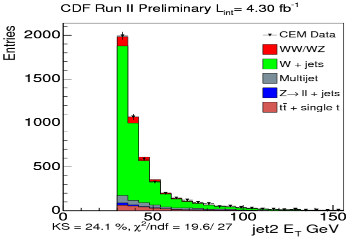



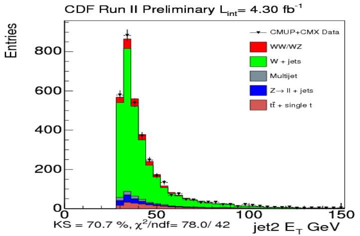

1. $M_{jj}$ of data compared to SM predictions
2. Want to investigate further $\rightarrow$ Need to check every background carefully. Public Webpage
   1. Look at Excess sidebands ($115 > M_{jj}$ or $M_{jj} > 175$ GeV/$^2$)
   2. QCD multijet shape modeling
   3. $t\bar{t}$ production
   4. $W + jet$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

# *Control sample:* $115 > M_{jj}$ *or* $M_{jj} > 175$ GeV/$^2$

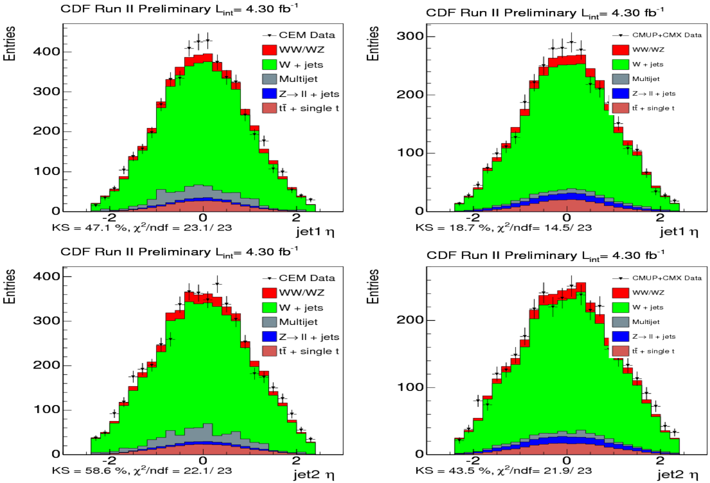


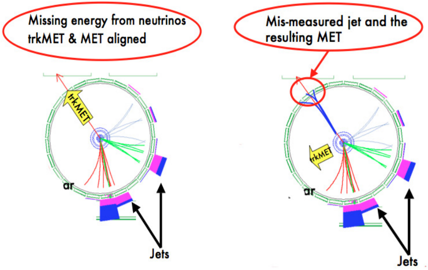



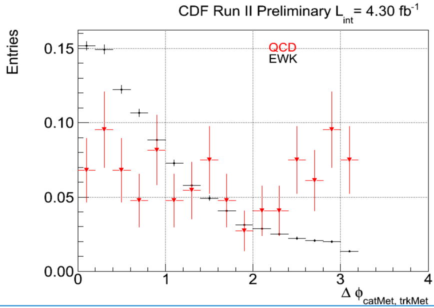
# *Control sample:* $115 > M_{jj}$ *or* $M_{jj} > 175\ GeV/^2$

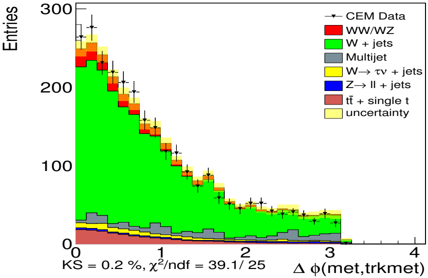


# QCD Multijet model

- Check shape and rate using independent ways:
  1. $\Delta\phi(\not{E}_T$, Missing Pt) $\rightarrow$ large for qcd events

