Update on $t\bar{t}t\bar{t}$ Searches in Single Lepton/OS Dilepton Channels Using 2016 Data

Denys Lontkovskyi [1] Freya Blekman [1] Long Wang [2]
Robert Clare [2] Steve Wimpenny [2]

[1]Vrije Universiteit Brussel

[2]University of California, Riverside

March 22, 2018

Overview

1. Current stage
2. Analysis setup
3. $t\bar{t}t\bar{t}$ Search Method
4. Re-blinded fit result
5. Conclusion

Current Status

- Aiming at re-preapproval, we are finalizing our documentation
  - Single-lepton: AN-17/146
  - Di-lepon: AN-17/143

- We have requested for the production of two new $t\bar{t}$ samples with 9M events each for the two channels(semi-lep and OS dilep), with dedicated cuts at generator level to increase MC stats by a factor of $\backsim$ 10 in high multiplicity/discriminant tails. Details on slide 24 to 30

- We are studying the effects of possible background from QCD multi jets with mis-identified leptons.

Data, MC and Objects

Data and MC

- Run2016 B-H, $35.9 pb^{-1}$
- Summer 16 MiniAOD MC for Morond 17
  - signal sample: $t\bar{t}t\bar{t}$ amc@NLO
  - background samples: $t\bar{t}$(+backup, mass, width, details on slide 23), single $t(\bar{t})$, DY, W+jets, $t\bar{t} + Z/H/W/diboson$

Objects

Single Lepton

- μ : tight ID, $p_T > 26$ GeV, $|\eta| < 2.1$, $RelIso < 0.15$
- e : tight ID, $p_T > 35$ GeV, $|\eta| < 1.4442$ or $1.566 < |\eta| < 2.1$
- jet : loose ID, $p_T > 30$ GeV, $|\eta| < 2.1$, $\Delta R > 0.4$

OS Dilepton

- μ : loose ID, leading(subleading) lep $p_T > 25(20)$ GeV, $|\eta| < 2.4$, $RelIso < 0.15$
- e : loose ID, leading(subleading) lep $p_T > 25(20)$ GeV, $|\eta| < 1.4442$ or $1.566 < |\eta| < 2.4$
- jet : loose ID, $p_T > 30$ GeV(25 GeV if tagged as b), $|\eta| < 2.4$, $\Delta R > 0.4$

Event Selection and MC Re-weighting

Event selection

Single Lepton

- $N_l^{tight}=1$
- $N_\mu^{loose}=0, N_e^{veto}=0$
- $N_j \geq 8(7)$ in $e(\mu)$ channel of which $N_{tags}^M \geq 2$
- $\not{E}_T > 50$ GeV
- $HT \geq 500$ GeV

OS Dilepton

- Exactly 2 OS leptons
- $M_{ll} \geq 106$ GeV or $76 \geq M_{ll} \geq 20$ GeV
- $N_j \geq 4$ of which $N_{tags}^M \geq 2$
- $HT \geq 500$ GeV

MC Re-weighting

- Trigger eff.
- Lepton scales
- Pileup Reweight
- JER/JEC
- b-tagging eff.
- top p_T *reweight*

$t\bar{t}t\bar{t}$ Search Method

Binned analysis fitting on event level BDT

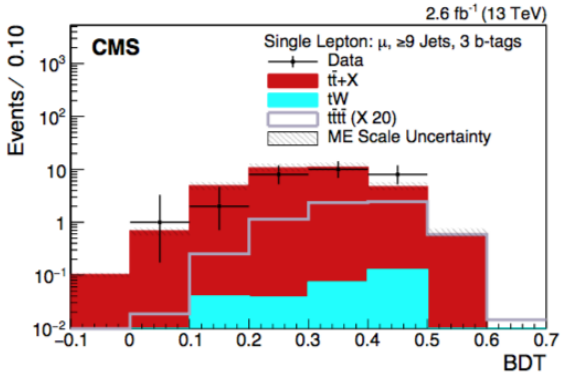


Figure: Single μ event level BDT in $\geq$9 jet 3 btag category

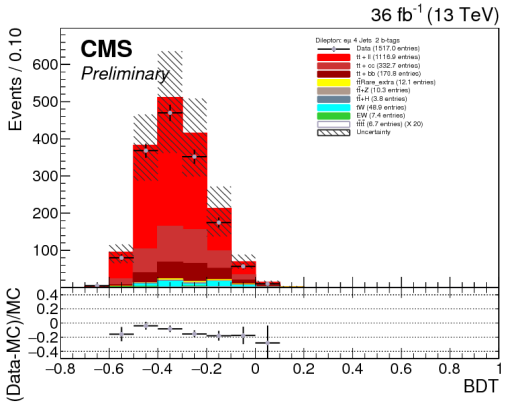


Figure: μe event level BDT in CR

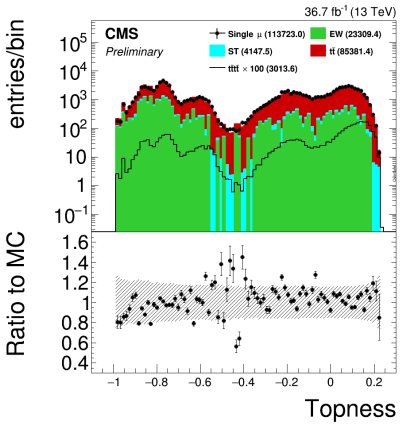
Event categorization in $N_j \otimes N_{tags}^M$ for limit fitting

