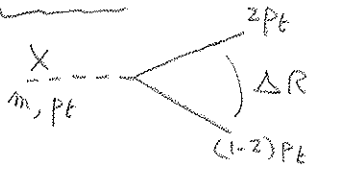
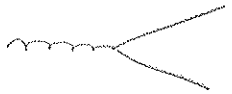
CHALLENGE



$$\Delta R \simeq \frac{m}{P_t \sqrt{z(1-z)}} \simeq \frac{2m}{P_t} - \frac{3m}{P_t} \Rightarrow$$ 1 single jet at large P_t.

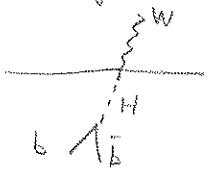
vs  QCD Background

examples:
- $W \to q\bar{q}$ [Seymour, 94; Butterworth, Cox, Forshaw, 2002; ...]
- $H \to b\bar{b}$ (WH/ZH) [Butterworth, Davison, Rubin, Salam, 2008]
- $t \to q\bar{q}b$ e.g [Kaplan, Rehermann, Schwartz, Tweedie, 2008]
- $\tilde{\chi}_0 \to qqq$ [Butterworth, J.Ellis, Raklev, Salam, 2009]

idea: look at jet substructure

Why: * $\frac{d\sigma}{dm}$ is not good enough (ex: QCD dijet is $\simeq 500 \times t\bar{t}$)

* boosted requirement reduces backgrounds
  ex: $t\bar{t}$ background to WH:



unlikely to have nearby $b\bar{b}$ from a $t\bar{t}$ decay.

Jet algs: k_t or Cambridge/Aachen

interparticle distance: $d_{ij}^{k_t} = \min(k_{ti}^2, k_{tj}^2)(\Delta\phi_{ij}^2 + \Delta y_{ij}^2)$
$d_{ij}^{C/A} = (\Delta\phi_{ij}^2 + \Delta y_{ij}^2)$

beam distance: $d_{iB}^{k_t} = R^2 k_{ti}^2$
$d_{iB}^{C/A} = R^2$

successively recombine particles/objects with minimal distance
(call it a jet for iB recomb.)

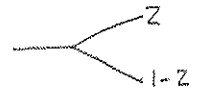
why these? provide a "cascade" view of the events

Filtering: cluster with C/A(R)
for each jet: • recluster with C/A(R_{filt})
• keep n filt subjets

idea: get rid of the UE ⇒ improve kinematic reconstruction.

Boosted tag : generic method

fat jet → find decay products of X

$X \to j_1 j_2$: * symmetric branching 

$z \ll 1 \Rightarrow$ mostly soft QCD emission

* "mass drop" : ----< >> small mass

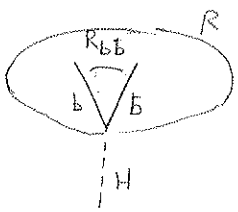
large mass $\qquad m_{j_1}, m_{j_2} \ll m$

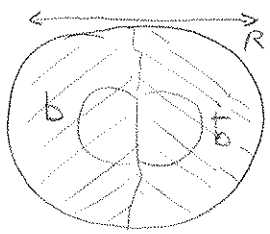
Method * 2 subjets : require 1 "symmetric mass drop"

* 3 " " " 2 " " " "

* method : successively unCluster the jet and look for sym mass-drop

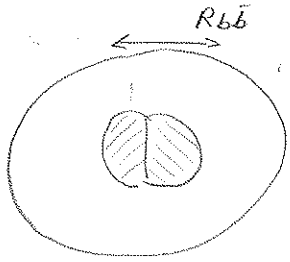
Comments : * C/A tends to be better than k_t





k_t

$H \to b\bar{b}$ likely last merging



C/A

first get down to $R_{b\bar{b}}$ before $H \to j_1 j_2$

$\Rightarrow$ lesser UE contents, more dynamic

* ≠ object $\Rightarrow$ ≠ techniques : mostly from scale knowledge

ex : • top : $m_t \simeq 175$ GeV $\Rightarrow$ "undo clustering until $m_{sub} \lesssim 40$ GeV" already not too bad

• $\tilde{\chi}_0$: need for something indep of the $\tilde{\chi}_0$ mass

• scale dependence more dangerous for busy events

* coloured objects : ex $t \to bW \to bq\bar{q}$

• tag : use R_{bW}

• reconstruct momentum : use R_t

Generic example : Pruning

[S. Ellis, Vermillon, Walsh, 09]

get rid of "soft radiation"

⓪ cluster with k_t or C/A

① rerun the clustering : for $j_1, j_2 \to j$

if $(z < z_{cut})$ and $(\Delta R_{1,2} > D_{cut})$ then discard softest

else merge

$D_{cut} \simeq m_j / p_{Tj}$ (m_j, p_{Tj} for the original jet)

$z_{cut} = 0.1$ for C/A

0.15 for k_t

Applicat° 1 : $H \to b\bar{b}$ just consider the method for illustration [Butterworth, Davison, Rubin, Salam, 08]

(0) Cluster with C/A ($R = 1.2$)

