Karlsruher Institut für Technologie

CMS Germany

BMBF FSP-CMS

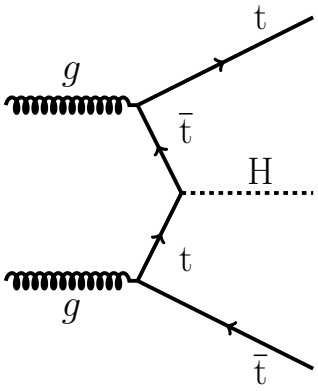
Boosted $t\bar{t}H$ with CMS

8th Annual Helmholtz Alliance Workshop “Physics at the Terascale”

Hannes Mildner | December 2nd, 2014

INSTITUT FÜR EXPERIMENTELLE KERNPHYSIK (IEKP)

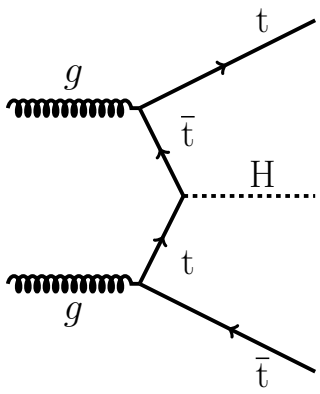
Introduction



$t\bar{t}H$ production

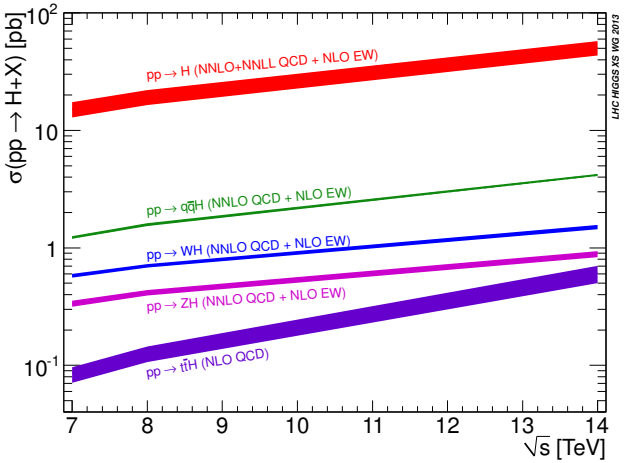
- Enables comparably model-independent measurement of top-Higgs coupling
- Three heavy-particle decays

Introduction



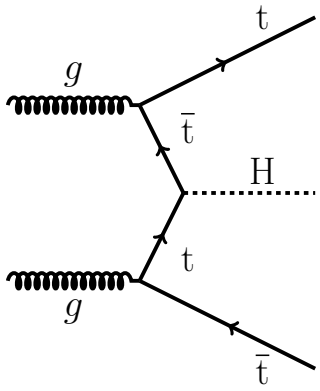
tt̄H production

- Enables comparably model-independent measurement of top-Higgs coupling
- Three heavy-particle decays



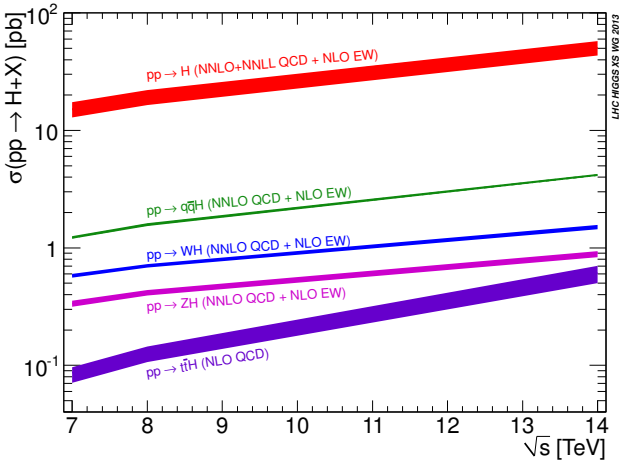
- Small cross section (130 fb @ 8 TeV, 510 fb @ 13 TeV)

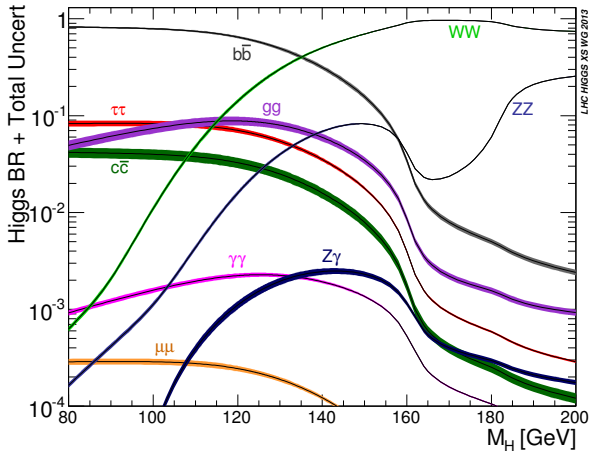
Introduction



tt̄H production

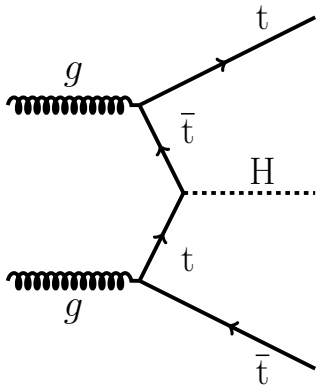
- Enables comparably model-independent measurement of top-Higgs coupling
- Three heavy-particle decays





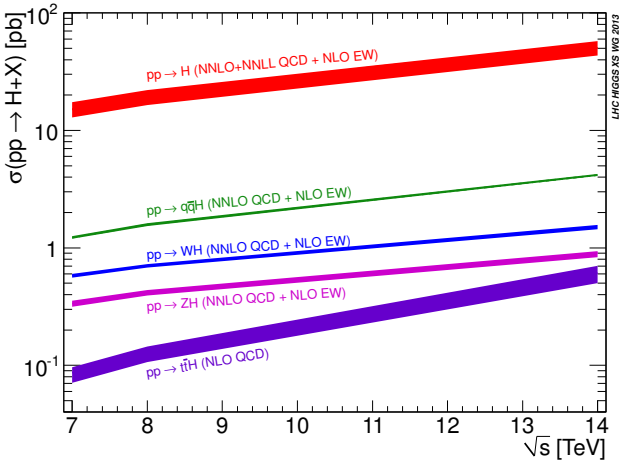
- Small cross section (130 fb @ 8 TeV, 510 fb @ 13 TeV)
- Many possible decays for Higgs boson

