Physics background at Tevatron

Andrea Messina
INFN - Roma

III Workshop Italiano sulla Fisica di ATLAS e CMS
Bari, 20-22 Ottobre 2005
Tevatron

- Proton-Antiproton collisions
- $\sqrt{s} = 1.96$ TeV
- 36 bunches: 396 ns crossing time
- Peak luminosity $1.4 \times 10^{32} \text{cm}^{-2}\text{s}^{-1}$

- 1 fb\(^{-1}\) per experiment on tape
- 1.3 fb\(^{-1}\) delivered luminosity
Run II upgrades
- New silicon tracking
- New drift chamber
- Upgraded muon chambers
- New plug calorimeters
- New TOF

Data taking efficiency 85 %
About 1 fb\(^{-1}\) on tape
Introduction

Larger data sample than ever

Measurements start being dominated by systematic errors

For high $p_T$ physics is crucial to have under control the backgrounds and the MonteCarlo tools for signal & backgrounds modeling
Introduction

**EWK Physics:**
- large cross sections (nb)
- small B/S (1-5%)
- limited by detector syst.

right place to develop, understand and test the algorithm for the background identification and subtraction to: e, \( \mu \) and Met signals

**QCD Physics:**
- large cross sections (nb)
- deeply falling with jet \( p_T \)
- limited by jet correction

understand jet energy scale, MI, UE. Now is crucial to tune MC on data to have a good description of UE & soft physic

**Top & Higgs:**
- small cross sections (pb)
- small S/B (O(1))
- limited by jet correction & physics bkgd modeling

- model Boson + jets bkgd
- measure HF fraction (W/Z+jet)
- understand non W background
**Electroweak Physics: W**

**W -> e\nu**

<table>
<thead>
<tr>
<th></th>
<th>Frac.</th>
<th>Meth.</th>
<th>Syst.</th>
</tr>
</thead>
<tbody>
<tr>
<td>W-&gt;\tau\nu</td>
<td>1.9%</td>
<td>MC</td>
<td>5%</td>
</tr>
<tr>
<td>QCD</td>
<td>1.6%</td>
<td>Data</td>
<td>50%</td>
</tr>
<tr>
<td>Z-&gt;ee</td>
<td>0.8%</td>
<td>MC</td>
<td>5%</td>
</tr>
<tr>
<td><strong>Tot.</strong></td>
<td>4.4%</td>
<td></td>
<td>20%</td>
</tr>
</tbody>
</table>

**W -> \mu\nu**

<table>
<thead>
<tr>
<th></th>
<th>Frac.</th>
<th>Meth.</th>
<th>Syst.</th>
</tr>
</thead>
<tbody>
<tr>
<td>W-&gt;\tau\nu</td>
<td>3.1%</td>
<td>MC</td>
<td>5%</td>
</tr>
<tr>
<td>QCD</td>
<td>0.7%</td>
<td>Data</td>
<td>50%</td>
</tr>
<tr>
<td>Z-&gt;\mu\mu</td>
<td>5.4%</td>
<td>MC</td>
<td>5%</td>
</tr>
<tr>
<td><strong>Tot.</strong></td>
<td>9.4%</td>
<td></td>
<td>5%</td>
</tr>
</tbody>
</table>

QCD bkgd to the W signal is the dominant bkgd to a number of analyses, need to be characterized where there is a good understanding of the other components.

Background dominated by \(\mu\) that escaping the detector mimic a Met signal.
Background from data: QCD

Missidentified leptons + large unbalanced $E_T$

Main technique: correlation between lept. Iso & Met

QCD uncorrelated on IsoMet plane

Extrapolate bkgd from low Met region

Correct for other boson bkgd:

- $Z\rightarrow ee$ (peaks reg B)
- $W\rightarrow tn$ (peaks reg C)

Signal contamination (reg B&C)
Systematic & alternatives

Systematic:
Vary the region definition

Fake method:
Di-jet events $E_T > 15$ GeV
$\text{Met} < 15\text{GeV} \rightarrow$ negligible EWK contamination.

