



BDT and Results for $1\tau 1l$

Search for Four top in Tau Final States

Huiling Hua¹

¹IHEP

September 4, 2024



Outline

- ① 2018
Old Theory Uncertainties
New Theory Uncertainties
- ② 2017
- ③ 2016preVFP
- ④ 2016postVFP
- ⑤ Combination

Updates

- BDT training with separate $t\bar{t}$ MC samples
- QCD renormalization and factorization scale uncertainties normalized to 1.

Section 1

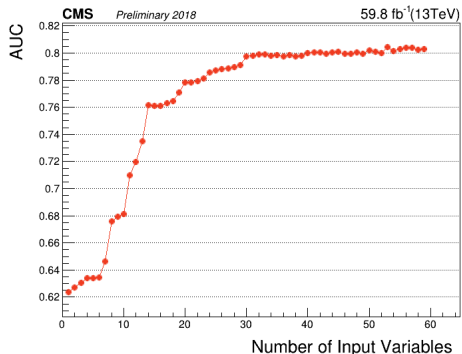
2018



BDT in 1tau1l

- Extra tt MC samples with different generator settings for BDT training
 - Only 2018 samples
 - TT extra samples: 39K
- Application with nominal tt MC samples, to avoid overtraining

Choosing input variables for 1tau1l

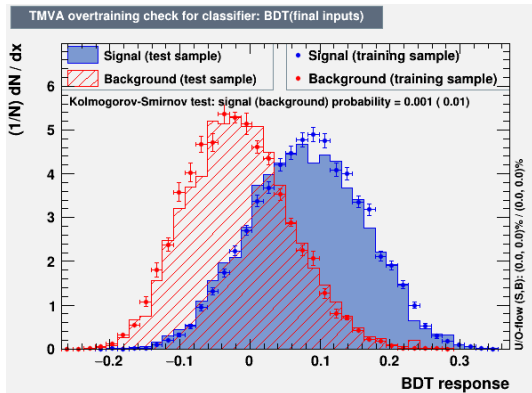


jets_num	n_{jet}
jets_rationHT_4toRest	$HT_{leading\ 4\ jets}/HT_{all\ jets}$
jets_centralty	centrality
jets_7pt	$p_T^{7th\ jet}$
jets_4pt	
jets_MHT	
jets_minDeltaR	
jets_1pt	
jets_tauST_invariantMass	
jets_METDivideHT	
bjetsM_num	
bjetsT_invariantMass	
bjetsT_minDeltaR	
bjetsT_2MET_stransMass	
bjetsT_leptons_minDeltaR	
bjetsM_invariantMass	
bjetsM_minDeltaR	
bjetsM_tauST_minDeltaR	
bjetsM_HT	
tausT_leptonl_charge	
tausT_leptonsT_invariantMass	
tausT_1pt	
tausT_invariantMass	
muonsTopMVAT_1pt	
tausT_1Met_transMass	
tausT_1leptonlMet1_stransMass	

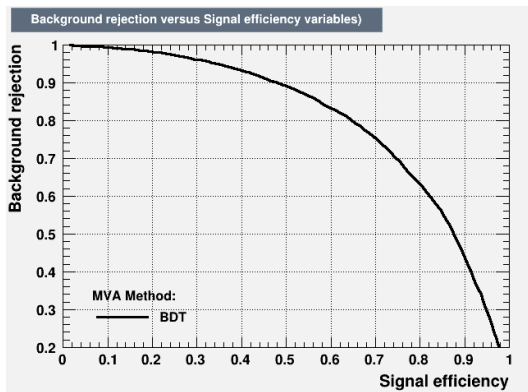
Table 1: Final 26 inputs for 1tau1l

- 1tau1l SR, with TT extra samples
- AUC with all variables: 0.806
- Input variables picked from only increase of AUC, remove the variables in the plateau: 27

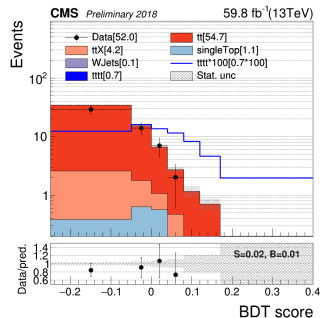
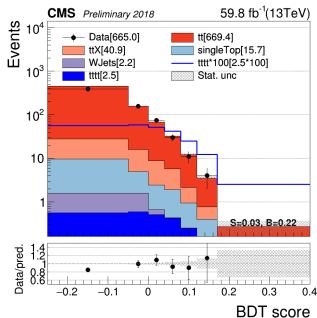
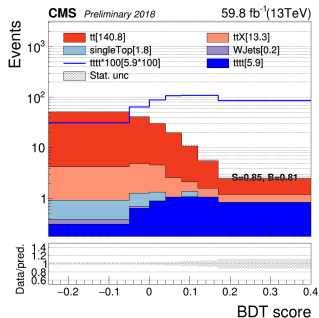
BDT performance with extra tt samples



