Top Tagging

Tilman Plehn

Universität Heidelberg

Fermilab 3/2015

Fat jets

Boosted particles at the LHC

1994 boosted $W \to 2$ jets from heavy Higgs [Seymour]

1994 boosted $t \to 3$ jets [Seymour]

2002 boosted $W \to 2$ jets from strongly interacting WW [YSplitter: Butterworth, Cox, Forshaw]

2006 boosted $t \to 3$ jets from resonances [Agashe, Belyaev, Krupovnickas, Perez, Virzi]

2008 boosted $H \to b\bar{b}$ [BDRS Higgs tagger: Butterworth, Davison, Rubin, Salam]

2008 boosted $t \to 3$ jets from resonances [JH/CMS tagger: Kaplan, Rehermann, Schwartz, Tweedie]

2009 boosted $t \to 3$ jets in Higgs production [HEPTopTagger: TP, Salam, Spannowsky]

2009 boosted $t \to 3$ jets from resonances [Template Tagger: Almeida, Lee, Perez, Sterman, Sung, Virzi]

. . .

2010 first meta analysis review [BOOST proceedings, Ed: Karagoz, Spannowsky, Vos]

. . .

2011 N-Subjettiness [Thaler, van Tilburg]

2011 Shower Deconstruction [Soper, Spannowsky]

2011 pedagogical review (partly basis for this talk) [TP, Spannowsky]

. . .

2014 BDRS tagger with 427 citations, BOOST 2009-2014 done

Jet Algorithms

Definition of jets

– jet–parton duality ⇔ what are partons in detector?

– need algorithm to reconstruct what was one parton [IR save recombination algos]

– crucial for any LHC analysis [ILC analyses without jets?]

– extension to b and t perturbative QCD problem

Different measures [FASTJET: Cacciari, Salam, Soyez]

– define jet–jet and jet–beam distance [exclusive with resolution y_{cut}]

$$\begin{array}{lll} k_T & y_{ij} = \dfrac{\Delta R_{ij}}{R} \min\left(p_{T,i}, p_{T,j}\right) & y_{iB} = p_{T,i} \\ \text{C/A} & y_{ij} = \dfrac{\Delta R_{ij}}{R} & y_{iB} = 1 \\ \text{anti-}k_T & y_{ij} = \dfrac{\Delta R_{ij}}{R} \min\left(p_{T,i}^{-1}, p_{T,j}^{-1}\right) & y_{iB} = p_{T,i}^{-1} \; . \end{array}$$

– (1) find minimum $y^{\text{min}} = \min_{ij}(y_{ij}, y_{iB})$
(2a) if $y^{\text{min}} = y_{ij}$ merge subjets i and j, back to (1)
(2b) if $y^{\text{min}} = y_{iB}$ remove i from subjets, go to (1)

– theoretical and experimental trade-off decisions

– fat jets: use clustering history

Analysis: $Z' \to t\bar{t}$

LHC target $m_{Z'} \gtrsim 1.5$ TeV

- purely leptonic decays rate limited
- semi-leptonic approximate reconstruction of neutrino 4-vector:

  massless neutrino
  2D missing energy vector
  top or W mass constraints
- purely hadronic decays deemed not useful

Many taggers [Hopkins/CMS tagger, HEPTopTagger, template tagger, shower deconstruction]

- hadronic top identification and reconstruction
- jet separation challenging for heavy Z'
- combination of calo and tracker great
- usually combined with b-tag

Analysis: $Z' \to t\bar{t}$

LHC target $m_{Z'} \gtrsim 1.5$ TeV

- purely leptonic decays rate limited
- semi-leptonic approximate reconstruction of neutrino 4-vector:
  massless neutrino
  2D missing energy vector
  top or W mass constraints
- purely hadronic decays deemed not useful

