Boosted ttH

Michele Selvaggi
Clement Helsens
David Jamin

Measurement

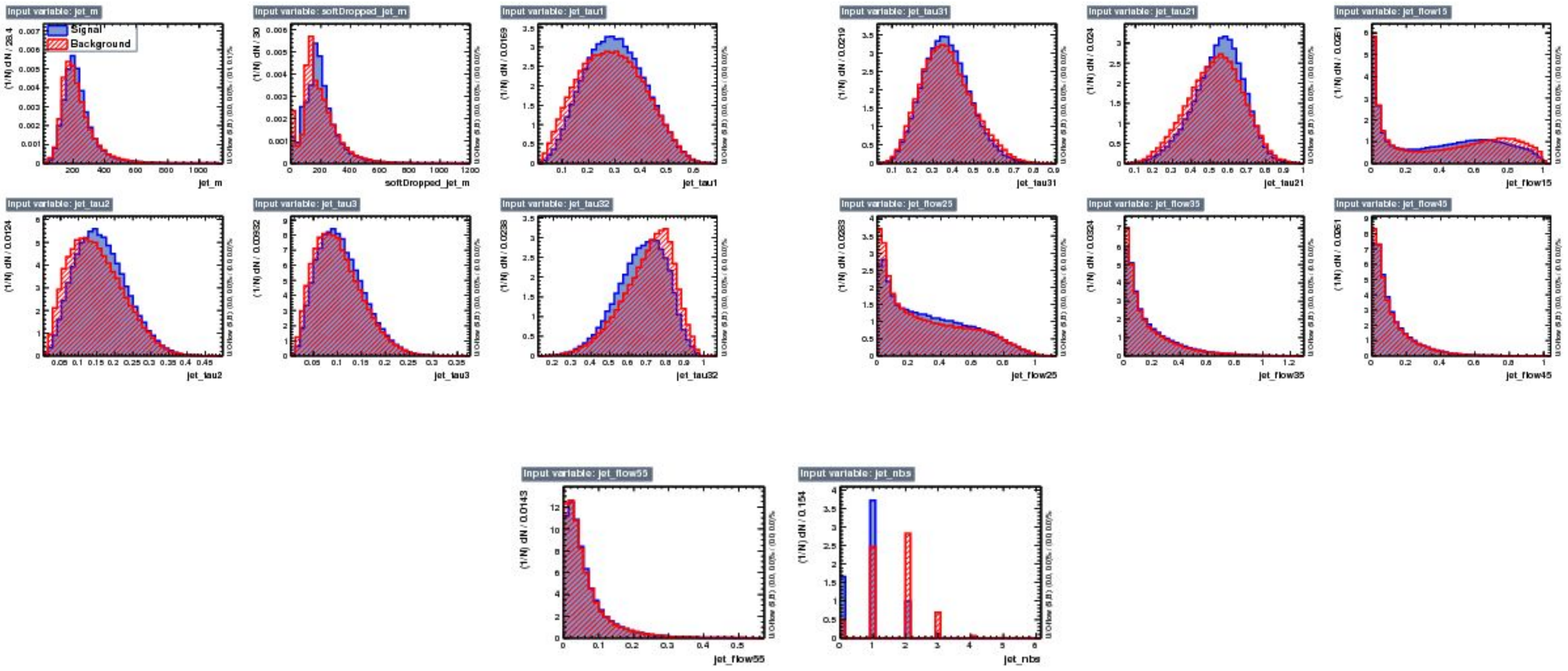
- Measure ttH/ttZ ratio, < 1% th. uncertainties
  - translate into 1% top yukawa measurement (1507.08169)
- with H->bb, Z->bb
- Final state:
  - boosted Higgs, H-> bb
  - boosted top hadronic
  - other top leptonic decay
  - signature: 2 fatjets, 1 lepton, MET, (+ 1 bjet)
- Backgrounds:
  - ttZ, tt+jets, tt+bb
  - W/Z+jets ignored for now (should be added)

