Top Tagging: Modern and More Modern Methods

Tilman Plehn

Universität Heidelberg

Bangalore 1/2017

Fat jets

Boosted particles at the LHC [Steve's talk]

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]

...

2014 first machine learning study [Cogan, Kagan, Strass, Schwartzman]

2017 machine learning vs tagger $t \to 3$ jets [Kasieczka, Plehn, Russell, Schell]

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

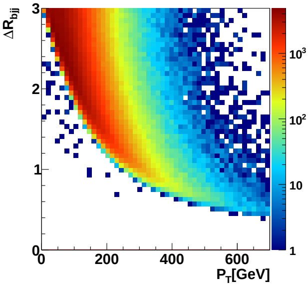
Many taggers [Hopkins/CMS tagger, HEPTopTagger, template tagger, shower deconstruction]

- 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

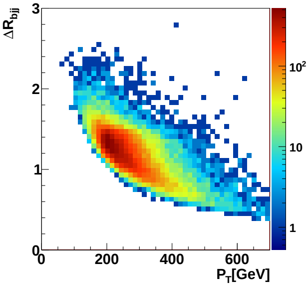
– usually combined with b-tag

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



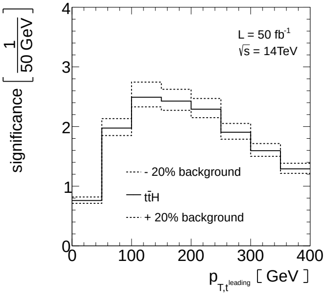
Analysis: $t\bar{t}H, H \to b\bar{b}$

Challenge of $t\bar{t}H, H \to b\bar{b}$ [CMS-TDR]

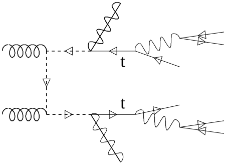
- trigger: $t \to bW^+ \to b\ell^+\nu$
  reconstruction and rate: $\bar{t} \to \bar{b}W^- \to \bar{b}jj$
- continuum background $t\bar{t}b\bar{b}, t\bar{t}jj$ [weighted by b-tag]
- 1– combinatorics: m_H in $pp \to 4b_{tag}\, 2j\, \ell\nu$
  2– kinematics: peak-on-peak
  3– systematics: $S/B \sim \mathcal{O}(0.1)$

Tagger [TP, Salam, Spannowsky; MadMax: TP, Schichtel, Wiegand]

- boost solving combinatorics [buckets]
- significance known to be in boosted regime
- side band in m_{bb} possible
- fun: fat Higgs jet + fat top jet



Analysis: top partners

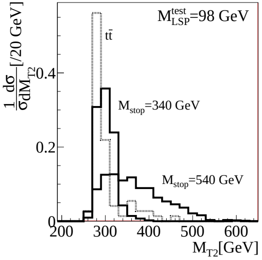


Stop pairs including reconstruction [TP, Spannowsky, Takeuchi, Zerwas]

- stop crucial for hierarchy problem
  comparison to other top partners
- hadronic: $\tilde{t}\tilde{t}^* \to t\tilde{\chi}_1^0\ \bar{t}\tilde{\chi}_1^0$
- stop mass from m_{T2} endpoint [like sleptons or sbottoms]

$$m_{T2}(\hat{m}_\chi) = \min_{\not{p}_T = q_1 + q_2} \left[\max_j \ m_{T,j}(q_j; \hat{m}_\chi) \right] \stackrel{!}{<} m_{\tilde{t}}$$

$\Rightarrow$ hadronic search as easy as $b\bar{b} + \not{E}_T$ [ATLAS Heidelberg]



N-Subjettiness

N-Jettiness to count subjets inside fat jet [Thaler, van Tilburg; Jesse's talk?]

– how many subjets do the calo entries correspond to?