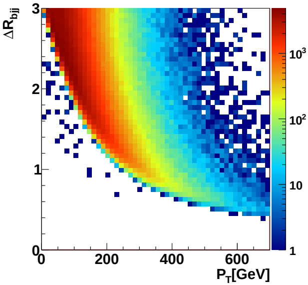
Many taggers [Hopkins/CMS tagger, HEPTopTagger, template tagger, shower deconstruction]

- hadronic top identification and reconstruction
- jet separation challenging for heavy Z'
- combination of calo and tracker great
- usually combined with b-tag

Validation and systematics

- tagging easier for higher boost, $p_{T,t} > 600$ GeV
- Standard Model events at lower $p_{T,t} < 400$ GeV

$\Rightarrow$ p_T range main challenge



Analysis: $Z' \to t\bar{t}$

LHC target $m_{Z'} \gtrsim 1.5$ TeV

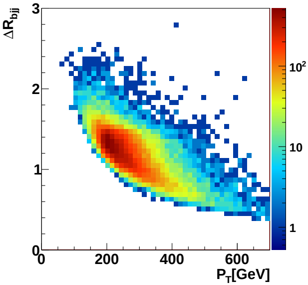
– purely leptonic decays rate limited
– semi-leptonic approximate reconstruction of neutrino 4-vector:
  massless neutrino
  2D missing energy vector
  top or W mass constraints
– purely hadronic decays deemed not useful

Many taggers [Hopkins/CMS tagger, HEPTopTagger, template tagger, shower deconstruction]

– hadronic top identification and reconstruction
– jet separation challenging for heavy Z'
– combination of calo and tracker great
– usually combined with b-tag

Validation and systematics

– tagging easier for higher boost, $p_{T,t} > 600$ GeV
– Standard Model events at lower $p_{T,t} < 400$ GeV
$\Rightarrow$ p_T range main challenge



HEPTopTagger

Mass drop algorithm [TP, Salam, Spannowsky, Takeuchi]

1– C/A fat jet, $R = 1.5$ and $p_T > 200$ GeV [FastJet limitation]

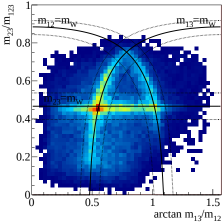
2– mass drop, cutoff $m_{\text{sub}} > 30$ GeV

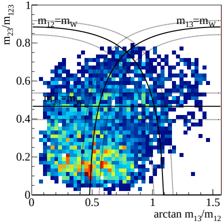
3– filtering leading to hard substructure triple

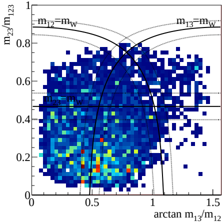
4– top mass window $m_{123} = [150, 200]$ GeV

5– A-shaped mass plane cuts as function of m_W / m_t

6– consistency condition $p_T^{(\text{tag})} > 200$ GeV







HEPTopTagger

Mass drop algorithm [TP, Salam, Spannowsky, Takeuchi]

1– C/A fat jet, $R = 1.5$ and $p_T > 200$ GeV [FastJet limitation]

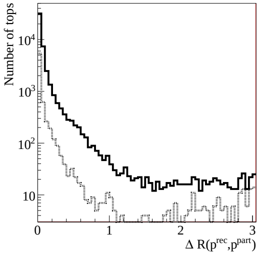
2– mass drop, cutoff $m_{\text{sub}} > 30$ GeV

3– filtering leading to hard substructure triple

4– top mass window $m_{123} = [150, 200]$ GeV

5– A-shaped mass plane cuts as function of m_W/m_t

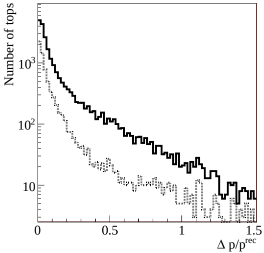
6– consistency condition $p_T^{(\text{tag})} > 200$ GeV