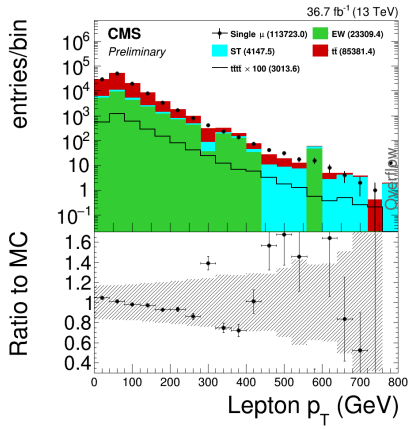
- Single lepton channel
  - μ: N_j: 7, 8, 9, 10+; N_{tags}^M: 2, 3, 4+
  - e: N_j: 8, 9, 10+; N_{tags}^M: 2, 3, 4+
- OS Dilepton channel: N_j: 4-5, 6-7, 8+; N_{tags}^M: 2, 3+

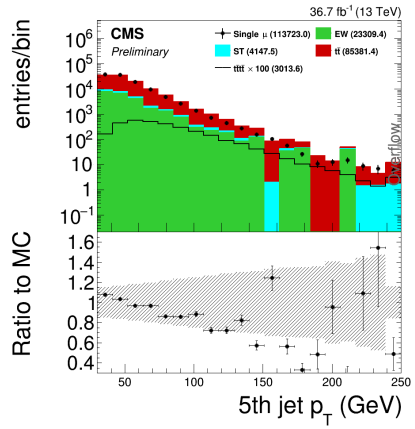
Control Plots (BDT Input Variables)

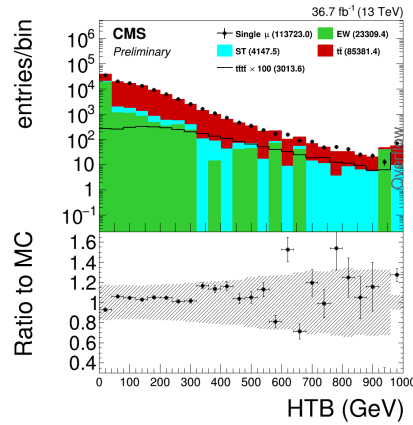
- Single lepton channel

$BDT_{...}$, H_T^b, H^{Rat}, p^{5thjet}, p^{6thjet}, M^H, HT_X, p^{lep}, $CSV_{...}$, $CSV_{...}$, $CSV_{...}$, $CSV_{...}$



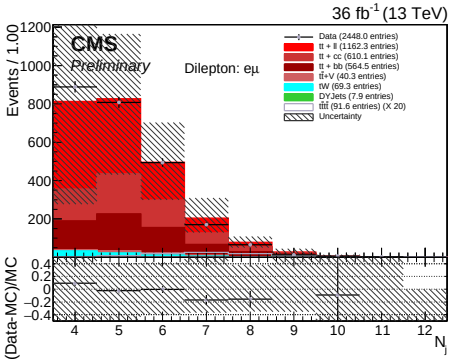


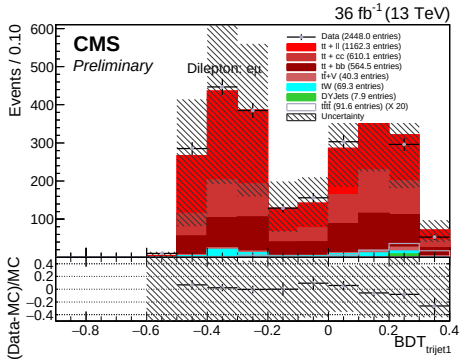


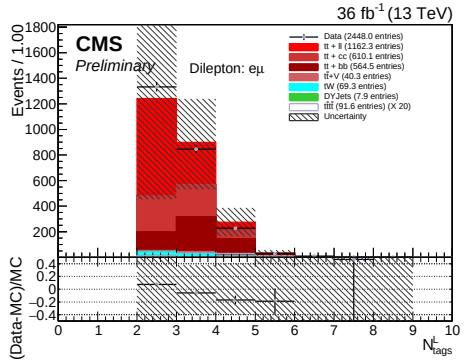


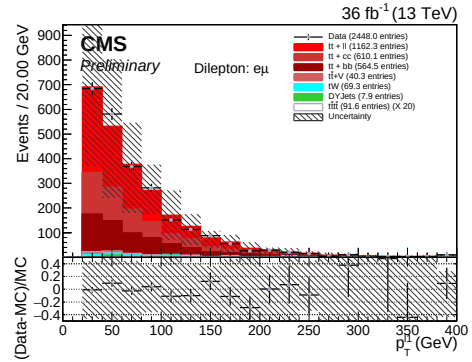
- OS Dilepton channel

N_j, $BDT_{trijet1}$, H_T^b, H_T^{2M}, HTH, S, H_T^{Rat}, p_T^{l1}, η^{l1}, ΔR_{ll}, ΔR_{bb}, N_{tags}^L, N_{tags}^M, p_T^{3rdjet}, p_T^{4thjet}









- Overall reasonable distributions agreement within uncertainties.