– event shape using N subjet directions $\hat{n}_j$ [$\beta > 0$]

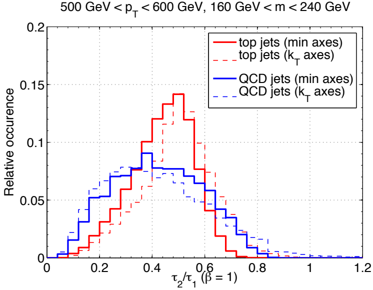
$$\tau_N = \frac{1}{\sum_{\alpha\in\text{jet}} p_{T,\alpha} R_0^\beta} \sum_{\alpha\in\text{jet}} p_{T,\alpha} \min_{k=1,\dots,N} (\Delta R_{k,\alpha})^\beta$$

– choice of reference axes

1- from subjet algorithm
2- from minimization of τ_N

– $\tau_N \to 1$ means many calo entries away from N axes
$\tau_N \to 0$ means perfect matching

– systematics cancelled in ratios
τ_{N+1}/τ_N with dip for $N+1$ subjets



N-Subjettiness

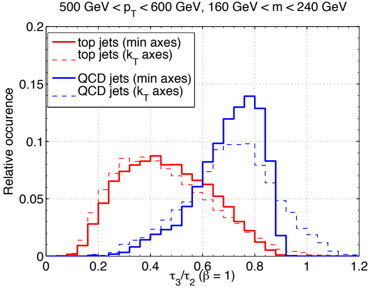
N-Jettiness to count subjets inside fat jet [Thaler, van Tilburg; Jesse's talk?]

- how many subjets do the calo entries correspond to?
- event shape using N subjet directions $\hat{n}_j$ [$\beta > 0$]

$$\tau_N = \frac{1}{\sum_{\alpha \in \text{jet}} p_{T,\alpha} R_0^\beta} \sum_{\alpha \in \text{jet}} p_{T,\alpha} \min_{k=1,\dots,N} (\Delta R_{k,\alpha})^\beta$$

- choice of reference axes

  1- from subjet algorithm
  2- from minimization of τ_N
- $\tau_N \to 1$ means many calo entries away from N axes
  $\tau_N \to 0$ means perfect matching
- systematics cancelled in ratios
  τ_{N+1}/τ_N with dip for $N+1$ subjets



N-Subjettiness

N-Jettiness to count subjets inside fat jet [Thaler, van Tilburg; Jesse's talk?]

– how many subjets do the calo entries correspond to?

– event shape using N subjet directions $\hat{n}_j$ [$\beta > 0$]

$$\tau_N = \frac{1}{\sum_{\alpha \in \text{jet}} p_{T,\alpha} R_0^\beta} \sum_{\alpha \in \text{jet}} p_{T,\alpha} \min_{k=1,\dots,N} (\Delta R_{k,\alpha})^\beta$$

– choice of reference axes

1- from subjet algorithm
2- from minimization of τ_N

– $\tau_N \to 1$ means many calo entries away from N axes
$\tau_N \to 0$ means perfect matching

– systematics cancelled in ratios
τ_{N+1}/τ_N with dip for $N+1$ subjets

Tagger

– simple selection

$$m_\text{fat jet} = 160 \dots 240 \;\; \text{GeV} \qquad\qquad \frac{\tau_3}{\tau_2} < 0.6$$

– multi-variate in N and β with some improvement

$\Rightarrow$ easily added to any other tagger

HEPTopTagger

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

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

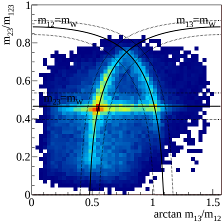
2– mass drop, cutoff $m_{\text{sub}} > 30$ GeV

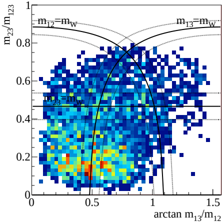
3– filtering leading to hard substructure triple

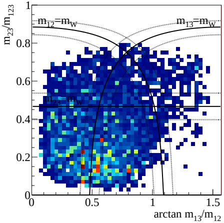
4– top mass window $m_{123} = [150, 200]$ GeV

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]

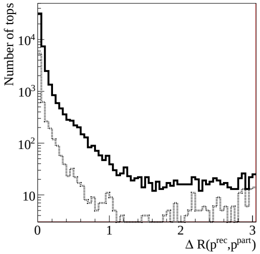
2– mass drop, cutoff $m_{\text{sub}} > 30$ GeV

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

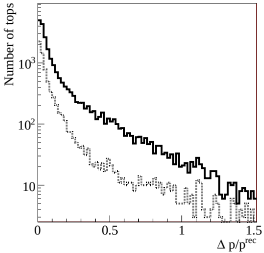
6– consistency condition $p_T^{(\text{tag})} > 200$ GeV