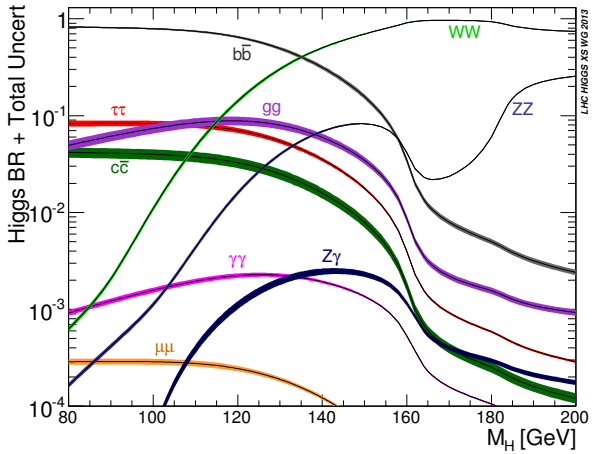
Introduction

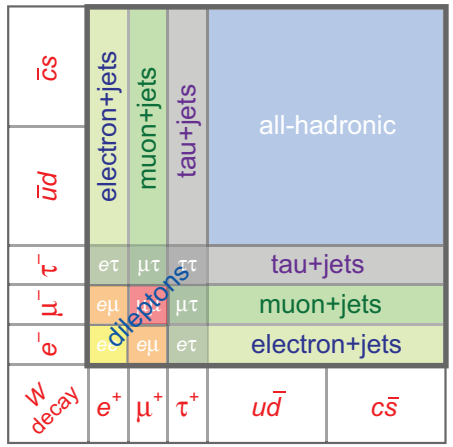


$t\bar{t}H$ production

- Enables comparably model-independent measurement of top-Higgs coupling
- Three heavy-particle decays



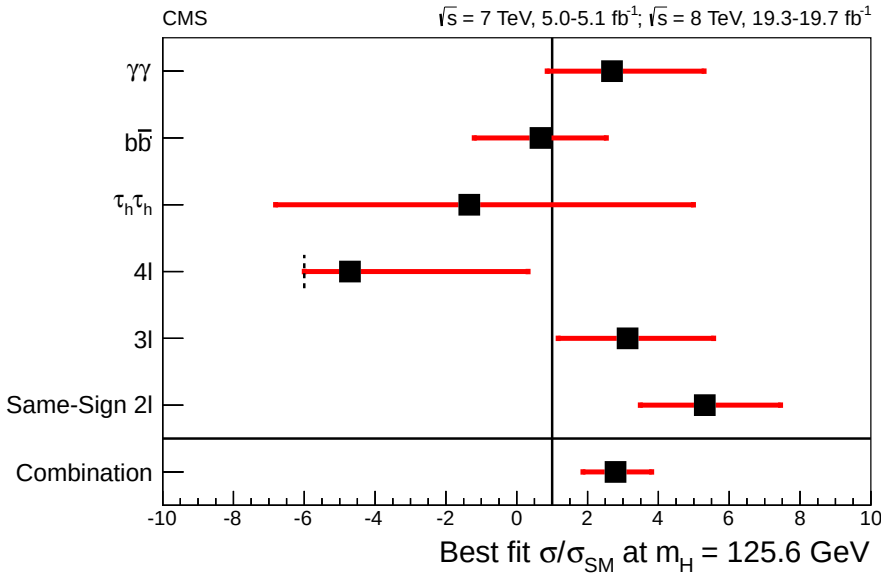




- Small cross section (130 fb @ 8 TeV, 510 fb @ 13 TeV)
- Many possible decays for Higgs boson
- In combination with different $t\bar{t}$ decay channels

Run I results

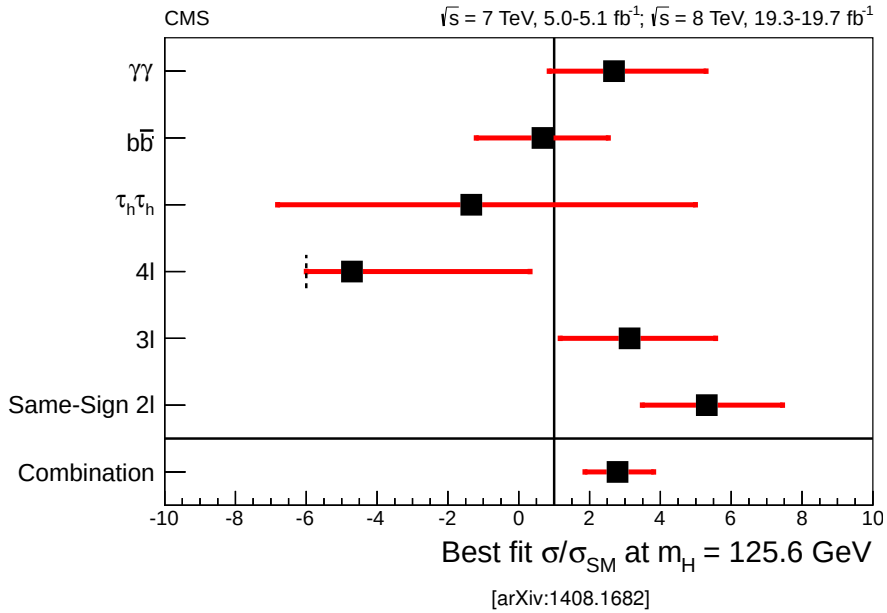
- Published results
  - Several channels analyzed
  - $t\bar{t}H$ with $H \to b\bar{b}$, $H \to \gamma\gamma$, and multi-lepton (e.g. $H \to WW$) channels most important ones
  - $\frac{\sigma^{\text{measured}}_{t\bar{t}H(125)}}{\sigma^{\text{SM}}_{t\bar{t}H(125)}} = \mu_{t\bar{t}H(125)} = 2.8^{+1.0}_{-0.9}$
  - Excess of $\mu^{\pm}\mu^{\pm} + X$ events

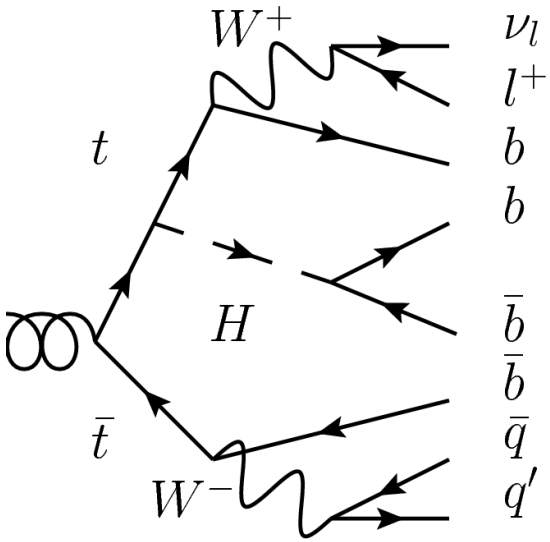


[arXiv:1408.1682]

Run I results

- Published results
  - Several channels analyzed
  - $t\bar{t}H$ with $H \to b\bar{b}$, $H \to \gamma\gamma$, and multi-lepton (e.g. $H \to WW$) channels most important ones
  - $\frac{\sigma^{\text{measured}}_{t\bar{t}H(125)}}{\sigma^{\text{SM}}_{t\bar{t}H(125)}} = \mu_{t\bar{t}H(125)} = 2.8^{+1.0}_{-0.9}$
  - Excess of $\mu^{\pm}\mu^{\pm} + X$ events