Sources of Systematic Uncertainties

Experimental Uncertainties

- Luminosity uncertainty
- Pileup $\pm 1\sigma$
- Lepton SFs uncertainty
- JER $\pm 1\sigma$
- JES(split)
  - SubTotalPileUp
  - SubTotalRelative
  - SubTotalPt
  - SubTotalScale
  - Jet flavor
- b-tag CSV $\pm 1\sigma$
- Heavy flavor fraction
- Top p_T reweight
- Jet normalization

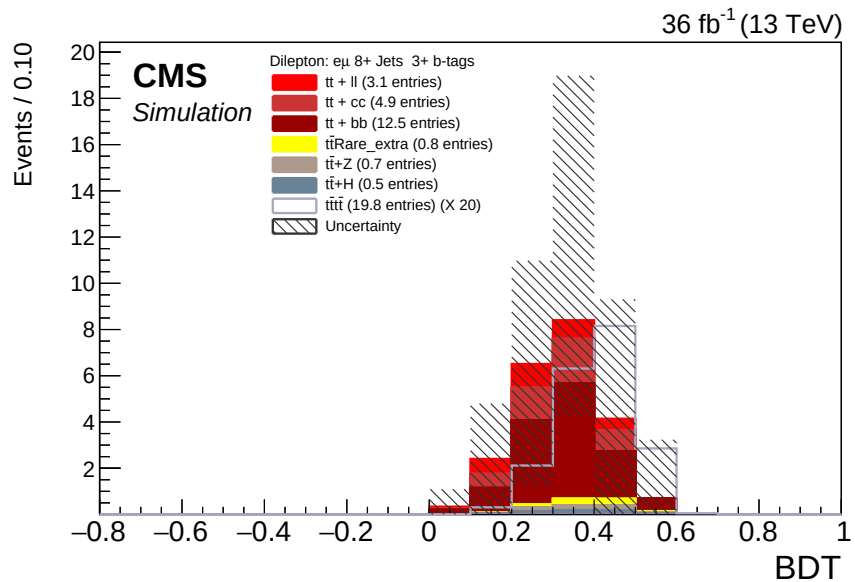
Theoretical Uncertainties

- ME scale
- MC cross sections
- UE tune
- PS scale
- ME-PS matching
- PDF

Fit Strategy

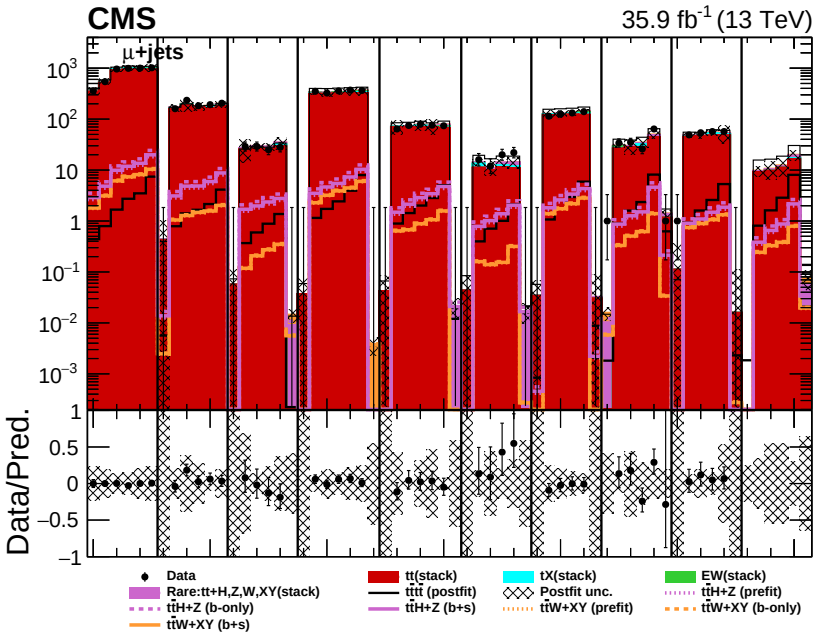
- Likelihood fit using Combine Tool
- Event level BDT output discriminator distributions for fit is performed simultaneously in different $N_j \otimes N_{tags}^M$ categories.
- Blind highest jet/tag multiplicity categories.
  - single lepton: blind 10+ jets & 3+ tags category
  - OS dilepton: blind 8+ jets & 3+ tags category
- Combine results from single lepton channel and OS dilepton channel.

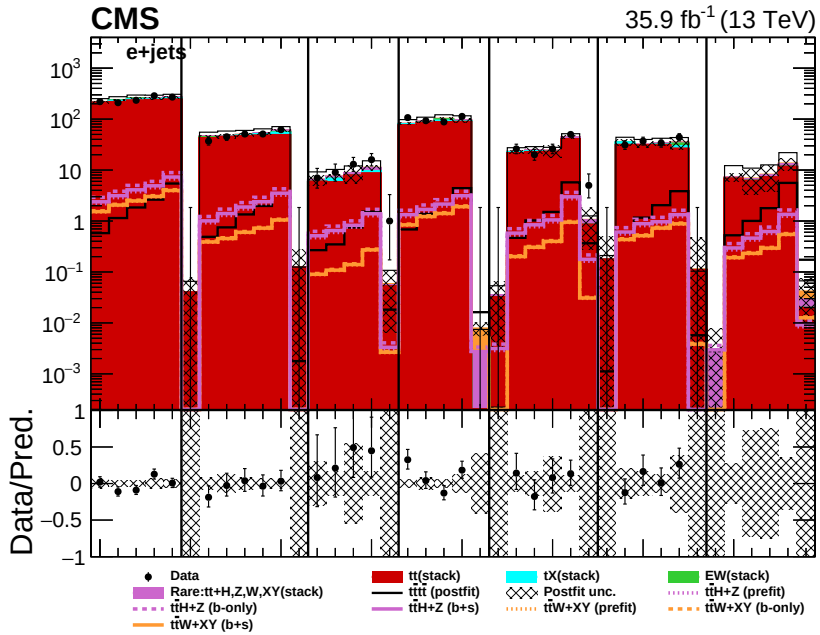
- Large uncertainties on MC in the signal regions.



