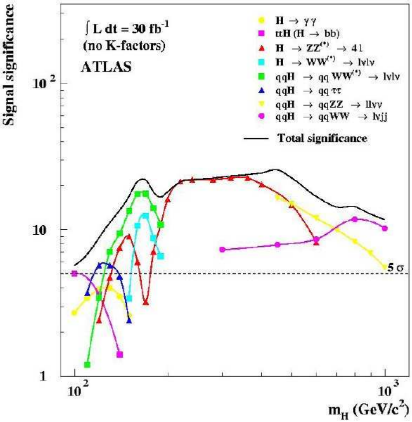
Jet substructure as a new Higgs search channel at the LHC

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, $H \to bb$, often swamped by backgrounds.

Various production processes

- $gg \to H\ (\to \gamma\gamma)$ — feasible
- $WW \to H \to \ldots$ — feasible
- $gg \to t\bar{t}H$ — v. hard
- $q\bar{q} \to WH, ZH$

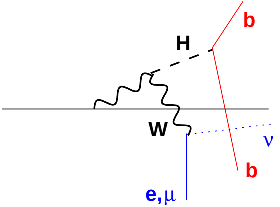
  small; but gives access to *WH* and *ZH* couplings
  Currently considered impossible

WH/ZH search channel @ LHC

- Signal is $W \to \ell\nu$, $H \to b\bar{b}$. Studied e.g. in ATLAS TDR
- Backgrounds include $Wb\bar{b}$, $t\bar{t} \to \ell\nu b\bar{b}jj$, ...

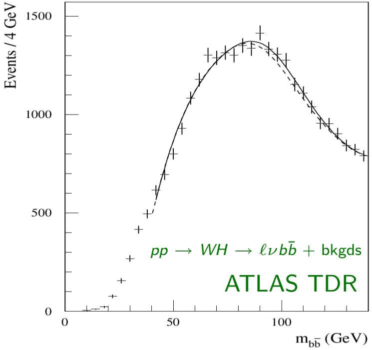
Difficulties, e.g.

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



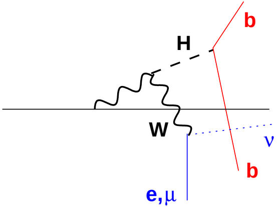
WH/ZH search channel @ LHC

- Signal is $W \to \ell\nu$, $H \to b\bar{b}$. Studied e.g. in ATLAS TDR
- Backgrounds include $Wb\bar{b}$, $t\bar{t} \to \ell\nu b\bar{b}jj$, ...



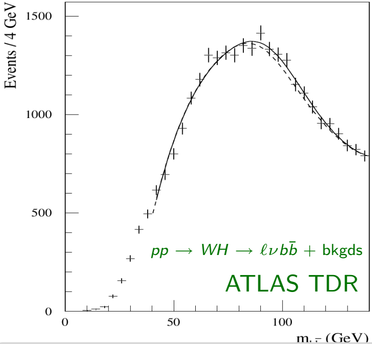
Difficulties, e.g.

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



WH/ZH search channel @ LHC

- Signal is $W \to \ell\nu$, $H \to b\bar{b}$. Studied e.g. in ATLAS TDR
- Backgrounds include $Wb\bar{b}$, $t\bar{t} \to \ell\nu b\bar{b}jj$, ...

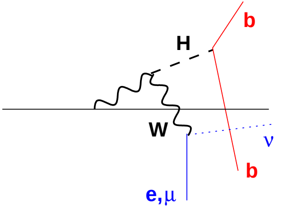


Difficulties, e.g.

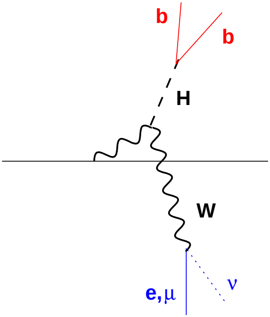
- Poor acceptance ($\sim 12\%$)
  Easily lose 1 of 4 decay products
- p_t cuts introduce intrinsic bkgd mass scale;
- $gg \to t\bar{t} \to \ell\nu b\bar{b}[jj]$ has similar scale
- small S/B
- Need exquisite control of bkgd shape

Conclusion (ATLAS TDR):

"The extraction of a signal from $H \to b\bar{b}$ 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 $Z \to \nu\bar{\nu}$ 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

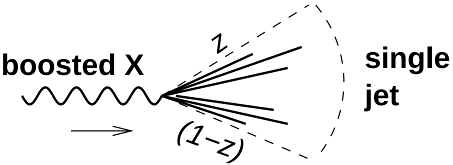
✗ Cross section will drop dramatically

By a factor of 20 for $p_{tH} > 200$ GeV

Will the benefits outweigh this?