EWK corrected
Raw method

Met for all jet weighted by e fake rate
QCD background composition

Electron identification

- Conv. removal (trident allowed)
- E/p < 2 if p<50 GeV
- Isol(0.4)< 0.1 (energy based)
- Had/EM < 0.055 + c(E)
- Good track
- Matching track EM cluster
- Lateral EM cluster shape

<table>
<thead>
<tr>
<th>Source</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conv.</td>
<td>60</td>
</tr>
<tr>
<td>Had. (ch.ex, overlap)</td>
<td>20</td>
</tr>
<tr>
<td>Heavy flavor</td>
<td>20</td>
</tr>
</tbody>
</table>
A jet is a composite object:

Complex detector properties
- non-linear detector properties
- non-instrumented regions

Complex underlying physics
- events contain spectator interaction
- processes connected via color
- hadron fragmentation
- different type of jets: q, g, b/c

Correct to particle level
- pile up
- detector efficiency/resolution

Model dep. correction
- underlying event
- parent parton energy

Data + Unfolding ↔ Jet Algorithms ↔ pQCD + Soft contribution
The Underlying Event

A typical Tevatron event consists of:

- hard interaction
- initial/final radiation
- interaction between remnants
- multiple parton interactions

Precise measurements at low Pt require good modeling of the underlying event.

Tune MonteCarlo to reproduce data in as many as possible variables sensitive to underlying physics.
Underlying Event Studies

“Transverse” region is very sensitive to the “underlying event”!

Important to model:
- Soft particle multiplicity
- Particle average energy

Study:
- $E_T$, $p_T$, charged particle density

Charged Particle Density: $dN/d\Delta\phi$

CDF Preliminary data uncorrected

30 < ET(jet#1) < 70 GeV

"Transverse" Region

Jet #1 Direction
- "Leading Jet"
- "Back-to-Back"

Jet #2 Direction
- "Toward" TransMAX
- "Away" TransMIN

Jet #3 Direction
- "Toward" TransMAX
- "Away" TransMIN

Study:
- $E_T$, $p_T$, charged particle density
Tune MonteCarlo on data

Plot shows the “Transverse” charged particle density versus $P_T(jet)$ compared to the QCD hard scattering predictions of two tuned versions of PYTHIA 6.206 (CTEQ5L, Set B (PARP(67)=1, less ISR) and Set A (PARP(67)=4), more ISR).

JIMMY was tuned to fit the energy density in the “transverse” region for “leading jet” events!

Shows the Etsum density in the “transMAX” and “transMIN” region versus $P_T(jet#1)$ for “Leading Jet” and “Back-to-Back” events.
Underlying event

We are making good progress in understanding and modeling the “underlying event”. We have PYTHIA Tune A and JIMMY tune A, however, we do not yet have a perfect fit to all the features of the “underlying event”. We are working on new improved Run 2 tunes!

Work in progress:
Look at the UE in other physics process Z, DY,...
Top physics

Top properties:
✓ Measure cross section
✓ Test QCD/SM
✓ Properties:
  ✓ different decay modes
  ✓ spin/W elicity
    ▼
      Larger sample
      Understand W + jet sample

Dilepton 5% : clean but small signal, 2ν
Lepton+jet 30%: 1ν, menageble background
Hadronic 44%: large background
τ + X 21%: τ ID is challenging

Top Mass:
Fundamental parameter SM
$M_{top} \approx$ EWSB scale
$M_{top}$ & $M_W$ constrain $M_{higgs}$
▼
Understand background(QCD/W+jet)
Understand jet energy scale
Xsec b-tagging (lepton+jet)

**Background**
- \(W_{bb}, W_{cc}, W_{c}\)
  - HF frac from MC
  - Normalized to \(W+\text{jets}\) data
- \(W+\text{light}\) (mistags)
  - Mistag from jet sample
  - Applied to \(W+\text{jets}\) data
- \(\text{Non-}W\)
  - From data
- Electroweak (\(WW, WZ, Z\rightarrow \tau \tau\))
  - From MC Small contribution