Template Fit in Single Lepton Channel

Postfit BDT distributions





- Equiprobable binning scheme
- Blind signal rich 10+/3+ category
- Reasonable description of the data in CRs

Template Fit in Single Lepton Channel

Table: Single lepton blinded fitting results

Channel	Expected limit $\times\sigma^{SM}_{t\bar{t}t\bar{t}}$	Expected xsec fb	Expected significance
e	$23.5^{+7.0}_{-6.5}$	216.2^{+64}_{-60}	0.09
μ	$16.0^{+7.0}_{-4.7}$	147.2^{+64}_{-43}	0.12
combined	$9.4^{+4.0}_{-2.7}$	86.5^{+37}_{-25}	0.25

Fit Diagnostic in Single Lepton Channel

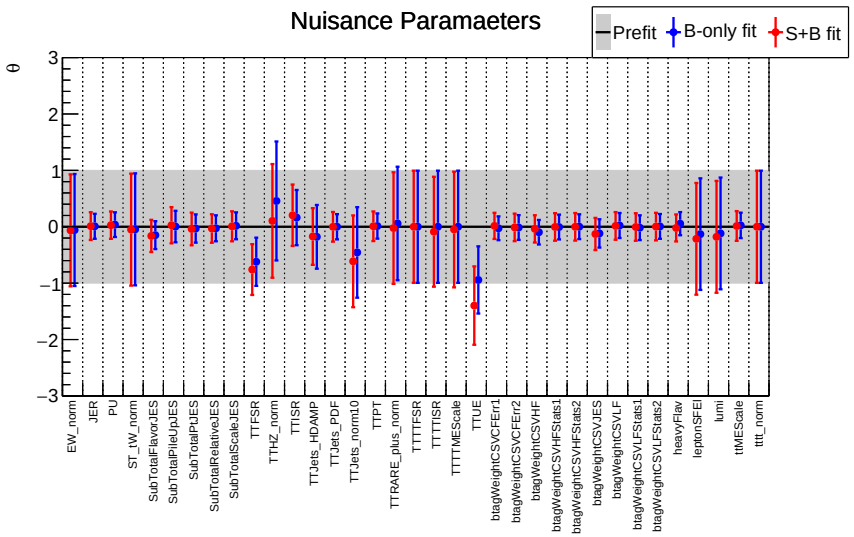


Figure: μ+jets channel

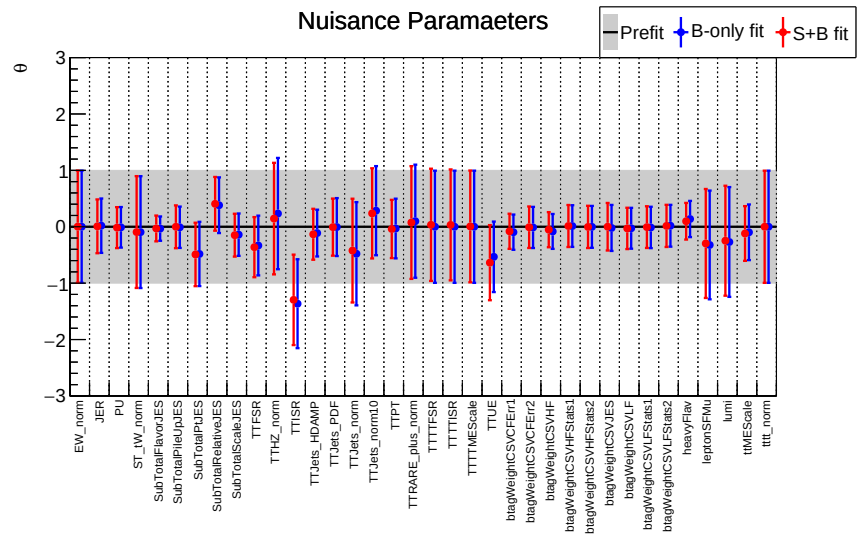
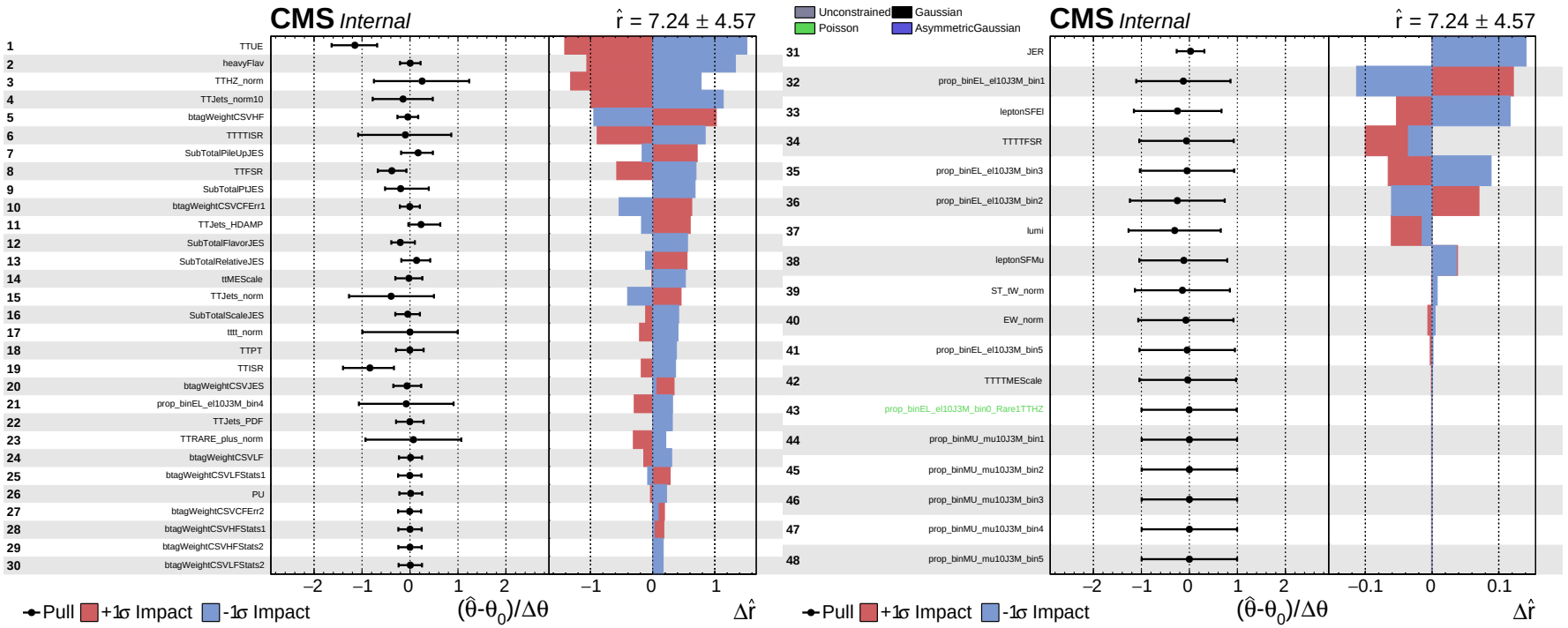


