Measurement of WW+WZ production in lvjj Final States

Gregorio Bernardi
(LPNE-Paris, CRNS/IN2P3)

On behalf of the

Do collaboration

Thanks to all my Dzero colleagues, in particular to J. Sekaric
**Motivation**

Probe of the EWK Symmetry Breaking mechanism
SM tests
Indirect searches for New Physics (NP)
(cross sections, kinematic distributions, gauge boson couplings)

**Important background** to Top, Higgs, SUSY
Good understanding is highly valuable

Proving ground for analysis techniques (MVA)
and statistical treatment used in the Tevatron
Higgs searches

Same final states, similar challenges
For SM tests, and $H \rightarrow WW$

Fully leptonic final states

$l = e/\mu$

- most analyzed
- small branching ratio
- clean signal (low backgrounds)
- observed at the Tevatron

**Dibosons**

Observation of Bosons at Hadron Colliders

- $W$
- $Z$
- $W\gamma$
- $Z\gamma$

1983 CERN

1995 Fermilab

WW 2005 Fermilab

WZ 2007 Fermilab

ZZ 2008 Fermilab

WH/ZH

(a) D0 8.6 fb$^{-1}$

(b) D0 8.6 fb$^{-1}$

(c) D0 8.6 fb$^{-1}$
To understand VH Production:

lepton+jets final states

\( l = e/\mu \) + jets

- 5-10 \times \) leptonic BR
- large background from jets
- significant systematic effects
- dijet mass resolution:
  \( |M_Z - M_W| < \text{D0 detector resolution} \)
- WW+WZ

Evidence by DØ (4.4\( \sigma \)) 2009
Observation by CDF in 2010

Cross Sections

<table>
<thead>
<tr>
<th>WW</th>
<th>WW+WZ</th>
<th>WW part</th>
</tr>
</thead>
<tbody>
<tr>
<td>( l\nu l\nu ) [pb]</td>
<td>( l\nu jj ) [pb]</td>
<td>( l\nu jj ) [pb]</td>
</tr>
<tr>
<td>4.3/fb</td>
<td>12.1 ± 1.8</td>
<td>18.1 ± 4.1</td>
</tr>
<tr>
<td>1.1/fb</td>
<td>11.5 ± 2.2</td>
<td>20.2 ± 4.6</td>
</tr>
</tbody>
</table>

\( \text{SM@NLO} \sigma(\text{WW}) = 11.7 ± 0.8 \text{ pb} \)
\( \text{SM@NLO} \sigma(\text{WZ}) = 3.5 ± 0.3 \text{ pb} \)
**Benchmarks**

- **WW+WZ→lνjj as a benchmark**
  First evidence of the dijet mass resonance from WW+WZ

\[ \sigma(WW + WZ) = 20.2 \pm 4.6 \text{ pb} \]

SM@NLO: \( \sigma(WW + WZ) = 16.1 \pm 0.9 \text{ pb} \)

SM@NLO: (l = e+μ)
\[ \sigma_{WW} \times \text{BR}(W \rightarrow l\nu; W \rightarrow jj) \approx 3.5 \text{ pb} \]
\[ \sigma_{WZ} \times \text{BR}(W \rightarrow l\nu; Z \rightarrow jj) \approx 0.6 \text{ pb} \]

- **WZ→lνjj as a benchmark**
  For WH, ZH (WZ→lνbb)

Cross Section Normalization

Components of the WH/ZH searches:
b-tagging, event selection, multivariate analysis, statistical treatment, systematic uncertainties, background modeling

\[ \sigma_{WZ} \times \text{BR}(W \rightarrow l\nu; Z \rightarrow bb) \approx 4 \sigma_{WH} \times \text{BR}(W \rightarrow l\nu; H@115 \rightarrow bb) \]
Event Selection, Signal and Background

Data: 4.3 fb\(^{-1}\) of Run lib Data
Trigger Selection: electron(+jets) and inclusive trigger (muon)