Boosted EW bosons

Hadronically decaying Higgs boson at high p_t = single massive jet?



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

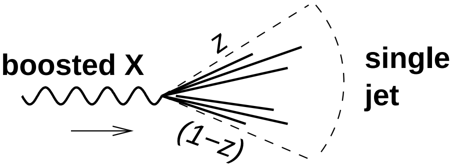
discussion of such problems: Seymour '93; Butterworth, Cox & Forshaw '02; Butterworth, Ellis & Raklev '07; Skiba & Tucker-Smith '07; Holdom '07; Baur '07; Agashe et al. '07; Lille, Randall & Wang '07; Contino & Servant '08; Brooijmans '08; Thaler & Wang '08; Kaplan et al '08 [. . .]

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 $\mathbf{p_t}$ = single massive jet?



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

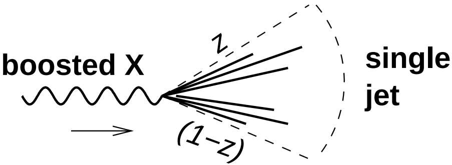
discussion of such problems: Seymour '93; Butterworth, Cox & Forshaw '02; Butterworth, Ellis & Raklev '07; Skiba & Tucker-Smith '07; Holdom '07; Baur '07; Agashe et al. '07; Lille, Randall & Wang '07; Contino & Servant '08; Brooijmans '08; Thaler & Wang '08; Kaplan et al '08 [...]

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 $\mathbf{p_t}$ = single massive 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 $Y_{ij} = \frac{\min(p_{ti}^2, p_{ti}^2)}{p_t^2} \Delta R_{ij}^2 > Y_{cut}$
- Look for peak in jet mass [Seymour '93]

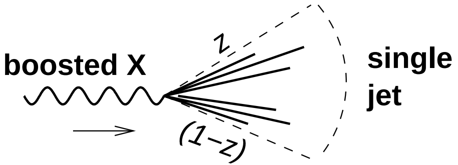
Butterworth, Cox & Forshaw '02; Butterworth, Ellis & Raklev '07

Drawbacks

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

Boosted EW bosons

Hadronically decaying Higgs boson at high $\mathbf{p_t}$ = single massive 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 $Y_{ij} = \frac{\min(p_{ti}^2, p_{ti}^2)}{p_t^2} \Delta R_{ij}^2 > Y_{cut}$
- Look for peak in jet mass

[Seymour '93]

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 $\sim \sqrt{Y_{cut}} \times$ jet p_t

#1: Our tool

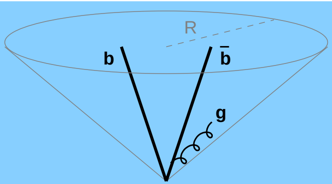
The Cambridge/Aachen jet alg.

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

#2: The jet analysis



Start with high-p_t jet

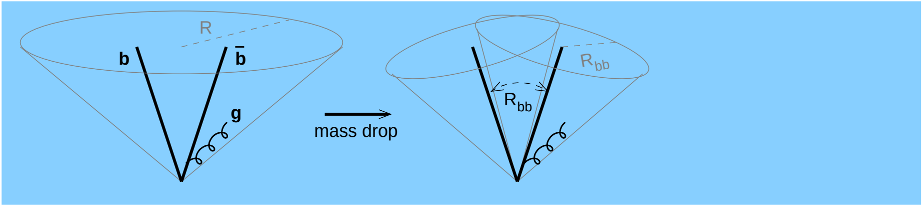
1. Undo last stage of clustering ($\equiv$ reduce R): $J \to J_1, J_2$

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

4. Require each subjet to have b-tag [else reject event]
   Correlate flavour & momentum structure

#2: The jet analysis



Start with high-p_t jet

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

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.

