Top and Higgs Tagging and Where it Helps

Tilman Plehn

Universität Heidelberg

DESY 01/2011

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 heavy resonances [Agashe, Belyaev, Krupovnickas, Perez, Virzi]

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

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

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

2010 boosted $t \to 3$ jets from top partners [HEPTopTagger: TP, Spannowsky, Takeuchi, Zerwas]

. . .

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

. . .

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 heavy resonances [Agashe, Belyaev, Krupovnickas, Perez, Virzi]

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

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

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

2010 boosted $t \to$ 3 jets from top partners [HEPTopTagger: TP, Spannowsky, Takeuchi, Zerwas]

...

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

...

Jet Algorithms

Definition of jets

- jet–parton duality $\Leftrightarrow$ what are partons in detector?
- need algorithm to reconstruct what was one parton
- stable w.r.t inclusion of soft radiation [IR save]
- crucial for any LHC analysis

Different measures [tool: FASTJET]

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

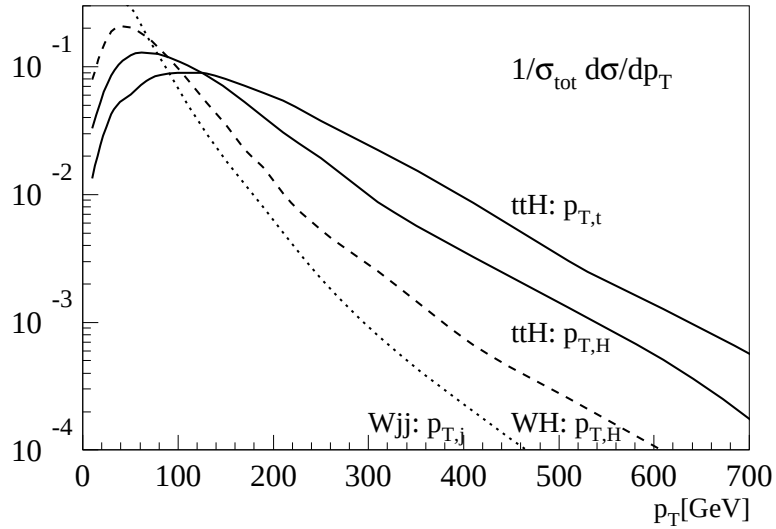
$$
\begin{array}{lll}
k_T & y_{ij} = \dfrac{\Delta R_{ij}}{D} \min\left(p_{T,i}, p_{T,j}\right) & y_{iB} = p_{T,i} \\
\text{C/A} & y_{ij} = \dfrac{\Delta R_{ij}}{D} & y_{iB} = 1 \\
\text{anti-}k_T & y_{ij} = \dfrac{\Delta R_{ij}}{D} \min\left(p_{T,i}^{-1}, p_{T,j}^{-1}\right) & y_{iB} = p_{T,i}^{-1}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

- (1) find minimum $y_{\text{min}} = \min_{kl}(y_{kl}, y_{kB})$
  (2a) if $y_{\text{min}} = y_{kl} < y_{\text{cut}}$ combine k and l, go to (1)
  (2b) if $y_{\text{min}} = y_{kB} < y_{\text{cut}}$ remove k, go to (1)
  (2c) if $y_{\text{min}} > y_{\text{cut}}$, done
- theoretical and experimental trade-off decisions
- fat jets: use clustering history

Example 1: $VH, H \to b\bar{b}$

New strategy for $H \to bb$ [Butterworth, Davison, Rubin, Salam]

- desperately needed [2/3 of all light Higgses; impact Dührssen & SFitter]
  but killed by continuum $Vb\bar{b}$ background
- S: large m_{bb}, boost-dependent R_{bb}
  B: large m_{bb} only for large R_{bb}
  S/B: go for large m_{bb} and small R_{bb}, so boost Higgs
- fat Higgs jet $R_{bb} \sim 2m_H/p_T \sim 0.8$ [like b tag for now]
- $q\bar{q} \to V_\ell H_b$ sizeable in boosted regime [$p_T \gtrsim$ 300 GeV, few % of total rate]



Example 1: $VH, H \to b\bar{b}$

New strategy for $H \to bb$ [Butterworth, Davison, Rubin, Salam]

- desperately needed [2/3 of all light Higgses; impact Dührssen & SFitter]
  but killed by continuum $Vb\bar{b}$ background
- S: large m_{bb}, boost-dependent R_{bb}
  B: large m_{bb} only for large R_{bb}
  S/B: go for large m_{bb} and small R_{bb}, so boost Higgs
- fat Higgs jet $R_{bb} \sim 2m_H/p_T \sim 0.8$ [like b tag for now]
- $q\bar{q} \to V_\ell H_b$ sizeable in boosted regime [$p_T \gtrsim$ 300 GeV, few % of total rate]

$\Rightarrow$ best performance: C/A algorithm

jet definition	σ_S/fb	σ_B/fb	$S/\sqrt{B}_{30}$
C/A, $R = 1.2$	0.57	0.51	4.4
$k_\perp$, $R = 1.0$	0.19	0.74	1.2
SISCone, $R = 0.8$	0.49	1.33	2.3

Example 1: $VH, H \to b\bar{b}$

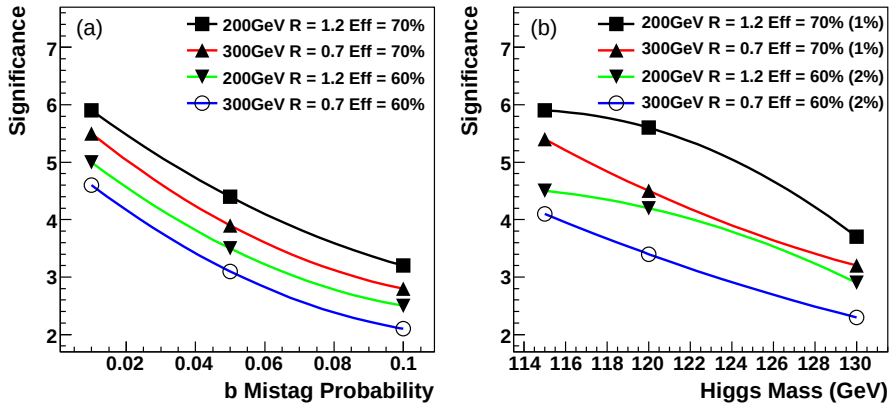
New strategy for $H \to bb$ [Butterworth, Davison, Rubin, Salam]