**Electron:**
- \(p_T \geq 15\) GeV, \(|\eta|_{\text{DET}} < 1.1\)

**Muon:**
- \(p_T \geq 20\) GeV, \(|\eta|_{\text{DET}} \leq 1.6\), 
- \(\Delta R_{\mu\text{-jet}} > 0.5\)

**Global**
- \(\text{MET} \geq 20\) GeV, \(M_T \geq 40 - 0.5 \cdot \text{MET}\)
- \(M_T (\mu\nu) < 200\) GeV, \(|\text{PV}_Z| < 60\) cm

**Jets (2 or 3 jet bin):**
- Minimum 2 vertex confirmed jets \(p_T^{\text{Jet1,Jet2}} \geq 20\) GeV, \(|\eta|_{\text{DET}} < 2.5\)
- calibrated jets (data and MC)
- Resolution calibration in MC
correction to jet/Z \(p_T\) imbalance and
energy resolution due to the different
quark/gluon sample composition

**Event Source** | **Generator** | \(\sigma(\text{SM}) / \sigma(\text{WW})\) | \(\sigma(\text{WW}) = 11.7\) pb
---|---|---|---
WW | Pythia | 1.0 | NLO
WZ | Pythia | 0.3 | NLO
ZZ | Pythia | 0.1 | NLO
W+light flavor jets | Alpgen | 850 | from FIT
W+heavy flavor jets | Alpgen +Pythia | 32 | from FIT
Z+light flavor jets | Alpgen | 32 | NNLO
Z+heavy flavor jets | Alpgen | 1.1 | NNLO
Double-Top | Alpgen +Pythia | 0.6 | NNLO
Single-Top | Comphep +Pythia | 0.2 | NNLO

**Events**
- Data: 100 k
- W/Z+light f. 70 k
- W/Z +Heavy f. 10 k
- Top 3 k
- Multijet 13 k
- Diboson 3.2 k
Muon channel, 2 jet bin, normalized with a scale factor to match the data (MC\times 1.04)
To separate bb-decays we use a Neural Network b-tagging algorithm.

NN b-tag output: $\geq$ L6 for two leading jets (12 identification points)

$<$ L6 or non-taggable: 0-tag bin

NN b-tag output as an input to a MVA (Random Forest)
Random Forest (RF)
Input: 15 well described variables to separate signal from background (including min. and max. of b-tag NN output, cf backup slides)

- **0-tag**: trained with all dibosons as signal
- **1-tag**: trained with WZ+ZZ as signal
- **2-tag**: trained with WZ+ZZ as signal

after a combined fit to data over all jet/tag sub-channels
Total Diboson cross section

Obtained from the fit to the **Dijet Mass Distributions** or to the Random Forest output, minimizing a $\chi^2$ function with respect to variations in the systematic uncertainties

(simultaneous fit of the electron and muon distributions in the 0, 1, and 2-tag sub-channels, and 2- and 3- jet bins)

- **W+jets and diboson (WW+WZ)** normalizations are free parameters

<table>
<thead>
<tr>
<th></th>
<th>Electron channel</th>
<th>Muon channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diboson signal</td>
<td>$1725 \pm 84$</td>
<td>$1465 \pm 67$</td>
</tr>
<tr>
<td>$W/Z+$light-flavor jets</td>
<td>$37200 \pm 1000$</td>
<td>$33520 \pm 710$</td>
</tr>
<tr>
<td>$W/Z+$heavy-flavor jets</td>
<td>$5370 \pm 610$</td>
<td>$4850 \pm 490$</td>
</tr>
<tr>
<td>$t\bar{t}$ + single top</td>
<td>$1750 \pm 130$</td>
<td>$1214 \pm 86$</td>
</tr>
<tr>
<td>Multijet</td>
<td>$10600 \pm 1000$</td>
<td>$1980 \pm 380$</td>
</tr>
<tr>
<td>Total predicted</td>
<td>$56700 \pm 720$</td>
<td>$43030 \pm 620$</td>
</tr>
<tr>
<td>Data</td>
<td>$56698$</td>
<td>$43044$</td>
</tr>
</tbody>
</table>

ZZ contribution to VV $\approx 1.5\%$

Measured $\sigma(WV)$ [pb]:

<table>
<thead>
<tr>
<th>RF Output</th>
<th>Measured $\sigma(WV)$ [pb]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dijet Mass</td>
<td>$18.3 \pm 1.5$ (stat) $^{+3.5}_{-3.3}$ (syst)</td>
</tr>
</tbody>
</table>

