Jet structures in New Physics and Higgs searches

Gavin P. Salam LPTHE, UPMC Paris 6 & CNRS

LHC New Physics Forum, IWH Heidelberg, Germany, 23–26 February 2009

Part based on work with Jon Butterworth, Adam Davison (UCL) & Mathieu Rubin (LPTHE) LHC will (should...) span two orders of magnitude in p_t :

$$\frac{m_{EW}}{2} \longleftrightarrow 50 m_{EW}$$

That's why it's being built

In much of that range, EW-scale particles are **light** [a little like *b*-quarks at the Tevatron]

This talk:

about reconstructing high- p_t EW-scale particles



Rules of thumb:

 $m = 100 \text{ GeV}, p_t = 500 \text{ GeV}$

 $R < \frac{2m}{p_t}$: always resolve two jetsR < 0.4 $R \gtrsim \frac{3m}{p_t}$: resolve one jet in 75% of cases $(\frac{1}{8} < z < \frac{7}{8})$ $R \gtrsim 0.6$

New heavy particles can decay to W, Z, top \rightarrow hadrons

- Need "taggers" for boosted hadronic SM particles
- ▶ To help extract new-physics signals; help identify their decays

Continue here: top-quark ID

New EW-scale particles may be *easier* to discover at high- p_t

- Some relevant fraction produced at high- p_t ($\sqrt{s} \gg m$)
- Jet combinatorics are easier at high p_t cleaner events
- Easier to organise cuts so as not to sculpt backgrounds

Start here: light Higgs-boson search

New heavy particles can decay to W, Z, top \rightarrow hadrons

- Need "taggers" for boosted hadronic SM particles
- ▶ To help extract new-physics signals; help identify their decays

Continue here: top-quark ID

New EW-scale particles may be *easier* to discover at high- p_t

- Some relevant fraction produced at high- p_t ($\sqrt{s} \gg m$)
- Jet combinatorics are easier at high p_t cleaner events
- Easier to organise cuts so as not to sculpt backgrounds

Start here: light Higgs-boson search

New heavy particles can decay to W, Z, top \rightarrow hadrons

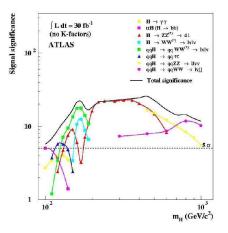
- Need "taggers" for boosted hadronic SM particles
- ▶ To help extract new-physics signals; help identify their decays

Continue here: top-quark ID

New EW-scale particles may be easier to discover at high- p_t

- Some relevant fraction produced at high- p_t ($\sqrt{s} \gg m$)
- Jet combinatorics are easier at high p_t cleaner events
- Easier to organise cuts so as not to sculpt backgrounds

Start here: light Higgs-boson search



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

Various production processes

	$gg \to H$	$(\rightarrow \gamma \gamma$)	feasible
--	------------	------------------------------	---	----------

- $WW \rightarrow H \rightarrow \dots$ 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

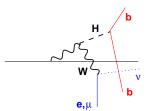
• Signal is $W \to \ell \nu$, $H \to b\bar{b}$.

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

Studied e.g. in ATLAS TDR

Difficulties, e.g.

- Poor acceptance (~ 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



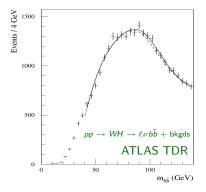
Jets, G. Salam, LPTHE (p. 6)

$\rm WH/ZH$ search channel @ LHC

• Signal is $W \to \ell \nu$, $H \to b \overline{b}$.

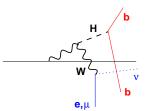
• Backgrounds include $Wb\bar{b}$, $t\bar{t}
ightarrow \ell
u b\bar{b} j j$, . . .

Studied e.g. in ATLAS TDR



Difficulties, e.g.

