Jet substructure as a new Higgs search channel at the LHC

Gavin P. Salam
LPTHE, UPMC Paris 6 & CNRS

SUSY08
Seoul, Korea, 16–21 June 2008

Work in collaboration with
Jon Butterworth, Adam Davison (UCL) & Mathieu Rubin (LPTHE)

arXiv:0802.2470, PRL in press
Low-mass Higgs search @ LHC: complex because dominant decay channel, $H \rightarrow bb$, often swamped by backgrounds.

Various production processes

- $gg \rightarrow H (\rightarrow \gamma\gamma)$ feasible
- $WW \rightarrow H \rightarrow \ldots$ feasible
- $gg \rightarrow t\bar{t}H$ v. hard
- $q\bar{q} \rightarrow WH, ZH$ small; but gives access to $WH$ and $ZH$ couplings

Currently considered impossible
WH/ZH search channel @ LHC

- Signal is $W \rightarrow \ell \nu$, $H \rightarrow b \bar{b}$. Studied e.g. in ATLAS TDR

- Backgrounds include $Wb \bar{b}$, $t \bar{t} \rightarrow \ell \nu b \bar{b}jj$, ...

Difficulties, e.g.

- Poor acceptance ($\sim 12\%$) Easily lose 1 of 4 decay products
- $p_t$ cuts introduce intrinsic bkgd mass scale;
- $gg \rightarrow t \bar{t} \rightarrow \ell \nu b \bar{b}[jj]$ has similar scale
- small S/B
- Need exquisite control of bkgd shape
Signal is $W \rightarrow \ell \nu$, $H \rightarrow b \bar{b}$. Studied e.g. in ATLAS TDR

Backgrounds include $Wb\bar{b}$, $t\bar{t} \rightarrow \ell \nu b\bar{b}jj$, ...

Difficulties, e.g.

- Poor acceptance ($\sim 12\%$)
  Easily lose 1 of 4 decay products
- $p_t$ cuts introduce intrinsic bkgd mass scale;
- $gg \rightarrow t\bar{t} \rightarrow \ell \nu b\bar{b}[jj]$ has similar scale
- small S/B
- Need exquisite control of bkgd shape
WH/ZH search channel @ LHC

- Signal is $W \rightarrow \ell \nu$, $H \rightarrow b \bar{b}$.
- Backgrounds include $Wb\bar{b}$, $t\bar{t} \rightarrow \ell \nu b\bar{b}jj$, ...

Difficulties, e.g.
- Poor acceptance ($\sim 12\%$)
  Easily lose 1 of 4 decay products
- $p_t$ cuts introduce intrinsic bkgd mass scale;
- $gg \rightarrow t\bar{t} \rightarrow \ell \nu b\bar{b}[jj]$ has similar scale
- small S/B
- Need exquisite control of bkgd shape

Conclusion (ATLAS TDR):
"The extraction of a signal from $H \rightarrow b\bar{b}$ decays in the WH channel will be very difficult at the LHC, even under the most optimistic assumptions [...]"
Study subset of WH/ZH with high $p_t$

- At high $p_t$:
  - ✓ Higgs and W/Z more likely to be central
  - ✓ high-$p_t$ $Z \rightarrow \nu\bar{\nu}$ becomes visible
  - ✓ Fairly collimated decays: high-$p_t$ $\ell^\pm, \nu, b$
    - Good detector acceptance
  - ✓ Backgrounds lose cut-induced scale
  - ✓ $t\bar{t}$ kinematics cannot simulate bkgd
    - Gain clarity and S/B
  - ✗ Cross section will drop dramatically
    - By a factor of 20 for $p_{tH} > 200$ GeV
    - Will the benefits outweigh this?
The method

**Boosted EW bosons**

**Hadronically decaying Higgs boson at high \( p_t = \text{single massive jet} \)?**

\[
R \gtrsim \frac{m}{p_t} \frac{1}{\sqrt{z(1-z)}}
\]

discussion of such problems: Seymour '93; Butterworth, Cox & Forshaw '02; Butterworth, Ellis & Raklev '07; Skiba & Tucker-Smith '07; Holdom '07; Baur '07; Agashe et al. '07; Lille, Randall & Wang '07; Contino & Servant '08; Brooijmans '08; Thaler & Wang '08; Kaplan et al '08 [...]  

