SiD Benchmarking Analyses With b/c
Flavour Tagging

Tomáš Laštovička, University of Oxford
2009 Linear Collider Workshop of the Americas
Albuquerque, NM
Overview

Credits
LCFI Package
Higgs Boson Decay Branching Ratios
Top Quark Analysis
Sbottom Production Within Sbottom Co-annihilation Scenario
Remarks and Future Plans
Summary
Credits

- Simulation/Reconstruction
  - Tim Barklow, Norman Graf, Jan Strube
- Higgs Branching Ratios
  - Yambazi Banda (Oxford)
- Top Analysis
  - Erik Devetak (Oxford)
- Sbottom Production
  - Alexander Belyaev (Southampton)

+ Andrei Nomerotski and myself (Oxford)
LCFI PACKAGE
LCFI Package

- Used for jet flavour tagging and secondary vertex reconstruction.
- Topological vertex finder ZVRES.
- Standard LCIO input/output
  - Marlin environment (used for both ILD/SiD)
- Flavour tagging based on Neural Nets.
  - Combine several variables.
LCFI Package Optimisation for SiD LoI

- Default LCFI Neural Nets performed poorly with the full sim/rec SiD data.
- LoI Solution:
  - NN retrained and a different approach chosen (1 larger NN per tag, instead of 3 nets)
  - Package parameters not optimised due to very limited time and manpower constraints.

SiD LoI, full sim/dig/rec

Dashed (LCFI default) vs. re-trained NNs.
HIGGS BOSON DECAY BRANCHING RATIOS
Motivation

- Measure the Higgs branching ratio to $c\bar{c}$ by looking at the following channels:
  - $Z \rightarrow \nu\nu$, $H \rightarrow c\bar{c}$
  - $Z \rightarrow q\bar{q}$, $H \rightarrow c\bar{c}$

- High quality $c$-tagging required.

- Extend analysis further to $H \rightarrow b\bar{b}$ and $H \rightarrow gg$.
  - Finished, not a part of LoI.
Data Samples

- For data samples the following is assumed:
  - Centre-of-mass = 250 GeV (peak xsec for higgstrahlung)
  - Integrated luminosity = 250 fb$^{-1}$
  - Signal Higgs mass = 120 GeV
  - +80% e$^-$ polarization, -30% e$^+$ polarization
  - ~ 7 Million Standard Model background events
  - ~ 200 000 inclusive ZH signal events
  - Full simulation and reconstruction
Event Selection

1) **Classification** in two Z-decay modes
   - Neutrino channel (2 jets) and Hadronic Channel (4 jets)
   - Visible energy and a number of leptons cut

2) **Basic Event Selection**
   - Kinematic and topological cuts

3) **Neural Net** event selection
   - Based on 2 Neural Nets: 1\textsuperscript{st} trained to separate SM and ZH, and 2\textsuperscript{nd} to separate ZH-background and ZH-signal.
   - Inputs: Jet tags, basic selection variables, ..
   - Then cut on both NN\textsubscript{1} and NN\textsubscript{2} outputs simultaneously.
Results

Leading to combined BR uncertainty of about 8.5%.

Similar approach yields

- 4.5% for BR(H→bB) and 11.1% for BR(H→gg)
- ZH cross sections uncertainty is dominant for BR(H→bB)

Analyses still being developed.
TOP QUARK ANALYSIS
Data Samples

- Standard Model background sample
  - About 7M events, weighted

- bBfFfF sample
  - $M_{\text{top}} = 174$ GeV, 250k events
  - Signal (bBqQqQ) plus remaining background
  - Six jets, at least two of them are b-jets.