Figure: e+jets channel

- Post-fit uncertainty reduction is under investigation

Impact of Nuisance Parameters



- Dominant sources of systematic uncertainty:
  - UE variation. Affects jet multiplicity spectrum (Sample has low statistics)
  - $t\bar{t}b\bar{b}$ normalization
  - Normalization of $t\bar{t}Z, H \to b\bar{b}$
  - Reweighting of HF component in CSV discriminant

Postfit BDT distributions

- high BDT region in signal sensitive regions is blinded
- Obvious excess in the distribution tails within all three channels, could be due to lack of MC stats

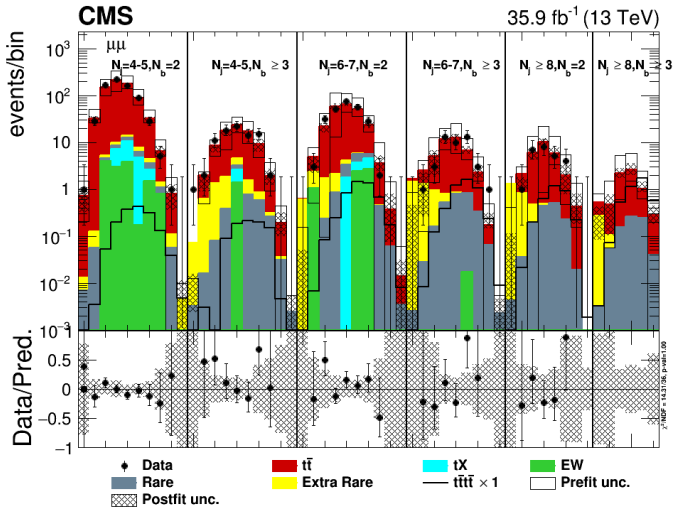


Figure: $\mu\mu$ channel

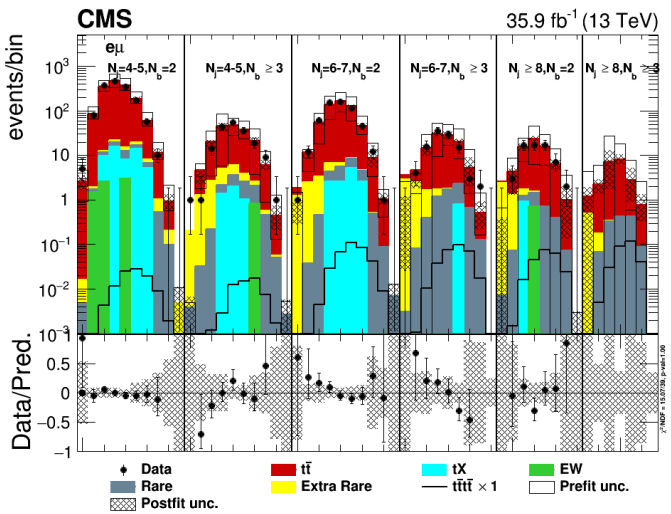


Figure: $e\mu$ channel

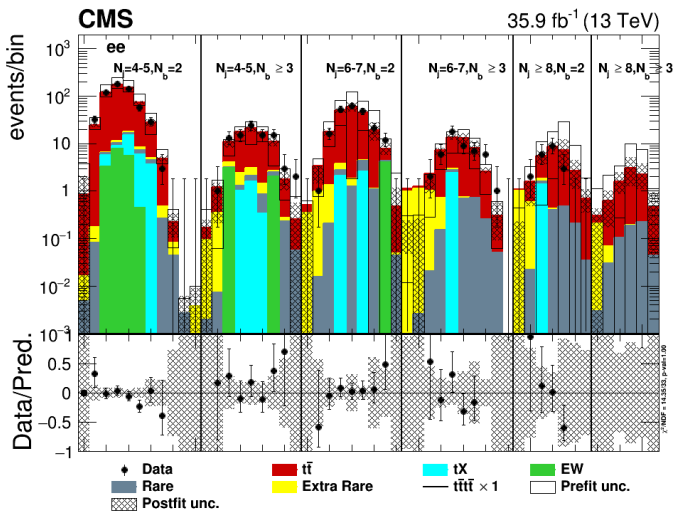


Figure: ee channel

Template Fit in OS Dilepton Channel

Table: OS dilepton blinded fitting results

Channel	Expected limit $\times\sigma^{SM}_{t\bar{t}t\bar{t}}$	Expected xsec *fb*	Expected significance
$\mu\mu$	$14.56^{+9.64}_{-5.24}$	134^{+89}_{-48}	0.19
$e\mu$	$9.88^{+6.53}_{-3.53}$	91^{+60}_{-32}	0.37
ee	$17.56^{+11.34}_{-6.19}$	162^{+104}_{-57}	0.29
combined	$6.88^{+4.44}_{-2.42}$	63^{+41}_{-22}	0.52