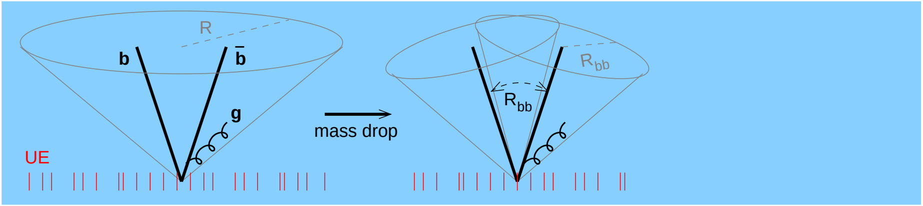
3. Require $y_{12} = \frac{\min(p_{t1}^2, p_{t2}^2)}{m_{12}^2} \Delta R_{12}^2 \simeq \frac{\min(z_1, z_2)}{\max(z_1, z_2)} > 0.09$ [else goto 1]

   dimensionless rejection of asymmetric QCD branching

4. Require each subjet to have b-tag [else reject event]

   Correlate flavour & momentum structure

#3: jet filtering



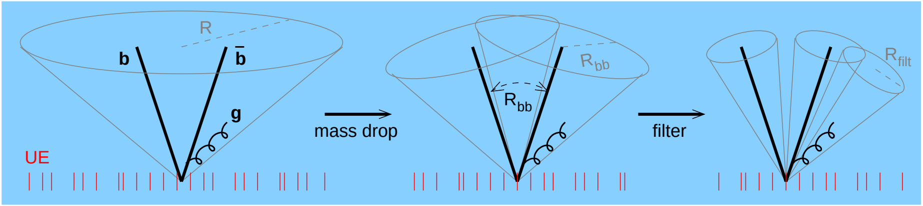
At moderate p_t, R_{bb} is quite large; *UE & pileup degrade mass resolution*

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

Filter the jet

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

#3: jet filtering



At moderate p_t, R_{bb} is quite large; *UE & pileup degrade mass resolution*

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

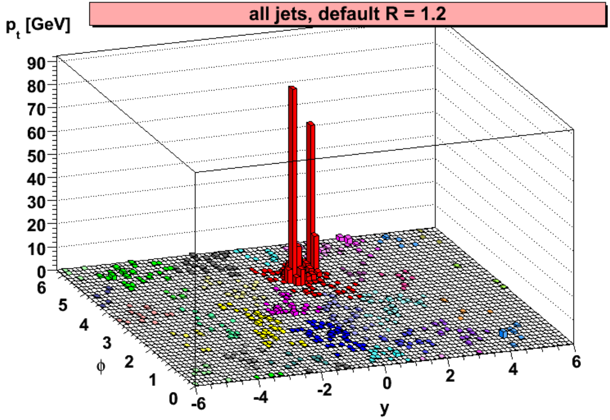
Filter the jet

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

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

SIGNAL

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



Zbb BACKGROUND

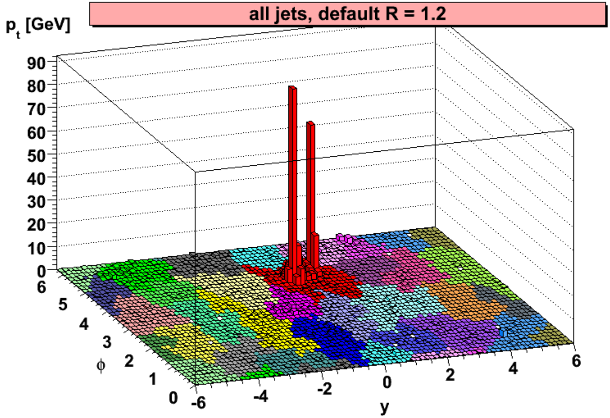
Cluster event, C/A, R=1.2

arbitrary norm.

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

SIGNAL

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



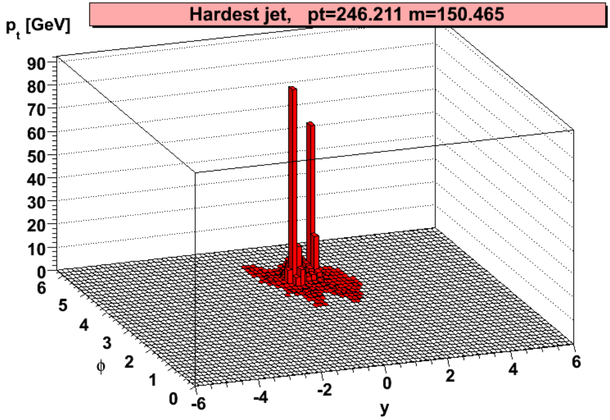
Zbb BACKGROUND

Fill it in, → show jets more clearly

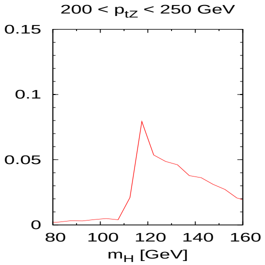
arbitrary norm.