**Drawbacks**

- Optimal \( R \) depends on \( m, p_t, z \) — hard to get single “best” choice
- \( Y_{ij} \) cut implicitly introduces mass scale \( \sim \sqrt{Y_{cut}} \times \text{jet } p_t \)
Hadronically decaying Higgs boson at high $p_t = $ single massive jet?

$$R \gtrsim \frac{m}{p_t} \frac{1}{\sqrt{z(1-z)}}$$

discussion of such problems: Seymour ’93; Butterworth, Cox & Forshaw ’02; Butterworth, Ellis & Raklev ’07; Skiba & Tucker-Smith ’07; Holdom ’07; Baur ’07; Agashe et al. ’07; Lille, Randall & Wang ’07; Contino & Servant ’08; Brooijmans ’08; Thaler & Wang ’08; Kaplan et al ’08 [...]

Drawbacks

- Optimal $R$ depends on $m$, $p_t$, $z$ — hard to get single “best” choice
- $Y_{ij}$ cut implicitly introduces mass scale $\sim \sqrt{Y_{cut}} \times$ jet $p_t$
**Boosted EW bosons**

**Hadronically decaying Higgs boson at high $p_t = \text{single massive jet?}$**

![Diagram showing boosted X decaying into single jet](image)

$R \gtrsim \frac{m}{p_t} \frac{1}{\sqrt{z(1-z)}}$

**Most powerful idea till 2007**

- Find jets with $k_t$ jet algorithm with given $R$.
- Uncluster last recomb. for jet and require $Y_{ij} = \frac{\min(p_{tij}^2, p_{ti}^2)}{p_t^2} \Delta R_{ij}^2 > Y_{cut}$ [Seymour '93]
- Look for peak in jet mass

**Butterworth, Cox & Forshaw '02; Butterworth, Ellis & Raklev '07**

**Drawbacks**

- Optimal $R$ depends on $m, p_t, z$ — hard to get single “best” choice
- $Y_{ij}$ cut implicitly introduces mass scale $\sim \sqrt{Y_{cut} \times \text{jet } p_t}$
The method

Boosted EW bosons

Hadronically decaying Higgs boson at high $p_t = \text{single massive jet}$?

$\text{boosted } \mathbf{X} \rightarrow \text{single jet}$

\[ R \gtrsim \frac{m}{p_t} \frac{1}{\sqrt{z(1-z)}} \]

Most powerful idea till 2007

- Find jets with $k_t$ jet algorithm with given $R$
- Uncluster last recomb. for jet and require $Y_{ij} = \frac{\min(p_{ti}^2, p_{ti}^2)}{p_t^2} \Delta R_{ij}^2 > Y_{cut}$ [Seymour '93]
- Look for peak in jet mass

Butterworth, Cox & Forshaw '02; Butterworth, Ellis & Raklev '07

Drawbacks

- Optimal $R$ depends on $m$, $p_t$, $z$ — hard to get single “best” choice
- $Y_{ij}$ cut implicitly introduces mass scale $\sim \sqrt{Y_{cut}} \times \text{jet } p_t$
The Cambridge/Aachen jet alg.

Dokshitzer et al ’97
Wengler & Wobisch ’98

Work out $\Delta R_{ij}^2 = \Delta y_{ij}^2 + \Delta \phi_{ij}^2$ between all pairs of objects $i, j$;
Recombine the closest pair;
Repeat until all objects separated by $\Delta R_{ij} > R$.

Provides a “hierarchical” view of the event;
work through it backwards to analyse a jet.
Start with high-$p_t$ jet

1. Undo last stage of clustering (≡ reduce $R$): $J \rightarrow J_1, J_2$

2. If $\max(m_1, m_2) \lesssim 0.67m$, call this a mass drop [else goto 1]
   Automatically detects correct $R \sim R_{bb}$ to catch angular-ordered radn.

3. Require $y_{12} = \frac{\min(p_{t1}^2, p_{t2}^2)}{m_{12}^2} \Delta R_{12}^2 \sim \frac{\min(z_1, z_2)}{\max(z_1, z_2)} > 0.09$ [else goto 1]
   dimensionless rejection of asymmetric QCD branching

4. Require each subjet to have $b$-tag [else reject event]
   Correlate flavour & momentum structure
Start with high-$p_t$ jet

1. Undo last stage of clustering (≡ reduce $R$): $J \rightarrow J_1, J_2$

2. If $\max(m_1, m_2) \lesssim 0.67 m$, call this a mass drop [else goto 1]
   Automatically detects correct $R \sim R_{bb}$ to catch angular-ordered radn.