Top reconstruction

– direction less critical

– energy requiring calibration



HEPTopTagger

Mass drop algorithm [TP, Salam, Spannowsky, Takeuchi]

1– C/A fat jet, $R = 1.5$ and $p_T > 200$ GeV [FastJet limitation]

2– mass drop, cutoff $m_{\text{sub}} > 30$ GeV

3– filtering leading to hard substructure triple

4– top mass window $m_{123} = [150, 200]$ GeV

5– A-shaped mass plane cuts as function of m_W/m_t

6– consistency condition $p_T^{(\text{tag})} > 200$ GeV

Improvements for upcoming run

– signal efficiency
– background sculpting
– p_T range
– resonance reconstruction

ATLAS Analysis

ATLAS resonance search [CERN-PH-EP-2012-291, ATLAS-CONF-2013-084]

– resonances decaying to $t\bar{t}$ [mass 1.0-1.5 TeV]

– Z' or KK gluon only different in width, $\Gamma_{Z'} < \Gamma_G$

– semi-leptonic searches done before

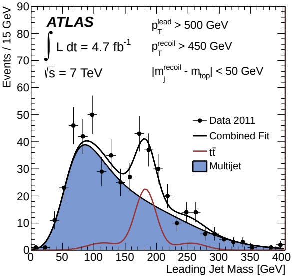
$\Rightarrow$ mostly test of top taggers

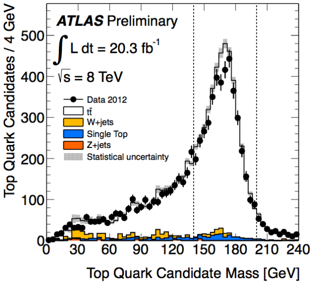
Test of subjet methods [Kasieczka, Schätzel, Anders, Schöning]

– starting with lots of jet calibration

– also add b-tag in/around fat jet

– fat jet and top masses in data [background region]





ATLAS Analysis

ATLAS resonance search [CERN-PH-EP-2012-291, ATLAS-CONF-2013-084]

– resonances decaying to $t\bar{t}$ [mass 1.0-1.5 TeV]

– Z' or KK gluon only different in width, $\Gamma_{Z'} < \Gamma_G$

– semi-leptonic searches done before

$\Rightarrow$ mostly test of top taggers

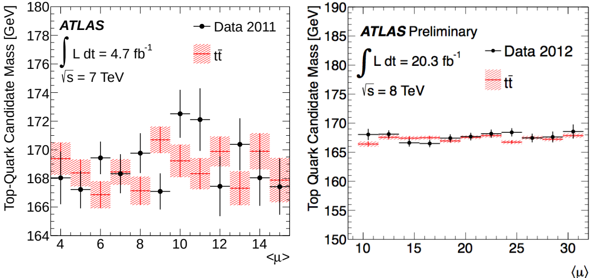
Test of subjet methods [Kasieczka, Schätzel, Anders, Schöning]

– starting with lots of jet calibration

– also add b-tag in/around fat jet

– fat jet and top masses in data [background region]

– pile-up dependence?



ATLAS Analysis

ATLAS resonance search [CERN-PH-EP-2012-291, ATLAS-CONF-2013-084]

– resonances decaying to $t\bar{t}$ [mass 1.0-1.5 TeV]

– Z' or KK gluon only different in width, $\Gamma_{Z'} < \Gamma_G$

– semi-leptonic searches done before

$\Rightarrow$ mostly test of top taggers

Test of subjet methods [Kasieczka, Schätzel, Anders, Schöning]

– starting with lots of jet calibration

– also add b-tag in/around fat jet

– fat jet and top masses in data [background region]

– pile-up dependence?