- bBfFfF template samples
  - $M_{\text{top}} = 174 (174.5, 173.5)$ GeV, each 1.1M events

All samples normalised to 500 fb$^{-1}$ and produced @ $\sqrt{s} = 500$GeV.
- Half of luminosity for -80/+30% polarisation, the other half for +80/-30%.
Event Selection

- Basic selection cuts:
  - 99.996% bkg rejection
  - 10% signal rejection eff.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Barrel</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{\text{tot}}$</td>
<td>&gt;</td>
<td>400 GeV</td>
</tr>
<tr>
<td>$\log(y_{56})$</td>
<td>&gt;</td>
<td>-8.5</td>
</tr>
<tr>
<td>number of particles in event</td>
<td>&gt;</td>
<td>80</td>
</tr>
<tr>
<td>number of tracks in event</td>
<td>&gt;</td>
<td>30</td>
</tr>
</tbody>
</table>

no isolated leptons

(a) log($y_{56}$) – Hadronic ttbar

(b) Total Energy – Hadronic ttbar

(a) Number of Particles – Hadronic ttbar

(b) Number of Tracks – Hadronic ttbar
Jet Flavour Tagging

- Good performance for six-jet events.
- Selection done based on a sum of NN outputs (b-tag only) of all jets.
Results – Top Quark Mass

- Kinematic fitting significantly improves the resolution.
Results – Top Mass Measurement Uncertainty

- Mass measurement uncertainty estimated using curve and template fits
  - Both give consistent numbers, around 50MeV
  - Template method preferred, stable and better $\chi^2$ behavior.

<table>
<thead>
<tr>
<th>Event Selection</th>
<th>Fit Range (GeV)</th>
<th>$\chi^2_{\text{min}}$/NDF</th>
<th>Mass (GeV)</th>
<th>$\sigma$ (GeV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Kinematic fit</td>
<td>120-200</td>
<td>148/159</td>
<td>174.135</td>
<td>0.090</td>
</tr>
<tr>
<td>No Kinematic fit</td>
<td>140-180</td>
<td>83/79</td>
<td>174.173</td>
<td>0.097</td>
</tr>
<tr>
<td>Kinematic fit</td>
<td>150-200</td>
<td>94/99</td>
<td>174.033</td>
<td>0.053</td>
</tr>
<tr>
<td>Kinematic fit</td>
<td>165-200</td>
<td>63/69</td>
<td>173.991</td>
<td>0.056</td>
</tr>
<tr>
<td>Kinematic fit</td>
<td>165-185</td>
<td>42/39</td>
<td>173.990</td>
<td>0.058</td>
</tr>
<tr>
<td>Probability &gt; 1%</td>
<td>150-200</td>
<td>101/99</td>
<td>174.018</td>
<td>0.049</td>
</tr>
<tr>
<td>Probability &gt; 1%</td>
<td>165-200</td>
<td>61/69</td>
<td>174.013</td>
<td>0.049</td>
</tr>
<tr>
<td>Probability &gt; 1%</td>
<td>165-185</td>
<td>41/39</td>
<td>174.010</td>
<td>0.053</td>
</tr>
<tr>
<td>Probability &gt; 5%</td>
<td>150-200</td>
<td>97/99</td>
<td>174.024</td>
<td>0.050</td>
</tr>
<tr>
<td>Probability &gt; 5%</td>
<td>165-200</td>
<td>61/69</td>
<td>174.017</td>
<td>0.050</td>
</tr>
<tr>
<td>Probability &gt; 5%</td>
<td>165-185</td>
<td>38/39</td>
<td>174.17</td>
<td>0.053</td>
</tr>
<tr>
<td>Probability &gt; 10%</td>
<td>150-200</td>
<td>100/99</td>
<td>174.012</td>
<td>0.050</td>
</tr>
<tr>
<td>Probability &gt; 10%</td>
<td>165-200</td>
<td>68/69</td>
<td>174.012</td>
<td>0.051</td>
</tr>
<tr>
<td>Probability &gt; 10%</td>
<td>165-185</td>
<td>40/39</td>
<td>174.14</td>
<td>0.052</td>
</tr>
<tr>
<td>Probability &gt; 20%</td>
<td>150-200</td>
<td>91/99</td>
<td>174.013</td>
<td>0.049</td>
</tr>
<tr>
<td>Probability &gt; 20%</td>
<td>165-200</td>
<td>68/69</td>
<td>174.010</td>
<td>0.050</td>
</tr>
<tr>
<td>Probability &gt; 20%</td>
<td>165-185</td>
<td>39/39</td>
<td>174.022</td>
<td>0.052</td>
</tr>
<tr>
<td>Probability &gt; 30%</td>
<td>150-200</td>
<td>98/99</td>
<td>174.021</td>
<td>0.049</td>
</tr>
<tr>
<td>Probability &gt; 30%</td>
<td>165-200</td>
<td>68/69</td>
<td>174.020</td>
<td>0.050</td>
</tr>
<tr>
<td>Probability &gt; 30%</td>
<td>165-185</td>
<td>47/39</td>
<td>174.027</td>
<td>0.052</td>
</tr>
</tbody>
</table>
Results – Cross Section and Production Asymmetry