Event Selection

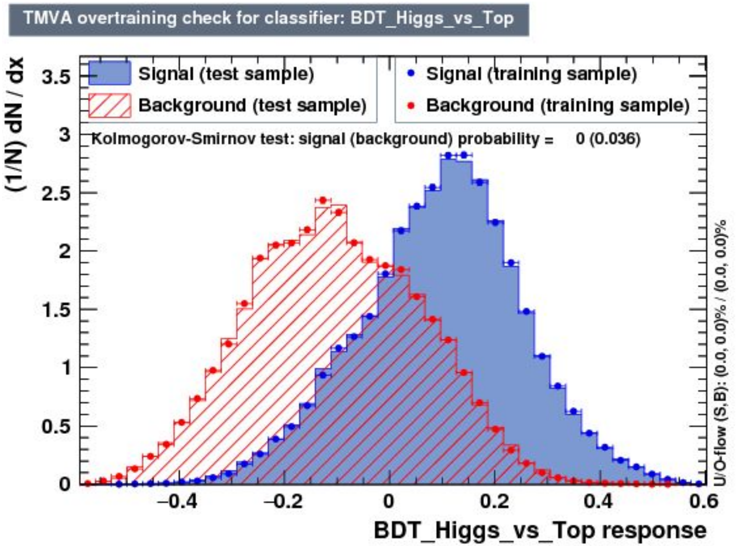
- Pre-selection:
    - 2 fat-jets R = 1.5 : pT > 250 GeV, |eta| < 3.0
    - 1 e/mu pT > 25 GeV, |eta| < 3.0
    - Also reconstruct R=0.4 jets (for b-tagging purposes)

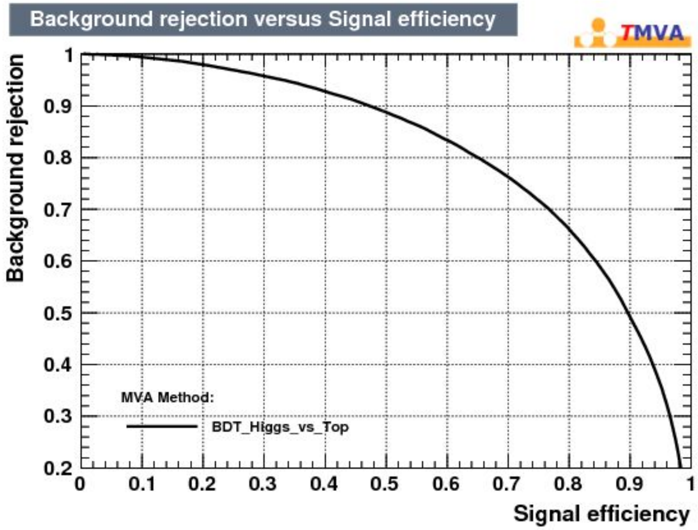
- Step 1: Decide which fatjet is top and which is Higgs
    - use a BDT with the following input variables
        - tau(i,j), mSD, m, EFlow, nbjets inside fatjet