$\Rightarrow$ subjet methods established

ATLAS Analysis

ATLAS resonance search [CERN-PH-EP-2012-291, ATLAS-CONF-2013-084]

– resonances decaying to $t\bar{t}$ [mass 1.0-1.5 TeV]

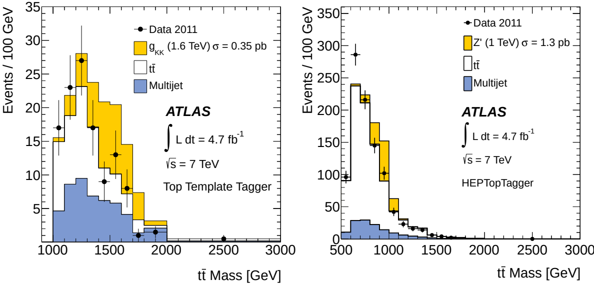
– Z' or KK gluon only different in width, $\Gamma_{Z'} < \Gamma_G$

– semi-leptonic searches done before

$\Rightarrow$ mostly test of top taggers

Resonance search

– m_{tt} from TemplateTagger and HEPTopTagger



ATLAS Analysis

ATLAS resonance search [CERN-PH-EP-2012-291, ATLAS-CONF-2013-084]

– resonances decaying to $t\bar{t}$ [mass 1.0-1.5 TeV]

– Z' or KK gluon only different in width, $\Gamma_{Z'} < \Gamma_G$

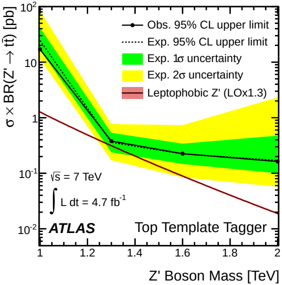
– semi-leptonic searches done before

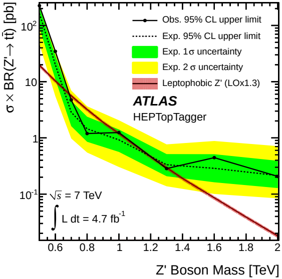
$\Rightarrow$ mostly test of top taggers

Resonance search

– m_{tt} from TemplateTagger and HEPTopTagger

– limit on Z'





ATLAS Analysis

ATLAS resonance search [CERN-PH-EP-2012-291, ATLAS-CONF-2013-084]

– resonances decaying to $t\bar{t}$ [mass 1.0-1.5 TeV]

– Z' or KK gluon only different in width, $\Gamma_{Z'} < \Gamma_G$

– semi-leptonic searches done before

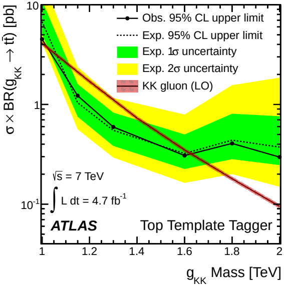
$\Rightarrow$ mostly test of top taggers

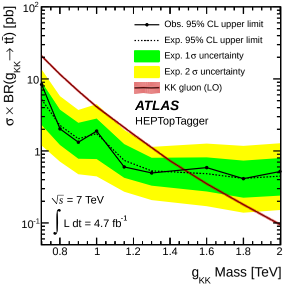
Resonance search

– m_{tt} from TemplateTagger and HEPTopTagger

– limit on Z'

– limit on KK gluon





ATLAS Analysis

ATLAS resonance search [CERN-PH-EP-2012-291, ATLAS-CONF-2013-084]

– resonances decaying to $t\bar{t}$ [mass 1.0-1.5 TeV]

– Z' or KK gluon only different in width, $\Gamma_{Z'} < \Gamma_G$

– semi-leptonic searches done before

⇒ mostly test of top taggers

Resonance search

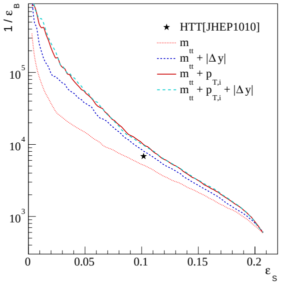
– m_{tt} from TemplateTagger and HEPTopTagger

– limit on Z'

– limit on KK gluon

⇒ it actually works!