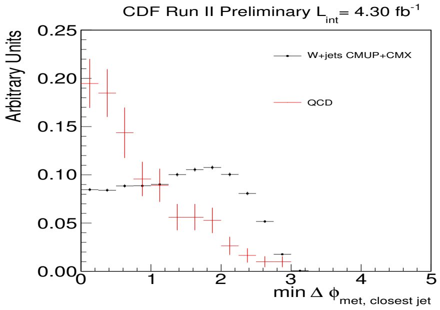


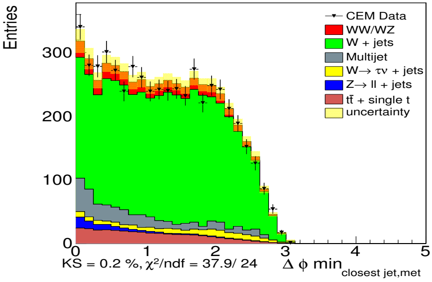



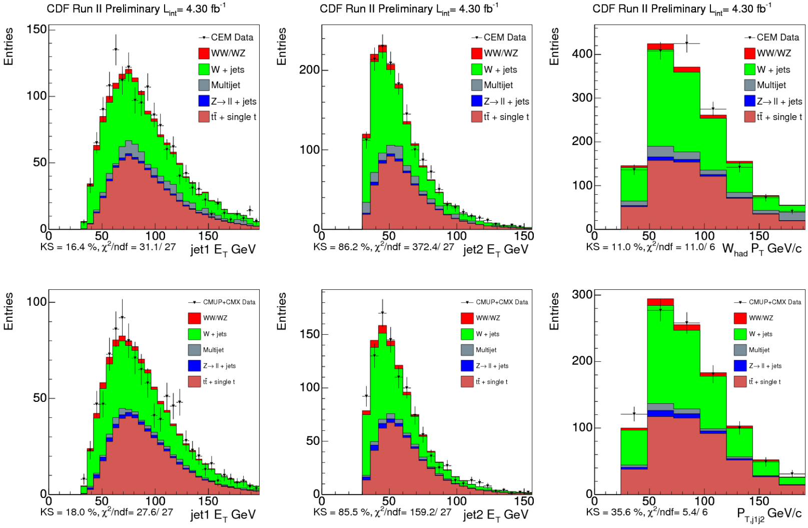
# QCD Multijet model

- Check shape and rate using independent ways:
  1. $\Delta\phi(\not{E}_T, \text{Missing Pt}) \rightarrow$ large for qcd events
  2. $\Delta\phi(\not{E}_T, \text{closest jet}) \rightarrow$ mis-measured jets tend to be aligned with the $\not{E}_T$



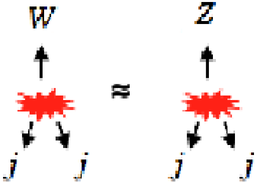



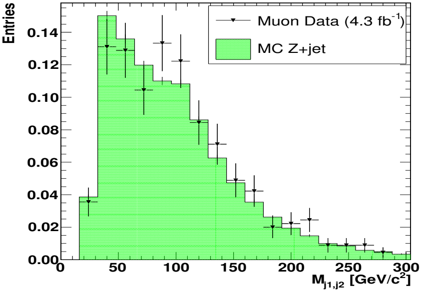
# $W + n \geq 3$ Jets: $t\bar{t}$ enhanced sample

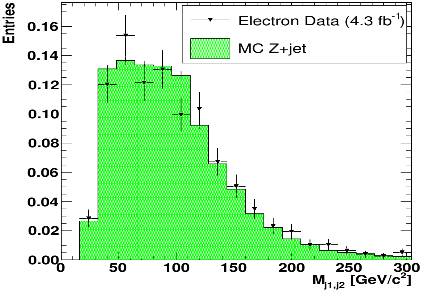


# $W + jets$ shape

- 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 + PYTHIA only
- Problem: 15 times less data