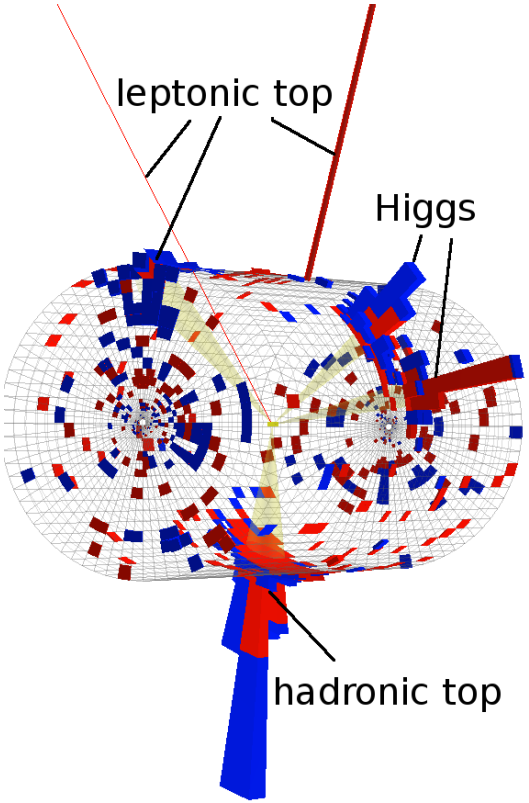
- KIT (with Ohio State and University of Virginia) is analyzing semileptonic $t\bar{t}$-decays and $H \to b\bar{b}$
  - \+ Large branching ratio
  - \+ Little QCD-multijet background thanks to lepton
  - – $\approx$ 6 (b-)jets – combinatorial problem, bad energy resolution
  - – Large $t\bar{t}$ + jets background ($\sigma_{tt} = 245$ pb @ 8 TeV)

$t\bar{t}H$ with boost

- Plehn, Salam, Spannowsky (2009): *Fat Jets for a light Higgs*
  Analyze $t\bar{t}H$-Events with Higgs/top with high p_T (*boost*)
  - Better S/B ratio
  - Simplified combinatorics (higher $p_T \Rightarrow$ collimated decay products)

$t\bar{t}H$ with boost

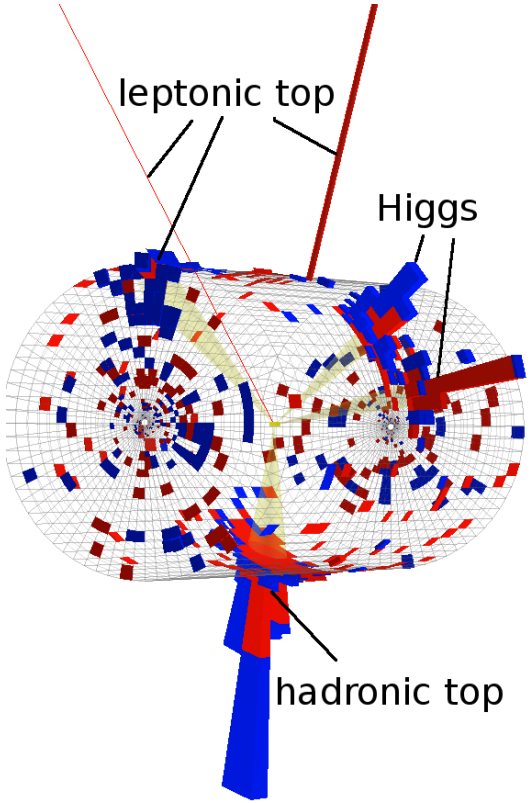
- Plehn, Salam, Spannowsky (2009): *Fat Jets for a light Higgs*
  Analyze $t\bar{t}H$-Events with Higgs/top with high p_T (*boost*)
  - Better S/B ratio
  - Simplified combinatorics (higher $p_T \Rightarrow$ collimated decay products)



- Simulated *boosted* $t\bar{t}H$-event
  - Fat jets with substructure from Higgs and hadronically decaying top
  - Lepton, MET, and b-jet from leptonic top

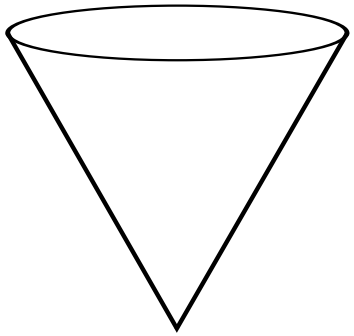
$t\bar{t}H$ with boost

- Plehn, Salam, Spannowsky (2009): *Fat Jets for a light Higgs*
  Analyze $t\bar{t}H$-Events with Higgs/top with high p_T (*boost*)
  - Better S/B ratio
  - Simplified combinatorics (higher $p_T \Rightarrow$ collimated decay products)



- Simulated *boosted* $t\bar{t}H$-event
  - Fat jets with substructure from Higgs and hadronically decaying top
  - Lepton, MET, and b-jet from leptonic top
- Collimated jets experimentally challenging
  - Using Subjet-filterjet algorithm for Higgs identification
  - HEP top-tagger is used for top-identification

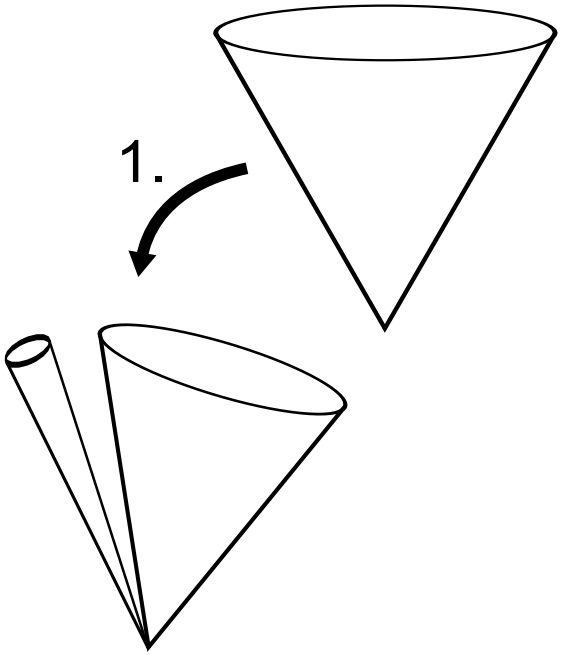
Subjet-filterjet algorithm



Clustering (Cambridge-Aachen 1.2)

- Cluster jet with large cone size (*fat jets*)

