Milestone for Tevatron Higgs Searches: First Observation of Diboson Production in Hadronic Final States

Sasha Pronko
Fermilab
Diboson Processes at the Tevatron

- \( \text{SU}(2)_L \otimes \text{U}(1)_Y \)
  - Electroweak (EW) group structure is central to Standard Model
- EW boson self-interactions are completely dictated by gauge symmetry
  - Sensitive to new physics

[Graph showing production cross section for different processes at Tevatron Run II, \( \sqrt{s} = 1.96 \text{ TeV} \)]
Diboson Final States

- Dibosons were discovered in $l\nu l\nu$, $lllv$, $llll$, and $ll\nu\nu$, modes
  - Small branching ratios, clean signatures, easy to trigger
- “Semileptonic” modes with at least one $W/Z \rightarrow jj$
  - ~40% branching fractions, ~1000× backgrounds, difficult to trigger
Recent Tevatron Results with Dibosons

- Dibosons (WW, WZ, and ZZ) were observed in fully leptonic mode

- Dibosons at Tevatron were not previously observed in $lujj$ or $vujj$ modes
  - CDF: 2.4σ in $WZ/WW\rightarrow lujj$
  - D0: 4.3σ in $WZ/WW\rightarrow lujj$
  - evidence
From Dibosons to Higgs Searches

- $H \rightarrow WW$ is dominant decay channel for $M_H > 135$ GeV/c$^2$
  - Direct WW production is largest non-reducible background
- Need to be well measured and understood
From Dibosons to Higgs Searches

- $HW \rightarrow \nu l + bb$ and $HZ \rightarrow \nu \nu + bb$ are leading channels for light Higgs ($M_H < 135 \text{ GeV/c}^2$) searches at the Tevatron
  - Similar signatures and challenges to $WW/WZ \rightarrow \nu l + jj$ and $ZZ/WZ \rightarrow \nu \nu + jj$
    - Small signal in a large background
    - Test of analysis techniques

- Recent CDF Higgs results for $M_H = 115 \text{ GeV/c}^2$
  - Observed limit
    - $HW: 5.3 \times \sigma_{SM}$ in 4.3 fb$^{-1}$
    - $HZ: 6.1 \times \sigma_{SM}$ in 3.6 fb$^{-1}$

Road to Higgs is paved with Dibosons
Tevatron is Running Very Well!

- Produced in 1 fb\(^{-1}\)
  \[\approx 6,200,000 \text{ } W \rightarrow l\nu + X\]
  \[\approx 2,600,000 \text{ } Z \rightarrow \nu\nu + X\]
  \[\approx 5,100 \text{ } WW \rightarrow jlj\nu\]
  \[\approx 1,300 \text{ } WZ \rightarrow jlj\nu + jj\nu\]
  \[\approx 420 \text{ } ZZ \rightarrow \nu\nu jj\]
  \[\approx 64 \text{ } H \rightarrow WW^* \rightarrow l\nu jj \ ?\]
  \[\approx 33 \text{ } WH \rightarrow l\nu bb \ ?\]
  \[\approx 13 \text{ } ZH \rightarrow \nu\nu bb \ ?\]

- \(l=e,\mu,\tau\); \(M_H=120\text{ GeV/}c^2\)

- \(\sim 7\text{ } fb^{-1}\) per experiment; \(\sim 1.9\text{ } fb^{-1}\) in FY09
- 55-60 pb\(^{-1}\) per week in FY09
- Ramping up speed after this summer shutdown
  - Already \(\sim 300\text{ } pb^{-1}\) since 09/15
- Running in 2011? Expect 10-12 fb\(^{-1}\) per experiment
CDF in Run II

- Multipurpose, classic design
- Operating well
  - 80-90% efficiency
- Broad physics program
  - QCD, EWK, top, B-physics, Higgs searches, new physics searches
How Do You Find Dibosons in jj+MET?

• **Strategy**
  
  – Select jj+MET events
    • Sensitive to $h$ and $\nu\nu$ decay modes of $W$ and $Z$
    • Need only Calorimeter & COT (tracking)
      – 10% more data!
  