- Poor acceptance (~ 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



Jets, G. Salam, LPTHE (p. 6) Intro

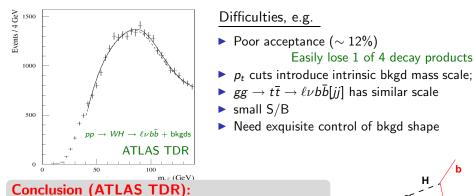
WH/ZH search channel @ LHC

▶ Signal is $W \rightarrow \ell \nu$, $H \rightarrow bb$.

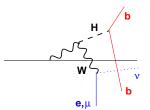
Backgrounds include $Wb\bar{b}, t\bar{t} \rightarrow \ell \nu b\bar{b} j j, \ldots$

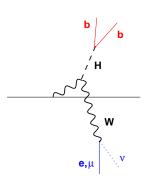
Studied e.g. in ATLAS TDR

Easily lose 1 of 4 decay products



"The extraction of a signal from $H \rightarrow bb$ decays in the WH channel will be very difficult at the LHC, even under the most optimistic assumptions [...]"





At high p_t :

- $\checkmark\,$ Higgs and W/Z more likely to be central
- ✓ high- p_t Z → $\nu \bar{\nu}$ becomes visible
- ✓ Fairly collimated decays: high- $p_t \ \ell^{\pm}, \nu, b$ Good detector acceptance
- $\checkmark\,$ Backgrounds lose cut-induced scale
- ✓ $t\overline{t}$ kinematics cannot simulate bkgd Gain clarity and S/B

X Cross section will drop dramatically By a factor of 20 for p_{tH} > 200 GeV Will the benefits outweigh this?

. . .

FTC

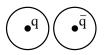
How do we find a boosted Higgs inside a single jet? Special case of general (unanswered) question: how do we best do jet-finding?

Various people have looked at boosted objects over the years

- ▶ Seymour '93 [heavy Higgs $\rightarrow WW \rightarrow \nu \ell \text{jets}$]
- ▶ Butterworth, Cox & Forshaw '02 [$WW \rightarrow WW \rightarrow \nu \ell j$ ets]
- Agashe et al. '06 [KK excitation of gluon $\rightarrow t\overline{t}$]
- ▶ Butterworth, Ellis & Raklev '07 [SUSY decay chains $\rightarrow W, H$]
- Skiba & Tucker-Smith '07 [vector quarks]
- Lillie, Randall & Wang '07 [KK excitation of gluon $\rightarrow t\bar{t}$]

Jets, G. Salam, LPTHE (p. 9) Boosted object finding







Select on the jet mass with one large (cone) jet Can be subject to large bkgds [high- p_t jets have significant masses]

Choose a small jet size (R) so as to resolve two jets Easier to reject background if you actually see substructure [NB: must manually put in "right" radius]

Take a large jet and split it in two Let jet algorithm establish correct division

Past methods

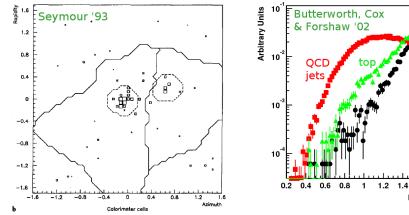


Fig. 2. A hadronic W decay, as seen at calorimeter level, a without, and b with, particles from the underlying event. Box sizes are logarithmic in the cell energy, lines show the borders of the sub-jets for infinitely soft emission according to the cluster (solid) and cone (dashed) algorithms

Use k_t jet-algorithm's hierarchy to split the jets

1.8 1.6 log (WPT $\times \sqrt{y}$)

$$d_{ij}^{k_t} = \min(p_{ti}^2, p_{tj}^2) \Delta R_{ij}^2$$

2

Past methods

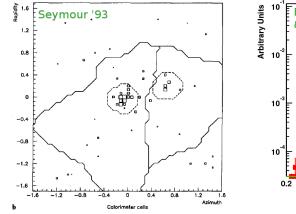
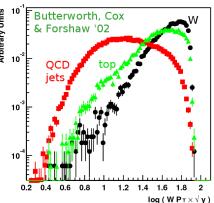