- AUC with 27 input variables: 0.799



BDT in SR and control regions

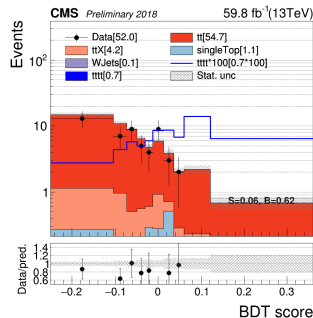
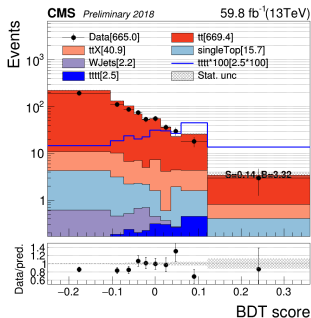
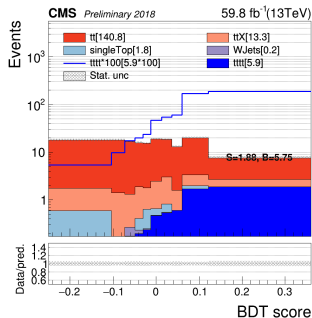


- With nominal tt MC samples
- Add systematic uncertainties

How to optimize binning?

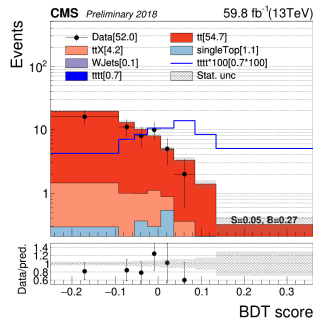
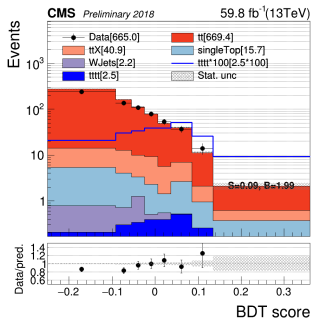
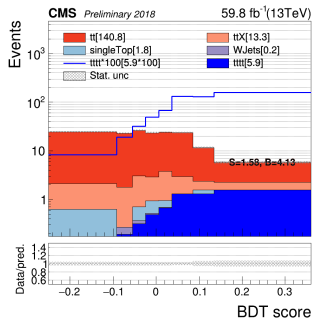
- Dance between statistic uncertainty and significance
- Adaptive binning for background
 - Binning background with roughly the same number of events in each bin except the last one
 - Even background to better contain nuisance parameters
 - Last bin with the highest significance
- Choose the binning with the highest significance

BDT in SR and control regions: **BinD**



- With nominal tt MC samples
- Add systematic uncertainties

BDT in SR and control regions: **BinE**



- With nominal $t\bar{t}$ MC samples
- Add systematic uncertainties

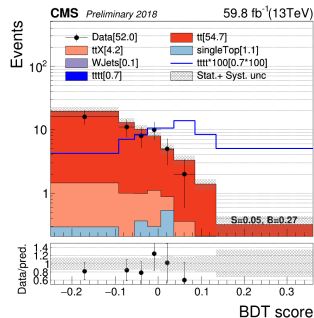
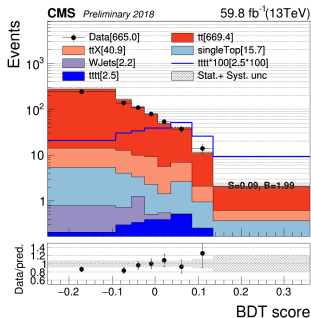
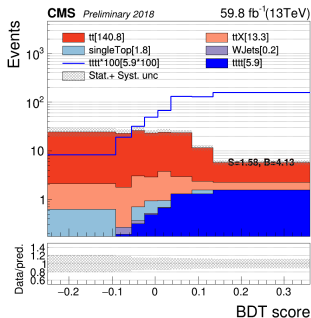
Section 1