- desperately needed [2/3 of all light Higgses; impact Dührssen & SFitter]
  but killed by continuum $Vb\bar{b}$ background
- S: large m_{bb}, boost-dependent R_{bb}
  B: large m_{bb} only for large R_{bb}
  S/B: go for large m_{bb} and small R_{bb}, so boost Higgs
- fat Higgs jet $R_{bb} \sim 2m_H/p_T \sim 0.8$ [like b tag for now]
- $q\bar{q} \to V_\ell H_b$ sizeable in boosted regime [$p_T \gtrsim$ 300 GeV, few % of total rate]

$\Rightarrow$ best performance: C/A algorithm

Bottom line, details later

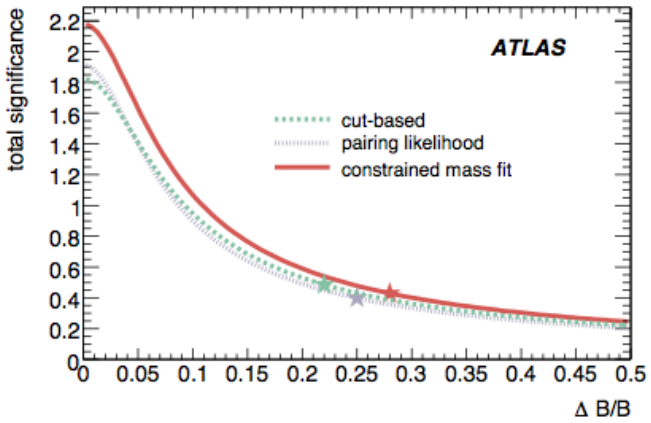
- combined channels $V \to \ell\ell, \nu\nu, \ell\nu$
- NLO rates [bbV notorious, not from data alone]
- Z peak as sanity check
- checked by Freiburg [Piquadio]
  subjet b tag excellent [70%/1%]
  charm rejection challenging
  $m_H \pm 8$ GeV tough



Example 2: $t\bar{t}H, H \to b\bar{b}$

Sad story of $t\bar{t}H, H \to b\bar{b}$ [Atlas-Bonn study, CMS-TDR even worse]

- 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]
- not a chance:
  1– combinatorics: m_H in $pp \to 4b_{tag}\ 2j\ \ell\nu$
  2– kinematics: peak-on-peak
  3– systematics: $S/B \sim 1/9$



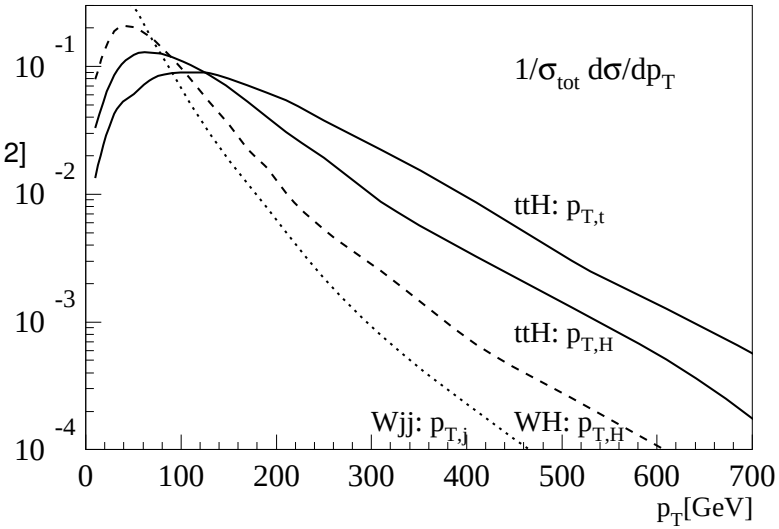
Example 2: $t\bar{t}H, H \to b\bar{b}$

Sad story of $t\bar{t}H, H \to b\bar{b}$ [Atlas-Bonn study, CMS-TDR even worse]

- 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]
- not a chance:
  1– combinatorics: m_H in $pp \to 4b_{tag}\ 2j\ \ell\nu$
  2– kinematics: peak-on-peak
  3– systematics: $S/B \sim 1/9$

Fat jets idea [TP, Salam, Spannowsky]

- $pp \to t_\ell t_h H_b$ even harder than VH
- S/B: $R_{bb} < 1.2$; $b\bar{b}$ pair boosted [solves 1]
- boosted regime different for S and B [solves 2]
- see how far we get... [watch S/B for 3]
- cool: fat Higgs jet + fat top jet



Example 2: $t\bar{t}H, H \to b\bar{b}$

Sad story of $t\bar{t}H, H \to b\bar{b}$ [Atlas-Bonn study, CMS-TDR even worse]

- 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]
- not a chance:
  1– combinatorics: m_H in $pp \to 4b_{tag}\ 2j\ \ell\nu$
  2– kinematics: peak-on-peak
  3– systematics: $S/B \sim 1/9$

Another bottom line

- require tagged top and Higgs
  trigger on lepton
- remove 'Higgs' as $t_\ell \to b$ plus QCD
  3rd b tag in continuum [costing $S/\sqrt{B}$]
  only continuum $t\bar{t}b\bar{b}$ left

per 1 fb^{-1}	signal	$t\bar{t}Z$	$t\bar{t}b\bar{b}$	$t\bar{t}$+jets
events after acceptance	24.1	6.9	191	4160
events with one top tag	10.2	2.9	70.4	1457
events with $m_{bb} = 110 - 130$ GeV	2.9	0.44	12.6	116
corresponding to subjet pairings	3.2	0.47	13.8	121
subjet pairings two b tags	1.0	0.08	2.3	1.4
including a third b tag	0.48	0.03	1.09	0.06