**Systematics**
- B-Tagging: 5.2%
- Lepton Isolation: 5%
- \(W+\text{Heavy Flavor}\): 4.4%
- Luminosity: 5.9%
- Total: 11%

**CDF Run II Preliminary (318 pb\(^{-1}\))**

\[ M_{\text{top}} = 178 \text{ GeV/c}^2 \]

\[ \sigma = 8.7^{+0.9}_{-0.9}^{+1.2}_{-0.9} \text{ pb} \]

\[ H_T > 200 \text{ GeV For } N_{jets} \geq 3 \]
Xsec no b tagging (lepton+jet)

Background:
ALPGEN+HERWIG $W + 3p$
and $W + 4p$ MC

$\mu$ and $W$-background float
QCD fixed to 4.6%

Seven input variables
$H_T$
Aplanarity
Maximum jet $\eta$
$\Sigma$ $E_T$ (Jets 3, 4, and 5)
$\Sigma p_T / \Sigma E_T$
Minimum dijet invariant mass
Minimum dijet separation ($\Delta R$)

<table>
<thead>
<tr>
<th>Sample</th>
<th>Events</th>
<th>Fitted $t\bar{t}$</th>
<th>$\sigma(t\bar{t})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$W + 3$ jets</td>
<td>936</td>
<td>148.2 ± 20.6</td>
<td>6.0 ± 0.8 ± 1.0 pb</td>
</tr>
<tr>
<td>$W + 4$ jets</td>
<td>210</td>
<td>80.9 ± 15.0</td>
<td>6.1 ± 1.1 ± 1.4 pb</td>
</tr>
</tbody>
</table>

CDF Preliminary (347 pb$^{-1}$)

$N_{jets} \geq 3$
signal: 148.2 ± 20.6 events
multijet: 43.2 ± 3.7 events
Wjets: 744.4 ± 32.8 events

Systematics | Total
---|---
Jet Et Scale | 8.3%
W+jets $Q^2$ Scale | 10.2%
t$\bar{t}$ PDF | 4.4%
Total | 16.4%

PYTHIA $t\bar{t}$ MC
**$M_{top}$ lepton+jet: 1D templates**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Jet $E_T$ cut [ GeV ]</th>
<th>S/B</th>
<th>$&lt;\text{bkg}&gt;$</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 b-tags</td>
<td>3 jets $E_T &gt; 15$, 4$^{th}$ jet $E_T &gt; 15$</td>
<td>18:1</td>
<td>0.7 ± 0.2</td>
<td>16</td>
</tr>
<tr>
<td>1 b-tag</td>
<td>(T) 4 jets $E_T &gt; 15$</td>
<td>4:1</td>
<td>7.6 ± 1.2</td>
<td>57</td>
</tr>
<tr>
<td>1 b-tag</td>
<td>(L) 3 jets $E_T &gt; 15$, 4$^{th}$ jet $8 &lt; E_T &lt; 15$</td>
<td>1:1</td>
<td>10.2 ± 1.7</td>
<td>25</td>
</tr>
<tr>
<td>0 b-tags</td>
<td>4 jets $E_T &gt; 21$</td>
<td>N/A</td>
<td>N/A</td>
<td>40</td>
</tr>
</tbody>
</table>

**$L_{\text{sample}} = L_{\text{shape}} \times L_{\text{bkg}}$**

- Background constrained to prediction (except for 0-tag)

**$M_{top} = 173.2^{+2.9}_{-2.8}$ (stat.) ± 3.1(JES) ± 1.5(syst.) GeV/c$^2$**
$M_{\text{top}}$ using $W \rightarrow jj$ template

- Sensitive to Jet Energy Scale (JES)
- Minimal sensitivity to $M_{\text{top}}$

$L_{\text{sample}} = L_{\text{shape}}^{M_{\text{top}}} \times L_{\text{shape}}^{m_{jj}} \times L_{\text{bkg}}$

