

News from the HEPTopTagger

arXiv:1312.1504 & work with C. Anders, C. Bernaciak, G. Kasieczka, T. Plehn, G. Salam, and T. Strebler

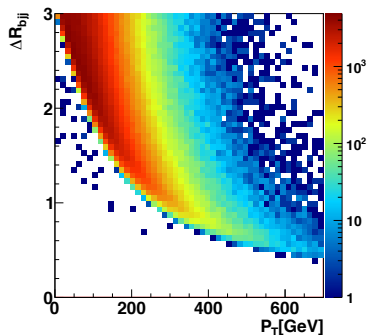
Torben Schell

Institute for Theoretical Physics, Heidelberg University

Boost 2014
August 21, 2014

HEPTopTagging

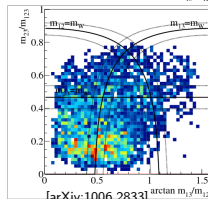
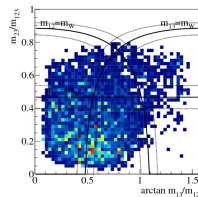
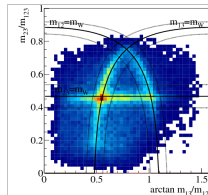
- reconstruction of boosted hadronic tops
- collimated decay products
 - fat jets
 - reduced combinatorial problems
- SM: number of top quarks vs. collimation
- substructure analysis based on subjet masses



HEPTopTagger – Steps

[arXiv:1006.2833]

- 0 **fat jet:** $C/A \ R = 1.5$, $p_T > 200$ GeV
- 1 **hard substructures:**
mass drop $f_{\text{drop}} = 0.8$, $m_i < m_{\text{sub}} = 30$ GeV
- 2 **filtering:**
filter a triple of hard substructures \rightarrow 3 jets (j_1, j_2, j_3)
- 3 **mass window:** $150 \text{ GeV} < m_{123} < 200 \text{ GeV}$
- 4 **mass plane cuts:** $0.85 \frac{m_W}{m_t} < \frac{m_{ij}}{m_{123}} < 1.15 \frac{m_W}{m_t}$
 $m_{23} \approx m_W$: $0.2 < \arctan \frac{m_{13}}{m_{12}} < 1.3$; else $\frac{m_{23}}{m_{123}} > 0.35$
- 5 **consistency:** $p_T^{(\text{tag})} > 200$ GeV



Needed improvements

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



Recent developments

- $R = 1.8$ with change in cut order ($3 \leftrightarrow 4$)

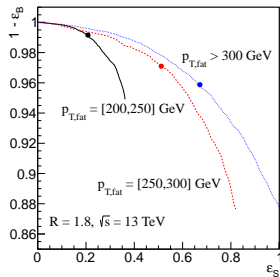
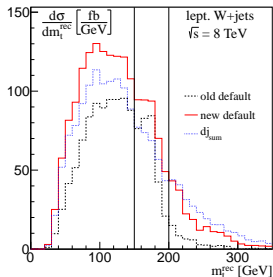
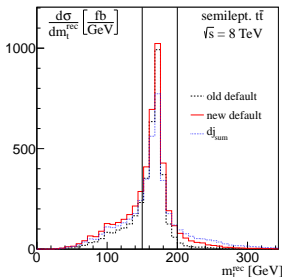
[arXiv:1312.1504]

- alternative triplet selections

- maximal $d_{j_{\text{sum}}} = \sum_{(ij)} d_{ij}$ with $d_{ij} = p_{T,i} p_{T,j} (\Delta R_{ij})^4$
- triplet of hardest subjects

- Boosted Decision Tree ROC curves

no window in m_{123} $\{\arctan(m_{13}/m_{12}), m_{23}/m_{123}, m_{123}, (m_w/m_t)_{\text{rec}}\}$



(Alpgen+Pythia6 MLM semilept. $t\bar{t} + \leq 2j$ vs. lept. $W + (2 - 4)j$)