**SM@NLQ $\sigma(WW+WZ) = 16.1 \pm 0.9$ pb**
Total Diboson cross section

Obtained from the fit to the Dijet Mass Distributions or to the Random Forest output, minimizing a $\chi^2$ function with respect to variations in the systematic uncertainties (simultaneous fit of the electron and muon distributions in the 0, 1, and 2-tag sub-channels, and 2- and 3-jet bins)

- $W$+jets and diboson ($WW+WZ$) normalizations are free parameters

<table>
<thead>
<tr>
<th></th>
<th>Electron channel</th>
<th>Muon channel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diboson signal</td>
<td>1725 ± 84</td>
<td>1465 ± 67</td>
</tr>
<tr>
<td>$W/Z$+light-flavor jets</td>
<td>37200 ± 1000</td>
<td>33520 ± 710</td>
</tr>
<tr>
<td>$W/Z$+heavy-flavor jets</td>
<td>5370 ± 610</td>
<td>4850 ± 490</td>
</tr>
<tr>
<td>$t\bar{t}$ + single top</td>
<td>1750 ± 130</td>
<td>1214 ± 86</td>
</tr>
<tr>
<td>Multijet</td>
<td>10600 ± 1000</td>
<td>1980 ± 380</td>
</tr>
<tr>
<td>Total predicted</td>
<td>56700 ± 720</td>
<td>43030 ± 620</td>
</tr>
<tr>
<td>Data</td>
<td>56698</td>
<td>43044</td>
</tr>
</tbody>
</table>

Measured $\sigma(WV)$ [pb]

- RF Output: $19.6 \pm 1.4$ (stat) $^{+2.9}_{-2.7}$ (syst)
- Dijet Mass: $18.3 \pm 1.5$ (stat) $^{+3.5}_{-3.3}$ (syst)

$\text{SM@NLO } \sigma(WW+WZ) = 16.1 \pm 0.9 \text{ pb}$
\[ \sigma_{RF}(WW + WZ) = 19.6 \pm 3.2 - 3.0 \text{ pb} \Rightarrow \text{significance} : 7.9\sigma \text{ (expected} : 5.9\sigma) \]
\[ \sigma_{MJJ}(WW + WZ) = 18.3 \pm 3.8 - 3.6 \text{ pb} \Rightarrow \text{significance} : 5.6\sigma \text{ (expected} : 4.6\sigma) \]

Significance gain from the Multi Variate Approach: 22%

**Significance:**

\[ \Delta \chi^2 = -2 \ln \left( \frac{L(D; S + B, \theta_k)}{L(D; B, \theta_k)} \right) = \left| \chi^2(D; S + B, \theta_k) - \chi^2(D; B, \theta_k) \right| \]

D: data or background-only
**WW & WZ cross sections**

From the fit to data of the Random Forest output (or the Dijet Mass Distributions), minimizing a $\chi^2$ function with respect to variations in the systematic uncertainties (simultaneous fit of the electron and muon distributions in the 0, 1, and 2-tag sub-channels, and 2- and 3-jet bins)
**WW & WZ cross sections**

From the fit to data of the Random Forest output (or the Dijet Mass Distributions), minimizing a $\chi^2$ function with respect to variations in the systematic uncertainties (simultaneous fit of the electron and muon distributions in the 0, 1, and 2-tag sub-channels, and 2- and 3- jet bins)

Taking $W+\text{jets}$, WW and WZ normalizations are free parameters:

$$\sigma_{RF}(WW) = 15.9 \pm 3.7 \text{ pb}$$

$$\sigma_{RF}(WZ) = 3.3 \pm 4.1 \text{ pb}$$

SM@NLO $\sigma(WW) = 11.7 \pm 0.8 \text{ pb}$

SM@NLO $\sigma(WZ) = 3.5 \pm 0.3 \text{ pb}$

Constraining WW:

$$\sigma_{WW \text{ fixed}}^{RF} (WZ) = 6.5 \pm 3.1 \text{ pb}$$

observed significance: 2.2$\sigma$ (expected: 1.2$\sigma$)
Summary

- DØ analyzed $l\nu jj$ final states ($e, \mu$) using 4.3 fb$^{-1}$ of RunIIb data.
- Kinematic distributions in good agreement with the SM predictions.
- Measured WW+WZ cross section in agreement with the SM, using similar techniques as in the Higgs WH analysis \( \Rightarrow \) yields 7.9σ observed significance.
- Extraction of the WZ signal (2.2 σ observed).