$L_{\text{JES}} = \exp\left(-\frac{(\text{JES} - \text{JES}_{\text{STD}})^2}{2\sigma_{\text{JES}_{\text{STD}}}^2}\right)$

Find best $(M_{\text{top}}, \text{JES})$ to fit data
# $M_{\text{top}}$ Systematics

## Systematic Source

<table>
<thead>
<tr>
<th>Systematic Source</th>
<th>2-D fit $\Delta M_{\text{top}}$ (GeV)</th>
<th>1-D fit $\Delta M_{\text{top}}$ (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JES</td>
<td>2.5</td>
<td>3.1</td>
</tr>
<tr>
<td>b-jet modeling</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>ISR</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>FSR</td>
<td>0.6</td>
<td>0.4</td>
</tr>
<tr>
<td>PDFs</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>Generators</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Bkg shape</td>
<td><strong>1.1</strong></td>
<td><strong>1.0</strong></td>
</tr>
<tr>
<td>b-tagging</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>MC statistics</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td>Method</td>
<td>0.5</td>
<td>-</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>3.0 = 2.5 + 1.7</strong></td>
<td><strong>3.4 = 3.1 + 1.5</strong></td>
</tr>
</tbody>
</table>

- Jet Energy Scale dominates
- In-situ calibration improves with stat.

Where we are now

Soon the background systematic will dominate
The systematic on the top cross section and mass measurement receive a substantial contribution from the background knowledge.

How do we plan to improve the W + jets background knowledge.
W/Z/photon + jets cross sections are important measurements:

Bkgd for a number of high $p_T$ analyses:

- top cross section & top mass
  - $W$+jet/$W$+HF main background ~ 40%
- higgs searches
  - $Z$+b $\bar{b}$$\bar{b}$
  - $W$+b $\bar{b}$$\bar{b}$

QCD-Wise:

- high $q^2$ interactions ($\sim M($boson$)$)
- smaller subset of diagrams compare to jet production
- specific color flaw paths
- proton PDF ($W/Z + HF$)
**MC issues**

Naïve: $W + n\ p \ (\text{ME})+(\text{PS}) \sim W + \geq N\ \text{jet}$

$W+(n+1)\ p \ (\text{ME})+(\text{PS}) \sim W + \geq N+1\ \text{jet}$

...$

Problems:

- how to generate the whole $n$ jet spectrum avoiding double counting?
- how be sure that the selected jet is coming from the ME and not PS?

Goal:

- have a prescription to safely merge different MC multiplicity samples
- reduce the dependence on parton lavel phase-space
Next step: tune the new tools on data by making measurement model independent (out of cone correction) and easily usable to test new models.
...new merging tools

Separate multijet phase-space
- Matrix element domain
- Parton shower domain

LO ME calculation interfaced with parton shower MonteCarlo

MLM's matching
Michelangelo Mangano

CKKW prescription
Catani, Krauss, Kuhn, Webber

Have to be validated & tuned on data
MC validation on data

- Measure cross sections for boson + jets (k-factors)
- Differential cross sections wrt (hard scale merging parameter)
  - Jet-$E_T$ spectrum
  - Angular correlation jet-jet
  - Di-jet invariant mass
  - Boson $p_T$ spectrum

Be as less as possible model dependent. Provide clear results to be compared with theory

$\sigma(W|E_{\text{ele}}^{\tau}>20\text{GeV}, \text{Met}>30\text{GeV})$ is well defined both theoretically & experimentally

$\delta \sigma \left[ \begin{array}{l} P_T^{\nu}>20, M_T^{\nu}>20 \\ P_T^e>30, \eta^e<1.1 \end{array} \right]$

$\delta E_T^j$

Restrict the W cross-section to the measurable decay phase space

Jets corrected hadron level

JETCLU 0.4 (Midpoint)