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

Limitations to watch

– signal efficiency

– background sculpting

– p_T range

– resonance reconstruction

A masterpiece in diplomacy

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

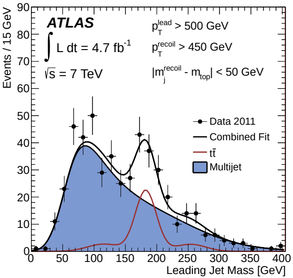
$\Rightarrow$ mostly test of top taggers

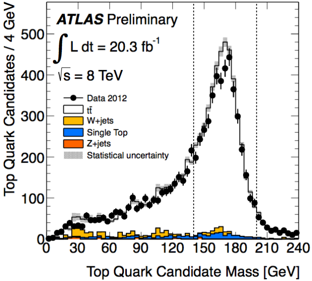
Subjet methods [Kasieczka, Schätzel, Anders (ATLAS Heidelberg)]

– starting with lots of jet calibration

– also add b-tag in/around fat jet

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





A masterpiece in diplomacy

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

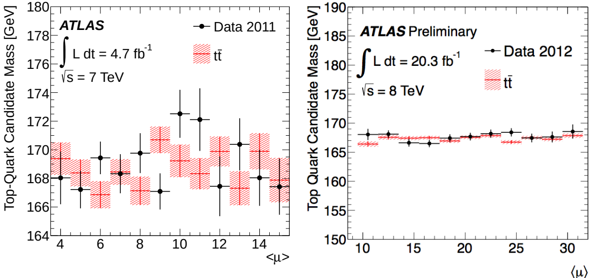
Subjet methods [Kasieczka, Schätzel, Anders (ATLAS Heidelberg)]

– 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?



A masterpiece in diplomacy

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

Subjet methods [Kasieczka, Schätzel, Anders (ATLAS Heidelberg)]

– 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?
⇒ top tagger basics established

A masterpiece in diplomacy

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

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

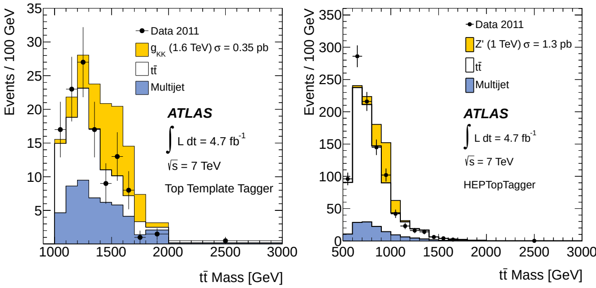
– Z' or KK gluon only different in width, $\Gamma_{Z'} < \Gamma_G$

– semi-leptonic searches done before

$\Rightarrow$ mostly test of top taggers

Resonance search

– m_{tt} from TemplateTagger and HEPTopTagger



A masterpiece in diplomacy

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, $\Gamma_{Z'} < \Gamma_G$

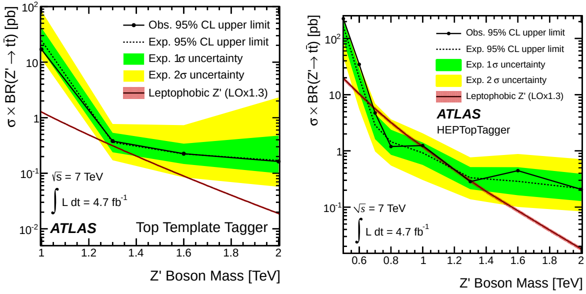
– semi-leptonic searches done before

$\Rightarrow$ mostly test of top taggers

Resonance search

– m_{tt} from TemplateTagger and HEPTopTagger

– limit on Z'



A masterpiece in diplomacy

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

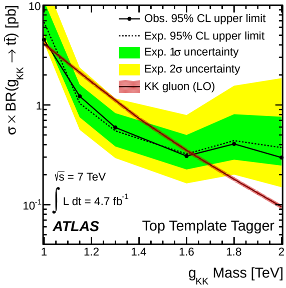
Resonance search

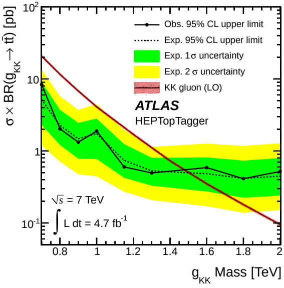
– m_{tt} from TemplateTagger and HEPTopTagger

– limit on Z'

– limit on KK gluon

$\Rightarrow$ it all works!