Fig. 2. A hadronic W decay, as seen at calorimeter level, a without, and b with, particles from the underlying event. Box sizes are logarithmic in the cell energy, lines show the borders of the sub-jets for infinitely soft emission according to the cluster (solid) and cone (dashed) algorithms

Use k_t jet-algorithm's hierarchy to split the jets



Use k_t alg.'s distance measure (rel. trans. mom.) to cut out QCD bkgd:

$$d_{ij}^{k_t} = \min(p_{ti}^2, p_{tj}^2) \Delta R_{ij}^2$$

Y-splitter

only partially correlated with mass

#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 $\Delta R_{ij} > R$. [in FastJet]

Gives "hierarchical" view of the event; work through it backwards to analyse jet

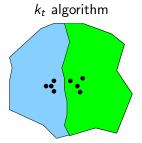
#1: Our tool

The Cambridge/Aachen jet alg.

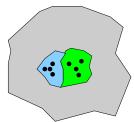
Dokshitzer et al '97 Wengler & Wobisch '98

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 $\Delta R_{ij} > R$. [in FastJet]

Gives "hierarchical" view of the event; work through it backwards to analyse jet

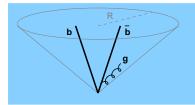


Cam/Aachen algorithm



Allows you to "dial" the correct R to keep perturbative radiation, but throw out UE

#2: The jet analysis



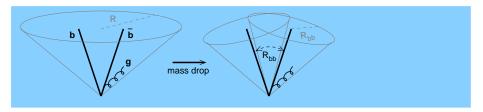
Start with high- p_t jet

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



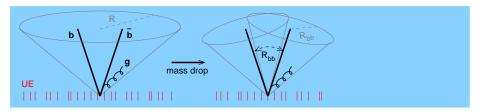
Start with high- p_t jet

- 1. Undo last stage of clustering (\equiv reduce *R*): $J \rightarrow J_1, J_2$
- 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 $y_{12} = \frac{\min(p_{t1}^2, p_{t2}^2)}{m_{12}^2} \Delta R_{12}^2 \simeq \frac{\min(z_1, z_2)}{\max(z_1, z_2)} > 0.09$ [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

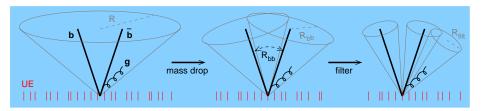


At moderate p_t , R_{bb} is quite large; UE & pileup degrade mass resolution $\delta M \sim R^4 \Lambda_{UE} \frac{p_t}{M}$ [Dasgupta, Magnea & GPS '07]

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]

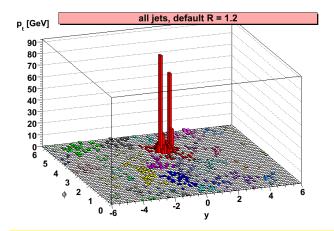
Filter the jet

- 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

Jets, G. Salam, LPTHE (p. 14) Boosted object finding $pp \rightarrow ZH \rightarrow \nu \bar{\nu} b \bar{b}$, @14TeV, $m_H = 115 \,\text{GeV}$

SIGNAL

Herwig 6.510 + Jimmy 4.31 + FastJet 2.3



Zbb BACKGROUND

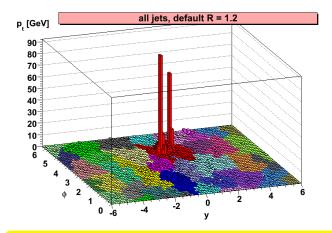
Cluster event, C/A, R=1.2

arbitrary norm.