3. Require $y_{12} = \frac{\min(p_{t1}^2, p_{t2}^2)}{m_{12}^2} \Delta R_{12}^2 \approx \frac{\min(z_1, z_2)}{\max(z_1, z_2)} > 0.09$ [else goto 1]
   dimensionless rejection of asymmetric QCD branching

4. Require each subjet to have $b$-tag [else reject event]
   Correlate flavour & momentum structure
At moderate $p_t$, $R_{bb}$ is quite large; **UE & pileup degrade mass resolution**

$$\delta M \sim R_{bb}^4 \Lambda_{UE} \frac{p_t}{M}$$ [Dasgupta, Magnea & GPS '07]

**Filter the jet**

- Reconsider region of interest at smaller $R_{filt} = \min(0.3, R_{bb}/2)$
- Take 3 hardest subjets $b, \bar{b}$ and leading order gluon radiation
At moderate $p_t$, $R_{bb}$ is quite large; \textit{UE & pileup degrade mass resolution}

$$\delta M \sim R^4 \Lambda_{\text{UE}} \frac{p_t}{M}$$ [Dasgupta, Magnea & GPS '07]

**Filter the jet**

- Reconsider region of interest at smaller $R_{\text{filt}} = \min(0.3, R_{bb}/2)$
- Take \textbf{3} hardest subjets $b, \bar{b}$ and leading order gluon radiation
The method

\[ pp \rightarrow ZH \rightarrow \nu \bar{\nu} b \bar{b}, @14 \text{ TeV, } m_H = 115 \text{ GeV} \]

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

Cluster event, C/A, R=1.2

arbitrary norm.
The method

\[ pp \rightarrow ZH \rightarrow \nu \bar{\nu} b \bar{b}, \ @14 \text{ TeV}, \ m_H = 115 \text{ GeV} \]

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

Fill it in, \( \rightarrow \) show jets more clearly
The method

\[ pp \rightarrow ZH \rightarrow \nu \bar{\nu} b \bar{b}, \ \& 14 \text{ TeV}, \ m_H = 115 \text{ GeV} \]

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

Consider hardest jet, \( m = 150 \text{ GeV} \)

\( {\text{SIGNAL}} \)

\( {\text{Zbb BACKGROUND}} \)

arbitrary norm.
The method

\[ pp \to ZH \to \nu\bar{\nu} b\bar{b}, \ @14 \text{TeV}, \ m_H = 115 \text{GeV} \]

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

\[ \text{split: } m = 150 \text{ GeV}, \ \frac{\max(m_1,m_2)}{m} = 0.92 \to \text{repeat} \]
The method

\[ pp \rightarrow ZH \rightarrow \nu \bar{\nu} b \bar{b}, \ @14 \text{ TeV}, \ m_H = 115 \text{ GeV} \]

**Herwig 6.510 + Jimmy 4.31 + FastJet 2.3**

\[ \text{Split: } m = 139 \text{ GeV}, \quad \frac{\max(m_1,m_2)}{m} = 0.37 \rightarrow \text{mass drop} \]
The method

\[ pp \rightarrow ZH \rightarrow \nu \bar{\nu} b \bar{b}, \ 14 \text{ TeV}, \ m_H = 115 \text{ GeV} \]

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

\[ y_{12} \approx \frac{p_{t2}}{p_{t1}} \approx 0.7 \rightarrow \text{OK + 2 } b\text{-tags (anti-QCD)} \]
The method

\[ pp \rightarrow ZH \rightarrow \nu\bar{\nu}b\bar{b}, \ O(14 \text{ TeV}), \ m_H = 115 \text{ GeV} \]

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

\[ R_{\text{filt}} = 0.3 \]

\[ \text{arbitrary norm.} \]
Jets, G. Salam, LPTHE (p. 9)

The method

\[ pp \rightarrow ZH \rightarrow \nu \bar{\nu} b \bar{b}, \quad @14 \text{ TeV}, \quad m_H = 115 \text{ GeV} \]

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

\[ R_{filt} = 0.3: \text{ take 3 hardest, } m = 117 \text{ GeV} \]
Consider $HW$ and $HZ$ signals: $H \rightarrow b\bar{b}$, $W \rightarrow \ell \nu$, $Z \rightarrow \ell^+\ell^-$ and $Z \rightarrow \nu\bar{\nu}$, 3 channels: $\ell^\pm + \not\!E_T; \ell^+\ell^-; \not\!E_T$