$pp \longrightarrow ZH \longrightarrow \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

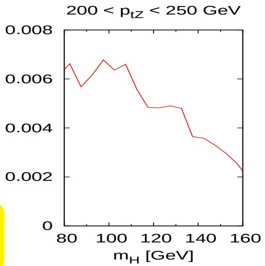
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



SIGNAL



Zbb BACKGROUND

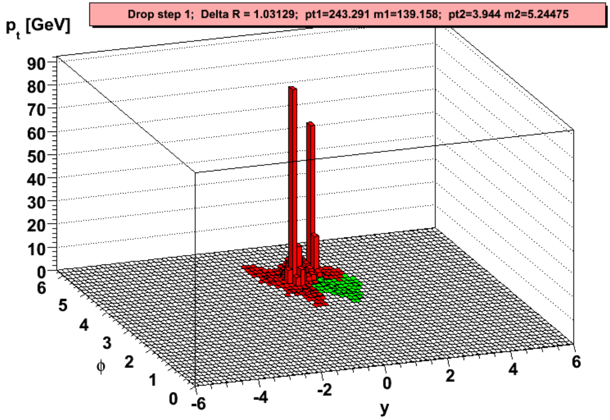


Consider hardest jet, $m = 150$ GeV

arbitrary norm.

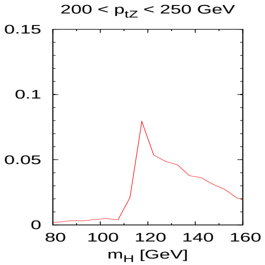
$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

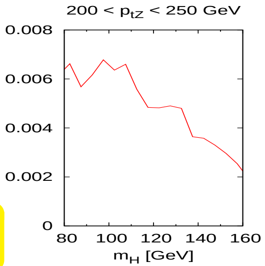


split: $m = 150$ GeV, $\frac{\max(m_1, m2)}{m} = 0.92 \to$ repeat

SIGNAL



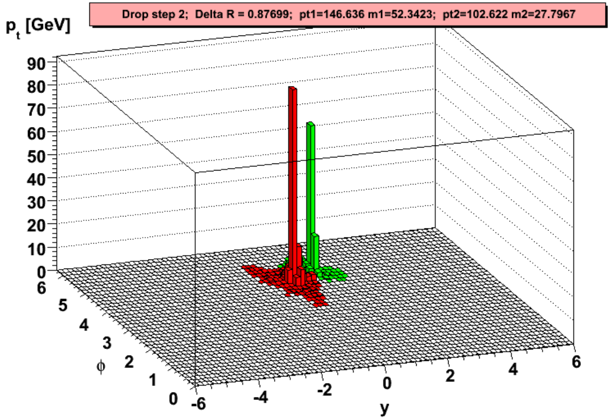
Zbb BACKGROUND



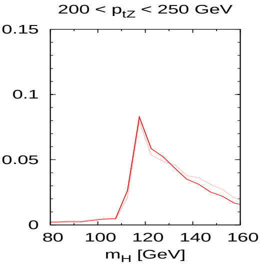
arbitrary norm.

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

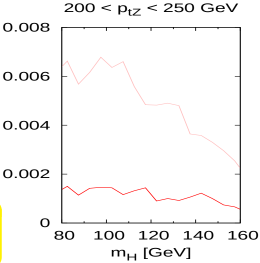
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