Jets, G. Salam, LPTHE (p. 14) Boosted object finding $pp \rightarrow ZH \rightarrow \nu \bar{\nu} b \bar{b}$, @14TeV, $m_H = 115 \,\text{GeV}$

SIGNAL

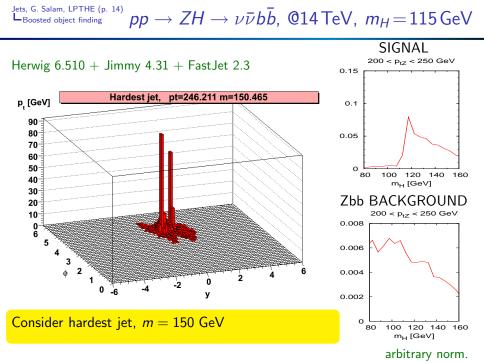
Herwig 6.510 + Jimmy 4.31 + FastJet 2.3

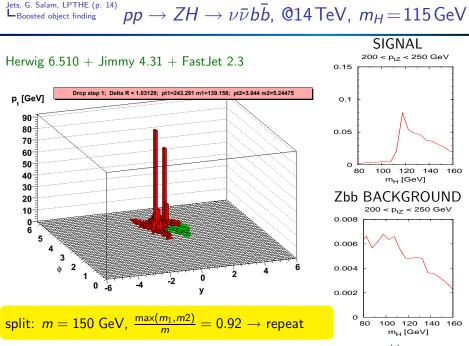


Zbb BACKGROUND

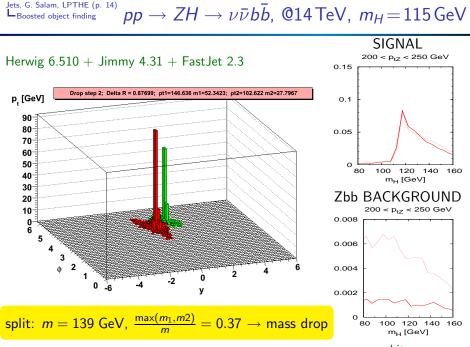
Fill it in, \rightarrow show jets more clearly

arbitrary norm.

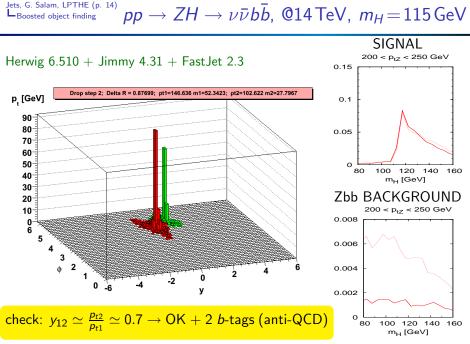




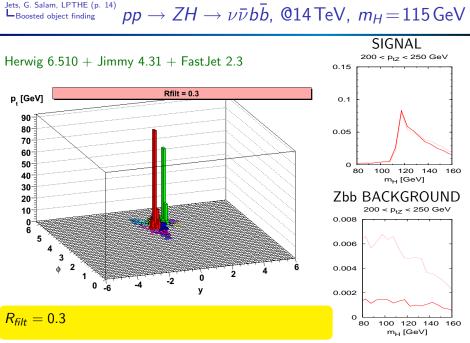
arbitrary norm.



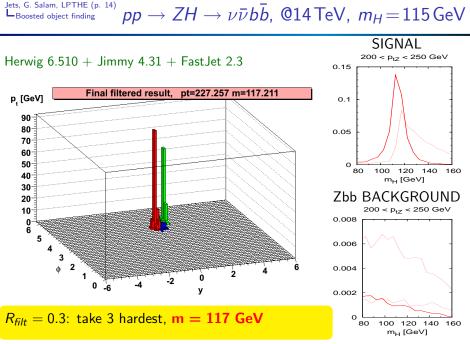
arbitrary norm.



arbitrary norm.



arbitrary norm.