  – Maximal use of data to estimate backgrounds
  
  – Simple but smart analysis techniques
    • Focus on deep understanding of backgrounds
    • It’s never late to add multivariate techniques
  
  – Do it fast!!!
How Do You Find Dibosons in jj+MET?

• **Challenges**
  – Need lots of data ✓
    • Analysis is based on 3.5 fb$^{-1}$ of data
  – High efficiency triggers at all luminosities
    • L2 trigger upgrade
  – Large backgrounds dominated by QCD multijet events with fake MET and Z/W+jets
    • Sophisticated technique to suppress QCD multijets and estimate systematics
  – Extracting small signal
Calorimeter Trigger Upgrade

- **Upgraded L2 trigger**
  - More sophisticated algorithm (almost same as in offline)
  - Better resolution and turn-on
  - Better performance at high luminosity
Trigger Efficiency

- Use all MET and MET+jets triggers
  - Every bit of extra data counts!
  - Complicates luminosity accounting
- Use $Z \rightarrow \mu\mu$ events (standard candle) with two jets from high $P_T$ triggers to find trigger efficiency
  - Integrated efficiency $96.4\% \pm 2.2\%$

$\mu$ is minimum ionizing particle
$\mu \equiv$ MET in calorimeter
Dibosons are Swamped with Backgrounds

- Triggered data dominated by QCD events with fake MET
Backgrounds: QCD Multijets

- Huge production rate
  - ~9 orders of magnitude above WW+WZ+ZZ
- Fake MET due to jet energy mis-measurement in the calorimeter
  - Rare fluctuations × huge rate = significant background
- Reject as much as possible
- Use data to model whatever remains

Jet 1  
Jet 2  
Jet n

Lost/mis-measured jet

Mis-measured jet tends to align with MET: Δϕ should help rejection
Backgrounds: EWK processes

- **Use MC to describe kinematics**
  - $W$+jets
    - $W \rightarrow e\nu, \mu\nu, \tau\nu$
  - $Z$+jets
    - $Z \rightarrow \nu\nu$ (looks like signal)
    - $Z \rightarrow ee, \mu\mu, \tau\tau$
  - Top quark production

**Dominant EWK backgrounds**

- $\tau \rightarrow$ hadrons
- $e, \mu, \tau$
- $Z \rightarrow \nu\nu$ (looks like signal)

---

11/30/09

A Pronko, SMU Video Seminar
MET Resolution Model (Metmodel)

Example of jet energy resolution

Mis-measurements in jet energy are leading source of fake MET

Obtain jet energy resolution as function of $E_{\text{jet}}$ & $\eta$

- Select events with true MET
  - Calculate MET-significance based on event configuration & known energy resolution
  - Use MET-significance to select with true MET

$E_{\text{det}}/E_{\text{true}} - 1$
Validation of Metmodel

• Use $W(\rightarrow e\nu)+\text{jet}$ data to validate MET-resolution

• Regions dominated by events with fake MET
  - Low MET-significance and small $\Delta\phi(\text{jet-MET})$
Diboson Candidate Selection: 44,910 events

- Analysis is based on 3.5 fb\(^{-1}\) of data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Cut values</th>
</tr>
</thead>
<tbody>
<tr>
<td>MET</td>
<td>&gt;60 GeV</td>
</tr>
<tr>
<td>Jet -1,2 (E_T)</td>
<td>&gt;25 GeV</td>
</tr>
<tr>
<td>Jet EmFr</td>
<td>&lt;0.9</td>
</tr>
<tr>
<td>Jet -1,2 (</td>
<td>\eta</td>
</tr>
<tr>
<td>(\Delta \phi_{\text{closest}})</td>
<td>&gt;0.4 rad</td>
</tr>
<tr>
<td>MET-significance</td>
<td>&gt;4</td>
</tr>
<tr>
<td>(\Delta R_{\text{lep-jet}})</td>
<td>&gt;0.2</td>
</tr>
<tr>
<td>(E^{EM}/E^{tot})</td>
<td>0.3-0.85</td>
</tr>
<tr>
<td>(M_{jj})</td>
<td>40 GeV/c(^2) – 160 GeV/c(^2)</td>
</tr>
<tr>
<td>Jet timing</td>
<td>&lt;4.5 ns</td>
</tr>
</tbody>
</table>
Modeling Remaining Multijet Background