SIGNAL



Zbb BACKGROUND

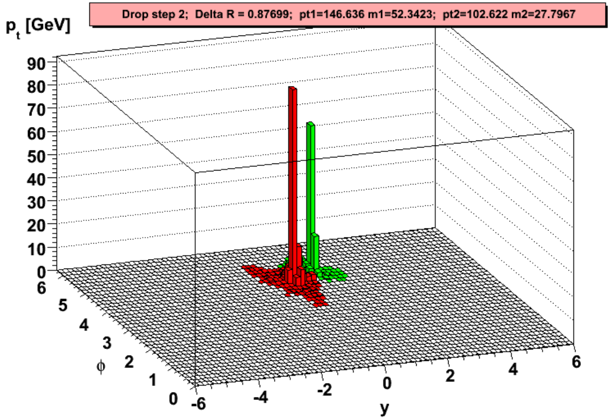


split: $m = 139$ GeV, $\frac{\max(m_1, m2)}{m} = 0.37 \to$ mass drop

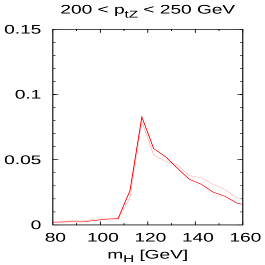
arbitrary norm.

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

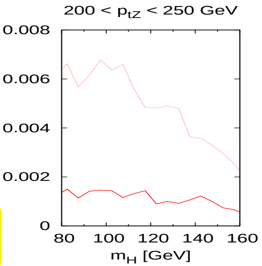
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



SIGNAL



Zbb BACKGROUND

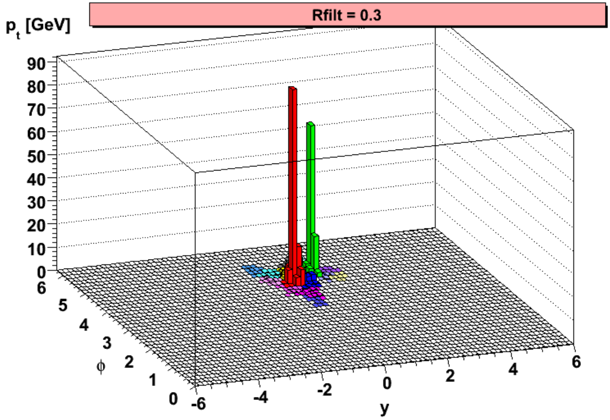


check: $y_{12} \simeq \frac{p_{t2}}{p_{t1}} \simeq 0.7 \to$ OK + 2 b-tags (anti-QCD)

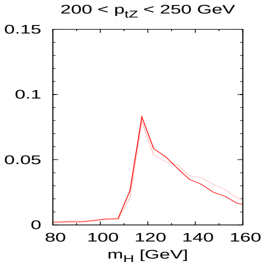
arbitrary norm.

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

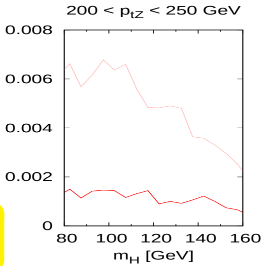
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



SIGNAL



Zbb BACKGROUND

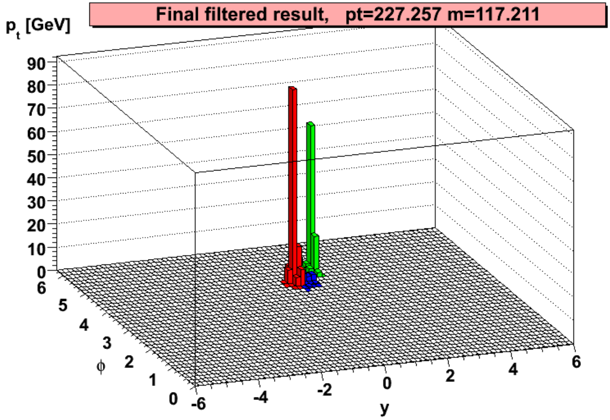


$R_{filt} = 0.3$

arbitrary norm.

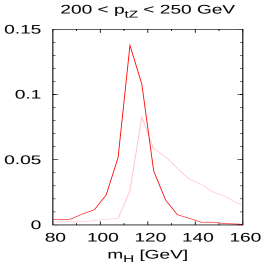
$pp \to ZH \to \nu\bar{\nu}b\bar{b}$, @14 TeV, $m_H = 115$ GeV

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

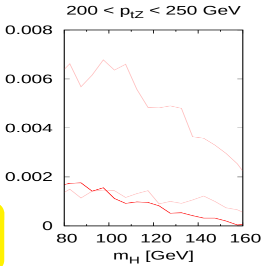


$R_{filt} = 0.3$: take 3 hardest, **m = 117 GeV**

SIGNAL



Zbb BACKGROUND



arbitrary norm.