<table>
<thead>
<tr>
<th>Cross Sections</th>
<th>WW</th>
<th>WW+WZ</th>
<th>WW part</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$l\nu l\nu$ [pb]</td>
<td>$l\nu jj$ [pb]</td>
<td>$l\nu jj$ [pb]</td>
</tr>
<tr>
<td>4.3/fb</td>
<td>12.1 ± 1.8</td>
<td>18.1 ± 4.1</td>
<td></td>
</tr>
<tr>
<td>1.1/fb</td>
<td>11.5 ± 2.2</td>
<td>20.2 ± 4.6</td>
<td></td>
</tr>
<tr>
<td>4.3/fb</td>
<td>19.6$^{+3.2-3.0}$</td>
<td>15.9$^{+3.7-3.2}$</td>
<td></td>
</tr>
</tbody>
</table>
Backup Slides
Additional MC Corrections

ALPGEN jet $\eta$ Reweighting

Correction to W/Z+jets MC samples [shape $\sim \frac{\text{data-QCD-nonVjets}}{Vjets}$]

Before correction

After correction
Backup: RF Input Variables (1)
Backup: RF Input Variables (II)
## Modeling of SM processes

### RunIIb1 MC

<table>
<thead>
<tr>
<th>Event Source</th>
<th>Generator</th>
<th>$\sigma$(SM) / $\sigma$(WW)</th>
<th>$\sigma$(WW) = 11.7 pb</th>
</tr>
</thead>
<tbody>
<tr>
<td>WW</td>
<td>Pythia</td>
<td>1.0</td>
<td>NLO</td>
</tr>
<tr>
<td>WZ</td>
<td>Pythia</td>
<td>0.3</td>
<td>NLO</td>
</tr>
<tr>
<td>ZZ</td>
<td>Pythia</td>
<td>0.1</td>
<td>NLO</td>
</tr>
<tr>
<td>W+light flavor jets</td>
<td>Alpgen</td>
<td>850</td>
<td>from FIT</td>
</tr>
<tr>
<td>W+heavy flavor jets</td>
<td>Alpgen</td>
<td>32</td>
<td>from FIT</td>
</tr>
<tr>
<td>Z+light flavor jets</td>
<td>Alpgen</td>
<td>+Pythia</td>
<td>32</td>
</tr>
<tr>
<td>Z+heavy flavor jets</td>
<td>Alpgen</td>
<td>+Pythia</td>
<td>1.1</td>
</tr>
<tr>
<td>Double-Top</td>
<td>Alpgen</td>
<td>+Pythia</td>
<td>0.6</td>
</tr>
<tr>
<td>Single-Top</td>
<td>Comphep</td>
<td>+Pythia</td>
<td>0.2</td>
</tr>
</tbody>
</table>

### Standard MC Corrections

(to account for differences from data)

- Reconstruction and Identification efficiencies of leptons/jets
- Trigger selection
- Z boson $p_T$ modeling (njet-dependent)
- W boson $p_T$ modeling (inclusive)
- Luminosity reweighting, beam z-position reweighting

### Multijet Background

(jet misidentified as a lepton)

- Estimated from (multijet enriched) data
  - Muon channel: Reverse muon isolation cuts
  - Electron channel: Loose electron quality criteria
- Corrected for contributions already accounted for by MC
- Normalization: template fit of $M_T(W\rightarrow l\nu)$
**Additional MC Corrections**