**Common cuts**

- $p_{tV}, p_{tH} > 200$ GeV
- $|\eta_{Higgs-jet}| < 2.5$
- $\ell = e, \mu$, $p_{t,\ell} > 30$ GeV, $|\eta_\ell| < 2.5$
- No extra $\ell, b$'s with $|\eta| < 2.5$

**Channel-specific cuts:** see next slide

**Assumptions**

- Real/fake $b$-tag rates: 0.7/0.01 optimistic, but not inconceivable
- $S/\sqrt{B}$ from 16 GeV window ATLAS jet-mass resln $\sim$ half this? cf. talk by Adam Davison in P6 @16:10

**Tools:** Herwig 6.510, Jimmy 4.31 (tuned), hadron-level $\rightarrow$ FastJet 2.3

**Backgrounds:** $VV, Vj, jj, t\bar{t}$, single-top, with $> 30$ fb$^{-1}$ (except $jj$)
Results combine HZ and HW, $p_t > 200$ GeV

Common cuts
- $p_t V, p_t H > 200$ GeV
- $|\eta_H| < 2.5$
- $[p_t, \ell > 30$ GeV, $|\eta_\ell| < 2.5$]
- No extra $\ell, b$’s with $|\eta| < 2.5$
- Real/fake $b$-tag rates: 0.7/0.01
- $S/\sqrt{B}$ from 16 GeV window

Leptonic channel

$Z \rightarrow \mu^+\mu^-, e^+e^-$
- $80 < m_{\ell^+\ell^-} < 100$ GeV

At 5.9$\sigma$ for 30 fb$^{-1}$ this looks like a possible new channel for light Higgs discovery. Deserves serious exp. study!
combine HZ and HW, $p_t > 200$ GeV

**Missing $E_T$ channel**

- $q\bar{q}$
- $V+$jets
- $VV$
- $V+$Higgs

**Common cuts**

- $p_{tV}, p_{tH} > 200$ GeV
- $|\eta_H| < 2.5$
- $[p_{t,\ell} > 30$ GeV, $|\eta_\ell| < 2.5]$
- No extra $\ell$, $b$'s with $|\eta| < 2.5$
- Real/fake $b$-tag rates: $0.7/0.01$
- $S/\sqrt{B}$ from 16 GeV window

**Missing-$E_T$ channel**

- $Z \rightarrow \nu\bar{\nu}, W \rightarrow \nu[\ell]$
- $E_T > 200$ GeV

At 5.9σ for 30 fb$^{-1}$ this looks like a possible new channel for light Higgs discovery. Deserves serious exp. study!
combine HZ and HW, $p_t > 200$ GeV

Common cuts

- $p_t V, p_t H > 200$ GeV
- $|\eta_H| < 2.5$
- $[p_t, \ell > 30$ GeV, $|\eta_\ell| < 2.5$
- No extra $\ell, b$'s with $|\eta| < 2.5$
- Real/fake $b$-tag rates: 0.7/0.01
- $S/\sqrt{B}$ from 16 GeV window

Semi-leptonic channel

- $W \rightarrow \nu\ell$
- $E_T > 30$ GeV (& consistent $W$.)
- no extra jets $|\eta| < 3, p_t > 30$

At $5.9\sigma$ for 30 fb$^{-1}$ this looks like a possible new channel for light Higgs discovery. Deserves serious exp. study!
Results

Combine HZ and HW, $p_t > 200$ GeV

3 channels combined

Common cuts
- $p_{tV}, p_{tH} > 200$ GeV
- $|\eta_H| < 2.5$
- $[p_{t,\ell} > 30$ GeV, $|\eta_\ell| < 2.5]$  
- No extra $\ell, b$'s with $|\eta| < 2.5$
- Real/fake $b$-tag rates: 0.7/0.01
- $S/\sqrt{B}$ from 16 GeV window

3 channels combined

Note excellent $VZ, Z \rightarrow b\bar{b}$ peak for calibration

NB: $q\bar{q}$ is mostly $t\bar{t}$

At 5.9$\sigma$ for 30 fb$^{-1}$ this looks like a possible new channel for light Higgs discovery. Deserves serious exp. study!
How can we be doing so well despite losing factor 20 in $X$-sct?