Example 2: $t\bar{t}H, H \to b\bar{b}$

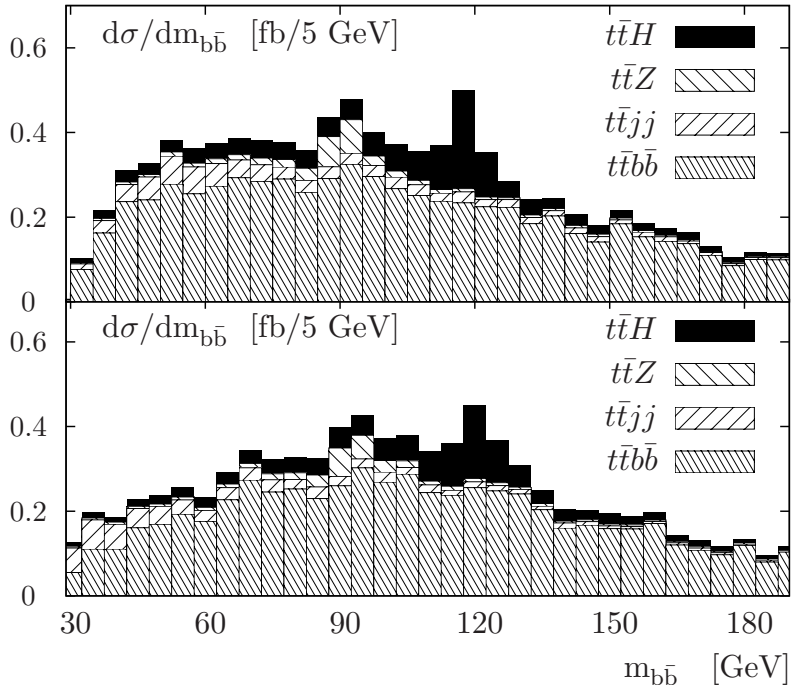
Sad story of $t\bar{t}H, H \to b\bar{b}$ [Atlas-Bonn study, CMS-TDR even worse]

- 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]
- not a chance:
  1– combinatorics: m_H in $pp \to 4b_{tag}\ 2j\ \ell\nu$
  2– kinematics: peak-on-peak
  3– systematics: $S/B \sim 1/9$

Another bottom line

- require tagged top and Higgs
  trigger on lepton
- remove 'Higgs' as $t_\ell \to b$ plus QCD
  3rd b tag in continuum [costing $S/\sqrt{B}$]
  only continuum $t\bar{t}b\bar{b}$ left

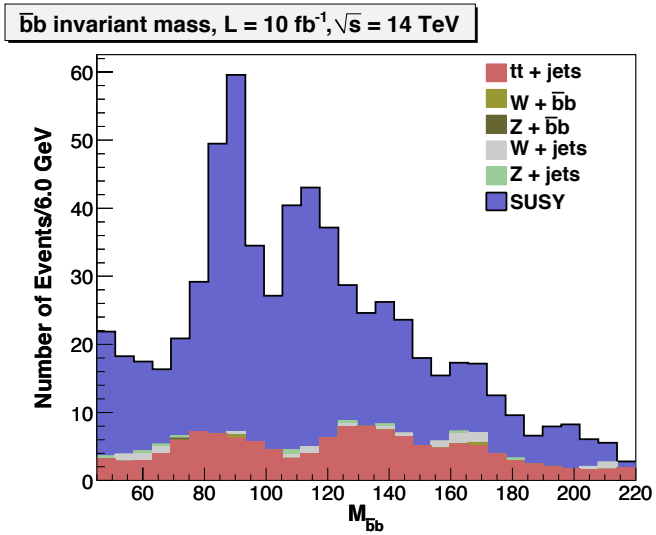
m_H	S	S/B	$S/\sqrt{B}_{100\ \text{fb}^{-1}}$
115	57	1/2.1	5.2 (5.7)
120	48	1/2.4	4.5 (5.1)
130	29	1/3.6	2.9 (3.0)



Example 3: $H \to b\bar{b}$ in SUSY cascades

Higgs in cascade decays [Kribs, Martin, Roy, Spannowsky]

- idea: find Higgs in cascade decays [Cambridge]
- BSM sample after missing energy or hard γ cut
- blind Higgs tag over remaining event [QCD rejection?]
- side bin analysis in m_{bb}
- more to follow...



Higgs tagger

Higgs tag for busy QCD environment [BDRS; TP, Salam, Spannowsky]

- uncluster one-by-one: $j \to j_1 + j_2$
  1– unbalanced $m_{j_1} > 0.8 m_j$ means QCD; discard j_2
  2– soft $m_{j_1} < 30$ GeV means QCD; keep j_1
- double b tag [possibly add balance criterion]
  three leading $J = p_{T,1} p_{T,2} (\Delta R_{12})^4$ vs m_{bb}
- no mass constraint — side bin
  typical mis-tag probability $< 10^{-5}$
- underlying event and pileup deadly
  filter reconstruction jets [Butterworth–Salam, cf pruning, trimming]
  zoomed-in C/A analysis with $R_{\text{filt}} = \min(0.3, R_{bb}/2)$
- reconstruct m_H w/ one QCD jet

Higgs tagger

Higgs tag for busy QCD environment [BDRS; TP, Salam, Spannowsky]

- uncluster one-by-one: $j \to j_1 + j_2$
  1– unbalanced $m_{j_1} > 0.8 m_j$ means QCD; discard j_2
  2– soft $m_{j_1} < 30$ GeV means QCD; keep j_1
- double b tag [possibly add balance criterion]
  three leading $J = p_{T,1} p_{T,2} (\Delta R_{12})^4$ vs m_{bb}
- no mass constraint — side bin
  typical mis-tag probability $< 10^{-5}$
- underlying event and pileup deadly
  filter reconstruction jets [Butterworth–Salam, cf pruning, trimming]
  zoomed-in C/A analysis with $R_{\text{filt}} = \min(0.3, R_{bb}/2)$
- reconstruct m_H w/ one QCD jet