- **Track MET (trkMET)**
  - Analogous to MET
- **True MET**
  - Small $\Delta\phi(\text{trkMET}-\text{MET})$
- **Fake MET**
  - Large $\Delta\phi(\text{trkMET}-\text{MET})$

MET from neutrinos

trkMET & MET aligned

Mis-measured jet and resulting fake MET

\(p\)
\(p\bar{p}\)
\(j\)
\(j\)
\(Z\)
\(\nu\)
\(\bar{\nu}\)
Modeling Remaining Multijet Background

- Subtract EWK from data in $\Delta\phi(\text{trkMET-MET})>1.0$ region
- Account for QCD background contribution in peak with dijet MC
- Address MC-data differences in resolution with $Z\rightarrow\mu\mu$ events
Checking Background Model

- Great agreement in distributions sensitive to fake MET
  - MET-significance
  - $\Delta \phi$ (closest)
- EWK background and signal have same shapes
Extracting Diboson Signal

- Fit $M_{jj}$ distribution using three templates
  - EWK, QCD, signal
- Minimize the unbinned extended negative log likelihood (ROOFIT)
- Nuisance parameters in the fit
  - EWK normalization
  - Jet energy scale (JES)
  - QCD shape & normalization
  - Signal normalization
M_{jj} Templates: Multijet Background

- Shape & normalization taken from data in the region $\Delta \phi(\text{trkMET-MET})>1.0$ after EWK subtraction
- Shape & normalization are constrained in $M_{jj}$ fit
- Uncertainties are driven by extrapolation into $\Delta \phi(\text{trkMET-MET})<1.0$ region
M_{jj} Templates: EWK Background

- Shapes taken from MC
- Total number of EWK events is unconstrained in fit

<table>
<thead>
<tr>
<th>Process</th>
<th>Expected % of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z→νν</td>
<td>28.9</td>
</tr>
<tr>
<td>Z→ττ</td>
<td>1.0</td>
</tr>
<tr>
<td>Z→μμ</td>
<td>0.7</td>
</tr>
<tr>
<td>Z→ee</td>
<td>0.0</td>
</tr>
<tr>
<td>W→τν</td>
<td>24.1</td>
</tr>
<tr>
<td>W→νν</td>
<td>14.4</td>
</tr>
<tr>
<td>W→μν</td>
<td>12.8</td>
</tr>
<tr>
<td>tt</td>
<td>0.9</td>
</tr>
<tr>
<td>Single top</td>
<td>0.5</td>
</tr>
<tr>
<td>Total</td>
<td>82.9</td>
</tr>
</tbody>
</table>

Expected 36,906 EWK events

11/30/09
A Pronko, SMU Video Seminar
M$_{jj}$ Templates: Signal

<table>
<thead>
<tr>
<th>Process</th>
<th>Expected % of sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>WW</td>
<td>2.3</td>
</tr>
<tr>
<td>WZ</td>
<td>0.7</td>
</tr>
<tr>
<td>ZZ</td>
<td>0.3</td>
</tr>
<tr>
<td>Total Signal</td>
<td>3.3</td>
</tr>
<tr>
<td>EWK</td>
<td>82.9</td>
</tr>
<tr>
<td>QCD</td>
<td>13.8</td>
</tr>
</tbody>
</table>

- Shape from MC (Gaussian + polynomial)
- Number of signal events is unconstrained in fit
- Jet energy scale has a Gaussian constraint in fit
  - Gaussian width depends linearly on JES
Expected Signal Significance

- Check with pseudo experiments (PE)
- PE’s input from expectations
  - EWK: 36,906
  - QCD: 6,144
  - Signal: 1,480
- Expected mean statistical significance \(\sim 6\sigma\)
Systematics