Better HEPTopTagger



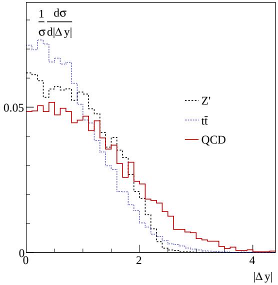


Figure : Left: ROC curves for the dominant QCD background vs. the Z' signal after including additional kinematic information shown in Eq.(2). Right: $|\Delta y|$ distribution of the reconstructed top quarks for signal and backgrounds.

Better HEPTopTagger

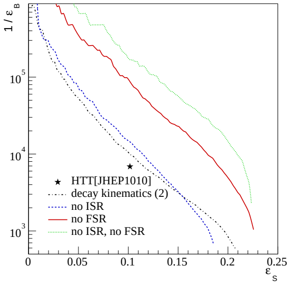


Figure : ROC curves for different combinations of initial state jet radiation (ISR) and final state jet radiation (FSR) in the Z' signal generation. The background is QCD with ISR and FSR for all curves.

Better HEPTopTagger

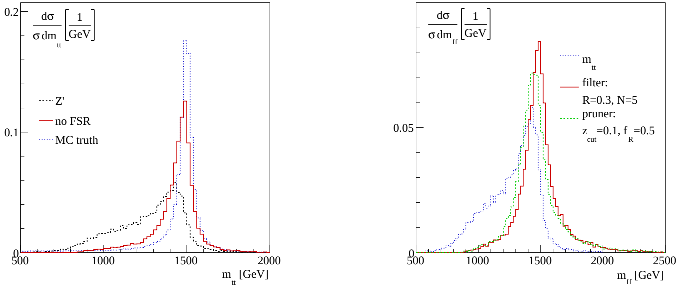


Figure : Effect of final state radiation on the invariant mass of the tagged and reconstructed $t\bar{t}$ system m_{tt} for the Z' signal (left), invariant mass $m_{\rm ff}$ of top-tagged fat jets (center), and different approaches to reconstruct the Z' mass peak (right). Monte Carlo truth is $\sqrt{p^2_{Z'}}$ with an assumed width of 65 GeV.

Better HEPTopTagger

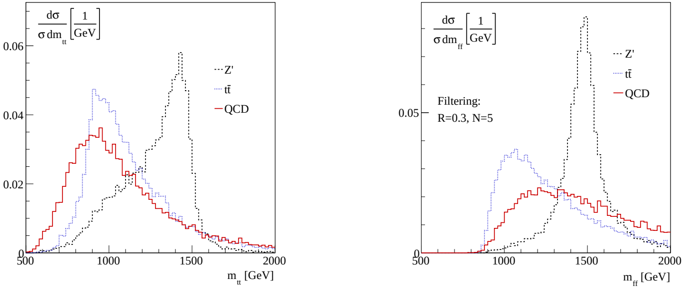


Figure : Reconstructed mass distribution of the Z' signal and the backgrounds based on the tagged tops (left) and the corresponding filtered fat jets (right).

Better HEPTopTagger

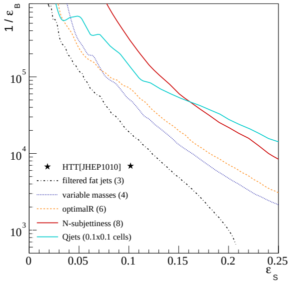


Figure : Performance of the optimalR mode based on the kinematic variables in Eq.(6) (red) and the same curve also including N-subjettiness variables as defined in Eq.(8) (orange). We only consider the dominant QCD background.