Better than traditional b jets

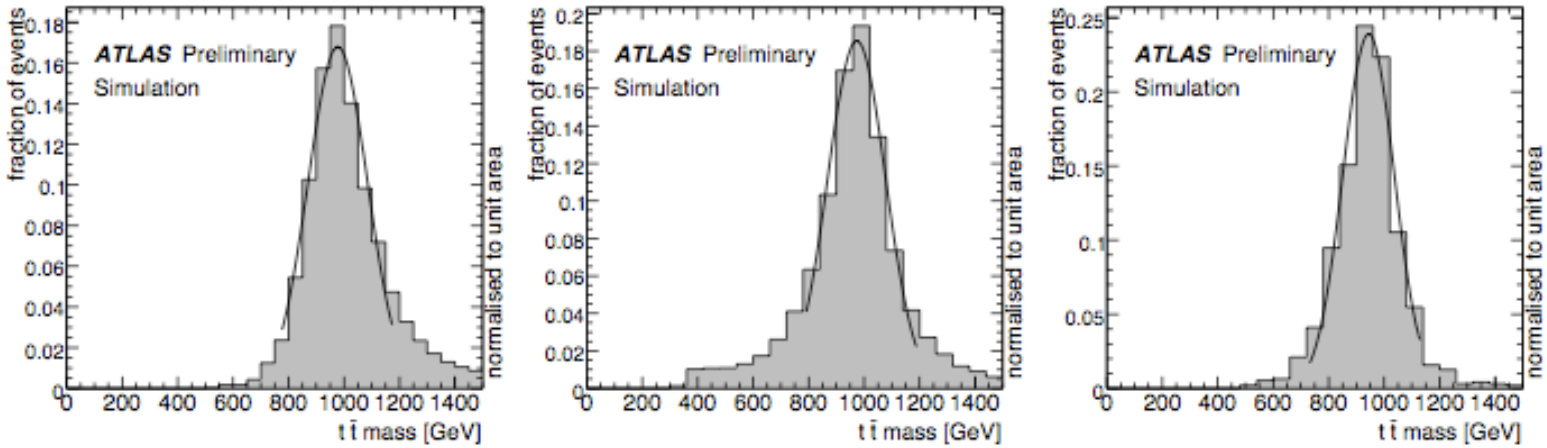
- no combinatorical choices
- more soft partons included in m_H
- b tagging easier than in continuum
- QCD features useful [Soper & Spannowsky]

...

Top tagger

Highly boosted top quarks [Kaplan, Rehermann, Schwartz, Tweedie; Princeton, Seattle...]

- identify hadronic tops with $p_T \gtrsim 800$ GeV
  isolation and b tagging challenging
- C/A algorithm with p_T drop criterion [Hopkins tagger, no b tag]
- top mass included, no sidebins
- ATLAS studies for semileptonic top pairs [adapted Y-splitter, full sim, ATLAS-2010-008]



(a) minimal (b) full reconstruction (c) mono-jet

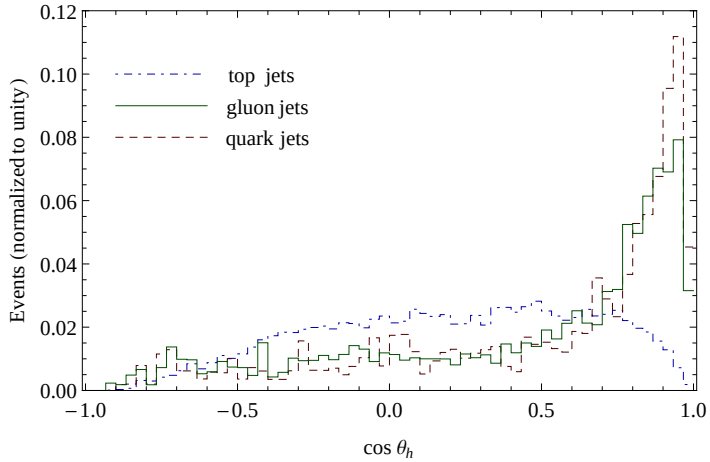
Top tagger

Highly boosted top quarks [Kaplan, Rehermann, Schwartz, Tweedie; Princeton, Seattle...]

- identify hadronic tops with $p_T \gtrsim 800$ GeV
  isolation and b tagging challenging
- C/A algorithm with p_T drop criterion [Hopkins tagger, no b tag]
- top mass included, no sidebins
- ATLAS studies for semileptonic top pairs [adapted Y-splitter, full sim, ATLAS-2010-008]

Out first top tag [TP, Salam, Spannowsky, Takeuchi]

- start with C/A jet $[R = 1.5]$ [Johns Hopkins]
- uncluster one-by-one: $j \to j_1 + j_2$
  1– unbalanced $m_{j_1} > 0.8 m_j$ means QCD; discard j_2
  2– soft $m_{j_1} < 30$ GeV means QCD; keep j_1
- top decay kinematics in relevant substructures
  reconstruct $m_W = 60...95$ GeV
  reconstruct $m_t = 150...200$ GeV
  helicity angle $\cos\theta_{t,j_1} > 0.7$
  no b tag needed



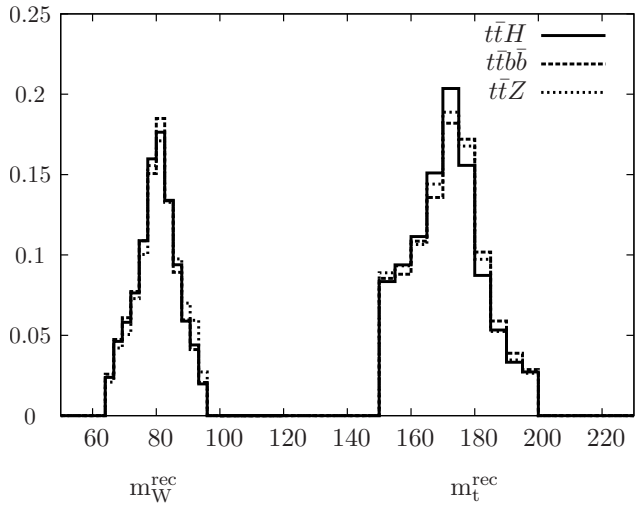
Top tagger

Highly boosted top quarks [Kaplan, Rehermann, Schwartz, Tweedie; Princeton, Seattle...]

