Measurement of the W boson helicity in top quark decay at DØ

Bertrand Martin (LPSC Grenoble) for the DØ collaboration

APS April meeting – 04/15/07
Motivations

Analysis guideline:
- dilepton
- lepton + jets

Analysis technique:
- $\cos\theta^*$ templates

Results of $f_+$ measurement
- Bayesian C.L.
Motivations

Standard Model:

Due to the observed Parity violation, charged current (W boson) only couples left handed particles 

V-A structure of the EW current in the S.M. Lagrangian

Physics observable sensitive to a possible V+A component: W boson helicity

Helicity is measured through the $\cos(\theta^*)$ distribution

$\theta^*$: angle between the top quark flight direction and the charged lepton momenta in the W rest frame:

3 components in the $\cos(\theta^*)$ distribution: 3 helicity states

Fraction \[
\begin{align*}
\text{Left handed:} & \quad f_- \sim 0.30 \\
\text{Longitudinal:} & \quad f_0 \sim 0.70 \\
\text{Right handed:} & \quad f_+ \sim 1.4 \times 10^{-3}
\end{align*}
\]

A non zero $f_+$ could sign new physics...
Select a data sample enriched in ttbar candidate events
(estimate physics and instrumental background contamination)

For each selected event,
- reconstruct the top quark & W boson leptonic decay
- compute $\cos(\theta^*)$

Compare the $\cos(\theta^*)$ distribution obtained in data to different signal hypotheses:
MC ttbar samples with:
- fixed $f_0$ (0.70)
- different $f_+$ values

Realize pseudo-experiments to estimate systematic uncertainties

which one is the most compatible with the observed data?
Event selection

Dilepton (ee, eμ, μμ):

Kinematics and topology
- 2 high $p_T$ leptons (opposite charge)
- $\geq 2$ high $p_T$ jets
- $M_\perp$ outside the $Z$ mass (ee, μμ)
- significant Missing $E_T$ (2 ν)
- sphericity (ee), $H_T$ (eμ)

Main backgrounds:
- Drell-Yan: Z/γ* + jets
- Diboson (WW, WZ, ZZ)
- Fake lepton

Lepton (e, μ) + jets:

Multivariate selection
- Only 1 high $p_T$ lepton
- $\geq 4$ high $p_T$ jets
- Missing $E_T$ (1 ν)
- Likelihood discriminant (to suppress W+jets)

Main backgrounds:
- W + jets
- QCD multijet production

See next slide
Likelihood discriminant ($L_t$) in $l+\text{jets}$

The variables used to discriminate signal (S) and background (B) must:

- be **well modeled** in the MC (K.S. proba > 5%)
- have **different shapes** between S and B

\[
L_t \sim \frac{S}{S + B} \sim \frac{\sum_{i=1}^{N \text{ var}} \ln \left( \frac{s}{b} \right)^{\text{fit}}}{e^{\sum_{i=1}^{N \text{ var}} \ln \left( \frac{s}{b} \right)^{\text{fit}}} + 1}
\]

Among $2^{11} - 1 = 2047$ possible $L_t$, the **best** one gives the **smallest error** on the measured $f_+$. 

11 “good” variables (kinematics, b-tagging...)

- Good MC modeling
- Discriminating?
- Fit ln(s/b)

Efficiency for best $L_t$

<table>
<thead>
<tr>
<th>Source</th>
<th>$\mu+$jets</th>
<th>$e+$jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>$tt$</td>
<td>$0.72 \pm 0.29$</td>
<td>$0.76 \pm 0.15$</td>
</tr>
<tr>
<td>$Wjjjj$</td>
<td>$0.04 \pm 0.004$</td>
<td>$0.07 \pm 0.02$</td>
</tr>
<tr>
<td>QCD</td>
<td>$0.12 \pm 0.17$</td>
<td>$0.10 \pm 0.02$</td>
</tr>
</tbody>
</table>
The \[\cos(\theta^*)\] distribution is built for:

- Data
- Signal for different \(f_+\) (\(V-A / V+A\))
- Background

**Likelihood maximization**: find which \(f_+\) value best reproduces the data distribution
Results with 370 pb$^{-1}$

Bayesian confidence level (CL %) : use a prior probability density

- flat for $f_+ \in [0, 0.30]$ and null in the non-physical region

Confidence interval $[x_{\text{min}}, x_{\text{max}}]$ based on the likelihood integral, such that:

- \[ f_{+}^{1+\text{jets}} = 0.11 \pm 0.09 \text{ (stat)} \]
- \[ f_{+}^{\text{dilepton}} = -0.09 \pm 0.15 \text{ (stat)} \]
- \[ f_{+}^{\text{comb}} = 0.056 \pm 0.080 \text{ (stat)} \]

The likelihood maximization does not guarantee $f_+ > 0$ !