The full analysis (scaled to 30 fb^{-1})

Consider HW and HZ signals: $H \to b\bar{b}$, $W \to \ell\nu$, $Z \to \ell^+\ell^-$ and $Z \to \nu\bar{\nu}$,

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 ℓ, b's with $|\eta| < 2.5$

Channel-specific cuts: see next slide

Assumptions

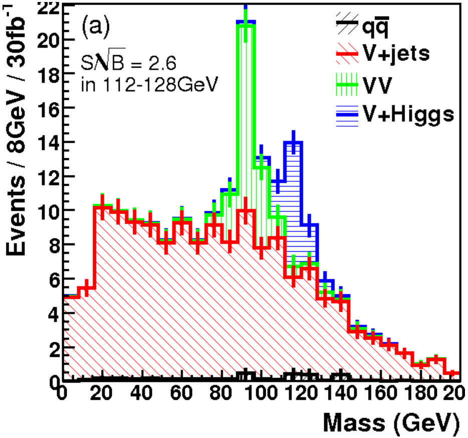
- Real/fake b-tag rates: 0.7/0.01 — optimistic, but not inconceivable
- $S/\sqrt{B}$ from 16 GeV window — ATLAS jet-mass resln $\sim$ half this?

cf. talk by Adam Davison in P6 @16:10

Tools: Herwig 6.510, Jimmy 4.31 (tuned), hadron-level $\to$ FastJet 2.3
Backgrounds: VV, Vj, jj, $t\bar{t}$, single-top, with > 30 fb^{-1} (except jj)

combine HZ and HW, $p_t > 200$ GeV

Leptonic channel



<u>Common cuts</u>

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

<u>Leptonic channel</u>

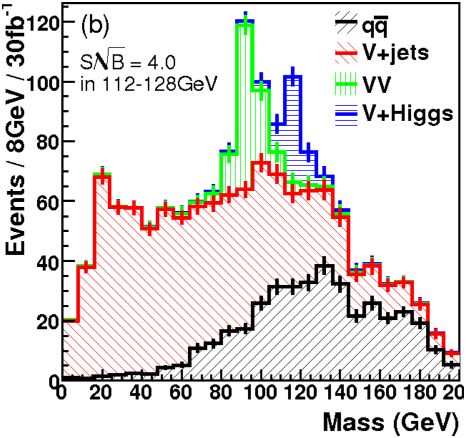
$Z \to \mu^+\mu^-, e^+e^-$

- $80 < m_{\ell^+\ell^-} < 100$ 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

Missing E_T channel



Common cuts

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

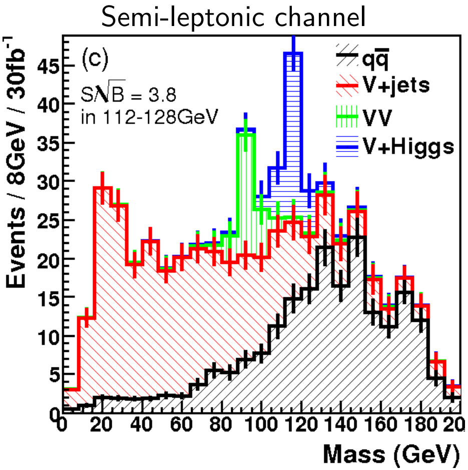
Missing-E_t channel

$Z \to \nu\bar{\nu}$, $W \to \nu[\ell]$

- $\not{E}_T > 200$ 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

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

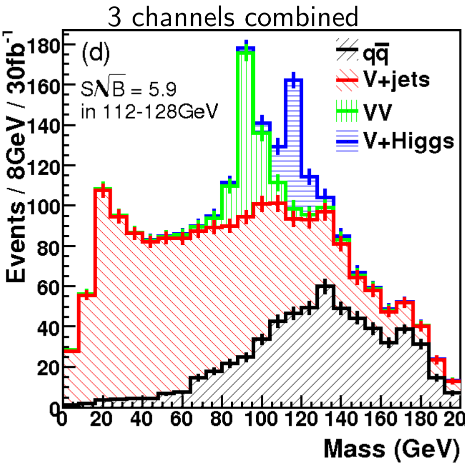
Semi-leptonic channel

$W \to \nu\ell$

- $\not{E}_T > 30$ GeV (& consistent W.)
- no extra jets $|\eta| < 3, p_t > 30$

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

- $p_{tV}, p_{tH} > 200$ GeV
- $|\eta_H| < 2.5$
- $[p_{t,\ell} > 30$ GeV, $|\eta_\ell| < 2.5]$
- No extra ℓ, 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, $Z \to b\bar{b}$ peak for calibration