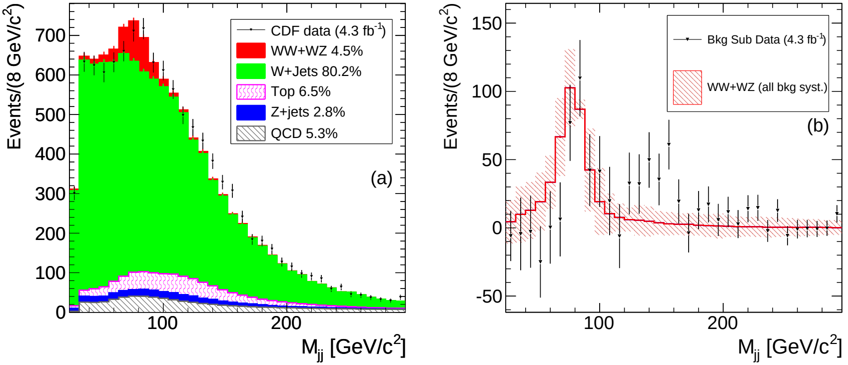


# Fitting procedure

- Combined $\chi^2$ fit to the dijet mass distribution in electron and muon samples.
- 5 templates:
  1. $W +$ jets (uncostrained, normalization determined from the fit)
  2. QCD (normalization constrained by the MET fit with 25 % error)
  3. $Z +$ 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)



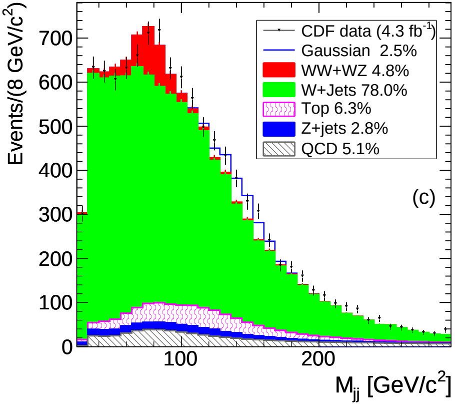
# Gaussian Assumption

- 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.

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

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

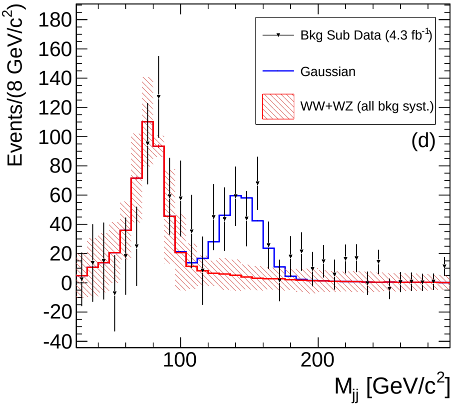
# Fit to data with SM templates + gaussian




- Data fitted with SM templates plus a gaussian.
- fit range 28-200 GeV/c$^2$
- $\Delta\chi^2$ observed 20.31 that corresponds to a statistical significance of 3.7$\sigma$ (including trial factor)

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

# *Fit to data with SM templates + gaussian*




- Data fitted with SM templates plus a gaussian.
- fit range 28-200 GeV/c$^2$
- $\Delta\chi^2$ observed 20.31 that corresponds to a statistical significance of 3.7$\sigma$ (including trial factor)

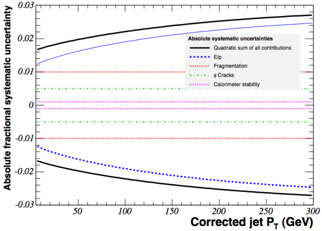
| | Electrons | Muons |
|---|---|---|
| Excess events | $156 \pm 42$ | $97 \pm 38$ |
| Excess events / expected diboson | $0.60 \pm 0.18$ | $0.44 \pm 0.18$ |
| Mean of the Gaussian component | $144 \pm 5$ GeV/c$^2$ | |

# Systematics

- 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 | $\pm 1.9$ |
| | $Q^2$ | $\pm 6.7$ |
| | JES | $\pm 6.1$ |
| | **Total** | $\pm \mathbf{9.3}$ |

- 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.