Max. likelihood

\[ \int_{x_{\text{min}}}^{x_{\text{max}}} L(x)dx = \frac{CL}{2} = \frac{\int_{0}^{0.30} L(x)dx}{\int_{0}^{0.30} L(x)dx} \]

Likelihood function

@ 95% of confidence level:

- $0 < f_{+}^{1+\text{jets}} < 0.264$
- $0 < f_{+}^{\text{dilepton}} < 0.239$
- $0 < f_{+}^{\text{comb}} < 0.226$
With 370 pb\(^{-1}\) of analyzed data, the combined lepton+jets and dilepton measurements of the right handed W fraction \(f_+\) is:

(assuming \(f_0 = 0.70\))

\[
f_+ = 0.056 \pm 0.080 \text{ (stat)} \pm 0.057 \text{ (syst)}
\]

\[
f_+ < 0.226 \text{ @95\%C.L.}
\]

This measurement is compatible with the predicted Standard Model value:

\[
f_+ = 1.36 \times 10^{-3}
\]

The analysis is currently updated with an integrated luminosity of \(\sim 1 \text{ fb}^{-1}\)

The CDF preliminary results for lepton + jets with \(\sim 1 \text{ fb}^{-1}\) are:

- \(f_+ = -0.03 \pm 0.06 \text{ (stat)} \pm 0.03 \text{ (syst)}\) assuming \(f_0 = 0.70\) (\(f_+ < 0.10 \text{ @ 95\% C.L.}\))
- \(f_0 = 0.59 \pm 0.12 \text{ (stat)} \pm 0.07 \text{ (syst)}\) assuming \(f_+ = 0\)
Backup slides
Results on the W helicity will be shown for 370 pb\(^{-1}\) of analyzed data, while D0 has more than 2 fb\(^{-1}\) on tape.
Likelihood maximization (w.r.t. $n_s$ and $n_{b,i}$):

$$L(f_+) = \prod_{i=1}^{N_{bkg}} e^{\frac{(n_{b,i} - \bar{n}_{b,i})^2}{2\sigma_{b,i}^2}} \times \prod_{j=1}^{N_{bins}} P(d_j; n_j)$$

- $\bar{n}_{b,i}$ and $\sigma_{b,i}$ obtained after final selection:
  - kinematics & topo: dilepton
  - $L_t$ discriminant cut: lepton + jets

- Poissonian probability to observe in the bin $j$:
  - $d_j$ data events with a predicted average of
  $$n_j(f_+) = n_s(f_+) + \sum_{i=1}^{N_{bkg}} n_{b,i}$$

- How well does this $f_+$ hypothesis match with the $\cos \theta^* \text{ data distribution}$

- Gaussian term for the background normalization
Ensemble tests

Test of the maximum likelihood performance

Create a “pseudo-dataset” of MC events with:

- the same number of MC events as observed in the data
- the signal/background composition can fluctuate according to a binomial distribution ($n_{bkg} = N_{tot}^{\text{observed}} - n_s$)

Compare the fitted $f_+$ to the known input $f_+$

Repeat the procedure 1000 times for each $f_+$ value

Evaluation of systematic uncertainties

- Varying parameters can affect both the data sample composition (different selection efficiency of the likelihood discriminant) and the shape of cos($\theta^*$) distributions.

- Effect on the fitted $f_+$: studied with pseudo-experiments (varying the parameters in the pseudo-dataset)

- Source: Jet Energy Scale, $M_{top}$, MC statistics, heavy flavor content (W+jets), …

$\Delta f_+ \sim 0.03$ to $0.04$ (for each one)