2018

1.1 Old Theory Uncertainties

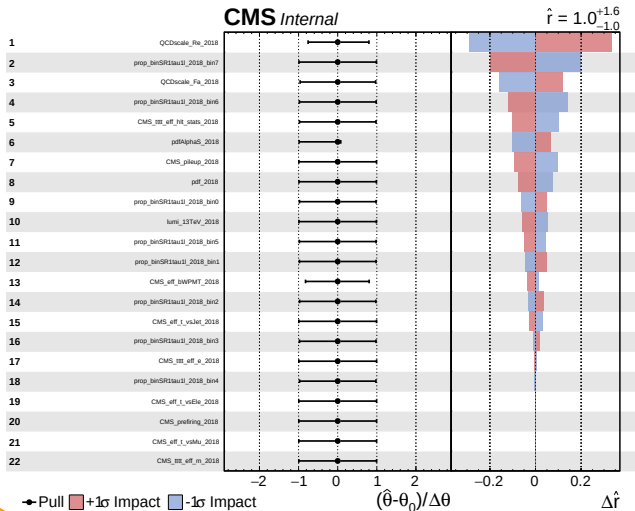


Results with systematics



- The results without systematics: sig = 0.87, limit = 2.53
- With all systematics: significance=0.78; limit = 3.10
- Seems problem with QCD renormalization and factorization scale

Impacts



- Seems some problems with QCD renormalization and factorization scale, and pdf

Section 1

2018

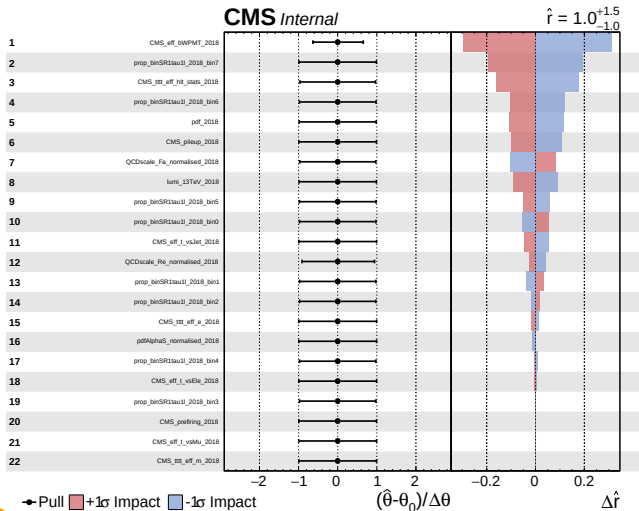
1.2 New Theory Uncertainties



BDT in SR and control regions: **BinE**; pdf and QCD scale renormalized

- QCD renormalization has around 30% variation on 4tops signal. Number 1 impact parameter
- Re-normalize QCD scale for all the samples:
$$\sum genWeight / \sum genWeight * QCQScale$$
- Results considering all systematics(except energy scale): significance=0.795; limit = 2.88

Impacts



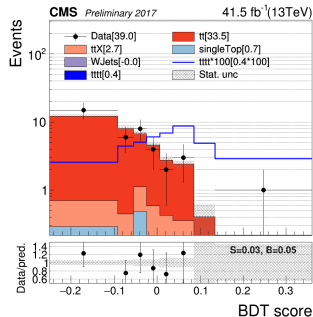
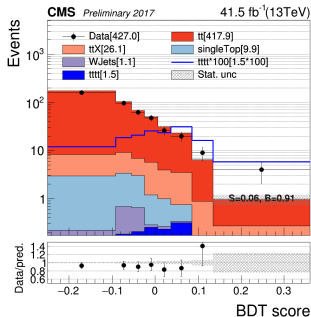
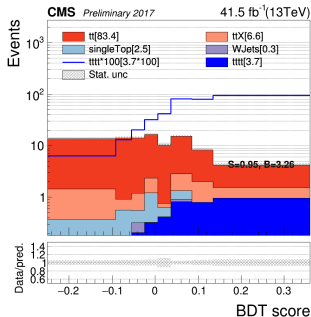
- Does it make sense that btag WP constrained?