$E_T^{\text{corr}} \geq 15 \text{ GeV}; \eta \leq 2.4$
Efficiency/Acceptance

- Defining cross-section phase space as W detector cut acceptance:
  - Acceptance will correct only for detector resolution effects - independent of theoretical model?
- Use W+np MC for acceptance and ID efficiency
  - Systematic on ID efficiency comparing Z MC and data
  - Systematic on acceptance from different MC models

Difference in the leptons kinematic model enhanced by correcting back to the whole W phase space

\[
A_{\text{Jet}} = \frac{N_{\text{Jet}}^{\text{Reco}}(E_T; P_T; E_X; WM_T; \eta^D; \text{fid})}{N_{\text{Jet}}^{\text{Gen}}(E_T; E_X; WM_T; \eta)}
\]
B/C Background Composition

$W \geq 1j$

$W \geq 2j$

$W \geq 3j$

$W \geq 4j$
For differential measurement a sample of QCD bkgd events is needed with a reliable shape.

Met.vs.Iso inadequate with jets:

- limited statistic for high jet multiplicity
- difficult to generalize for jet kinematic
- some correlation between anti-isolated ele & jet (for example in 1jet event)
- Met is sensitive to the event jet multiplicity
QCD background - part II

✓ form a fake electron sample by requiring passes “kinematic” cuts but fails at least 2 “ID” variables
  ✓ kinematic ET, PT, h, E/P, isolation
  ✓ pure ID Had/Em, chi2 track, matching track-cal

✓ ID variables will not introduce kinematic bias: the fake ele sample will be kinematically identical to fakes in data

✓ Correct the Met for jets by looking to the Z events
✓ Make the Met correction a function of the Z recoil energy than the jet multiplicity

✓ Bkgd rate by fitting the Met shape of antielectron & W-like bkgd to data
✓ Antielectron sample does contain EWK contamination 5%
Background Method

Max. likelihood fit of antielectron and MC model shapes to data - integrals above 30GeV give inclusive bkgds

QCD and MC models give excellent agreement with data in WM_T and electron E_T kinematics
**W/Z + HF jets**

Use the secondary vertex invariant mass to discriminate between b, c, light quark.

### W+ b jet

<table>
<thead>
<tr>
<th></th>
<th>$\sigma(W + b) / \sigma(W + \text{jet})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 jet</td>
<td>0.0169 ±0.0101</td>
</tr>
<tr>
<td>2 jet</td>
<td>0.0169 ±0.0101</td>
</tr>
</tbody>
</table>

### Z+ b jet

| $E_T^{\text{jet}}>20\text{GeV, }|p_T^{\text{jet}}|<1.5$ | CDF Prelim. Data | PYTHIA |
|---------------------------------|------------------|-------------|
| $\sigma(Z + b \text{ jet})$    | 0.96±0.32 ±0.14pb| 0.83pb      |
| $\sigma(Z + b \text{ jet})/\sigma(Z + \text{jet})$ | 0.0237 ±0.0078 ±0.033 | 0.0207      |
Conclusions

• To make precise measurement and to set stringent limits a good knowledge of the backgrounds is needed.
  - QCD, boson + jets and soft underlying physics are some of the high $p_T$ relevant backgrounds

• We are making progress in the MonteCarlo tuning to have a reliable description of relevant backgrounds
Backup slides
Separate multijet phase-space

- Matrix element domain
- Parton shower domain

1. Reweight the ME parton level inclusive N-jet rate to reproduce the probability of N-jet exclusive final state

2. Veto the showers with hard emission

---

1. Generate parton: $p_T > p_{T_{\text{Min}}}; \Delta R_{jj} > R_{\text{Min}}$

2. Perform parton showering

3. Reconstruct event with a cone algorithm ($E_{\text{T_{min}}}, R_{\text{jet}}$), before had.

4. Match partons with jets in $R$

5. If all partons are matched keep the event, else discard it

6. Define exclusive N-jet samples by requiring $N_{\text{jet}} = N_{\text{part}}$
Correct to parton level

MidPoint Cone Algorithm ($R = 0.7$, $f_{\text{merge}} = 0.75$)

Data corrected to the parton level

$0.1 < |y_{\text{jet}}| < 0.7$

Compared with NLO QCD ($\text{JetRad}$, $R_{\text{sep}} = 1.3$)

Contributions resulting from the underlying event and hadronization effects to the overall hadron to parton correction.