- identify hadronic tops with $p_T \gtrsim 800$ GeV
  isolation and b tagging challenging
- C/A algorithm with p_T drop criterion [Hopkins tagger, no b tag]
- top mass included, no sidebins
- ATLAS studies for semileptonic top pairs [adapted Y-splitter, full sim, ATLAS-2010-008]

Out first top tag [TP, Salam, Spannowsky, Takeuchi]

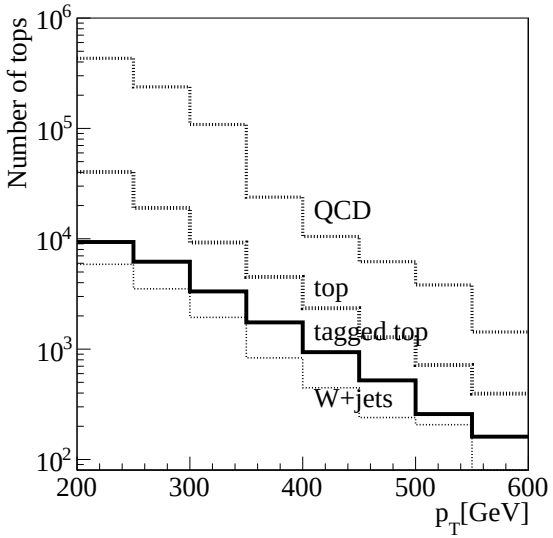
- start with C/A jet [$R = 1.5$] [Johns Hopkins]
- uncluster one-by-one: $j \to j_1 + j_2$
  1– unbalanced $m_{j_1} > 0.8 m_j$ means QCD; discard j_2
  2– soft $m_{j_1} < 30$ GeV means QCD; keep j_1
- top decay kinematics in relevant substructures
  reconstruct $m_W = 60...95$ GeV
  reconstruct $m_t = 150...200$ GeV
  helicity angle $\cos\theta_{t,j_1} > 0.7$
  no b tag needed
- filtering w/ 2 QCD jets



HEPTopTagger

Applicable: HEPTopTagger [TP, Salam, Spannowsky, Takeuchi]

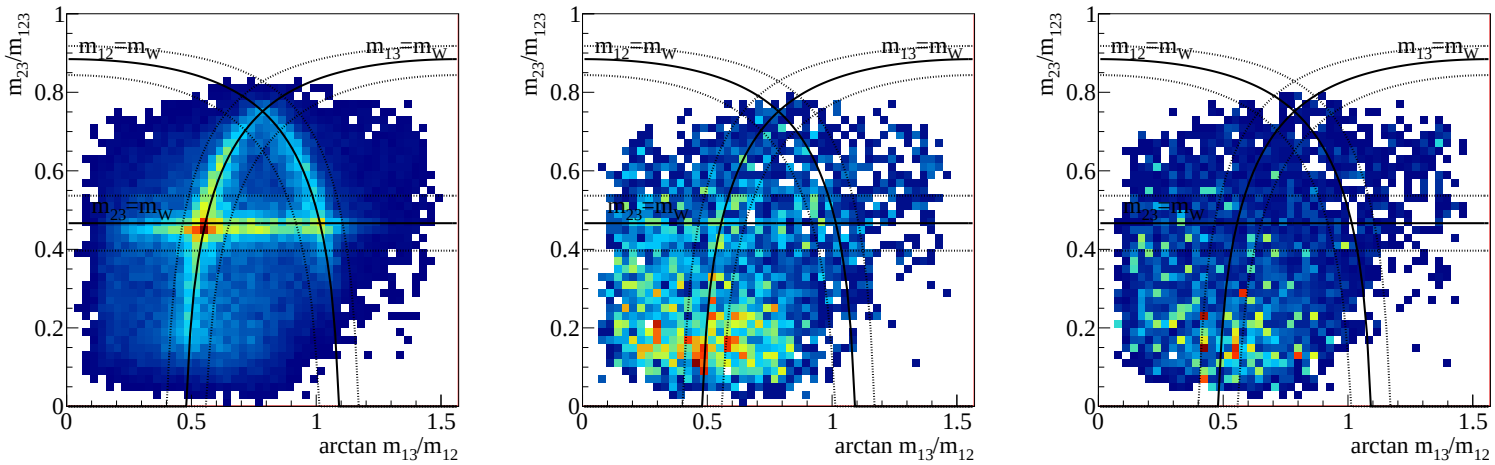
– extend lower $p_T \gtrsim 250$ GeV
testable in Standard Model $t\bar{t}$ events



HEPTopTagger

Applicable: HEPTopTagger [TP, Salam, Spannowsky, Takeuchi]

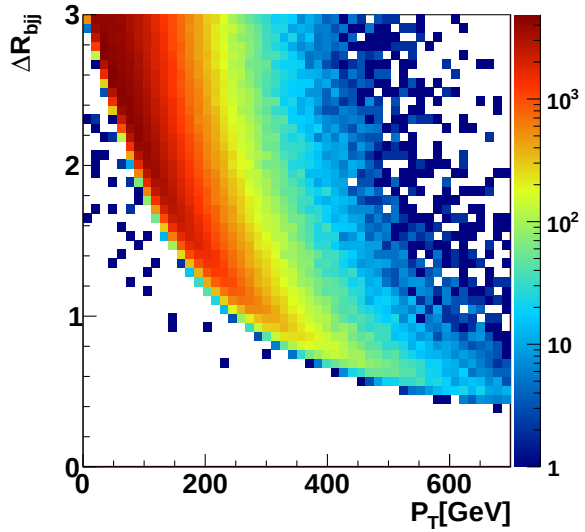
- extend lower $p_T \gtrsim 250$ GeV
  testable in Standard Model $t\bar{t}$ events
- new kinematic selection: m_{jjj}, $m_{jj}^{(1)}$, $m_{jj}^{(2)}$ [no boost]