- Uncertainties associated with nuisance parameters are folded into fit statistical uncertainty
- Remaining systematic uncertainties on signal extraction
  - EWK shape (next slide)
  - Jet energy resolution (JER)
    - Smear signal template according to JER uncertainty

<table>
<thead>
<tr>
<th>Signal Extraction</th>
<th>% uncertainty</th>
<th># of signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>EWK shape</td>
<td>7.7</td>
<td>117</td>
</tr>
<tr>
<td>Resolution</td>
<td>5.6</td>
<td>85</td>
</tr>
<tr>
<td><strong>TOTAL EXTRACTION</strong></td>
<td><strong>9.5</strong></td>
<td><strong>144</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acceptance</th>
<th>% uncertainty</th>
<th># of signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>JES</td>
<td>8</td>
<td>121</td>
</tr>
<tr>
<td>JER</td>
<td>0.7</td>
<td>11</td>
</tr>
<tr>
<td>Met Model</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Trigger Efficiency</td>
<td>2.2</td>
<td>33</td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>2.5</td>
<td>38</td>
</tr>
<tr>
<td>PDF</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td><strong>TOTAL ACCEPTANCE</strong></td>
<td><strong>9.0</strong></td>
<td><strong>136</strong></td>
</tr>
<tr>
<td><strong>LUMI</strong></td>
<td><strong>6</strong></td>
<td><strong>91</strong></td>
</tr>
<tr>
<td><strong>TOTAL SYSTEMATICS</strong></td>
<td><strong>14.4</strong></td>
<td><strong>218</strong></td>
</tr>
</tbody>
</table>
Systematics on Shape of EWK Background

- Use data $\gamma$+jets as alternative template
  - Many uncertainties eliminated
- Basic idea: kinematics of $V$+jets $\approx \gamma$+jets, $V=W,Z$

$$V + jets(data) = \frac{V + jets(MC)}{\gamma + jets(MC)} \times [\gamma + jets(data)]$$

MT

Photon

MET (MET+jets)

MET+photon = “MET”

CDF RunII Preliminary
Signal Extraction

• Fit result
  – Signal: $1516 \pm 239\,\text{(stat)} \pm 144\,\text{(syst)}$
    • Expected from PE: $1398 \pm 243$
    – JES: $0.985 \pm 0.015$

• Significance
  – Naively $1516/\sqrt{(239^2+144^2)}=5.4\sigma$
  – Consider parameter variations for all sources of systematics
    • Compare likelihood of background only fit with full fit result
    • Convert difference into probability
  – Lowest significance returned: $5.3\sigma$

$\chi^2/\text{ndf}$ has 37% probability
Cross Section

\[ \sigma = \frac{N_{VV}(extracted)}{\varepsilon \cdot A \cdot L} \]

- \( N_{VV}(extracted) = 1516 \)
- Efficiency, \( \varepsilon \)
  - Trigger: 96%
  - Cosmics removal: 99%
- Luminosity, \( L \): 3,450 pb\(^{-1}\)
- Acceptance is weighted by WW, WZ, ZZ cross sections
- Cross section
  - Measured: 18.0 ± 2.8(stat) ± 2.4(syst) ± 1.1(lumi) pb
  - Theory: 16.8±0.5 pb
Back to Higgs Searches...

- On the road to Higgs
  - Need more data
    - Tevatron running until 2011?
  - Increase acceptance
  - Need better analysis techniques
    - Metmodel, track MET, …
  - Reduce systematics
    - Use data to model backgrounds
  - Need smart techniques to extract small signals
    - ME, neural net, decision trees
  - Combine analyses
  - Explore new channels
Summary

• First observation of vector boson pair production in hadronic final state at the Tevatron
  – Published in PRL 103, 091803 (2009)
  – Milestone in search for low mass Higgs
• Measured diboson production cross section
  – Measured: $18.0 \pm 2.8\text{(stat)} \pm 2.4\text{(syst)} \pm 1.1\text{(lumi)} \text{ pb}$
  – SM prediction: $16.8\pm0.5 \text{ pb}$
• Developed and tested new effective techniques
  – Metmodel to remove QCD
  – Track MET to estimate remaining QCD
  – Used $\gamma+jj$ events to estimate shape systematics of EWK template
• Next goal before Higgs is WZ and ZZ observation in final state with two b-jets
We are almost at the summit of SM!!!
Backup Slides
Cosmic Removal