(1) undo the last clustering $j \to j_1 + j_2$

if mass drop: $m_{j_1} < \mu m_j$, sufficiently symmetric $y = \frac{\min(p_{t j_1}^2, p_{t j_2}^2)}{m_j^2} \Delta R_{12}^2 > y_{cut}$

Then j_1, j_2 are $b, \bar{b}$, j is the H ($R_{b\bar{b}} = R_{12}$)

else continue with the hardest of j_1, j_2

in practice $\mu = 0.67$, $y_{cut} \approx 0.09$

(2) filter with $R_{filt} = \min(0.3, R_{b\bar{b}}/2)$

$n_{filt} = 3$ ($b, \bar{b}$ + one g)

(3) b-tag for the 2 hardest subjets

results : b-tag efficiency: 60%
mis-tagging: 2%
$p_t \geq 200$ GeV
WH with $W \to l\nu$
ZH with $Z \to e^+e^-, \mu^+\mu^-$ or $\nu\bar{\nu}$
+ misc. cuts

$\Rightarrow$ $m_H = 115$ GeV
$30\ fb^{-1} \Rightarrow S/\sqrt{B} = 4.5$ in $[112-128]$ GeV
$100\ fb^{-1} \Rightarrow S/\sqrt{B} = 8.2$

($\frac{S}{\sqrt{B \cdot fb}} \simeq 0.8$ compared to 0.42 for a SISCone-based ($R=0.8$) analysis without subjets
0.22 for k_t ($R=1.0$) without subjets)

Application 2 : $t\bar{t}H$ $H \to b\bar{b}$, $t \to qq\bar{b}$, $t \to l\nu b$ [Plehn, Salam, Spannowsky, 09]

Method (0) cluster with C/A, $R = 1.5$, kinematic cuts:
at least 2 jets with $p_t \geq 200$ GeV, $|y_t| < 4$, for b-tagging $|y_{tag}| < 2.5$
1 lepton with $p_t \geq 15$ GeV, $|y_e| < 2.5$

(i) tag a top jet

(1) undo one step of the clustering $j \to j_1 j_2$ ($m_{j_1} > m_{j_2}$)

(2) if $m_{j_1} > 0.8\, m_j$, then only keep j_1 (j_2 is UE or soft QCD)
else keep j_1 and j_2

(3) for each kept subjet, if $m_{j_i} \leq 30$ GeV, store it
else, go back to (1) with $j = j_i$ (recurse for j_i)

(4) from the stored subjets, reconstruct the W:
$\to$ pair of stored subjets s.t. after filtering, $m_W^{rec} \in [65, 95]$ GeV
$R_{filt} = \min(0.3, R_{ij}/2)$; $n_{filt} = 3$

(5) reconstruct the top: $W + 3^{rd}$ subjet s.t., after filtering, $m_t^{rec} \in [150, 200]$ GeV
$\cos(\theta_W) < 0.7$

(6) if multiple candidates, keep the one with smaller $|m_W^{rec} - m_W| + |m_t^{rec} - m_t|$

efficiency: 43% of the signal
mis-tag: 5% (W + jets)

(ii) tag a Higgs : one wants an approach independent of the Higgs mass

①-③ same as the top with 0.9 instead of 0.8 for the cut of step ②

40 GeV " " 30 " " " " ③

④ order the (stored) subjet pairs by

$$J_{ij} = P_{ti} P_{tj} (\Delta R_{ij})^4 \quad (\simeq \text{Jade distance shifted towards large } \Delta R)$$

keep (and filter) the 3 hardes pairs as the Higgs candidate

⑤ 2 b tags matching the 2 Higgs subjets

⑥ $m_{bb} \in [m_H - 10, m_H + 10\ \text{GeV}]$ for S/B, S/√B computation

(iii) optional : 3rd b tag (not in the H & t jets)

Signal and backgrounds : (MadEvent + Herwig++ @) AlpGen + Herwig 6.5)

- $t\bar{t}H$: 702 fb (for $m_H = 120$ GeV) ($K = 1.57$)
- $t\bar{t}Z$: 1.1 pb reduced by boosted cuts, H-tag, b-tags
- $t\bar{t}b\bar{b}$: $P_{t,b} \geq 20$ GeV, $|\eta_b| < 2.5$, $R_{bb} > 0.8$: 2.6 pb ($K = 2.3$)
  reduced by H-tag, b-tags (matching !)
- $t\bar{t}$+jets : 3 topologies : (i) 2 jets mistagged as b
  (ii) $t \to Wb \to \ell\nu b$, b+jet to fake the Higgs
  (main motivation for 3rd b tagging)
  (iii) $t \to Wb \to bq\bar{q}$: QCD jet faking the "b in the top"
  b looking into the Higgs.

- Wjj : $P_{tj} \geq 15$ GeV : 15 nb
  boosted : $\times 10^{-3}$
  t mis-tagging : $\times 0.05$ } → 100 fb
  Higgs mass : $\times 0.1$
  b mis-tagging $\times 10^{-4}$ → 0.01 fb
  include b from shower : 0.1 fb i.e. < 10% of $t\bar{t}$+jets.

Params : $m_t = 172.3$ GeV
b-tag : 70% efficiency, 1% mis-tagging
$m_H = 115, 120, 130$ GeV

Note ideas for improvement in the paper.