Better HEPTopTagger

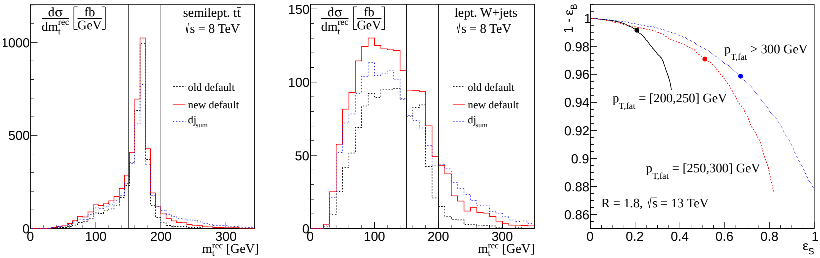
Multivariate BDT setup [TP, Spannowsky, Takeuchi; Anders, Bernaciak, Kasieczka, TP, Schell]

– increase size $R = 1.8$

– change in cut order $3 \leftrightarrow 4$

– alternative triplet selection
maximal $dj_{\text{sum}} = \sum p_{T,i} p_{T,j} (\Delta R_{ij})^4$
triplet of hardest subjets

– $\{m_{13}, m_{23}, m_{123}\}$



Better HEPTopTagger

Multivariate BDT setup [TP, Spannowsky, Takeuchi; Anders, Bernaciak, Kasieczka, TP, Schell]

– increase size $R = 1.8$

– change in cut order $3 \leftrightarrow 4$

– alternative triplet selection
maximal $dj_{\text{sum}} = \sum p_{T,i} p_{T,j} (\Delta R_{ij})^4$
triplet of hardest subjets

– $\{ m_{13}, m_{23}, m_{123} \}$

Moderate boost [Anders, Bernaciak, Kasieczka, TP, Schell]

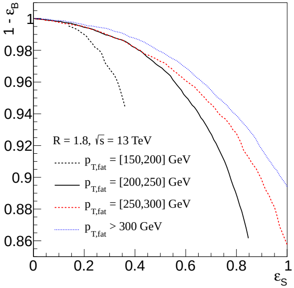
– focus on $p_T = [150, 200]$ GeV

– target type-2 tags [two hardest subjets matched]

– correlations via Fox–Wolfram moments

$$H_\ell^U = \sum_{i,j} \frac{1}{N^2} P_\ell(\cos \Omega_{ij})$$

– $\{ m_{13}, m_{23}, m_{123}, \text{FWMs} \}$



HEPTopTagger2

OptimalR [Kasieczka, TP, Salam, Schell, Strebler]

– optimal fat jet size R_{min} [large to include decay jets, small to avoid combinatorics]

– reduce R until we leave jet mass plateau

$$|m_{123} - m_{123}^{(R_{\text{max}})}| < 0.2\, m_{123}^{(R_{\text{max}})} \quad \Rightarrow \quad R_{\text{min}}$$

– estimate $R_{\text{min}}^{(\text{calc})}$ from kinematics

– $\{ m_{123}^{(R_{\text{min}})}, f_W^{(R_{\text{min}})}, R_{\text{min}} - R_{\text{min}}^{(\text{calc})} \}$ [$f_W = \min |m_{ij}/m_{123} - m_W/m_t|$]

HEPTopTagger2

OptimalR [Kasieczka, TP, Salam, Schell, Strebler]

– optimal fat jet size R_min [large to include decay jets, small to avoid combinatorics]

– reduce R until we leave jet mass plateau

$$|m_{123} - m_{123}^{(R_\text{max})}| < 0.2\, m_{123}^{(R_\text{max})} \quad \Rightarrow \quad R_\text{min}$$

– estimate $R_\text{min}^\text{(calc)}$ from kinematics

– $\{m_{123}^{(R_\text{min})}, f_W^{(R_\text{min})}, R_\text{min} - R_\text{min}^\text{(calc)}\}$ [$f_W = \min|m_{ij}/m_{123} - m_W/m_t|$]

N-Subjettiness [Thaler, van Tilburg]