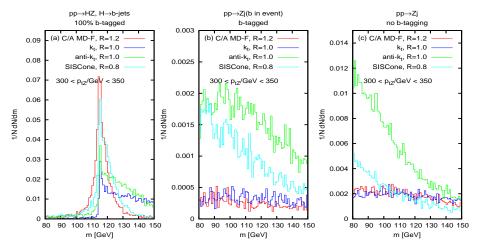


arbitrary norm.

Jets, G. Salam, LPTHE (p. 15) Boosted object finding

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

Jets, G. Salam, LPTHE (p. 16) Results

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

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
- \blacktriangleright No extra $\ell,~b{\rm 's}$ with $|\eta|<2.5$

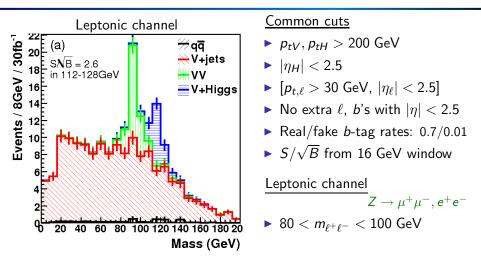
Channel-specific cuts: see next slide

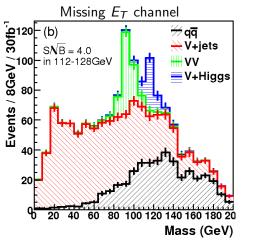
Assumptions

- ▶ Real/fake *b*-tag rates: 0.7/0.01
- S/\sqrt{B} from 16 GeV window

optimistic, but not inconceivable ATLAS jet-mass resln \sim half this?

<u>Tools:</u> 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*)

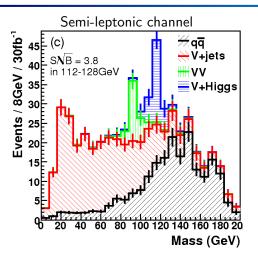




Common cuts

- ▶ p_{tV}, p_{tH} > 200 GeV
- ► $|\eta_H| < 2.5$
- $[p_{t,\ell} > 30 \text{ GeV}, |\eta_\ell| < 2.5]$
- 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



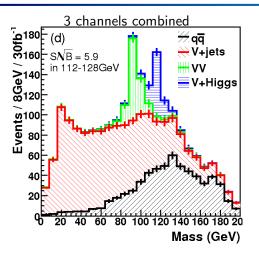
Common cuts

- ▶ $p_{tV}, p_{tH} > 200 \text{ GeV}$
- ► $|\eta_H| < 2.5$
- $[p_{t,\ell} > 30 \text{ GeV}, |\eta_\ell| < 2.5]$
- 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

 $W \to \nu \ell$

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



Common cuts

- ▶ p_{tV}, p_{tH} > 200 GeV
- ► $|\eta_H| < 2.5$
- $[p_{t,\ell} > 30 \text{ GeV}, |\eta_\ell| < 2.5]$
- \blacktriangleright 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>3 channels combined</u> Note excellent $VZ, Z \rightarrow b\bar{b}$ peak for calibration NB: $q\bar{q}$ is mostly $t\bar{t}$

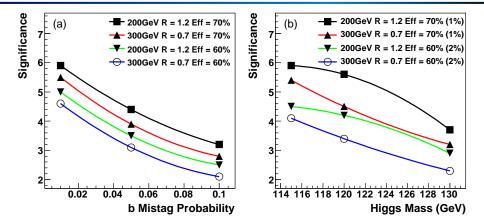
How can we be doing so well despite losing factor 20 in X-sct?