Subjet-filterjet algorithm



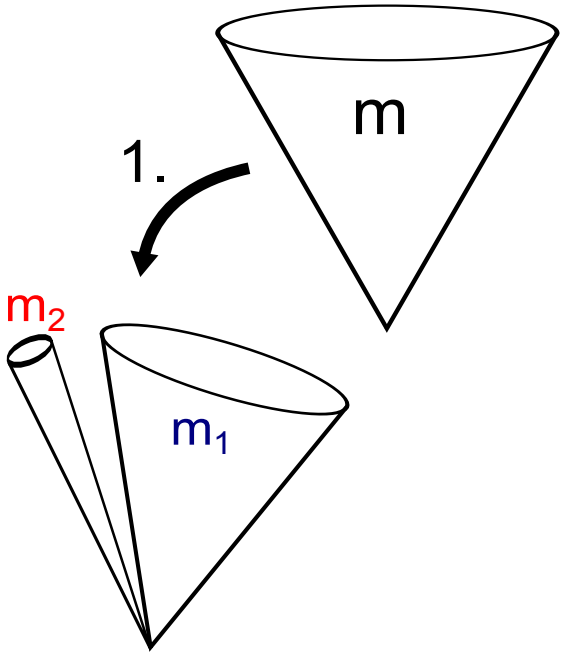
Clustering (Cambridge-Aachen 1.2)

- Cluster jet with large cone size (*fat jets*)

Declustering (variable-sized subjets)

- Undo last clustering (1.)

Subjet-filterjet algorithm



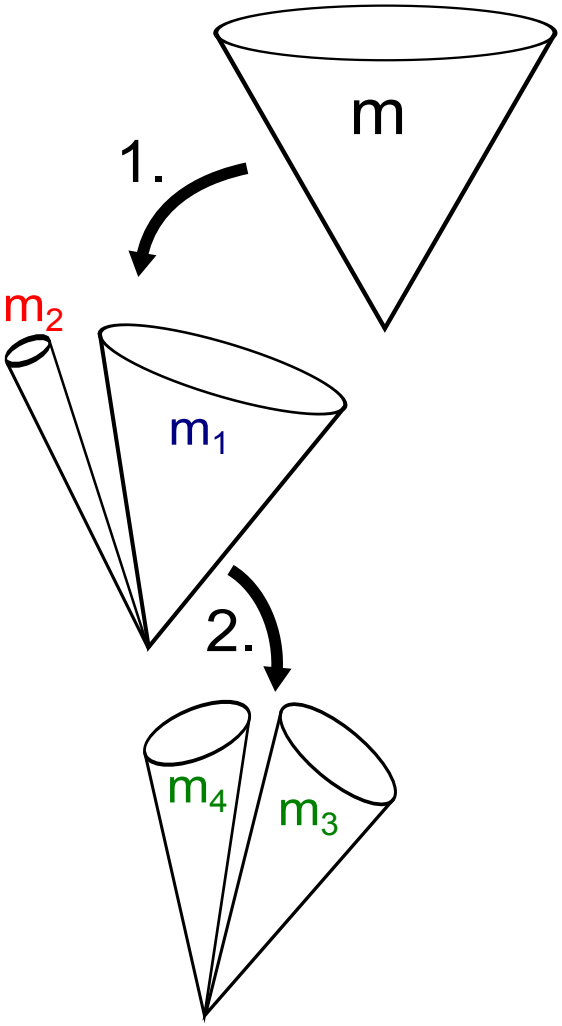
Clustering (Cambridge-Aachen 1.2)

- Cluster jet with large cone size (*fat jets*)

Declustering (variable-sized subjets)

- Undo last clustering (1.)
- Ratio of inv. masses m_1/m large $\Rightarrow$ Subjet 2 unimportant radiation, continue *declustering* Subjet 1 (2.)

Subjet-filterjet algorithm



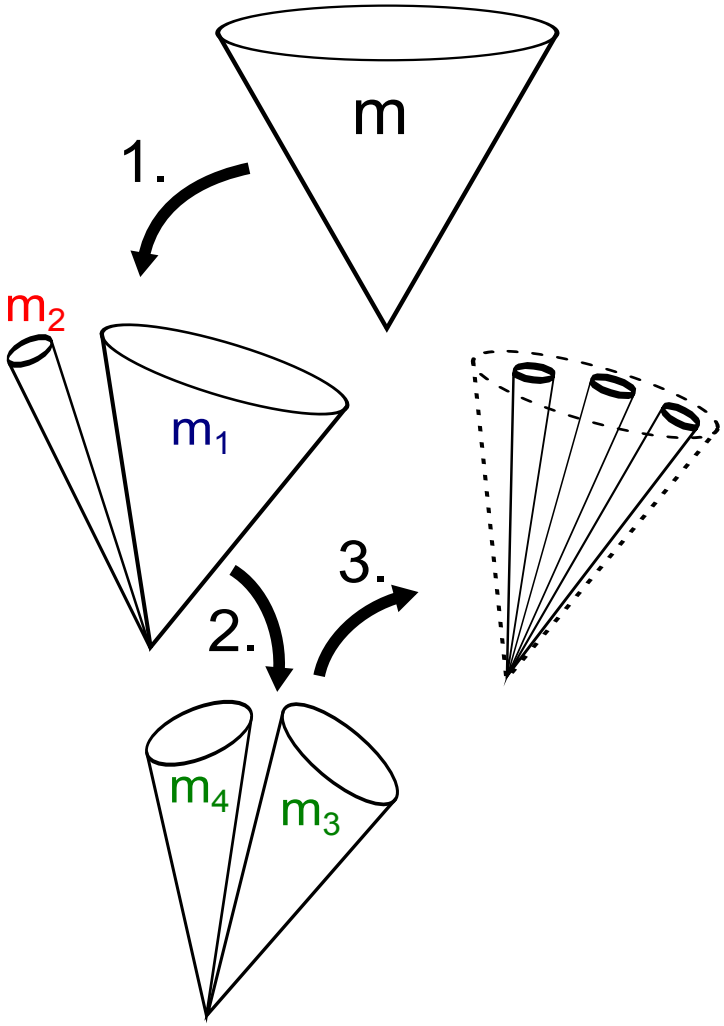
Clustering (Cambridge-Aachen 1.2)

- Cluster jet with large cone size (*fat jets*)

Declustering (variable-sized subjets)

- Undo last clustering (1.)
- Ratio of inv. masses m_1/m large $\Rightarrow$ Subjet 2 unimportant radiation, continue *declustering* Subjet 1 (2.)
- Until ratio m_3/m_1' and m_4/m_1' small: found decay

Subjet-filterjet algorithm



Clustering (Cambridge-Aachen 1.2)

- Cluster jet with large cone size (*fat jets*)

Declustering (variable-sized subjets)

- Undo last clustering (1.)
- Ratio of inv. masses m_1/m large $\Rightarrow$ Subjet 2 unimportant radiation, continue *declustering* Subjet 1 (2.)
- Until ratio m_3/m_1' and m_4/m_1' small: found decay

Filtering (Cambridge-Aachen 0.3)

- Cluster particles of Subjet 3 and 4 to slim jets (3.), ignore soft jets, analyze hard jets (invariant masses, b-tags)

HEP top-tagger

Clustering, declustering, filtering

- Finds filterjets similar to Subjet-filterjet algorithm
- Returns three subjets made of filterjets close to top mass