Top vs Higgs BDT



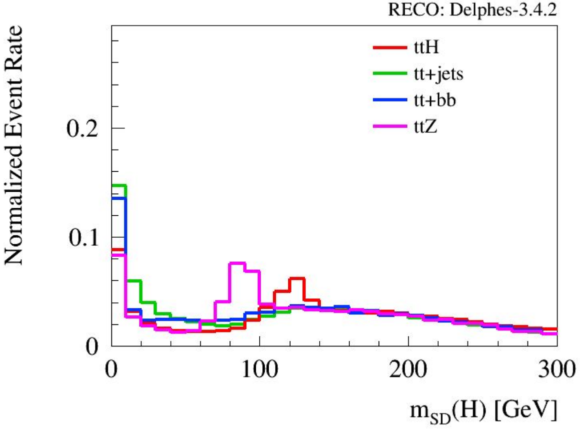
Top vs Higgs BDT

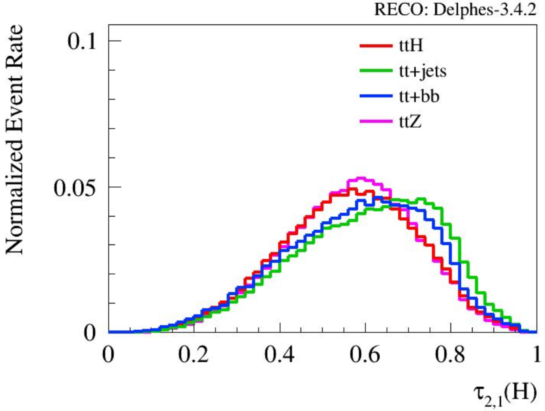


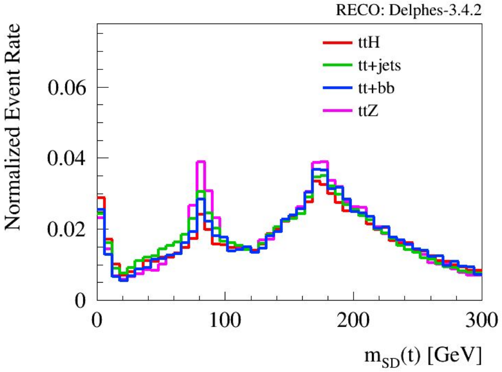


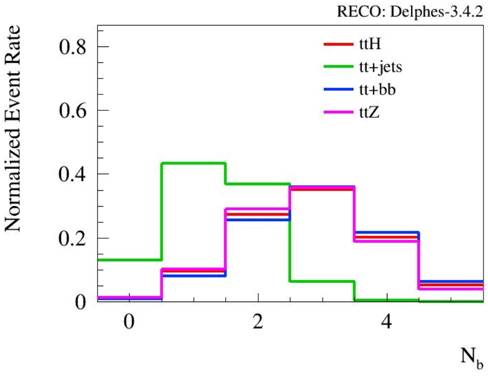
this is the first try, completely un-optimized ...

Validation Distributions (normalized)





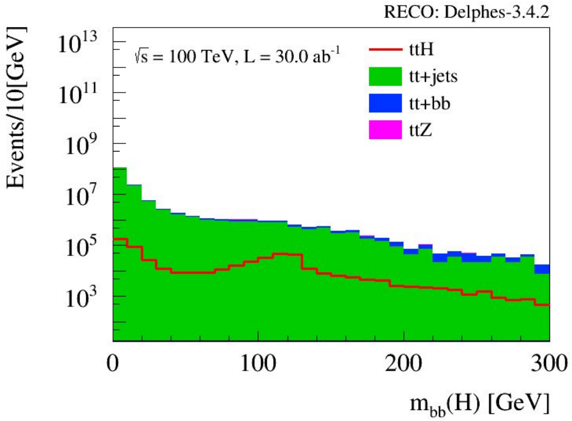


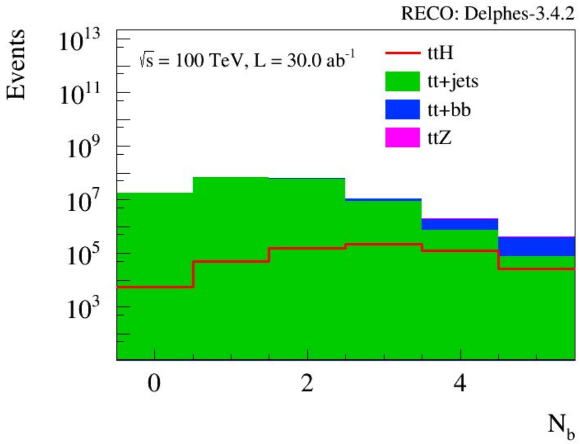


Event Selection

- Pre-selection:
  - 2 fat-jets R = 1.5 : pT > 250 GeV, |eta| < 3.0
  - 1 e/mu pT > 25 GeV, |eta| < 3.0
  - Also reconstruct R=0.4 jets (for b-tagging purposes)
- Step 1: Decide which fatjet is top and which is Higgs
  - train a BDT with the following input variables
    - tau(i,j), mSD, m, EFlow, nbjets inside fatjet
  - **use BDT output to tag “top” or “Higgs” jet**
- Step 2: Top/Higgs tagging:
  - Top jet: 120 < mSD < 250 GeV
  - Top jet: tau(3,2) < 0.8
  - Higgs jet: tau(2,1) < 0.6
  - no mass cut on Higgs jet (because we fit it)