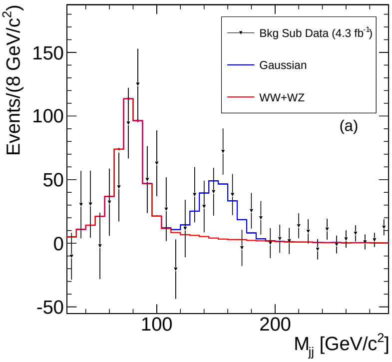
# *Shape systematic: $W + jets$*

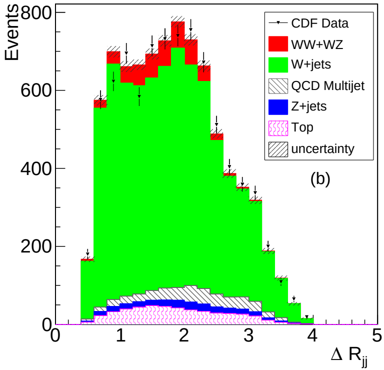
- Alpgen MC depends on some parameters:
  1. Factorization and renormalization scale $Q^2 = M_W^2 + \sum p_{T,j}^2$ which can be varied by a constant factor on an event by event basis
  2. $k_T$ Scale Factor: Alpgen's scale factor for $\alpha_s$ at each decay vertex.
  3. Parton matching cluster $p_T$ 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
- 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

# Significance

- 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 $\chi^2$: for each toy sample, the minimum $\chi^2$ of the scan is used in the $\Delta\chi^2$ evaluation.
- The largest p-value is $7.6 \times 10^{-4}$, corresponding to a significance of 3.2 standard deviations.

# Best Fitting Systematics

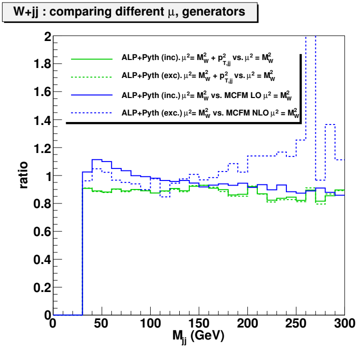


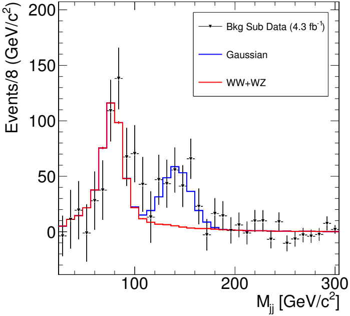




- Fit performed combination of systematic that fits data best: lowest $\chi^2$
- Returns a p-value intermediate between the largest and statistical only
- $\Delta R_{jj}$ ($M_{jj} < 115$ and $M_{jj} > 175$ GeV/c$^2$) shown with the same combination of systematic. $\chi^2/ndf$=26.6/18.

# NLO effects

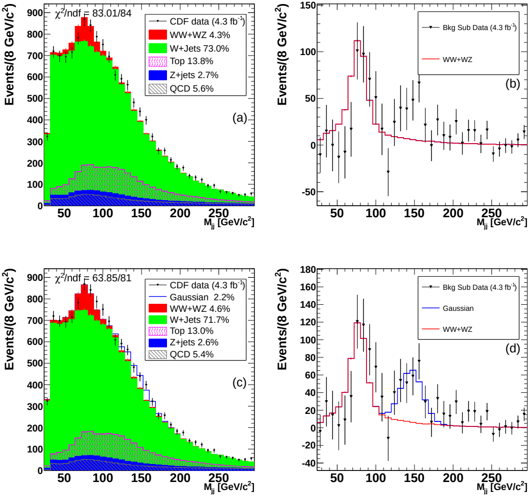
- In order to test Next to Leading Order contributions to the $W$+2 partons prediction, we compare (private communication with J.Campbell, E. Eichten, K.Lane, A.Martin) ALPGEN and interfaced to PYTHIA for showering to a sample of $W$+2 partons simulated using the MCFM.
- We extract a correction as a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



# Inclusive selection

- Tried also what happens if we don't veto on the number of jets
- Top normalization increases as expected but excess still in place



# *Production Cross Section ?*

- 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 GeV/c$^2$ and decaying isotropically into two jets.
- In this model, we estimate a cross section * BR(jet-jet) of the order of 4 pb
  - $\rightarrow$ not compatible with Standard Model $WH$ $\sigma \times BR(b\bar{b})$ = 39 fb