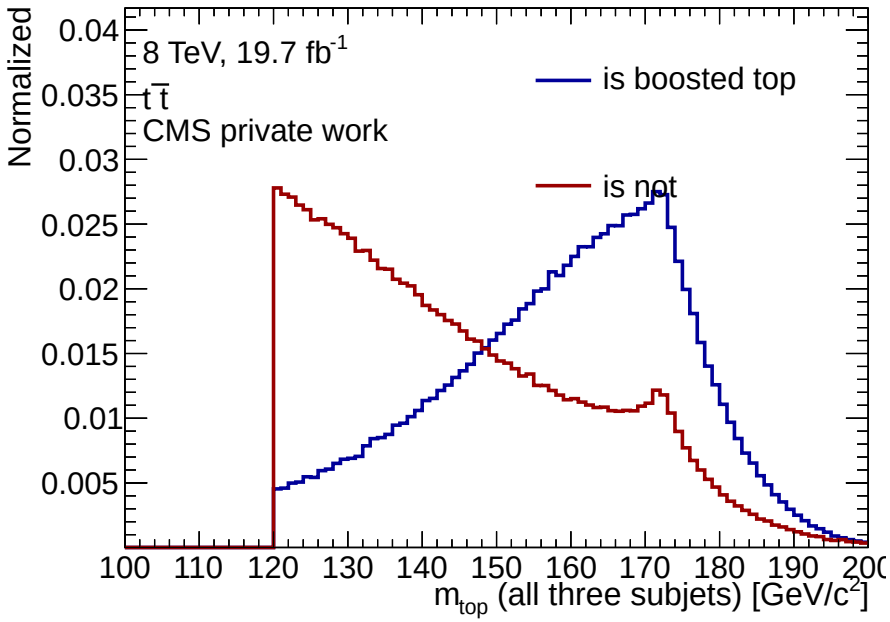
HEP top-tagger

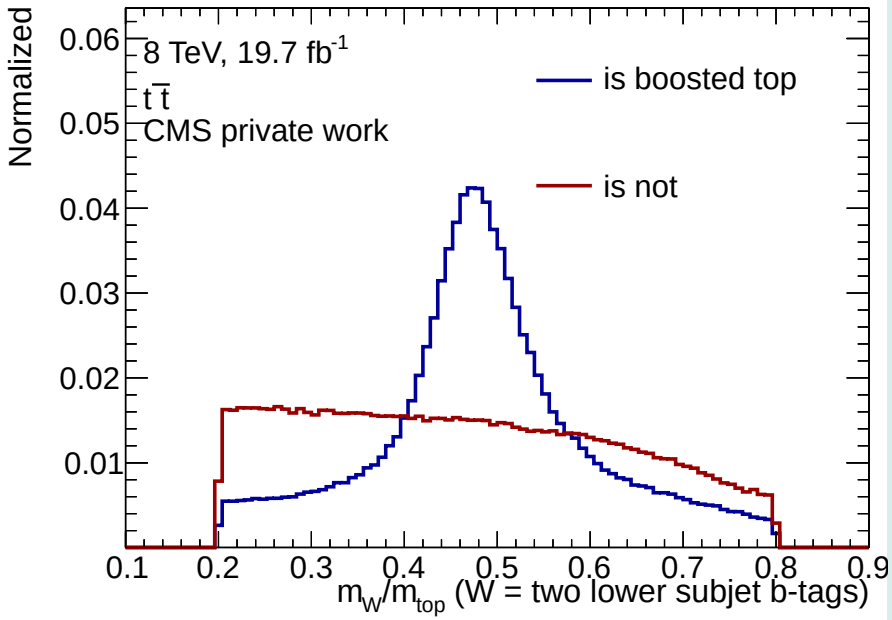
Clustering, declustering, filtering

- Finds filterjets similar to Subjet-filterjet algorithm
- Returns three subjets made of filterjets close to top mass

Further discrimination using invariant masses of subjets

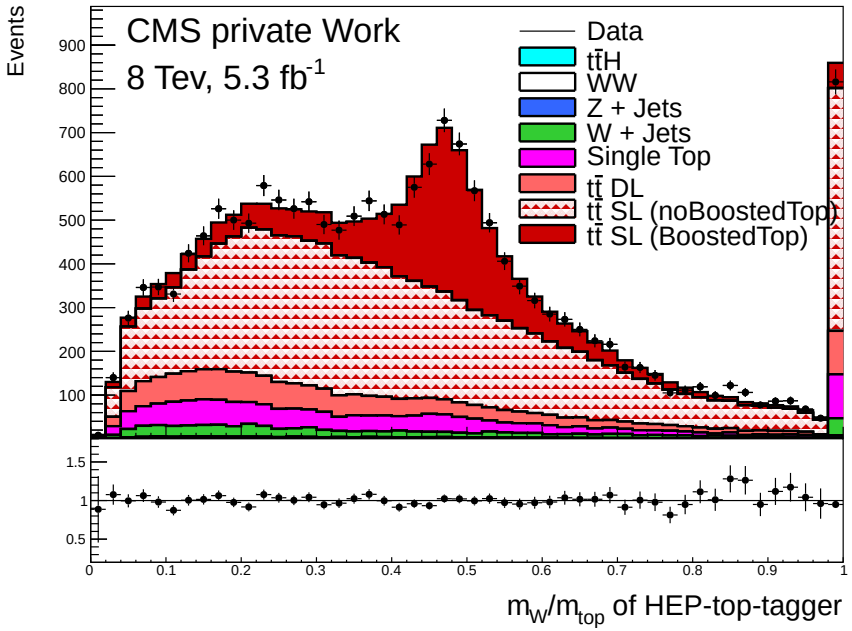
- Combining invariant masses of subjet-combinations to likelihood-ratio