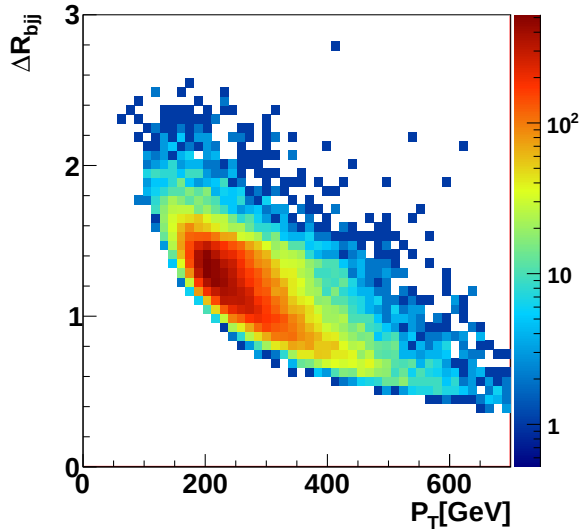


HEPTopTagger

Applicable: HEPTopTagger [TP, Salam, Spannowsky, Takeuchi]

- extend lower $p_T \gtrsim 250$ GeV
  testable in Standard Model $t\bar{t}$ events
- new kinematic selection: $m_{jjj}, m_{jj}^{(1)}, m_{jj}^{(2)}$ [no boost]

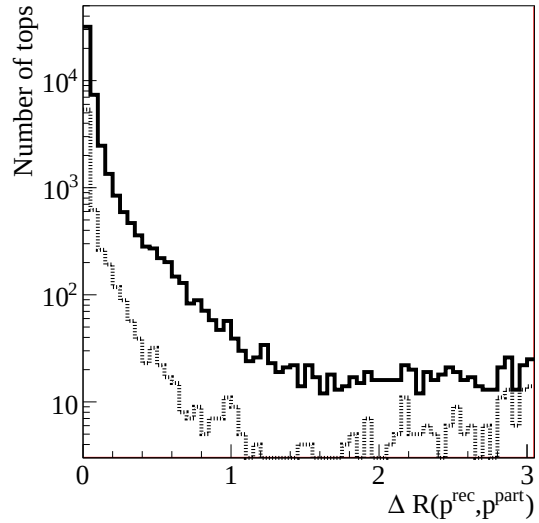


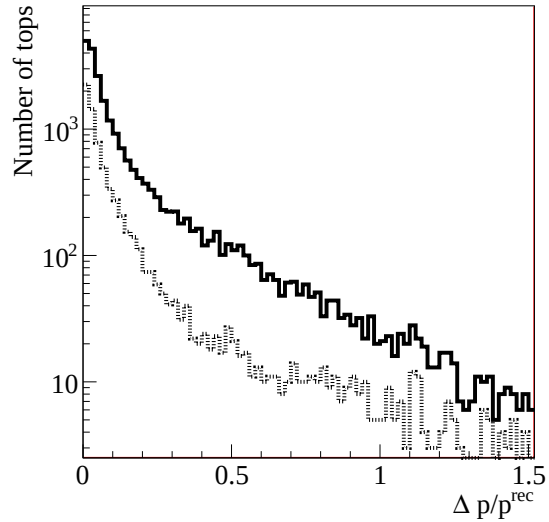


HEPTopTagger

Applicable: HEPTopTagger [TP, Salam, Spannowsky, Takeuchi]

- extend lower $p_T \gtrsim 250$ GeV
  testable in Standard Model $t\bar{t}$ events
- new kinematic selection: $m_{jjj}, m_{jj}^{(1)}, m_{jj}^{(2)}$ [no boost]
- momentum reconstruction for free

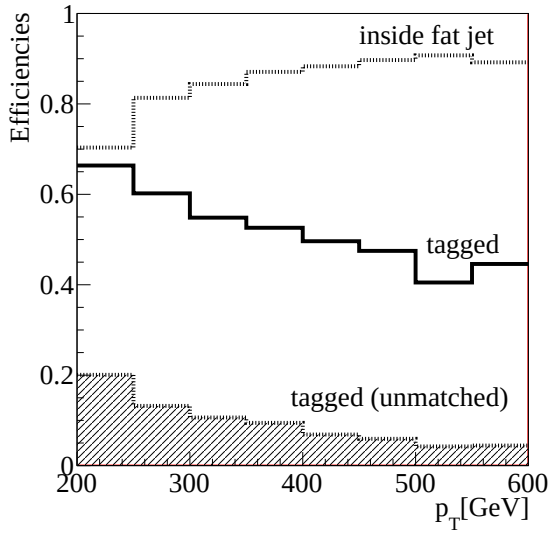




HEPTopTagger

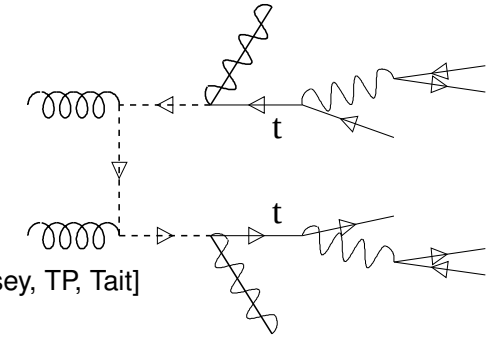
Applicable: HEPTopTagger [TP, Salam, Spannowsky, Takeuchi]

- extend lower $p_T \gtrsim 250$ GeV
  testable in Standard Model $t\bar{t}$ events
- new kinematic selection: $m_{jjj}, m_{jj}^{(1)}, m_{jj}^{(2)}$ [no boost]
- momentum reconstruction for free
- use color activity [Baryakthar, Hook, Janowiak, Wacker]
- tested by ATLAS Heidelberg [Kasieczka & Schätzel]
- hadronic top like tagged b