Fit Diagnostic in OS Dilepton Channel

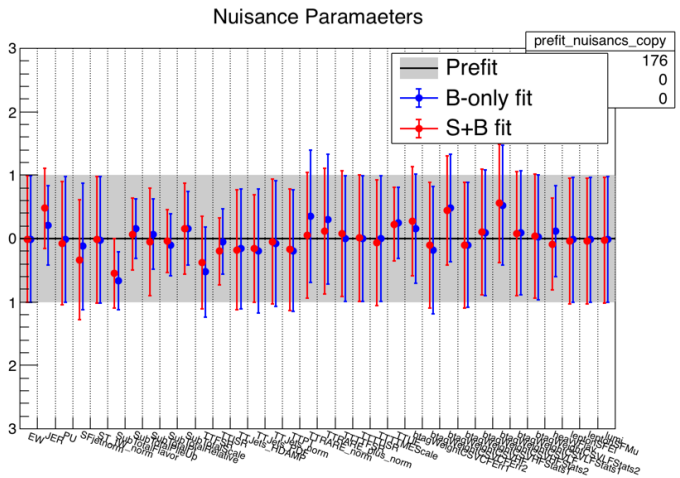


Figure: nuisance pulls in $\mu\mu$

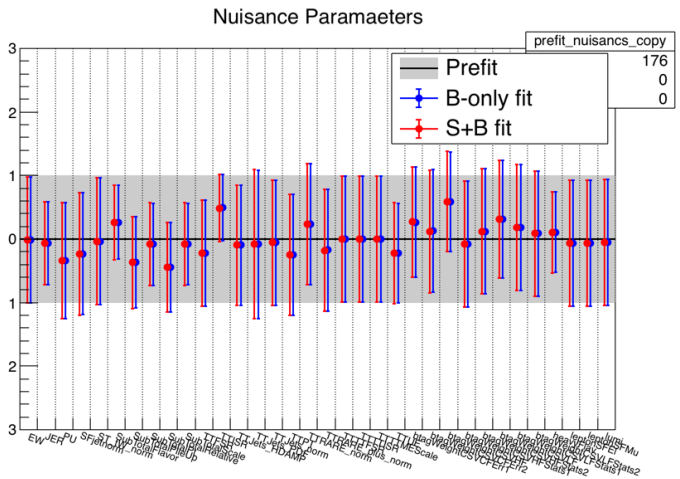


Figure: nuisance pulls in $e\mu$

- Most signal sensitive region is blinded
- No extreme pulls or constraints.
- Reasonable behavior for all the NPs
- Results are consistent between three sub-channels

Fit Diagnostic in OS Dilepton Channel

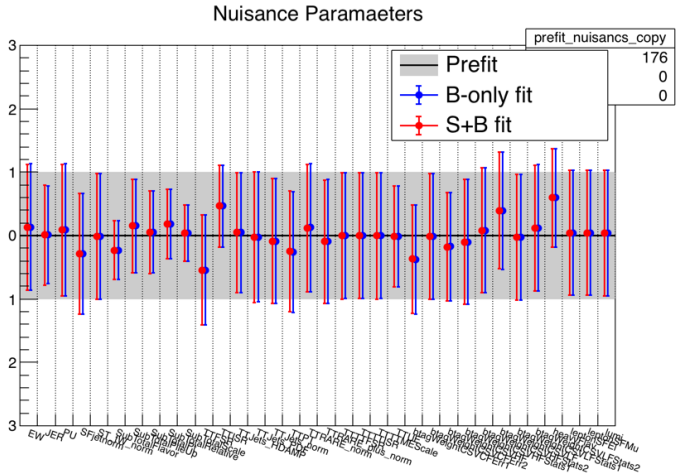


Figure: nuisance pulls in *ee*

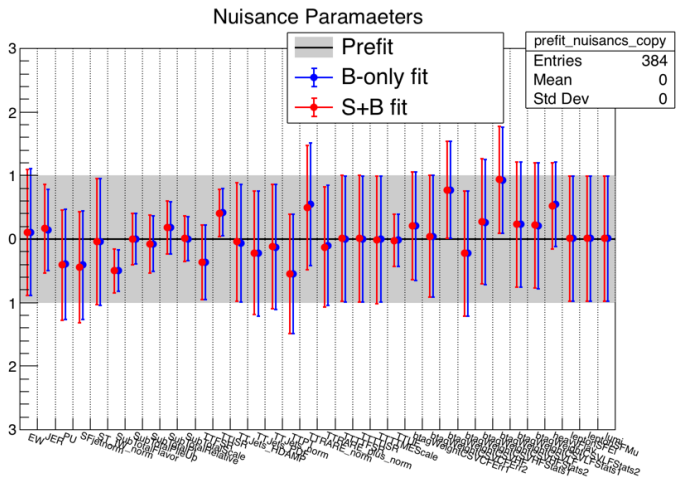


Figure: nuisance pulls in OS dilep

- Mainly the sub-components of JES are constrained
  - Statistic fluctuation
  - Correlated with other NPs