– also include rejected events [$R_\text{filt} = 0.2, N_\text{filt} = 5$]

– $\{m_{123}^{(R_\text{min})}, f_W^{(R_\text{min})}, R_\text{min} - R_\text{min}^\text{(calc)}, \tau_j, \tau_j^\text{(filt)}\}$

HEPTopTagger2

OptimalR [Kasieczka, TP, Salam, Schell, Strebler]

– optimal fat jet size R_min [large to include decay jets, small to avoid combinatorics]

– reduce R until we leave jet mass plateau

$$|m_{123} - m_{123}^{(R_\text{max})}| < 0.2\, m_{123}^{(R_\text{max})} \quad \Rightarrow \quad R_\text{min}$$

– estimate $R_\text{min}^\text{(calc)}$ from kinematics

– $\{ m_{123}^{(R_\text{min})}, f_W^{(R_\text{min})}, R_\text{min} - R_\text{min}^\text{(calc)} \}$ $[f_W = \min |m_{ij}/m_{123} - m_W/m_t|]$

N-Subjettiness [Thaler, van Tilburg]

– also include rejected events $[R_\text{filt} = 0.2, N_\text{filt} = 5]$

– $\{ m_{123}^{(R_\text{min})}, f_W^{(R_\text{min})}, R_\text{min} - R_\text{min}^\text{(calc)}, \tau_j, \tau_j^\text{(filt)} \}$

Qjets [Ellis, Hornig, Roy, Krohn, Schwartz]

– more than one clustering history, weighted by

$$\omega_{ij} = \exp \left[-\alpha \frac{y_{ij} - y_{ij}^\text{min}}{y_{ij}^\text{min}} \right]$$

then using distributions like $\langle m^2 \rangle - \langle m \rangle^2$

– $\{ m_{123}^{(R_\text{min})}, f_W^{(R_\text{min})}, R_\text{min} - R_\text{min}^\text{(calc)}, \tau_j, \tau_j^\text{(filt)}, \{ m_{123}^\text{Qjets} \} \}$

HEPTopTagger2

OptimalR [Kasieczka, TP, Salam, Schell, Strebler]

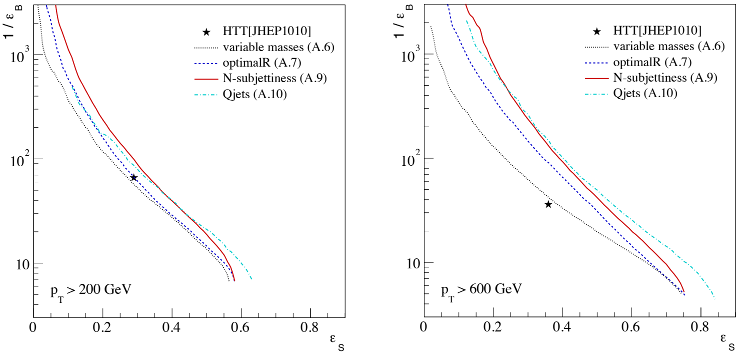
– $\{m_{123}^{(R_{\text{min}})}, f_W^{(R_{\text{min}})}, R_{\text{min}} - R_{\text{min}}^{(\text{calc})}\}$ $[f_W = \min|m_{ij}/m_{123} - m_W/m_t|]$

N-Subjettiness [Thaler, van Tilburg]

– $\{m_{123}^{(R_{\text{min}})}, f_W^{(R_{\text{min}})}, R_{\text{min}} - R_{\text{min}}^{(\text{calc})}, \tau_j, \tau_j^{(\text{filt})}\}$

Qjets [Ellis, Hornig, Roy, Krohn, Schwartz]

– $\{m_{123}^{(R_{\text{min}})}, f_W^{(R_{\text{min}})}, R_{\text{min}} - R_{\text{min}}^{(\text{calc})}, \tau_j, \tau_j^{(\text{filt})}, \{m_{123}^{\text{Qjets}}\}\}$



Non-improvements

Tagging without fat jets [Lapsien, Kogler, Haller; FastJetContrib]