# Flavour composition

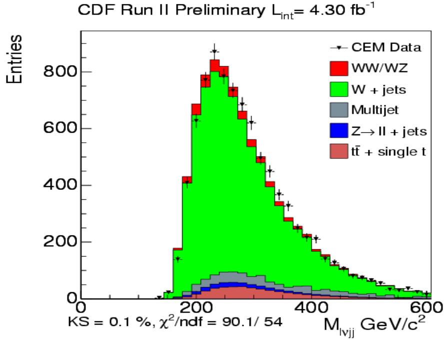
- Tried to look at the flavour composition of these jets.
  - Assuming a branching ratio in $b\bar{b}$ of 100% and considering tagging efficiency, we expect to be able to see a much stronger signal in the tagged sample
- Compared the fraction of events with b-jets in the sidebands (120$> M_{jj}$ or $M_{jj}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

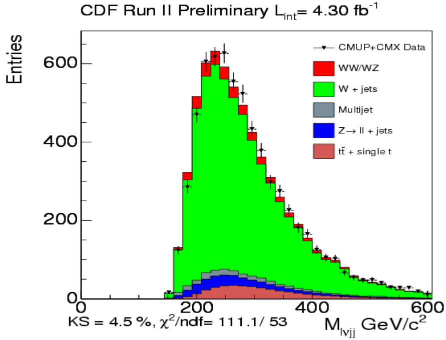
| Tag requirement | Excess region | Sideband region |
|---|---|---|
| Muons | | |
| 1 tag | $0.1027 \pm 0.0112$ | $0.0813 \pm 0.0096$ |
| 2 tag | $0.0078 \pm 0.0030$ | $0.0084 \pm 0.0030$ |
| Electrons | | |
| 1 tag | $0.0897 \pm 0.0088$ | $0.0945 \pm 0.0087$ |
| 2 tag | $0.0110 \pm 0.0030$ | $0.0095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

No significant difference is observed

# Resonant production

- Finally, to investigate the possibilities of a parent resonance or other quasi-resonant behavior, we consider the $M_{(\mathrm{lepton},\nu,jj)}$ and the $M_{(\mathrm{lepton},\nu,jj)}$-$M_{jj}$ distributions for events with $M_{jj}$ in the range 120-160 GeV/c$^2$and to investigate the Dalitz structure of the excess events, the distribution of $M_{(\mathrm{lepton},\nu,jj)}$-$M_{jj}$, in bins of $M_{jj}$.
- The distributions are compatible in shape with the background-only hypothesis in all cases.



# Conclusions

- 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 $7.6 \times 10^{-4}$, corresponding to a significance of 3.2 standard deviations, after accounting for all statistical and systematic uncertainties.
- Published in PRL 106, 171801 (2011)

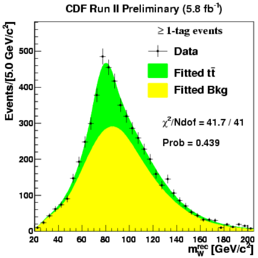
# Acknowledgments

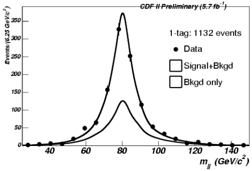
- Would like to thank J.Campbell, E. Eichten, R. K.Ellis, C.Hill, K.Lane, M. Mangano, A. Martin for the useful suggestions and discussions

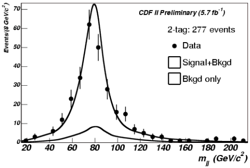
Backup

# JES from $t\bar{t}$ events

- Pure samples of light quark jets
- A few 1000s hadronic Ws, from three different ttbar decay modes.
- All right on the spot!
- JES for light quarks known to 1% level