$B = 97 \pm 6$

$\varepsilon = 98.9 \pm 0.2\%$

11/30/09
A Pronko, SMU Video Seminar
Trigger Efficiencies

- Final integral efficiency is 96.2%±2.2%
  - Take 90% efficiency for MET>120
    $\rightarrow$ 2% effect $\rightarrow$ assign additional 2% uncertainty

\[
\text{Eff} = \frac{c}{a-x} \frac{1 + e^{-b}}{a-x}
\]
“Effective” Sample Lumi: 3,450 pb$^{-1}$

- Use muon trigger to find “effective” lumi of all MET triggers
  - Muon sample: $N = \sigma A \epsilon L_{HiPt} = 357$, $L_{HiPt} = 3,483$ pb$^{-1}$, $\epsilon = 100\%$
  - MET sample: $N^{MET} = \sigma A \epsilon^{MET} L_{MET} = 339$, $\epsilon^{MET} = 96\%$
  - $L_{MET} = L_{HiPt} N^{MET}/(N \epsilon^{MET}) = 3,450$ pb$^{-1}$
  - Method also x-checked with MET40 & MET45 triggers only
Rejecting Fake MET in W+jet Events: $M_T$ plot

- Metsig>5 effectively removes QCD fakes
  - Only a small fraction remains in the region $M_T<10$ GeV
Rejecting Fake MET in W+jet Events: $\Delta \phi_{\text{closest}}$

- **Left plot:** clearly see QCD contribution at small $\Delta \phi$
- **Right plot:** QCD is gone if MET-sig>5

11/30/09

A Pronko, SMU Video Seminar
Signal Template for Final Fit

- Allow JES to float in the fit
  - From Final fit: $0.985 \pm 0.019$
- Parameterize width (Gaussian $\sigma$) as a function of JES
JER uncertainty

- JER uncertainty
  - \( N_{\text{vx}} = 1 \) vs \( N_{\text{vx}} > 1 \)
  - Fit function
  - \( \Delta \phi_{jj} \) cut
  - \( E_T(\text{jet3}) \) cut
“Sideband” Kinematics: $40 < M_{jj} < 60$, $110 < M_{jj} < 160$
“Sideband” Kinematics: $40 < M_{jj} < 60$, $110 < M_{jj} < 160$
“Sideband” Kinematics: $40<\text{M}_{jj}<60$, $110<\text{M}_{jj}<160$
No $\Delta \phi_{\text{closest}}$ Cut
No $\Delta \varphi_{\text{closest}}$ Cut
Significance Part-I

• MINUIT reports 1516/239=6.34σ
  – PEs (s+b fit to s+b generated) imply a 6.45σ
  – -2*ΔL=LogL(s+b)-LogL(b)=42
    • TMath::Prob(42,1) = 9.1x10^{-11} → 6.48σ

• Naïve approach:
  – stat^2+syst^2=234^2+144^2=275^2
  – 1516/275 = 5.5σ → 3.8x10^{-8}
Significance Part-II

• Try to estimate the degradation of all systematic uncert. on the significance
  - Fix all parameters except $N_{ewk}$ and $N_{sig}$
    • $\Delta L = 22 \rightarrow TMath::Prob(44,1)=3.3\times10^{-11} \rightarrow >6\sigma$
  - Use alternative JER and repeat:
    • $\Delta L = 22 \rightarrow TMath::Prob(44,1)=3.3\times10^{-11} \rightarrow >6\sigma$
  - Use alternative $\gamma+$jets and repeat:
    • $\Delta L = 14 \rightarrow TMath::Prob(28,1)=1.2\times10^{-7} \rightarrow 5.3\sigma$

• The smallest significance corresponds to $5.3\sigma$
  - Good agreement with Naïve approach: $5.5\sigma$