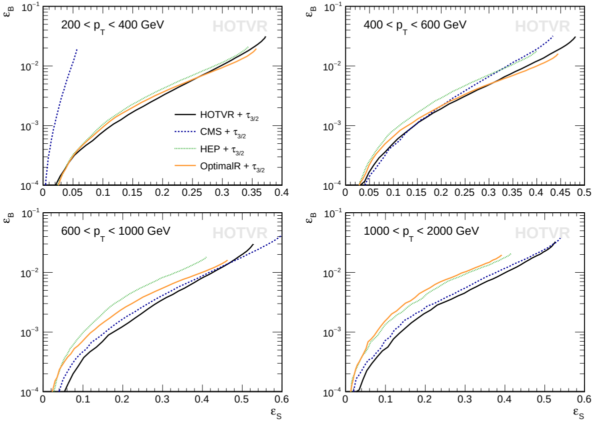
1– variable-R: jet size depending on p_T [Krohn, Thaler, Wang]

2– mass jump: jet mass condition in C-A jet algorithm [Stoll]
specifically $m_{ij} \lesssim 30$ GeV and $m_{ij} \gtrsim 1.4 \max(m_{ij})$

– possible advantage: not relying on fixed fat jet

– comparison with HEPTopTagger, CMSTagger in working points
pseudo-ROC curve by adding $\tau_{2,3}$ [why???]

– improvement?



Non-improvements

Tagging without fat jets [Lapsien, Kogler, Haller; FastJetContrib]

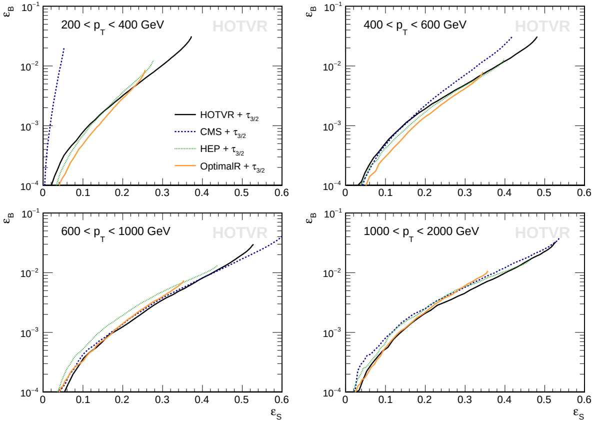
1– variable-R: jet size depending on p_T [Krohn, Thaler, Wang]

2– mass jump: jet mass condition in C-A jet algorithm [Stoll]
specifically $m_{ij} \lesssim 30$ GeV and $m_{ij} \gtrsim 1.4 \max(m_{ij})$

– possible advantage: not relying on fixed fat jet

– comparison with HEPTopTagger, CMSTagger in working points
pseudo-ROC curve by adding $\tau_{2,3}$ [why???]

– not in v2!



Non-improvements

Tagging without fat jets [Lapsien, Kogler, Haller; FastJetContrib]

1– variable-R: jet size depending on p_T [Krohn, Thaler, Wang]

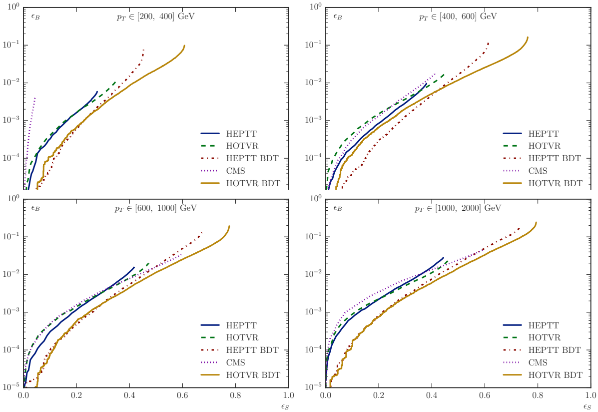
2– mass jump: jet mass condition in C-A jet algorithm [Stoll]
specifically $m_{ij} \lesssim 30$ GeV and $m_{ij} \gtrsim 1.4 \max(m_{ij})$

– possible advantage: not relying on fixed fat jet

– comparison with HEPTopTagger, CMSTagger in working points
pseudo-ROC curve by adding $\tau_{2,3}$ [why???]