	Signal	Background	
Eliminate $t\bar{t}$, etc.	_	×1/3	
$p_t > 200 { m GeV}$	imes 1/20	imes 1/60	[bkgds: <i>Wbb</i> , <i>Zbb</i>]
improved acceptance	$\times 4$	$\times 4$	
twice better resolution	_	imes 1/2	
add $Z ightarrow u ar{ u}$	imes1.5	imes1.5	
total	×0.3	×0.017	

much better S/B; better S/\sqrt{B} [exact numbers depend on analysis details]

Jets, G. Salam, LPTHE (p. 19)

Impact of *b*-tagging, Higgs mass

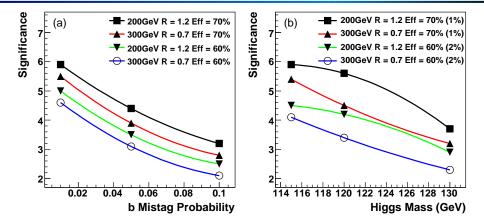


Most scenarios above 3σ

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

Jets, G. Salam, LPTHE (p. 19) Lesults

Impact of *b*-tagging, Higgs mass



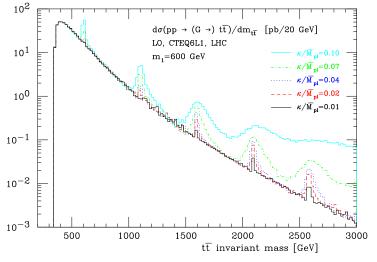
Most scenarios above 3σ

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

Boosted top [hadronic decays]

Jets, G. Salam, LPTHE (p. 21) $\mathbf{L}_{t\bar{t}}$

$X \to t \overline{t}$ resonances of varying difficulty



RS KK resonances $\rightarrow t\bar{t}$, from Frederix & Maltoni, 0712.2355

NB: QCD dijet spectrum is \sim 500 times $t\bar{t}$

High- p_t top production often envisaged in New Physics processes. ~ high- p_t EW boson, but: top has 3-body decay and is coloured.

4 papers on top tagging in '08 (at least). All use the jet mass + something extra.

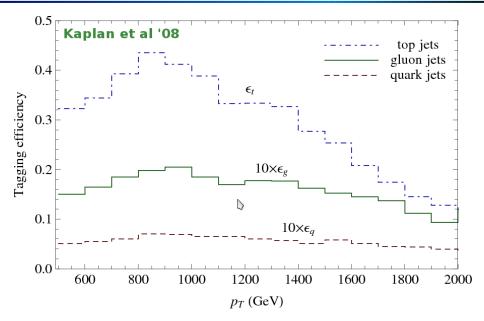
Questions

- What efficiency for tagging top?
- What rate of fake tags for normal jets?

Rough results for top quark with $p_{t} \sim 1$ TeV				
	"Extra"	eff.	fake	
[from T&W]	just jet mass	50%	10%	
Brooijmans	3,4 k_t subjets, d_{cut}	45%	5%	
Thaler & Wang	2,3 k_t subjets, z_{cut} + various	40%	5%	
Kaplan et al.	2,3 k_t subjets, z_{cut} + various 3,4 C/A subjets, z_{cut} + θ_h	40%	1%	
Almeida et al.	predict mass dist ⁿ , use jet-shape	-	_	



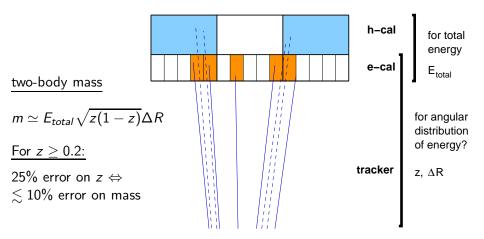
Efficiency v. p_t with calo (0.1×0.1)



Jets, G. Salam, LPTHE (p. 24) $t\bar{t}$ Boosted top

Fair assumptions for detector?