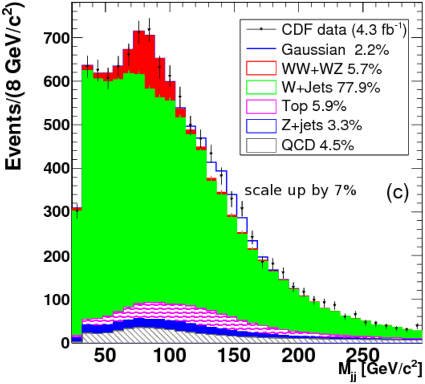



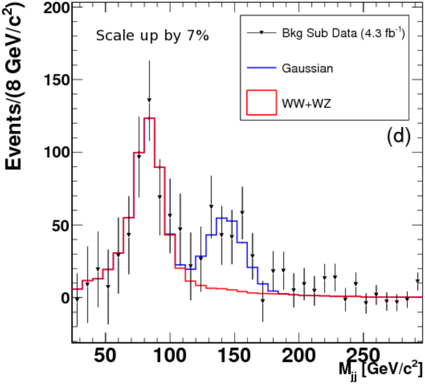




More plots/infos on http://www-cdf.fnal.gov/physics/new/top/public_mass.html

# *Extreme Scenario*

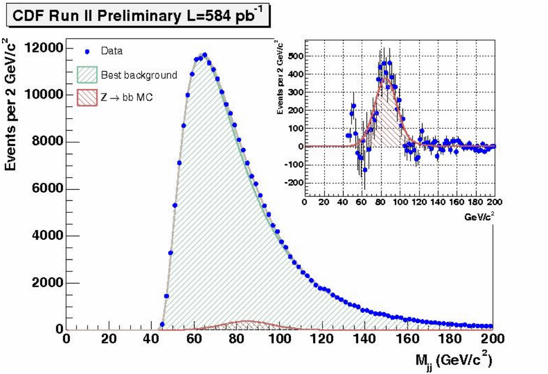
- N.B. it is NOT correct to simply shift the plot: one needs to rescale first, then redo cuts and fits.
- Here we tried an unreasonably large JES shift : +7%.
- The significance of the excess is unaffected: 3.2 sigma.



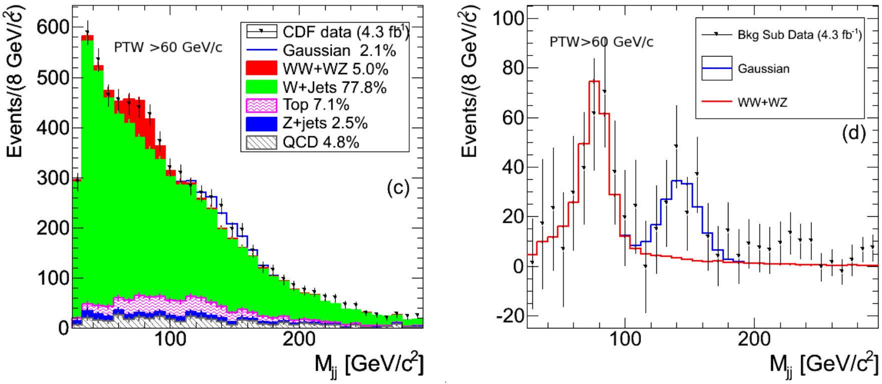
# JES from $Z^0 \to b\bar{b}$ events

- JES correction factor for b-quarks
  $k = 0.974 \pm 0.011(\mathrm{stat})^{+0.017}_{-0.014}(\mathrm{syst})$



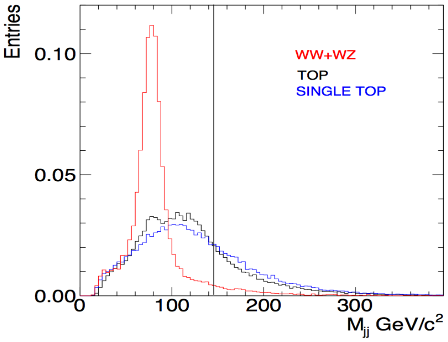

More plots/infos from
http://www-cdf.fnal.gov/physics/new/qcd/abstracts/zbb_07.html