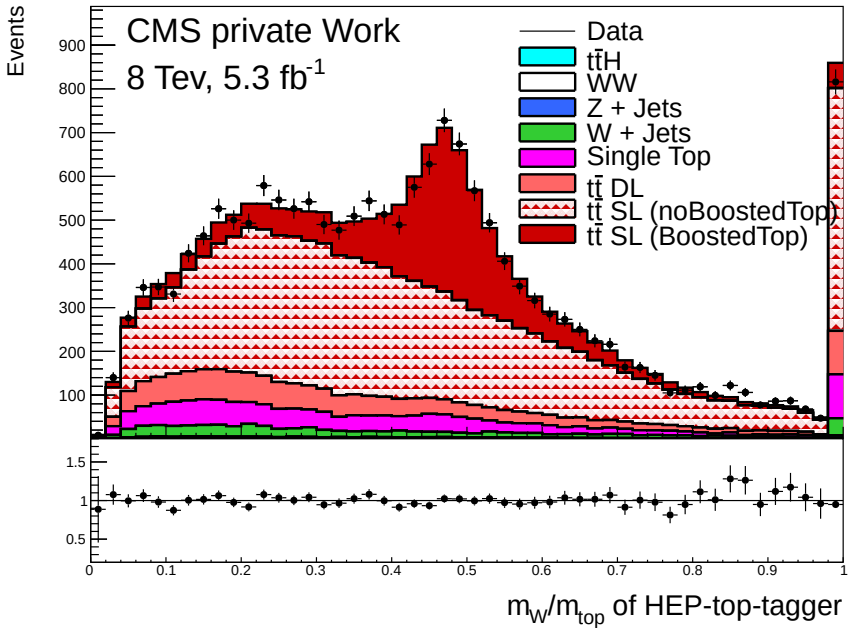
Data vs. MC

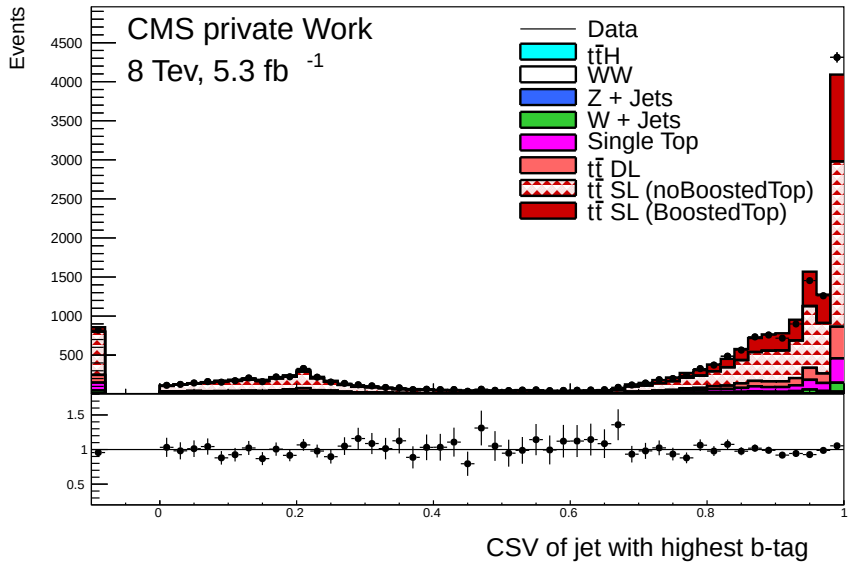
- We study modeling of subjet-tools in $t\bar{t}$ events
  - Properties of Cambridge-Aachen fat/sub/filter jets mostly well-modeled



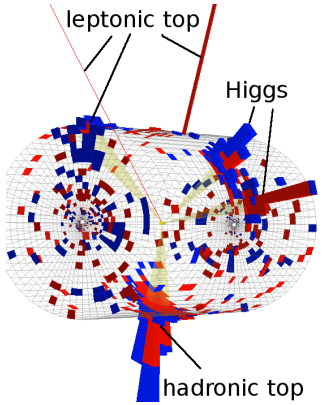
Data vs. MC

- We study modeling of subjet-tools in $t\bar{t}$ events
  - Properties of Cambridge-Aachen fat/sub/filter jets mostly well-modeled
  - B-tagging calibrations of anti-k_T-0.5-jets can be used for subjets, too





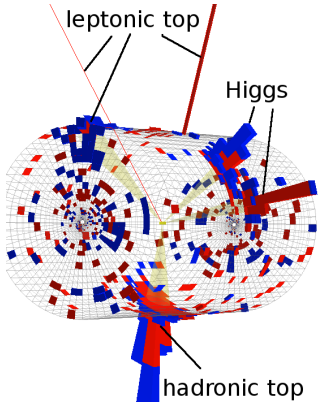
Event selection



“boosted” event selection

- 30 GeV isolated lepton
- 200 GeV CA-1.5 jet
- 200 GeV CA-1.2 jet with 2 b-tagged filterjets

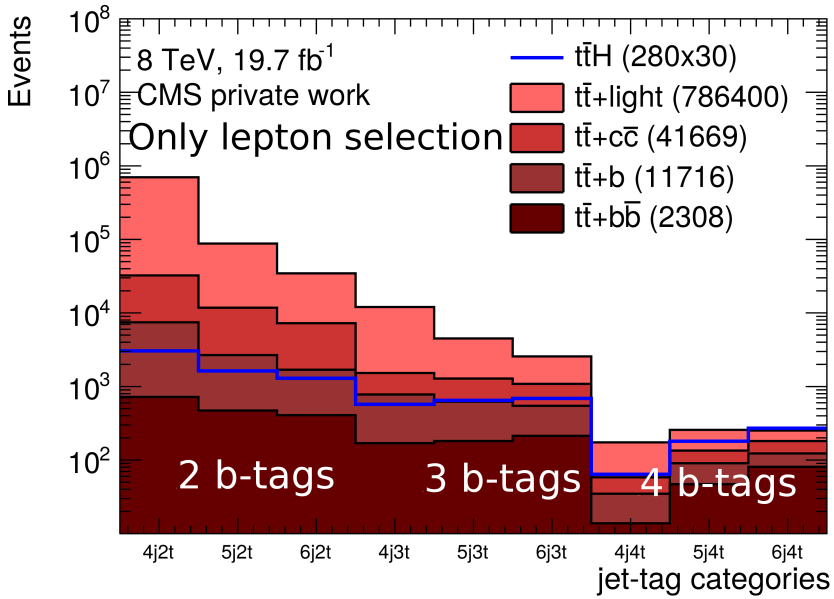
Event selection



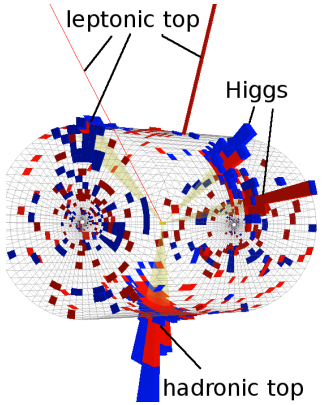
“boosted” event selection

- 30 GeV isolated lepton
- 200 GeV CA-1.5 jet
- 200 GeV CA-1.2 jet with 2 b-tagged filterjets