### Inclusive Muon Trigger Modeling

- Inclusive trigger gives ~50% increase in statistics on Single Muon OR (additional data collected by jet triggers mainly)
- Modeling: bootstrap from SMOR parameterized in HT ($p_{T}^{j1}+p_{T}^{j2}$)
- Correction: data driven, as a function of HT

\[
\varepsilon^{\text{Inclusive}} = \varepsilon^{\text{SMOR}} + \varepsilon^{\text{JETS (correction)}} \quad (\text{max } \varepsilon^{\text{JETS}} < 1 - \varepsilon^{\text{SMOR}})
\]

\[
\varepsilon^{\text{JETS}} = \left[ \frac{N_{\text{data}}^{\text{Inclusive}} - QCD_{\text{Inclusive}} - N_{\text{SMOR}}^{\text{data}} - QCD_{\text{SMOR}}}{N_{\text{MC}}^{\text{Inclusive}}} \right]
\]

- **muon, 2 jet bin**
- **muon, 3+4 jet bin**
Additional MC Corrections

Inclusive Muon Trigger Modeling

- Inclusive trigger gives ~50% increase in statistics on Single Muon OR (additional data collected by jet triggers mainly)
- Modeling: bootstrap from SMOR parameterized in HT ($p_T^{j1}+p_T^{j2}$)
- Correction: data driven, as a function of HT

$$\epsilon^{\text{Inclusive}} = \epsilon^{\text{SMOR}} + \epsilon^{\text{JETS \ (correction)}} \ (\text{max} \ \epsilon^{\text{JETS}} < 1 - \epsilon^{\text{SMOR}})$$

$$\epsilon^{\text{JETS}} = \left[ \frac{N_{\text{data}}^{\text{Inclusive}} - \text{QCD}_{\text{Inclusive}} - N_{\text{SMOR}}^{\text{data}} - \text{QCD}^{\text{SMOR}}}{N_{\text{MC}}^{\text{Inclusive}}} \right]$$

muon, 2 jet bin

muon, 3+4 jet bin
(1) Unclustered Energy (UE) Reweighting

- Sum of the energy deposits in EM, HAD, ICD not associated to a jet
- Overlaid MB, low $p_T$ objects (jets, muons) not accounted for in missing $E_T$
- Correction to all MC samples [shape $\sim (data-QCD)/MC$]
(2) ALPGEN jet $\eta$ Reweighting
To separate bb-decays we use NN b-tagging algorithm.

Taggable jets are good vertex confirmed jets ($\varepsilon_{\text{data}} < \varepsilon_{\text{MC}}$)

MC correction for each $\eta_{\text{jet}}$ in different $z_{\text{PV}}$ bins:

- $(z \leq -30)$
- $(-30 < z \leq 0)$
- $(0 < z \leq 30)$
- $(30 < z)$

$$SF_{\text{taggable}}(z,\eta) = \frac{\varepsilon_{\text{data}}(z,\eta)}{\varepsilon_{\text{MC}}(z,\eta)}; \quad SF_{\text{non-taggable}}(z,\eta) = \frac{1 - \varepsilon_{\text{data}}(z,\eta)}{1 - \varepsilon_{\text{MC}}(z,\eta)}$$
Electron channel, 2 jet bin, normalized with a scale factor to match the data (MC×0.95)

Muon channel, 2 jet bin, normalized with a scale factor to match the data (MC×1.04)
(2) ALPGEN jet $\eta$ Reweighting

- Correction to W/Z+jets MC samples [shape $\sim$ (data-QCD-nonVjets)/Vjets]

Also derived for $\sigma(VV)\times0$ and $\sigma(VV)\times2$ (fit systematics)
Additional MC Corrections (cont’d)

(4) ALPGEN parton-jet $p_T$ Matching Reweighting

- minimum $p_T$ for jet clusters that are used for the MLM jet-parton matching procedure
- recommended: the generator level jet $p_T$ cut + 20% (or 5 GeV if larger)

(5) Diboson Modeling

- LO-to-NLO correction at the generator level using MC@NLO+Herwig
- 2D reweighting ($p_T$ VV - leading V) applied to WW and WZ events
- Systematic uncertainty is half the size of the correction

I. Each reweighting preserves the total normalization of the MC that is being reweighted

II. Two systematic uncertainties for each reweighting:

- From the fit uncertainty; vary the fit parameter that causes the most change in shape, using the covariance matrix to change the other parameters accordingly
- The uncertainty on the diboson cross section (0 and $2 \times \sigma_{SM}$), taken as correlated between reweightings