Low p_T mode

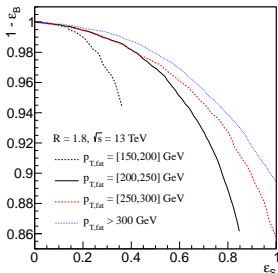
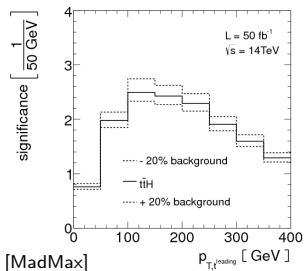
- target rejected candidates with $p_T \in [150, 200]$ GeV
- focus on **type-2** tags
- angular correlations \rightarrow **Fox-Wolfram moments**

$$H_I^x = \sum_{i,j=1}^N W_{ij}^x P_I(\cos \Omega_{ij}) \quad \text{with} \quad W_{ij}^U = \frac{1}{N^2}$$

- BDT: $\{\arctan(m_{13}/m_{12}), m_{23}/m_{123}, m_{123}, (m_w/m_t)_{\text{rec}}, \text{FWMs}\}$

	default		low- p_T mode	
	(mis)tags [fb]	fraction	(mis)tags [fb]	fraction
type-1	5309	57%	5967	52%
type-2	1283	14%	1863	16%
type-3	2712	29%	3601	32%
ε_S	0.287		0.353	
W +jets	1200		1663	
ε_B	0.007		0.010	

example working point

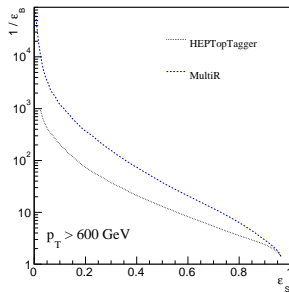
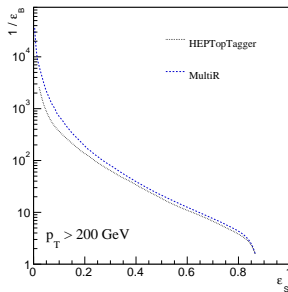


MultiR Tagger

- there is an optimal $R^{(\text{opt})}$
- reduce R until leaving top mass plateau

$$|m_{123}^{(R)} - m_{123}^{(R_{\max})}| < 0.2 m_{123}^{(R_{\max})} \rightarrow R_{\min}$$

- estimate as $R_{\min}^{(\text{calc})} \rightarrow$ additional variable $R_{\min} - R_{\min}^{(\text{calc})}$
- BDT: $\{ m_{123}^{(R_{\min})}, f_W^{(R_{\min})}, R_{\min} - R_{\min}^{(\text{calc})} \}$, $f_W = \min |m_{ij}/m_{123} - (m_W/m_t)_{\text{true}}|$



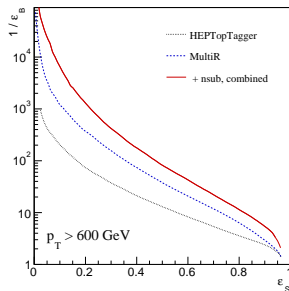
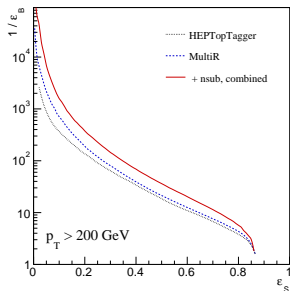
($\sqrt{s} = 13$ TeV, Pythia8 $t\bar{t}$ vs. QCD di-jet, BDT(HTT): $\{ m_{123}, f_W \}$)

N -Subjettiness HEPTopTagger

- MultiR working point + rejected events + N -Subjettiness [Thaler, Van Tilburg]
- two different filterings and BDT analyses

passed: $R_{\text{filt}} = 0.3$, $N_{\text{filt}} = 3$ rejected: $R_{\text{filt}} = 0.2$, $N_{\text{filt}} = 5$

$$\{ m_{\text{fat}}^{(\text{filt})}, m_{123}^{(R_{\text{min}})}, f_W^{(R_{\text{min}})}, R_{\text{min}} - R_{\text{min}}^{(\text{calc})}, \tau_i, \tau_i^{(\text{filt})} \} \quad i \leq 3$$