- All events sorted by number of anti-k_T-0.5 jets and corresponding b-tags



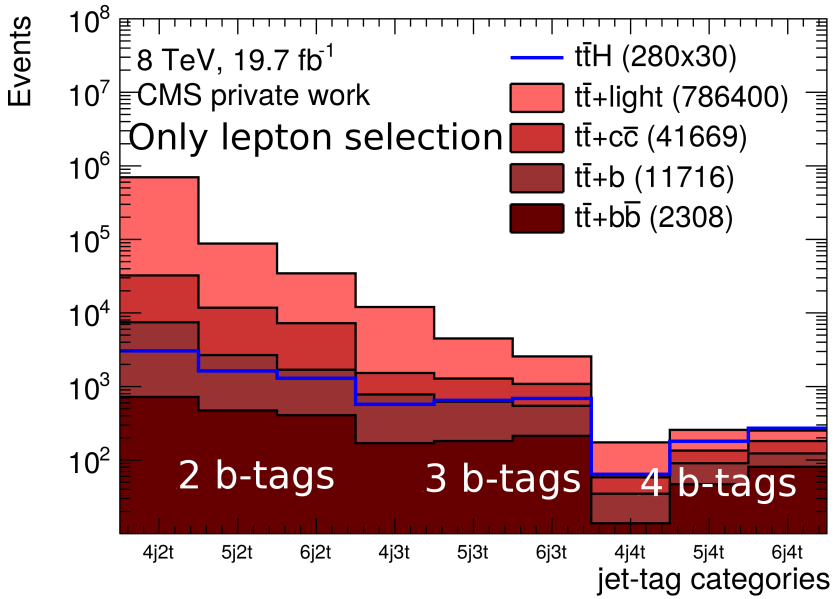
Event selection

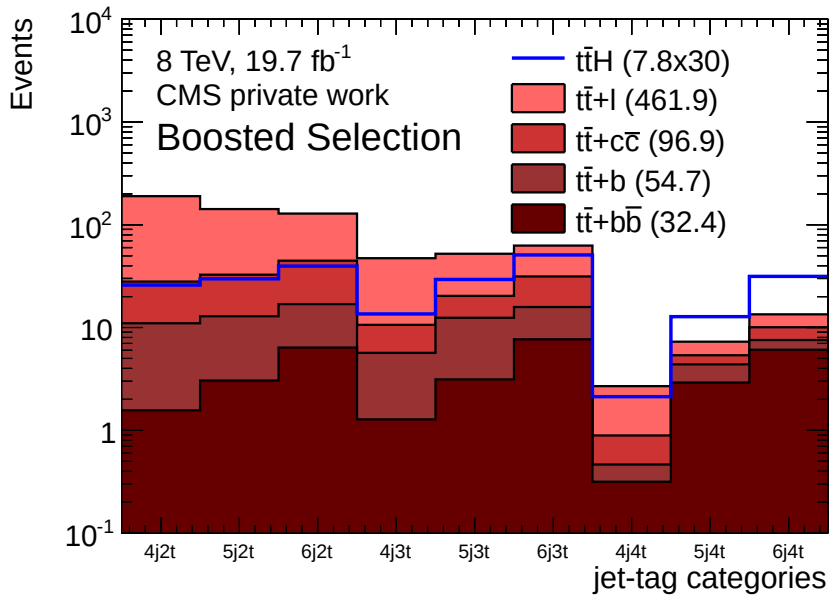


“boosted” event selection

- 30 GeV isolated lepton
- 200 GeV CA-1.5 jet
- 200 GeV CA-1.2 jet with 2 b-tagged filterjets

- All events sorted by number of anti-k_T-0.5 jets and corresponding b-tags
- After “boosted” selection: S/B improved



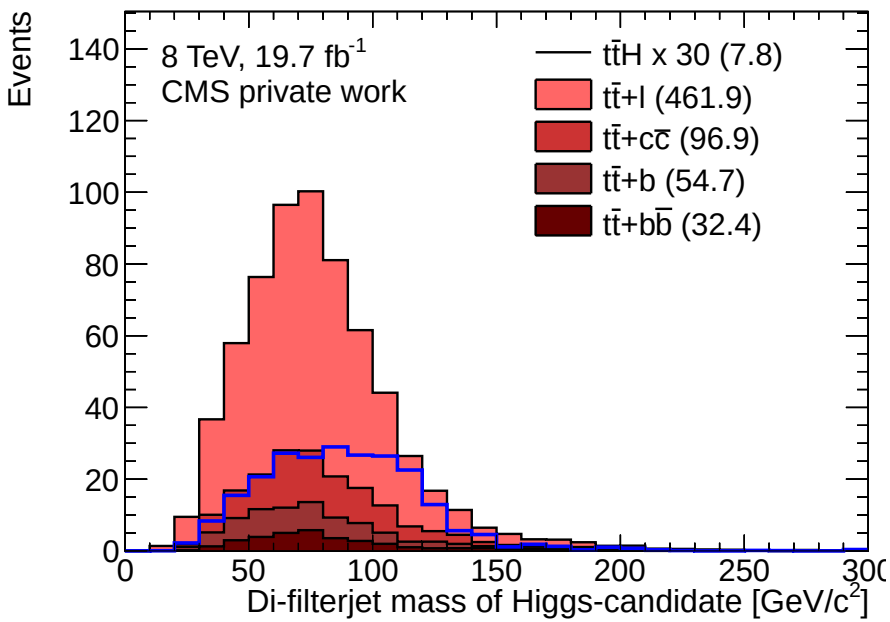


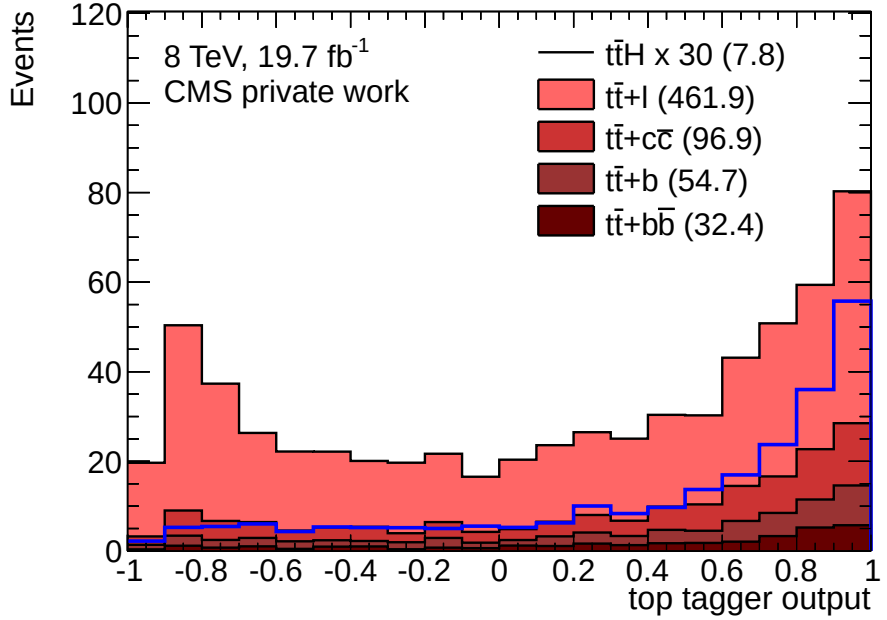
Multivariate analysis

- Further analysis of top and Higgs candidates with high p_T
  - Fat jets determine assignment of jets to Higgs boson and top quark (40% correct, instead of $< 20\%$ with default method)

Multivariate analysis

- Further analysis of top and Higgs candidates with high p_T
  - Fat jets determine assignment of jets to Higgs boson and top quark (40% correct, instead of $< 20\%$ with default method)
  - Invariant masses of Higgs boson and top-tag important variables for signal-background discrimination (especially against $t\bar{t}b\bar{b}$)