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

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

Rough impact of going to high-p_t

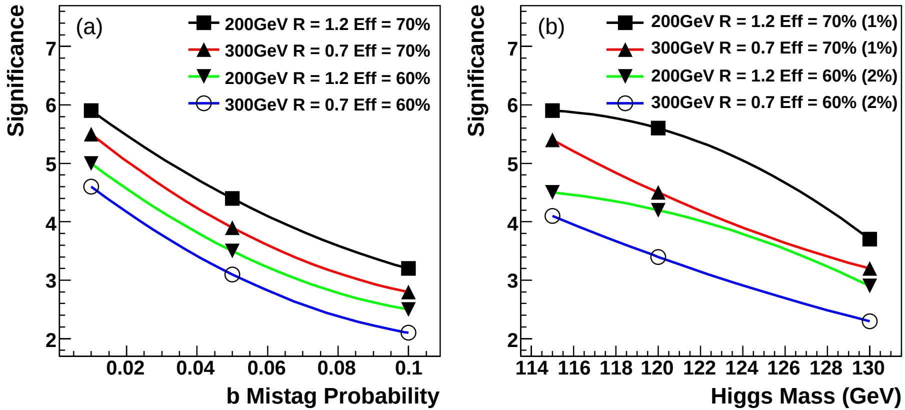
How can we be doing so well despite losing factor 20 in X-sct?

	Signal	Background	
Eliminate $t\bar{t}$, etc.	–	$\times 1/3$	[very approx.]
$p_t > 200$ GeV	$\times 1/20$	$\times 1/60$	[bkgds: $Wb\bar{b}, Zb\bar{b}$]
improved acceptance	$\times 4$	$\times 4$	
twice better resolution	–	$\times 1/2$	
add $Z \to \nu\bar{\nu}$	$\times 1.5$	$\times 1.5$	
total	$\times 0.3$	$\times 0.017$	

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

[exact numbers depend on analysis details]