The corrections are derived using PYTHIA. The uncertainty of the correction was estimated by taking the difference between HERWIG and PYTHIA and is shown as the band in the plot.
### W mass systematic

<table>
<thead>
<tr>
<th>Systematic [MeV]</th>
<th>Electrons (Run 1b)</th>
<th>Muons (Run 1b)</th>
<th>Common (Run 1b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lepton Energy Scale and Resolution</td>
<td>70 (80)</td>
<td>30 (87)</td>
<td>25</td>
</tr>
<tr>
<td>Recoil Scale and Resolution</td>
<td>50 (37)</td>
<td>50 (35)</td>
<td>50</td>
</tr>
<tr>
<td>Backgrounds</td>
<td>20 (5)</td>
<td>20 (25)</td>
<td></td>
</tr>
<tr>
<td>Production and Decay Model</td>
<td>30 (30)</td>
<td>30 (30)</td>
<td>25 (16)</td>
</tr>
<tr>
<td>Statistics</td>
<td>45 (65)</td>
<td>50 (100)</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>105 (110)</strong></td>
<td><strong>85 (140)</strong></td>
<td><strong>60 (16)</strong></td>
</tr>
</tbody>
</table>
Higgs limits
Top Mass

Best Tevatron Run 2 Preliminary

- **D0 Dilepton**
  
  $L = 230 \text{ pb}^{-1}$
  
  $155.0 \pm 14.0 \pm 7.0$

- **CDF Dilepton**
  
  $L = 340 \text{ pb}^{-1}$
  
  $165.3 \pm 6.3 \pm 3.6$

- **D0 Lepton+Jets**
  
  $L = 320 \text{ pb}^{-1}$
  
  $169.5 \pm 3.0 \pm 3.6$

- **CDF Lepton+Jets**
  
  $L = 318 \text{ pb}^{-1}$
  
  $173.5 \pm 2.7 \pm 3.0$

**New TeV Average (prel.)**

$172.7 \pm 1.7 \pm 2.4$

**(Run1 + Run2)**

**Run 1 World Average**

**(Run 1 only)**

$178.0 \pm 2.7 \pm 3.3$
Top Cross Section

- Dilepton: Combined (L= 200pb⁻¹) 7.0 ± 2.4 ± 1.6 ± 0.4
- Lepton+Jets: Kinematic ABN (L= 347pb⁻¹) 6.3 ± 0.8 ± 0.9 ± 0.3
- Lepton+Jets: Soft Muon Tag (L= 193pb⁻¹) 5.3 ± 3.3 ± 1.3 ± 0.3
- Lepton+Jets: Vertex Tag (L= 318pb⁻¹) 8.9 ± 0.9 ± 1.1 ± 0.5
- MET+Jets: Vertex Tag (L= 311pb⁻¹) 6.1 ± 1.2 ± 1.3 ± 0.4
- All-hadronic: Vertex Tag (L= 311pb⁻¹) 8.0 ± 1.7 ± 3.3 ± 0.5
- Combined (L= 350pb⁻¹) 7.1 ± 0.6 ± 0.7 ± 0.4

σ(p¯p → t¯t) (pb)
b-tagging

Displaced tracks $\Rightarrow$ sec. vtx
Tag $\varepsilon$ per jet from physics MC (Pythia top)
Scale factor from di-jet Data/MC

SecVtx Tag Efficiency for Top $b$-Jets

- Tight SecVtx
- Loose SecVtx

Top MC scaled to match data
Only $b$-jets with $|\eta|<1$