Selection step 2





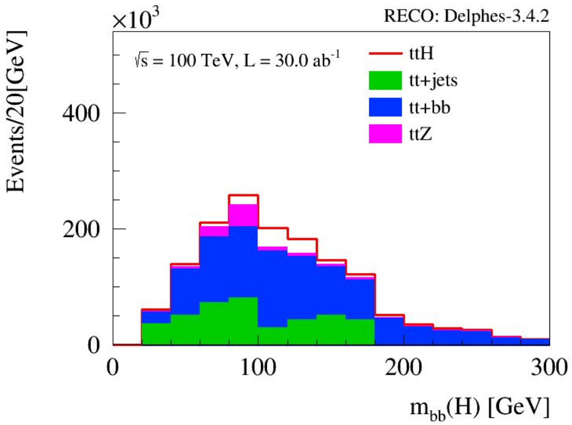
process	yield (30.0 ab-1)	stat. error	raw
ttH	1349501.2	8115.1	27654
tt+jets	483299579.7	1883511.2	65853
tt+bb	12292826.6	104963.4	13716
ttZ	1094289.4	6941.9	24849
signal	1349501.155	90.084	
background	496686695.645	1886446.38	

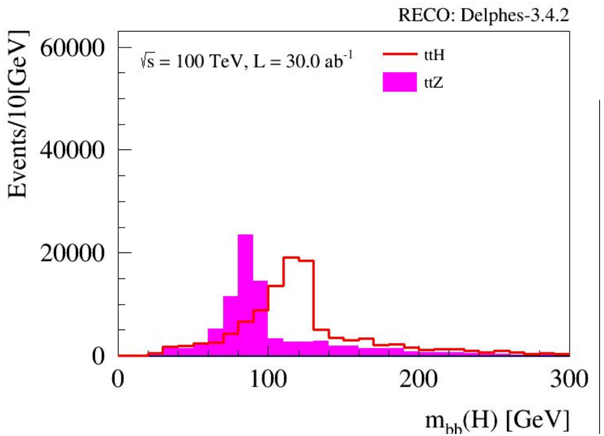
(sig_s, sig_b) (%)		S/B	significance	dmu/mu (%)
0.0	0.0	0.003	60.47	1.65
1.0	0.0	0.003	51.75	1.93
1.0	0.1	0.003	2.71	36.86

Event Selection

- Pre-selection:
  - 2 fat-jets R = 1.5 : pT > 250 GeV, |eta| < 3.0
  - 1 e/mu pT > 25 GeV, |eta| < 3.0
  - Also reconstruct R=0.4 jets (for b-tagging purposes)
- Step 1: Decide which fatjet is top and which is Higgs
  - train a BDT with the following input variables
    - tau(i,j), mSD, m, EFlow, nbjets inside fatjet
  - **use BDT output to tag "top" or "Higgs" jet**
- Step 2: Top/Higgs tagging:
  - Top jet: 120 < mSD < 250 GeV
  - Top jet: tau(3,2) < 0.8
  - Higgs jet: tau(2,1) < 0.6
  - no mass cut on Higgs jet (because we fit it)
- Step 3: bTagging:
  - >= 4 bjets, 2 bjets being close Dr = 1.1 to Higgs jet

Selection step 3





process	yield (30.0 ab-1)	stat. error	raw
ttH	174458.2	2917.8	3575
tt+jets	675255.7	70400.3	92
tt+bb	1407992.9	35523.2	1571
ttZ	138938.5	2473.6	3155
signal	174458.184	54.017	
background	2222187.095	78893.702	

(sig_s, sig_b) (%)		S/B	significance	dmu/mu (%)
0.0	0.0	0.079	112.69	0.89
1.0	0.0	0.079	74.8	1.34
1.0	0.1	0.079	54.15	1.85