Nuisance Parameters impact in OS dilep combined fit

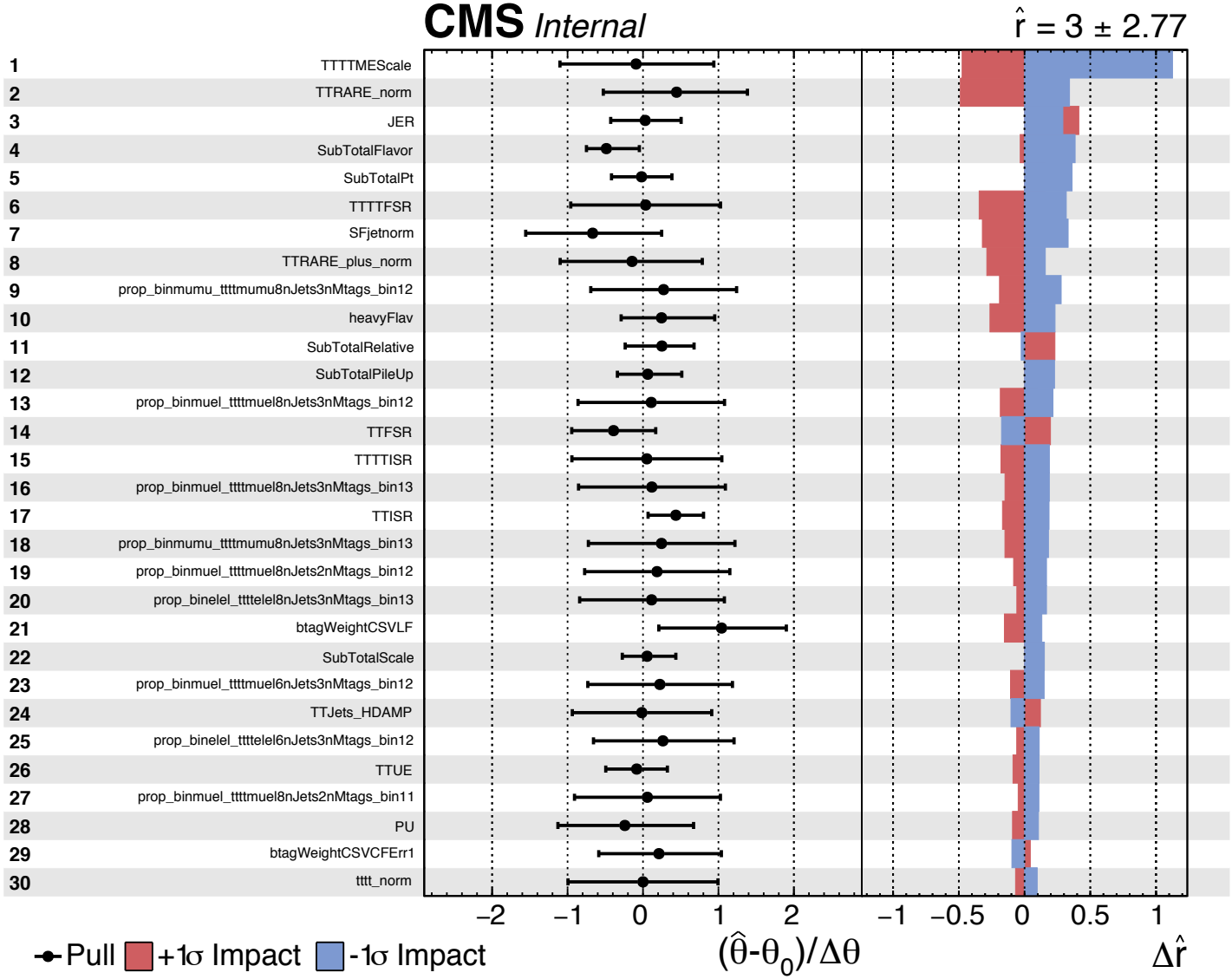


Figure: Impact of nuisance parameters on the parameter of interest

Nuisance Parameters impact in OS dilep combined fit

- Signal systematics have the largest impacts, as we are very close to the expected signal strength.
- TTRare resembles our signal, but under control.
- Jet energy scale uncertainties and MC stats in signal enriched bins dominate.
- All nuisance parameters behave reasonably.
- Full list of nuisance impacts are on backup slides 31 to 32

Combined Results

Table: Single lepton + OS dilepton blinded combined results

Channel	Expected limit $\times\sigma^{SM}_{t\bar{t}t\bar{t}}$	Expected limit *fb*	Expected significance
l+jets	$9.4^{+4.0}_{-2.7}$	86.5^{+37}_{-25}	0.25
OS ll+jets	$6.9^{+4.4}_{-2.4}$	63^{+41}_{-22}	0.52
combined	$5.2^{+2.6}_{-1.7}$	48^{+24}_{-16}	0.58

- Signal sensitivity is driven by the OS *ll*+jets channel

Conclusion

Current status:

- Analysis re-blinded in order to optimize search regions/binning taking into account limited MC statistics
- Expect $\times 10$ more $t\bar{t}$ simulated events using dedicated filters in signal rich search regions that have large statistical uncertainty
  - New high jet multiplicity, high-H_T filter was optimized for the analysis
  - MC requests are added to the system

Plans for coming weeks:

- Include new MC samples when they are ready
- Finalize QCD background estimation
- Update ANs and paper draft