	$t\bar{t}$			QCD	W+jets	
$p_{T,t}^{\min}$[GeV]	0	200	300			
one fat jet	100%	100%	100%	100%	100%	
two fat jets	44%	57%	70%	53%	50%	rel to one fat jet
one top tag	23%	37%	51%	2.0%	3.9%	rel to one fat jet
two top tags	2.0%	4.5%	8.5%	0.027%	0.07%	rel to one fat jet
	4.5%	8.0%	12%	0.05%	0.15%	rel to two fat jets

Stops pairs

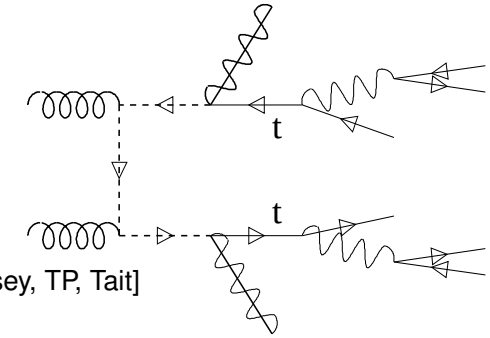


Stop pairs as the first application [TP, Spannowsky, Takeuchi, Zerwas]

- stop most important for hierarchy problem [review: Morrissey, TP, Tait]
  comparison to other top partners [Meade & Reece]
- dark matter means difficult semi-leptonic channel [possibly impossible]
- hadronic: $\tilde{t}\tilde{t}^* \to t\tilde{\chi}_1^0\, \bar{t}\tilde{\chi}_1^0$ [CMS: leptons as spontaneous life guards; Meade & Reece overly optimistic]

events in 1 fb^{-1}	$\tilde{t}_1\tilde{t}_1^*$						$t\bar{t}$	QCD	W+jets	Z+jets	S/B	$S/\sqrt{B}_{10\text{ fb}^{-1}}$
$m_{\tilde{t}}$[GeV]	340	390	440	490	540	640					340	
$p_{T,j} > 200$ GeV, ℓ veto	728	447	292	187	124	46	87850	$2.4 \cdot 10^7$	$1.6 \cdot 10^5$	n/a	$3.0 \cdot 10^{-5}$	
$\not{E}_T > 150$ GeV	283	234	184	133	93	35	2245	$2.4 \cdot 10^5$	1710	2240	$1.2 \cdot 10^{-3}$	
first top tag	100	91	75	57	42	15	743	7590	90	114	$1.2 \cdot 10^{-2}$	
second top tag	15	12.4	11	8.4	6.3	2.3	32	129	5.7	1.4	$8.3 \cdot 10^{-2}$	
b tag	8.7	7.4	6.3	5.0	3.8	1.4	19	2.6	$\lesssim 0.2$	$\lesssim 0.05$	0.40	5. 9
$m_{T2} > 250$ GeV	4.3	5.0	4.9	4.2	3.2	1.2	4.2	$\lesssim 0.6$	$\lesssim 0.1$	$\lesssim 0.03$	0.88	6. 1

Stops pairs

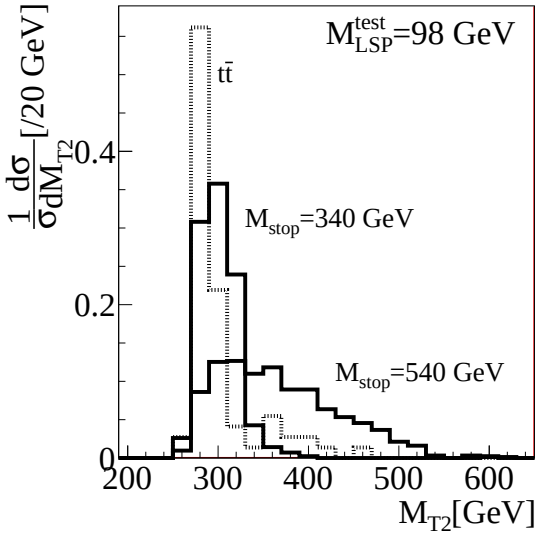


Stop pairs as the first application [TP, Spannowsky, Takeuchi, Zerwas]

- stop most important for hierarchy problem [review: Morrissey, TP, Tait]
  comparison to other top partners [Meade & Reece]
- dark matter means difficult semi-leptonic channel [possibly impossible]
- hadronic: $\tilde{t}\tilde{t}^* \to t\tilde{\chi}_1^0\, \bar{t}\tilde{\chi}_1^0$ [CMS: leptons as spontaneous life guards; Meade & Reece overly optimistic]
- stop mass from m_{T2} endpoint [like sleptons or sbottoms]

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

- as easy as $b\bar{b} + \not{E}_T$



Leptonic top tag

Leptonic tag [Thaler & Wang; Rehermann & Tweedie; TP, Spannowsky, Takeuchi]

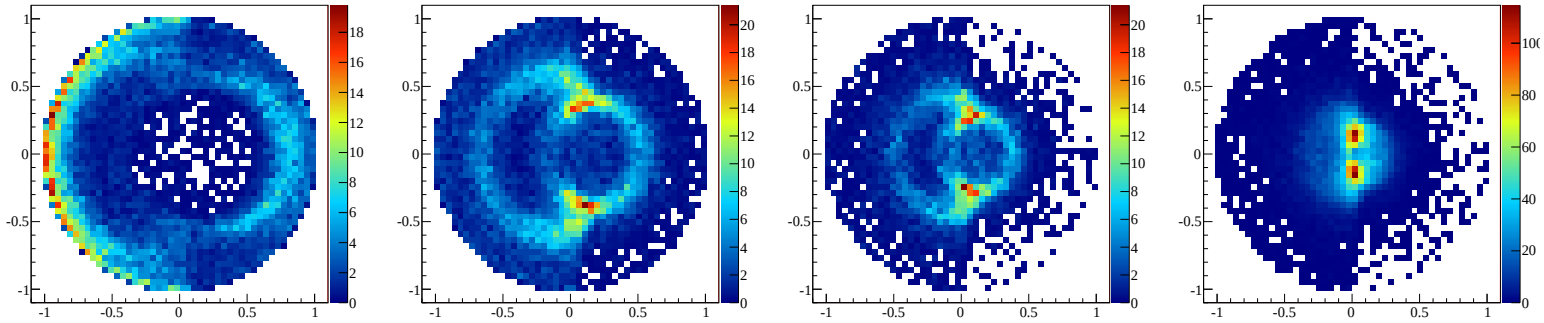
- known: masses of top decay products
  unknown: 3-momentum of neutrino
  measured: $E_b, E_\ell, m_{b\ell}$ [rest frame]
- W and t mass constraints
  third parameter elsewhere
  do not use measured $\not{p}_T$ vector