– proper job

[no Qjets, $p_T > 30$ GeV]



Non-improvements

Tagging without fat jets [Lapsien, Kogler, Haller; FastJetContrib]

1– variable-R: jet size depending on p_T [Krohn, Thaler, Wang]

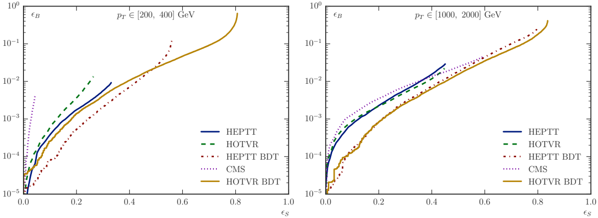
2– mass jump: jet mass condition in C-A jet algorithm [Stoll]
specifically $m_{ij} \lesssim 30$ GeV and $m_{ij} \gtrsim 1.4 \max(m_{ij})$

– possible advantage: not relying on fixed fat jet

– comparison with HEPTopTagger, CMSTagger in working points
pseudo-ROC curve by adding $\tau_{2,3}$ [why???]

– proper job

[no Qjets, all p_T]



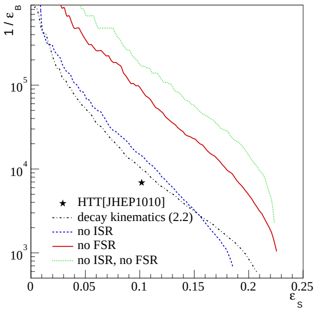
$\Rightarrow$ tagger performance stable and plateau'd

$\Rightarrow$ find other ways to improve?

Resonance search

Final state radiation

– Z' search without ISR/FSR

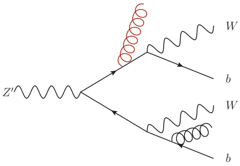


Resonance search

Final state radiation

– Z' search without ISR/FSR

– FSR major problem

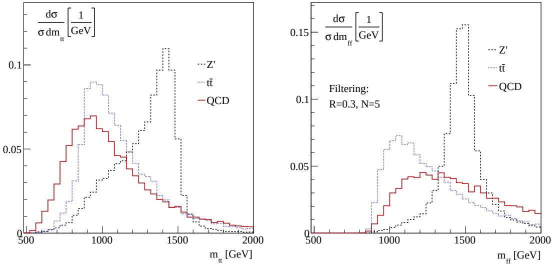


Resonance search

Final state radiation

- Z' search without ISR/FSR
- FSR major problem
- using reconstructed 4-momentum of tagged top
- tail in $m_{tt}^{(\text{rec})}$ from tagged momentum and FSR

$\Rightarrow$ add 4-momentum of fat jet: $\{m_{tt}, p_{T,t}, m_{jj}^{(\text{filt})}, p_{T,j}^{(\text{filt})}\}$



Resonance search

Final state radiation

– Z' search without ISR/FSR

– FSR major problem

– using reconstructed 4-momentum of tagged top

– tail in $m_{tt}^{(\text{rec})}$ from tagged momentum and FSR

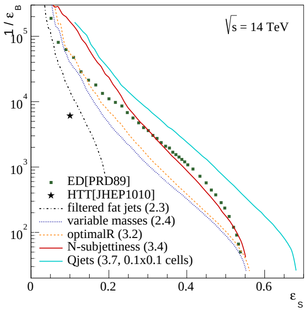
$\Rightarrow$ add 4-momentum of fat jet: $\{m_{tt}, p_{T,t}, m_{jj}^{(\text{filt})}, p_{T,j}^{(\text{filt})}\}$

Full blast HEPTopTagger2

– $Z' \to t\bar{t}$ signal

– high efficiencies remaining a challenge

$\Rightarrow$ ball in experimenalists' field



Neural networks

Recent developments in image regonition [convolutional networks]

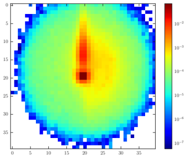
– wavelet transformation [Rentala, Shepherd, Tait; Monk]

– W-tagging with image recognition [Cogan etal, Oliveira etal, Baldi etal]

– top-tagging attempt [Almeida, Backovic, Cliche, Lee, Perelstein]

– QCD and shower study [Barnard etal]

$\Rightarrow$ does image recognition beat taggers?