- Cross section measurement
  - Estimated to about 0.5% precision
- Quark charge and forward backward asymmetries
  - Vertex charge, momentum weighted vertex and jet charges (LCFI)
  - For both t-quarks and they decay products b-quarks
  - Precision of about 0.008 reached for $A_{fb}$
SBOTTOM PRODUCTION
Motivation

- **Neutralino** is a very attractive CDM candidate.

- Cold Dark Matter favours some particular SUSY scenarios
  - one of them is co-annihilation scenario, when neutralino effectively co-annihilates with others quasi-degenerate SUSY particles into SM ones.

- Neutralino-sbottom co-annihilation scenario has not been studied previously.
  - This scenario is virtually impossible for LHC while feasible but challenging at the ILC.
  - The small mass split between neutralino and sbottom leads to small energy release and softness of the visible particles.
Why are Soft b-jets Difficult to Analyse?

i. Tagging efficiency is dropping down quickly at low energies.

ii. Jet finding algorithms begin to break.

iii. Large gamma-gamma and gamma-e backgrounds.
Data Samples

- $\sqrt{s} = 500$ GeV; 1000 fb$^{-1}$ luminosity; ~ 200k events /sample (CalcHEP)
- Five points close to ILC limits
  - $(M_{\text{NE1}}, M_{\text{sbottom}}) = (220,210), (230,220)$ - mass difference 10 GeV
  - $(M_{\text{NE1}}, M_{\text{sbottom}}) = (230,210), (240,220)$ - mass difference 20 GeV
  - $(M_{\text{NE1}}, M_{\text{sbottom}}) = (240,220)$ - mass difference 30 GeV
Analysis

- Events are pre-selected using few basic quantities
  - $E_{visible} < 80\text{ GeV}$, $\Delta R_{\eta\phi} < 3.0$, $10 \leq N_{particles} \leq 60$, $\max(|\eta_1|, |\eta_2|) < 2.0$
  - Veto on electrons or photons in forward detectors (>10 mrad)
- For the final selection Neural Net is trained with additional inputs.
- Example plots for point (230,210) – signal (line) was multiplied by $10^5$
The measurement is interpreted in terms of signal significance calculated as \( \frac{S}{\sqrt{S + B}} \) and depending on a particular neural net output cut.

Points (230,210) and (220,210) both reach above 4\( \sigma \) level.

- Other points are more difficult (low x-section, jet softness) but they all can be excluded @ 95% CL.
Remarks and Future Plans

- **Higgs self-coupling (ZHH) analysis**
  - Not included in the SiD LoI.
  - Uncertainty too large, after having FSR and full sim/rec samples.

- **Work in progress for TeV Linear Collider**
  - Tuning of the LCFI package for CLIC and physics/tagging/vertexing studies.
  - The package was never used in $\sqrt{s} = 3$ TeV environment before.
Summary

H → cC branching ratio uncertainty from $e^+e^- \rightarrow ZH$ estimated to ~8.5%. Analysis extended to $H \rightarrow bB$ and $H \rightarrow gg$.

Top mass uncertainty about 50 MeV on the tree level. Cross section and production asymmetry addressed.

We study a new cosmologically motivated sbottom co-annihilation scenario which can be uniquely probed at the ILC. Challenge is due to very soft jets and large $\gamma\gamma$ bkgr.

Higgs self-coupling analysis delivered large errors.

Work in progress for TeV LC in both SiD and CLIC geometries.