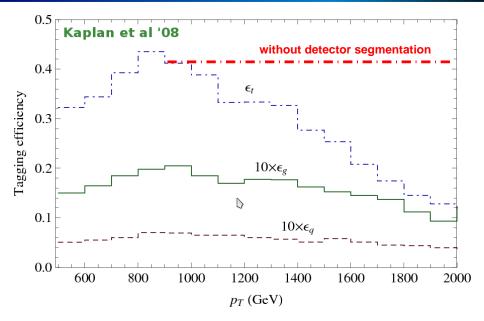
Theory $t\bar{t}$ studies use $\eta - \phi$ segmentation of 0.1. Limiting when $\Delta R \sim 0.1$ But charged tracks and EM-calo provide much better angular resolution.



Even rough info from tracks & e-cal very valuable



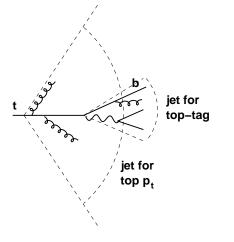
Efficiency v. p_t (ideal detector)



Jets, G. Salam, LPTHE (p. 26) $L_{t\bar{t}}$ Boosted top

If you want to use the tagged top (e.g. for $t\bar{t}$ invariant mass) QCD tells you:

the jet you use to tag a top quark \neq the jet you use to get its p_t



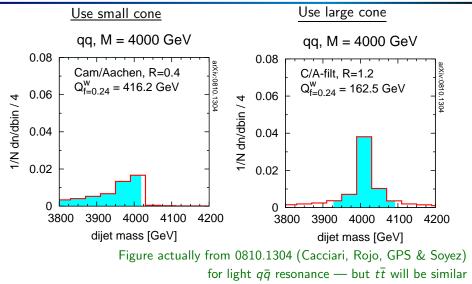
Within inner cone $\sim \frac{2m_t}{p_t}$ (dead cone) you have the top-quark decay products, but no radiation from top ideal for reconstructing top mass

Outside dead cone, you have radiation from top quark

> essential for top p_t Cacciari, Rojo, GPS & Soyez '09

Jets, G. Salam, LPTHE (p. 27) $t\bar{t}$ Boosted top

Impact of using small cone angle



How you look at your event matters: http://quality.fastjet.fr/

<u>General</u>

- Boosted EW-scale particles can be found in jets
- ► Cambridge/Aachen alg. is very powerful (flexible, etc.) tool for this
- General two-body eff/fake is 60% v. 3-7%

Higgs discovery

- ▶ high-p_t limit recovers WH and ZH channel at LHC
- Separately see WH, ZH couplings
- Deserves & needs in-depth experimental study

ongoing/starting within ATLAS/CMS

Тор

- Efficiencies/fake rates: up to 40/1(2)%
- Only get this if detector can resolve fine structure
- Top-quark at decay (the one you tag) and top-quark at production are different objects need different R for them

EXTRAS

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	$\sigma_{\mathcal{S}}/fb$	$\sigma_B/{ m fb}$	$S/\sqrt{B\cdot \mathrm{fb}}$
C/A, <i>R</i> = 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

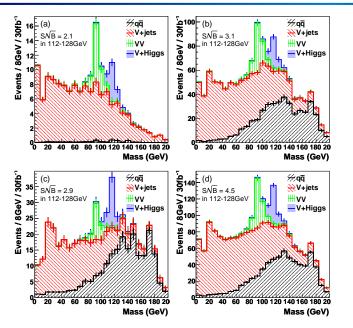
Analysis shown without K factors. What impact do they have?

Determined with MCFM, MC@NLO

- ▶ Signal: K ~ 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

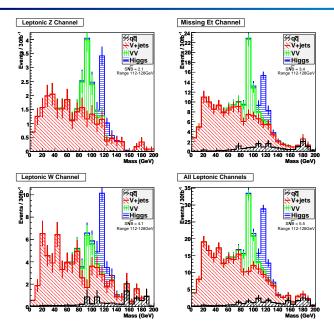
Conclusion: S/\sqrt{B} should not be severely affected by NLO contributions

Worsen *b*-tagging: 60%/2%



Jets, G. Salam, LPTHE (p. 33)

Raise p_t cut to 300 GeV



NB: kills $t\bar{t}$ background