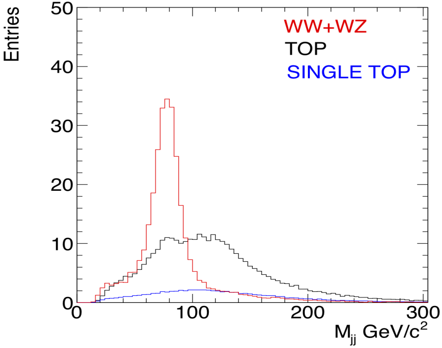
# Fit requiring $p_{T,jj} > 60$ GeV/c



# $t\bar{t}$ + single top Shape

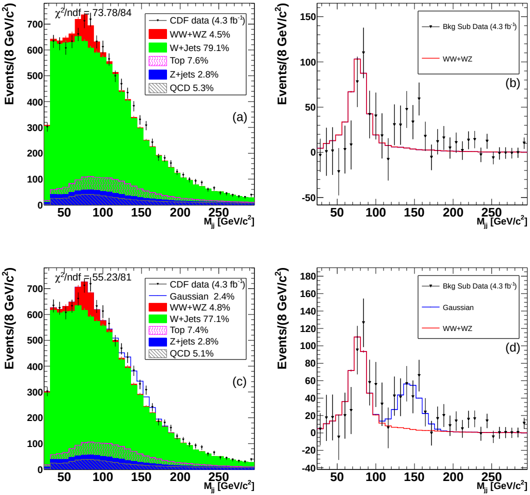
- $W \to jj$ peak clearly visible in $t\bar{t}$.
- Excess (black vertical line corresponds to gaussian mean) not located where $t\bar{t}$ + single top peak



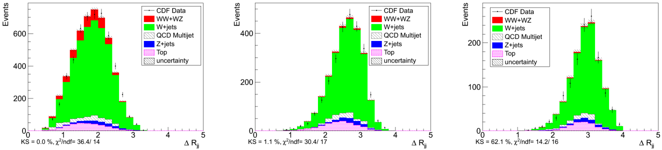
# Increasing single top cross section

- If we artificially double the single top contribution we obtain a statistical significance of 3.5 $\sigma$.



# $\Delta R$ studies

- If we do reweigh to the sidebands we observe that the significance drops to $2.3\sigma$
- Does it really make 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 $M_{jj}$.

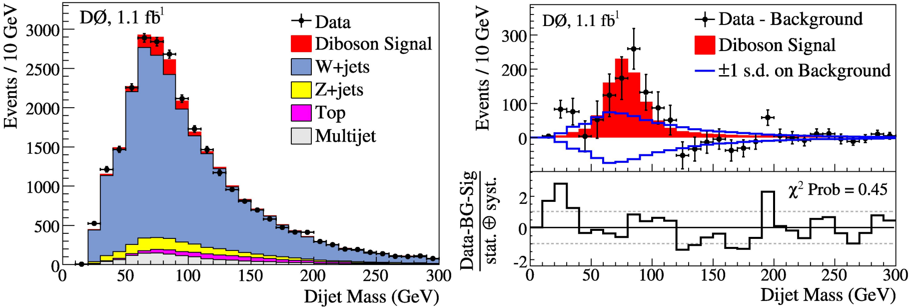



  2. reweighing to the sum of the two sidebands, since the low one has more statistics, is not completely right because we artificially make the mjj look more like the low sideband.
  3. The $M_{jj}$ distribution is highly corralated to the $\Delta R_{jj}$ one. We compare background predictions to $M_{jj}$ and $\Delta R_{jj}$.

# $W$+jets (mis)modeling $\rightarrow \Delta R_{jj}$

- 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 dijet $p_T < 40$ GeV/c.
- We use these regions to derive a correction as a function of $\Delta R_{jj}$ 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 $\Delta R_{jj}$ distribution is strongly correlated to $M_{jj}$ and the control regions both have significantly different distributions of $\Delta R_{jj}$.
- Reweighing our $W$+jets sample may be a correction to $\Delta R_{jj}$ mismodeling or may introduce bias in the $M_{jj}$ distribution.
- In addition, the $\Delta R_{jj}$ distribution is consistent within the one sigma variation of the systematic uncertainties for events outside the excess mass region.

# *Results from D0 using 1.1 $fb^{-1}$*




D0 WW+WZ paper