Resonance Reconstruction

Benchmark: Event Deconstruction [Soper, Spannowsky, arXiv:1402.1189]

Event generation:

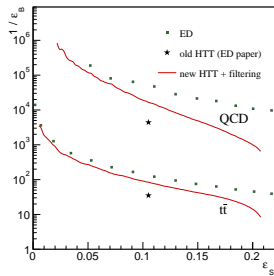
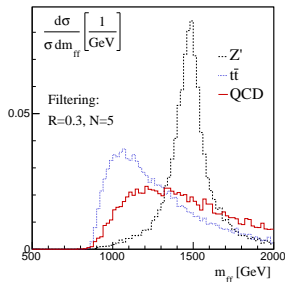
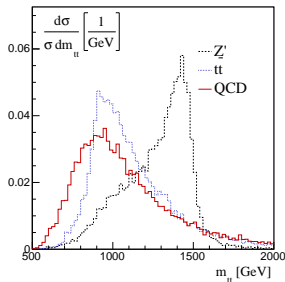
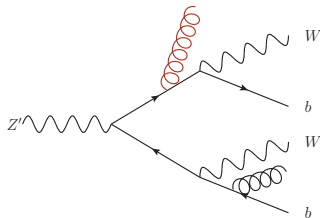
- Pythia8, LHC $\sqrt{s} = 13$ TeV
- signal: $Z' \rightarrow t_h \bar{t}_h$, $m_{Z'} = 1500$ GeV, $\Gamma(Z') = 65$ GeV
- background: QCD-dijet & $t_h \bar{t}_h$, both $p_T > 400$ GeV
- no detector simulation

Event selection:

- 2 hardest C/A, $R = 1.5$ fat jets (FastJet)
 - require $p_{T,\text{fat}} > 400$ GeV and $|y_{\text{fat}}| < 2.5$
- dominant background: QCD-dijet

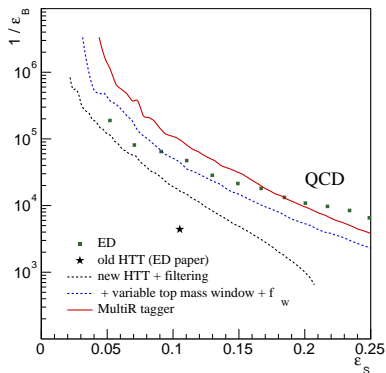
Final State Radiation

- HTT working point + **BDT** $\{ m_{tt}, p_{T,j} \}$
- HTT reconstructs on-shell tops
 → misses **final state radiation**
 → sizeable tail in m_{tt} distribution
- BDT: $\{ m_{tt}, p_{T,j}, m_{ff}^{(filt)}, p_{T,f_j} \}$



Further optimization

- beyond working point: double top mass window and A-band width
 $\{m_{tt}, p_{T,j}, m_{ff}^{(filt)}, p_{T,fj}^{(filt)}, \min(m_{123}), \max(m_{123}), f_W\}, f_W = \max(f_{W,1}, f_{W,2})$
- MultiR** $\{m_{tt}, p_{T,j}, m_{ff}^{(filt)}, p_{T,fj}^{(filt)}, \min(m_{123}^{(R_{min})}), \max(m_{123}^{(R_{min})}), f_W^{(R_{min})}, \max(R_{min} - R_{min}^{(calc)})\}$
- N -Subjettiness to come



Summary

- signal efficiency ✓
→ $R = 1.8$, inverted cut order, BDTs
- background sculpting ✓
→ alternative triplet selections
- p_T range ✓
→ low- p_T mode
→ high p_T : MultiR, N -Subjettiness HEPTopTagger
- resonance reconstruction ✓
→ account for final state radiation, MultiR

HEPTopTagger ready for LHC run II