Impact of *b*-tagging, Higgs mass

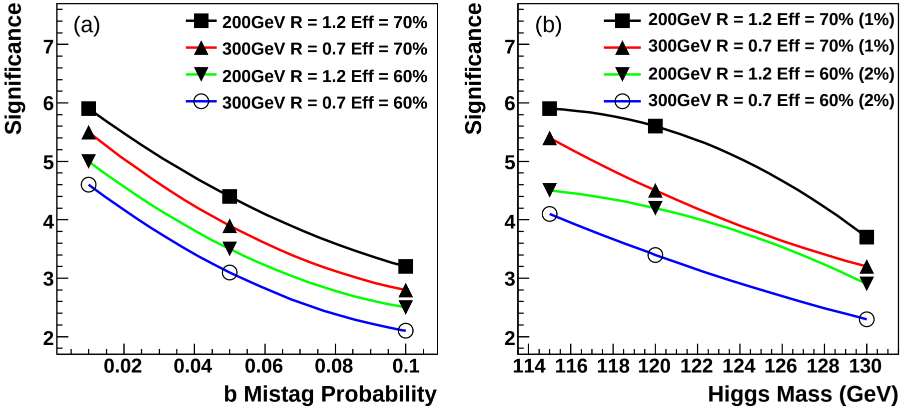


Most scenarios above 3σ

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 3σ

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

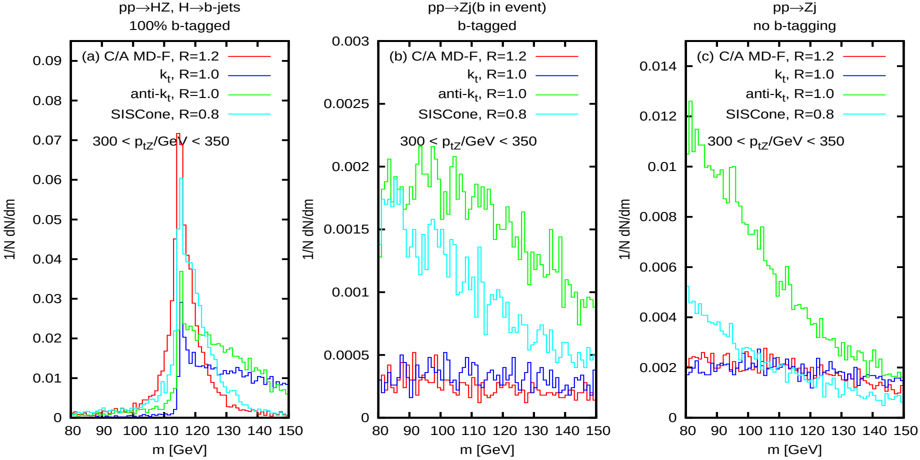
  3rd generation ≡ interact with the event structure

  Applied also to high-p_t top, Kaplan et al, arXiv:0806.0848

EXTRAS

Compare with "standard" algorithms

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



Cambridge/Aachen (C/A) with mass-drop and filtering (MD/F) works best

Jet-alg comparison

Cross section for signal and the Z+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.

Jet definition	σ_S/fb	σ_B/fb	$S/\sqrt{B \cdot \mathrm{fb}}$
C/A, $R = 1.2$, MD-F	0.57	0.51	0.80
k_t, $R = 1.0$, y_{cut}	0.19	0.74	0.22
SISCone, $R = 0.8$	0.49	1.33	0.42
anti-k_t, $R = 0.8$	0.22	1.06	0.21

K-factors

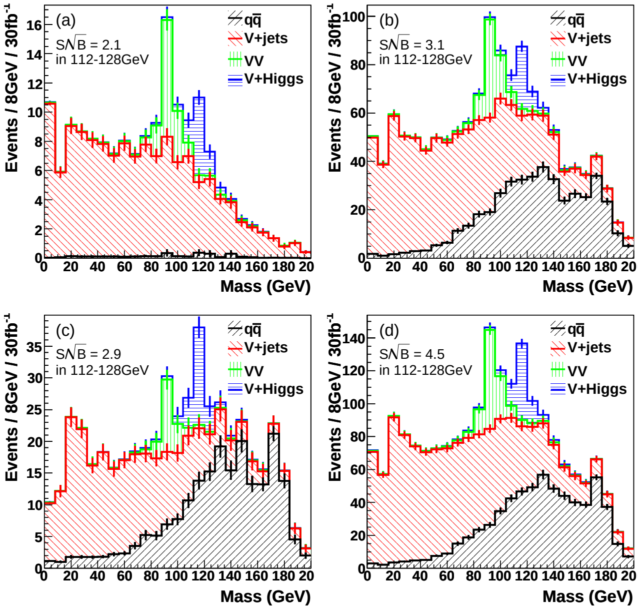
Analysis shown without K factors. What impact do they have?

Determined with MCFM, MC@NLO

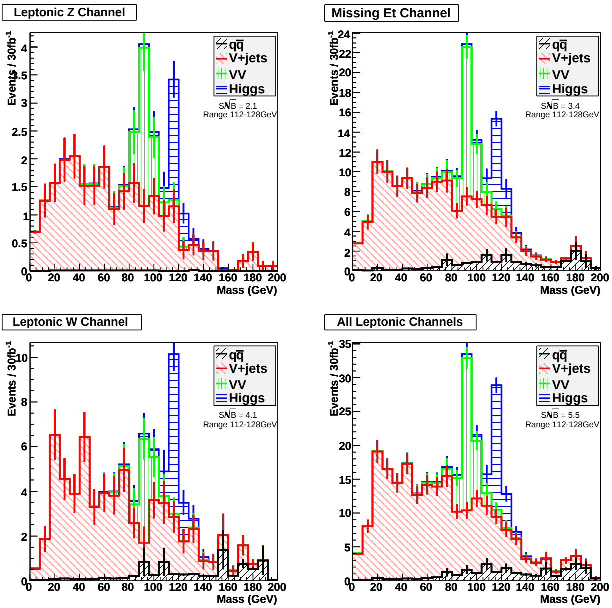
- Signal: $K \sim 1.6$
- Vbb backgrounds: $K \sim 2 - 2.5$
- $t\bar{t}$ backgrounds: $K \sim 2$ for total; not checked for high-p_t part

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 UAx, 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 reaonsable dream?