Section 2

2017



BDT in SR and control regions: **BinE**



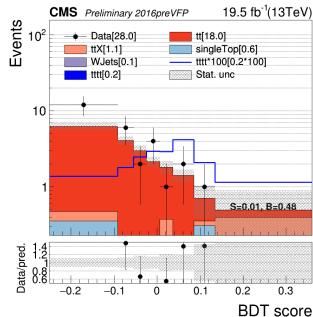
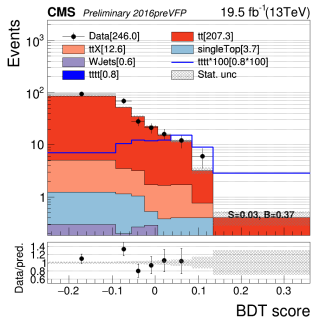
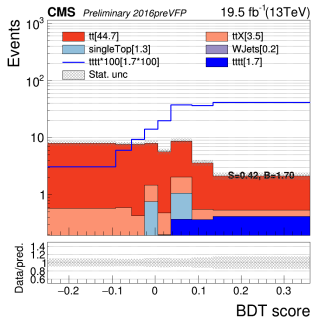
- With nominal tt MC samples
- With 2018 training of extra tt samples
- Might have to rebinning with more bg in each bin, because of the lower lumi in 2016preVFP and 2016postVFP

Section 3

2016preVFP



BDT in SR and control regions: **BinE**



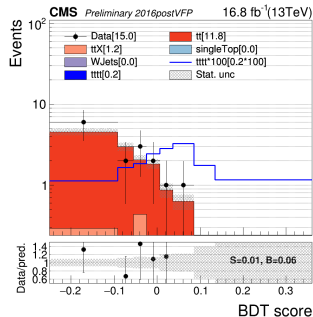
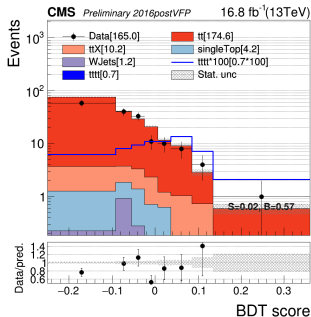
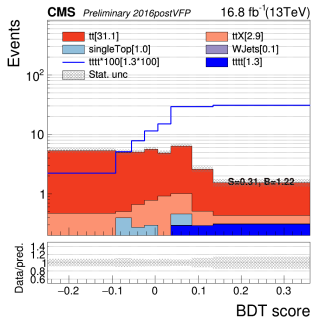
- With nominal tt MC samples
- With 2018 training of extra tt samples
- Might have to rebinning with more bg in each bin, because of the lower lumi in 2016preVFP and 2016postVFP

Section 4

2016postVFP



BDT in SR and control regions: **BinE**



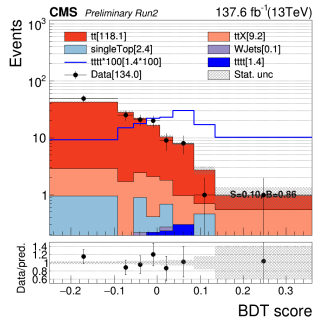
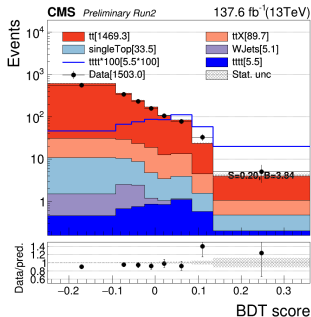
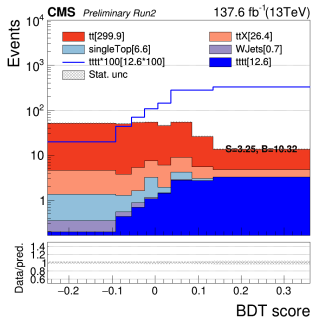
- With **nominal** tt MC samples
- With 2018 training of extra tt samples
- Might have to rebinning with more bg in each bin, because of the lower lumi in 2016preVFP and 2016postVFP

Section 5

Combination



Combination of 3 years



- Add data hists: not using quadratic sum of the statistical uncertainties;
- Add signal and background hists: using quadratic sum of the systematic uncertainties;
- Careful with correlation of the systematics

Section 6

Back up



jets_num jets_rationHT_4toRest jets_centrality	
jets_7pt jets_4pt jets_MHT jets_minDeltaR jets_1pt jets_tausT_invariantMass jets_METDivideHT	
bjetsM_num bjetsT_invariantMass bjetsT_minDeltaR bjetsT_2MET_stransMass bjetsT_leptons_minDeltaR bjetsM_invariantMass bjetsM_minDeltaR bjetsM_tausT_minDeltaR	