<table>
<thead>
<tr>
<th></th>
<th>Signal</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eliminate $t\bar{t}$, etc.</td>
<td>$-$</td>
<td>$\times 1/3$</td>
</tr>
<tr>
<td>$p_t &gt; 200$ GeV</td>
<td>$\times 1/20$</td>
<td>$\times 1/60$ [bkgds: $Wb\bar{b}$, $Zbb$]</td>
</tr>
<tr>
<td>improved acceptance</td>
<td>$\times 4$</td>
<td>$\times 4$</td>
</tr>
<tr>
<td>twice better resolution</td>
<td>$-$</td>
<td>$\times 1/2$</td>
</tr>
<tr>
<td>add $Z \rightarrow \nu\bar{\nu}$</td>
<td>$\times 1.5$</td>
<td>$\times 1.5$</td>
</tr>
<tr>
<td>total</td>
<td>$\times 0.3$</td>
<td>$\times 0.017$</td>
</tr>
</tbody>
</table>

much better $S/B$; better $S/\sqrt{B}$

[exact numbers depend on analysis details]
Impact of $b$-tagging, Higgs mass

Most scenarios above $3\sigma$

For it to be a significant discovery channel requires decent $b$-tagging, lowish mass Higgs [and good experimental resolution]

In nearly all cases, looks feasible for extracting $WH$, $ZH$ couplings
Impact of $b$-tagging, Higgs mass

Most scenarios above $3\sigma$

For it to be a significant discovery channel requires decent $b$-tagging, lowish mass Higgs [and good experimental resolution]

In nearly all cases, looks feasible for extracting $WH$, $ZH$ couplings
Conclusions

Specific

- New promising Higgs search channel
- Unique at LHC in terms of separately seeing $WH$, $ZH$ couplings
- Deserves & needs in-depth experimental study starting within ATLAS

General

- Clarity & simplicity of high-$p_t$ final state outweighed large X-sct loss
  
  Might this hold in other cases?
- 3rd generation jet-finding tools play a key role here
  
  $3^{rd}$ generation $\equiv$ interact with the event structure
  
  Applied also to high-$p_t$ top, Kaplan et al, arXiv:0806.0848
EXTRAS
Compare with “standard” algorithms

Check mass spectra in HZ channel, $H \rightarrow b\bar{b}$, $Z \rightarrow \ell^+\ell^-$

Cambridge/Aachen (C/A) with mass-drop and filtering (MD/F) works best
Cross section for signal and the $Z+\text{jets}$ background in the leptonic $Z$ channel for $200 < p_{TZ}/\text{GeV} < 600$ and $110 < m_J/\text{GeV} < 125$, with perfect $b$-tagging; shown for our jet definition (C/A MD-F), and other standard ones close to their optimal $R$ values.

<table>
<thead>
<tr>
<th>Jet definition</th>
<th>$\sigma_S/\text{fb}$</th>
<th>$\sigma_B/\text{fb}$</th>
<th>$S/\sqrt{B} \cdot \text{fb}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C/A, $R = 1.2$, MD-F</td>
<td>0.57</td>
<td>0.51</td>
<td>0.80</td>
</tr>
<tr>
<td>$k_t$, $R = 1.0$, $y_{cut}$</td>
<td>0.19</td>
<td>0.74</td>
<td>0.22</td>
</tr>
<tr>
<td>SISCone, $R = 0.8$</td>
<td>0.49</td>
<td>1.33</td>
<td>0.42</td>
</tr>
<tr>
<td>anti-$k_t$, $R = 0.8$</td>
<td>0.22</td>
<td>1.06</td>
<td>0.21</td>
</tr>
</tbody>
</table>
**K-factors**

Analysis shown without $K$ factors. What impact do they have?

Determined with MCFM, MC@NLO

- **Signal**: $K \sim 1.6$
- **$Vbb$ backgrounds**: $K \sim 2 - 2.5$
- **$t\bar{t}$ backgrounds**: $K \sim 2$ for total; not checked for high-$p_t$ part

Conclusion: $S/\sqrt{B}$ should not be severely affected by NLO contributions
Worsen $b$-tagging: 60%/2%
Raise $p_t$ cut to 300 GeV

NB: kills $t\bar{t}$ background
Jet algorithm generations

- 1st generation: the original UA1, Tevatron jet algorithms
  all IR or collinear unsafe

- 2nd generation: sequential recombination algorithms (JADE, $k_t$, Cambridge), and IR safe cones (SISCone, anti-$k_t$)
  All IR safe; some give jet substructure

- 3rd generation(?): algorithms and jet-analysis procedures whose behaviour adapts itself to the specific event under consideration.
  Not yet systematic reality; but reasonable dream?