Leptonic top tag

Leptonic tag [Thaler & Wang; Rehermann & Tweedie; TP, Spannowsky, Takeuchi]

- known: masses of top decay products
  unknown: 3-momentum of neutrino
  measured: $E_b, E_\ell, m_{b\ell}$ [rest frame]
- W and t mass constraints
  third parameter elsewhere
  do not use measured $\not{p}_T$ vector
- neutrino coordinates
  leading in $b - \ell$ direction
  sub-leading in $b - \ell$ decay plane
  sub-leading orthogonal to decay plane

  components $(p_\nu^{\parallel}, p_\nu^{\perp})$

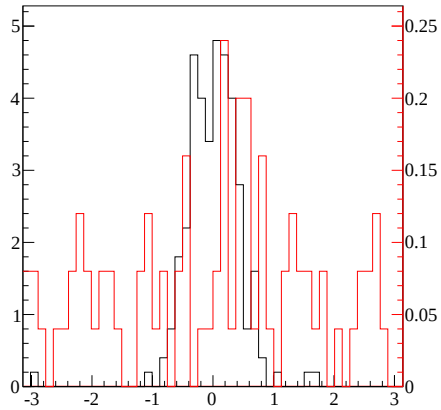


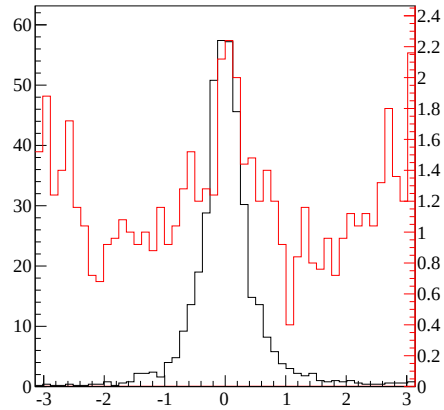
Leptonic top tag

Leptonic tag [Thaler & Wang; Rehermann & Tweedie; TP, Spannowsky, Takeuchi]

- known: masses of top decay products
  unknown: 3-momentum of neutrino
  measured: $E_b, E_\ell, m_{b\ell}$ [rest frame]
- W and t mass constraints
  third parameter elsewhere
  do not use measured $\not{p}_T$ vector
- neutrino coordinates
  leading in $b-\ell$ direction
  sub-leading in $b-\ell$ decay plane
  sub-leading orthogonal to decay plane
- approximations for $\Delta\Phi(\not{p}_T, \hat{p}_t)$ and m_{T2}

  decay plane: $p_\nu^\perp = 0$
  orthogonal: $p_\nu^\parallel = 0$



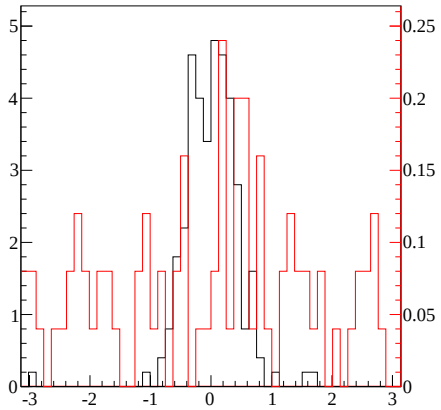


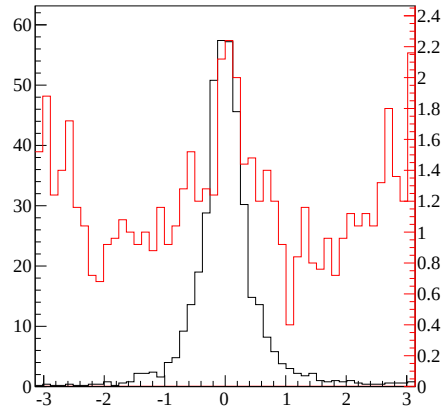
Leptonic top tag

Leptonic tag [Thaler & Wang; Rehermann & Tweedie; TP, Spannowsky, Takeuchi]

- known: masses of top decay products
  unknown: 3-momentum of neutrino
  measured: $E_b, E_\ell, m_{b\ell}$ [rest frame]
- W and t mass constraints
  third parameter elsewhere
  do not use measured $\not{p}_T$ vector
- neutrino coordinates
  leading in $b-\ell$ direction
  sub-leading in $b-\ell$ decay plane
  sub-leading orthogonal to decay plane
- approximations for $\Delta\Phi(\not{p}_T, \hat{p}_t)$ and m_{T2}

  decay plane: $p_\nu^\perp = 0$
  orthogonal: $p_\nu^\parallel = 0$





- scoop us and you will get killed in a really ugly manner...

Outlook

Fat jets — Aspirin of LHC phenomenology

- VH: bringing back 2/3 of light Higgses
- $t\bar{t}H$: curing combinatorics and backgrounds
- SUSY cascades: curing lack of analysis idea

  ...

- Z' etc: improving mass resolution
- $\tilde{t}\tilde{t}^*$: curing backgrounds

  ...

- HEPTopTagger code as FASTJET add-on [www.thphys.uni-heidelberg.de/~plehn/HEPTopTagger]
  leptonic tagger on the way [keep fingers crossed]

  LHC lecture notes arXiv:0910.4182
  BOOST review arXiv:1012.5412