Making calorimeter images easy to read [η vs ϕ energy bins]

– find maxima identify characteristic points

– shift move image to center the global maximum

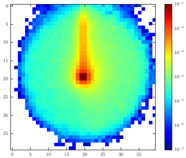
– rotation rotate the second maximum to 12 o'clock

– flip ensure third maximum is in the right half-plane

– crop crop the image to 40×40 pixels

– decide on E vs E_T for rapidities $\eta \gtrsim 2$

$\Rightarrow$ all not strictly necessary, but obviously CPU-saving



Neural networks

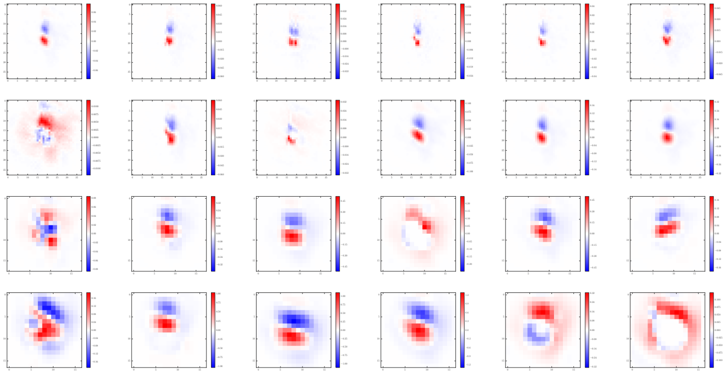
Getting binned calorimeter images ready [$n_\eta \times n_\phi$ energy bins]

– Convolutional layers to transform into optimal basis with kernel matrix W

– Activation to weigh the feature maps $\vec{x}$, for example step function

– Pooling to reduce the size of picture

– ConvNet output 6 kernels W, 4 layers/iterations



Neural networks

Getting binned calorimeter images ready [$n_\eta \times n_\phi$ energy bins]

- Convolutional layers   to transform into optimal basis with kernel matrix W
- Activation   to weigh the feature maps $\vec{x}$, for example step function
- Pooling   to reduce the size of picture

- DNN output   to link to data

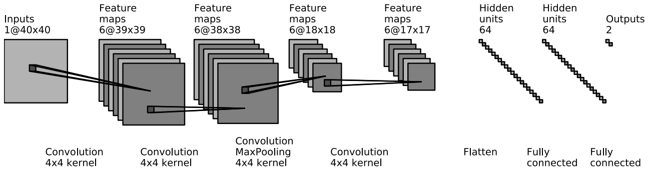
$$y_i = \max\left(0, \sum_{j=1}^{n^2} W_{ij} x_j + b_i\right) \qquad \text{for} \quad i \in [1, n_{\text{d-node}}]$$

- loss function   to train $\{W, b\}$ by minimization

$$L(\theta) = \frac{1}{N} \sum_{i=0}^{N} \left(y(W, b; x_i) - y_i\right)^2$$

- minimization   with learning rate, momentum

$$\theta_{n+1} = \theta_n - \eta \nabla L(\theta_n) + \alpha(\theta_n - \theta_{n-1})$$



Performance

Compare ConvNet with QCD tagger — preliminary! [including N-subjettiness, no Qjets]

1– state of the art HEPTopTagger2

2– ConvNet trained and applied on same simulation environment

⇒ QCD is safe, thinking and understanding hard to beat

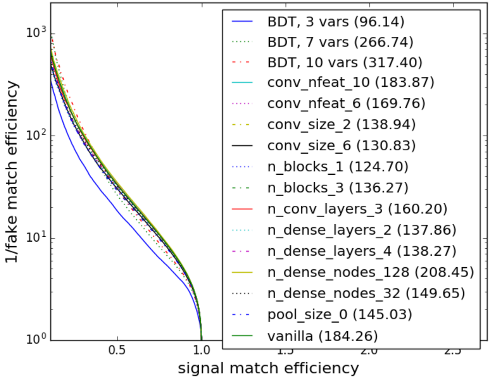
Performance

Compare ConvNet with QCD tagger — preliminary!

[including N-subjettiness, no Qjets]

1– state of the art HEPTopTagger2

2– ConvNet trained and applied on same simulation environment



⇒ QCD is safe, thinking and understanding hard to beat

Outlook

I am probably over time, so...

...QCD-based taggers are stable

...additional information beyond jets is needed

...some progress does not survive v2

...imagine recognition is cool, but QCD at least as cool

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/`