Selection step 4

100 < mH < 135 GeV

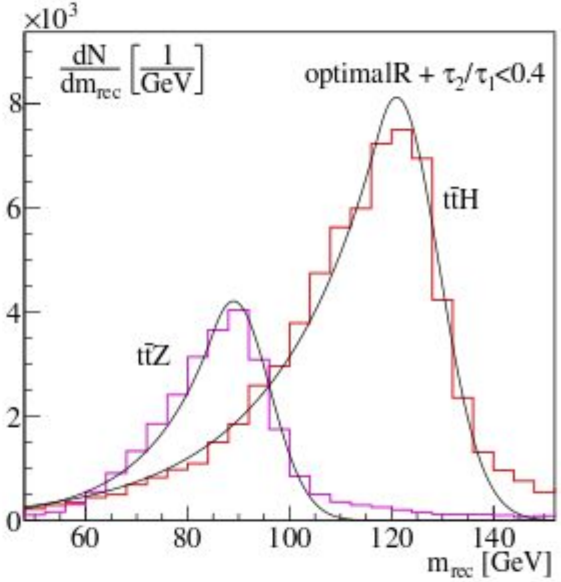
process	yield (30.0 ab-1)	stat. error	raw
ttH	53923.4	1622.2	1105
tt+jets	58288.5	20608.1	8
tt+bb	223163.7	14142.4	249
ttZ	10260.8	672.2	233
signal	53923.439	40.276	
background	291712.994	25003.076	

(sig_s, sig_b) (%)		S/B	significance	dmu/mu (%)
0.0	0.0	0.185	91.72	1.09
1.0	0.0	0.185	67.59	1.48
1.0	0.1	0.185	63.48	1.58

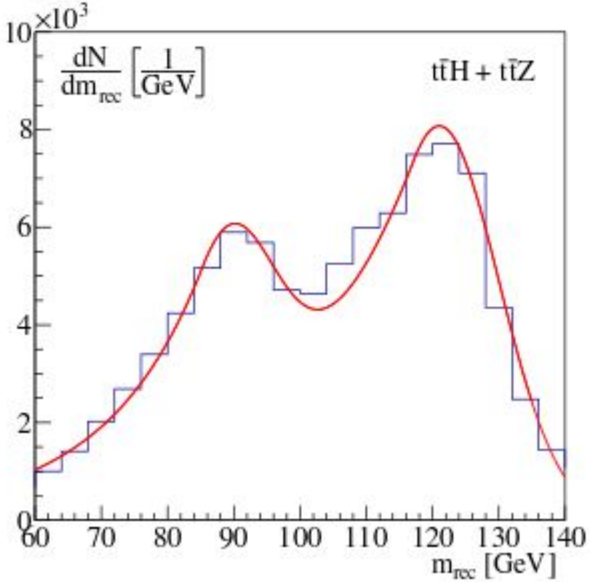
S/B ~ ⅕
dmu/mu ~ 1%

Gives good statistical precision, but this would be a standalone ttH measurement

Alternative - fit ttH/ttZ



1507.08169



- Fit with a double Crystall Ball function (NOT DONE HERE YET)
- In the paper, claimed N(H)/N(Z) = 2.80 +/- 0.03 < 1%

Conclusion

- To be addressed:
  - update K-factors (right now everything is LO)
  - generate more backgrounds, mostly (tt+jets, W+jets)
  - optimize BDT Higgs vs Top (more events, more parameters)
  - optimize selection (Top vs QCD, and loose Higgs vs. QCD taggers)
  - use more sub-jet information (non b-tagged subjets) for better discrimination
  - do the 2-crystal ball fit and make precise estimation of background stat. uncertainty from control region at high mass mbb.
- Open questions:
  - why do we need such tight btag requirements to kill tt+jets (in the paper 1507.08169 seems that 3 btag are sufficient)?
  - why relative peak height is switched? Broader mH distribution?
  - Before mH cut we have NH/NZ ~ 1.3 (vs. 2.8 in 1507.08169)
    - to be investigated ...