Better HEPTopTagger

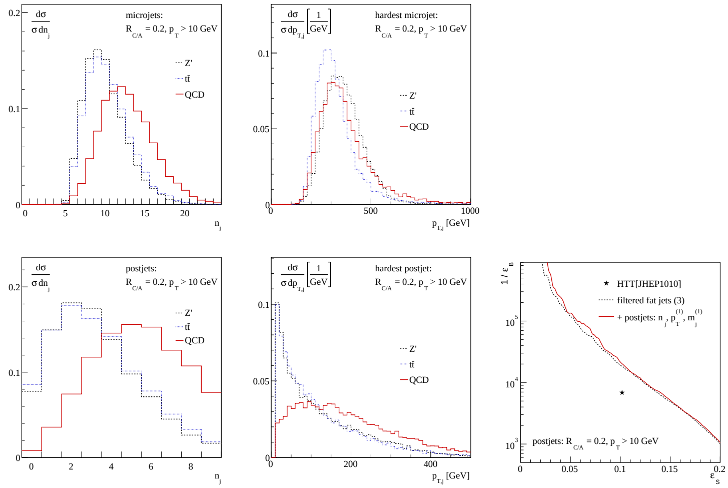


Figure : Information on the hardest jet before top tagging (upper row) and the hardest jet left over after top tagging (lower row). For the jets defined with $R = 0.2$ and $p_T > 10$ GeV we show the number of jets, the hardest jet's transverse momentum, and its mass in Z' candidate events (left to right).

Better HEPTopTagger

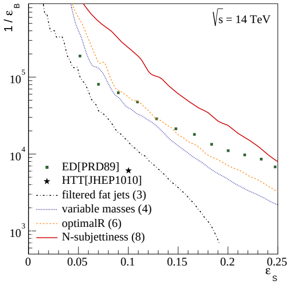


Figure : Comparison of the multivariate HEPTopTagger2 analysis presented in this paper with event deconstruction. This comparison in the absence of an experimental validation should be taken as first estimate.

Better HEPTopTagger

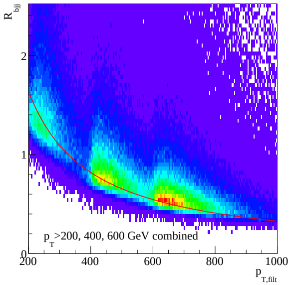


Figure : $R_{\text{opt}}^{(\text{calc})}$ fit based on Standard Model $t\bar{t}$ samples with $p_{T,t} > 200, 400, 600$ GeV for the parton level distance of decay products R_{bjj}. The fat jets are filtered with $R = 0.2$, $N = 10$. The functional form of the fit curve is given in Eq.(14).

Better HEPTopTagger

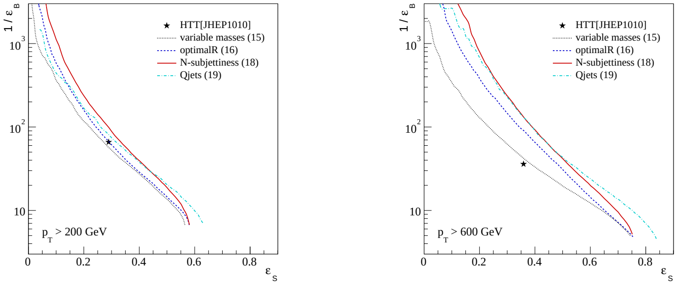


Figure : Performance of the HEPTopTagger2 for $t\bar{t}$ production in the Standard Model. We show the incremental improvements from the extended multivariate analyses for top quarks with $p_{T,t} > 200$ GeV and $p_{T,t} > 600$ GeV.

Outlook

Fat Jets are...

...turning jet physics into a cool topic

...a fast–moving and happy field

...bringing together experiment and theory

...an opportunity for young people to have impact

...always honoring good ideas

Sorry for not discussing...

...template taggers

...wavelet taggers

...your favorite tagger

Thank you to...

...former and current ATLAS–Heidelberg

...all groups working with and on taggers

Lectures on LHC Physics, Springer, arXiv:0910.4182 updated under `www.thphys.uni-heidelberg.de/~plehn/`