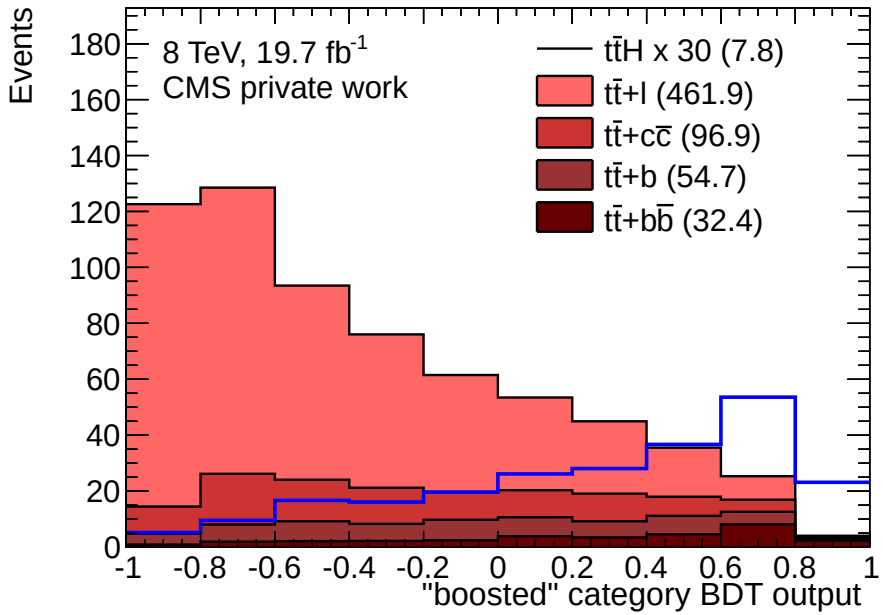
Multivariate analysis

- Training BDT for events accepted by boosted selection
- Using mixture of
  - Subjet information (most importantly di-filterjet mass of Higgs candidate)
  - B-tag information (most importantly 3rd and 4th highest tag of anti-k_T-0.5 jets)
  - General event variables (H_T, Sphericity, ...)

Multivariate analysis

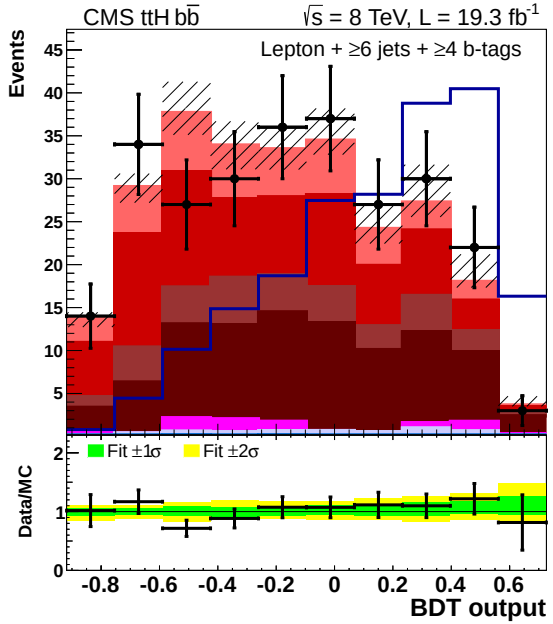
- Training BDT for events accepted by boosted selection
- Using mixture of
  - Subjet information (most importantly di-filterjet mass of Higgs candidate)
  - B-tag information (most importantly 3rd and 4th highest tag of anti-k_T-0.5 jets)
  - General event variables (H_T, Sphericity, ...)

Discriminant in boosted regime



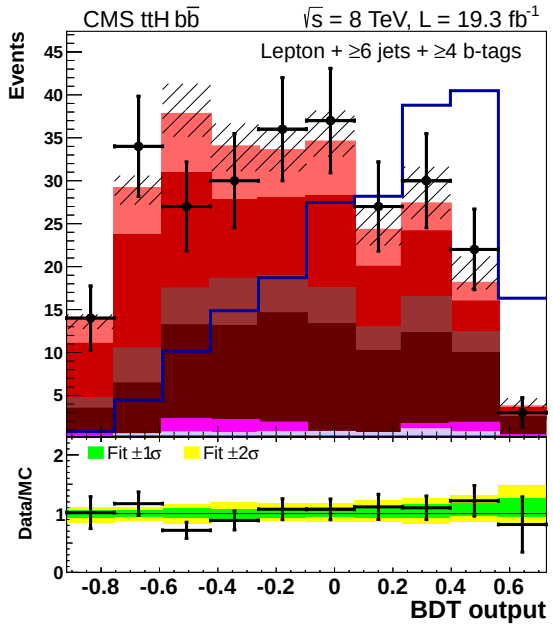
Comparison to existing analysis

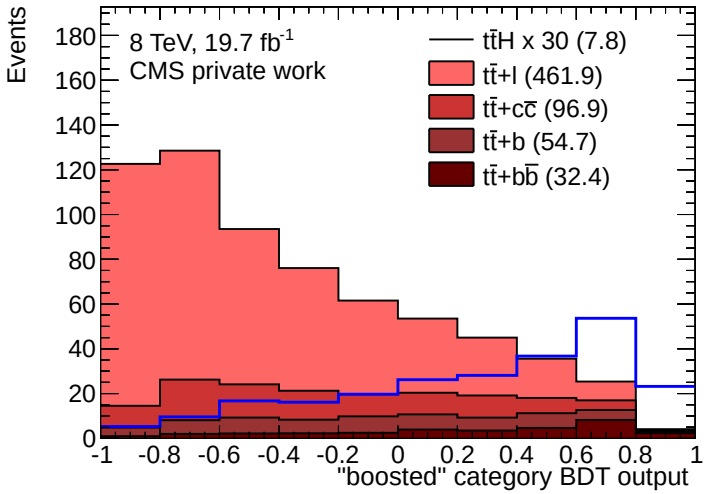
- So far: events categorized according to number of jets and b-tags
  - Best category (of 7): $\geq$6 Jets, $\geq$4 b-Tags
  - BDTs trained in all categories
  - Fit background and signal-model from MC-simulation



Comparison to existing analysis

- So far: events categorized according to number of jets and b-tags
  - Best category (of 7): $\geq$6 Jets, $\geq$4 b-Tags
  - BDTs trained in all categories
  - Fit background and signal-model from MC-simulation





- New: additional *boosted* category
  - S/B comparable to best jet-tag categories
  - New BDT with additional variables performs significantly better than existing BDTs in this category
  - Planning to create more categories with well-defined topologies to train dedicated MVAs

Conclusion and outlook

- $t\bar{t}H$-events with top quark and Higgs boson with high p_T
  - Good S/B
  - Simplified combinatorics
  - Require specialized techniques (subjet-algorithms)
- Introducing new analysis category for boosted events
  - Events in category have unique features
  - Dedicated MVA performs better than existing MVAs
- Outlook
  - Validating new techniques with existing data set
  - Creating more categories
  - Run 2: cross sections increase by a factor of 4