Backups

MC samples

Table: Simulated event samples, cross sections and event yields used in this analysis

Sample	σ (pb)	No of events
/TT_TuneCUETP8M2T4_13TeV-powheg-pythia8(& backup)	831.8	154652276
/TT_TuneCUETP8M2T4_mtop1665_13TeV-powheg-pythia8	831.8	19379727
/TT_TuneCUETP8M2T4_mtop1695_13TeV-powheg-pythia8	831.8	58540996
/TT_TuneCUETP8M2T4_mtop1715_13TeV-powheg-pythia8	831.8	19578294
/TT_TuneCUETP8M2T4_mtop1735_13TeV-powheg-pythia8	831.8	19323035
/TT_TuneCUETP8M2T4_mtop1755_13TeV-powheg-pythia8	831.8	59252096
/TT_TuneCUETP8M2T4_mtop1785_13TeV-powheg-pythia8	831.8	16376678
/TT_widthx0p2_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.8	19937607
/TT_widthx0p5_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.8	19909415
/TT_widthx0p8_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.8	18546839
/TT_widthx2_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.8	13505281
/TT_widthx4_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.8	19593910
/TT_widthx8_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.8	19524579
total		436M

Current status

- Existing inclusive $t\bar{t}$ sample, including mass and width variations: 436M events
  - Nevertheless, MC statistics is scarce
    - 200 MC events in 10j4b l+jets category
    - 140 MC events in 8j4j l^+l^-+jets category
  - Need factor of 10 increase in statistics to ensure fit stability in the sensitive high multiplicity/high discriminant region
  - This requires the development of a gen level filter for efficient MC production
- Various filter configurations have been studied and we have a preferred configuration
- MC production can be started once new release is available
- Backport and genfragment are available (Thanks Javier Fernandez!)
- CMSSW modification are signed by generator conveners (Thanks L.Perrozzi!)
- Currently waiting for new cmssw build containing the backport
  - New patch release has to be issued by release managers

Two scenarios for new filtered samples

- Two independent samples with different lepton and jet multiplicity cuts:
  - Optimal for the two channels
  - Uses gen level cuts which are probably not useful for other analyses
- Common sample with no fully hadronic $t\bar{t}$ decays
  - Potentially useful for other analyses
  - Worse filter efficiency

- Filter efficiency and acceptance loss were calculated on /MINIAODSIM level
- Efficiency values were verified using actual filter implementation in small private /GEN-SIM production

Filters summary (l+jets and OS)

- Preferred configuration: $HT > 500$, $\mathrm{nJets} + \mathrm{nLep} \geq 8$, $\mathrm{nLep} \geq 1$ (26M)

		Acceptance loss in different jet multiplicity regions[1]				
Filter cuts	Filter eff. $\times 10^{-2}$	OS $N_J = 7$	OS $N_J \geq 8$	SL $N_J = 9$	SL $N_J > 9$	Ext ($\times 10$)
HT>500 nJets+nLep≥6 ≥ 1 lepton	2.3 ± 0.03	—	—	—	—	100 M
HT>500 nJets+nLep≥7 ≥ 1 lepton	1.4 ± 0.02	0.04	0.02	—	—	61 M
HT>500 nJets+nLep≥8 ≥ 1 lepton	0.6 ± 0.02	0.14	0.04	0.11	0.08	26 M
HT>500 nJets+nLep≥9 ≥ 1 lepton	0.2 ± 0.01	0.43	0.14	0.19	0.1	9M
HT>500 nJets+nLep≥10 ≥ 1 lepton	0.08 ± 0.007	—	—	—	0.19	3.5M

[1] Fraction of events passing offline cuts but rejected by gen filter

Filter optimized for l+jets channel

- Preferred configuration: $HT > 500$, $\mathrm{nJets} + \mathrm{nLep} \geq 9$, $\mathrm{nLep} = 1$ (9M)

		Acceptance loss in different jet multiplicity regions[2]		
Filter cuts	Filter eff.	SL ($N_J^{rec} = 9$)	SL ($N_J^{rec} > 9$)	Ext ($\times 10$)
HT>500 nJets+nLep $\geq$8 1 lepton	0.005 ± 0.0002	0.11	0.08	21.8 M
HT > 500 nJets+nLep $\geq$ 9 1 lepton	0.002 ± 0.0001	0.19	0.10	8.7M
HT>500 nJets+nLep>=10 1 lepton	$0.0007 \pm 6.5 \times 10^{-5}$	—	0.19	3M

[2] Fraction of events passing offline cuts but rejected by gen filter

Filter optimized for OS dilpeton channel

- Preferred configuration: $HT > 500$, $\mathrm{nJets} + \mathrm{nLep} \geq 7$, $\mathrm{nLep} = 2$ (9M)

		Acceptance loss in different jet multiplicity regions[3]		
Filter cuts	Filter eff.	OS ($N_J^{rec} = 7$)	OS ($N_J^{rec} \geq 8$)	Ext ($\times 10$)
HT>500 nJets+nLep>=5 2 lepton	$0.0046 \pm 5.5e-05$	0.21	0.23	20 M
HT>500 nJets+nLep>=6 2 lepton	$0.0033 \pm 4.7e-05$	0.21	0.23	14.4 M
HT>500 nJets+nLep>=7 2 lepton	$0.0020 \pm 3.6e-05$	0.23	0.24	8.7 M
HT>500 nJets+nLep>=8 2 lepton	$0.0009 \pm 2.4e-05$	0.31	0.25	3.9 M

[3] Fraction of events passing offline cuts but rejected by gen filter

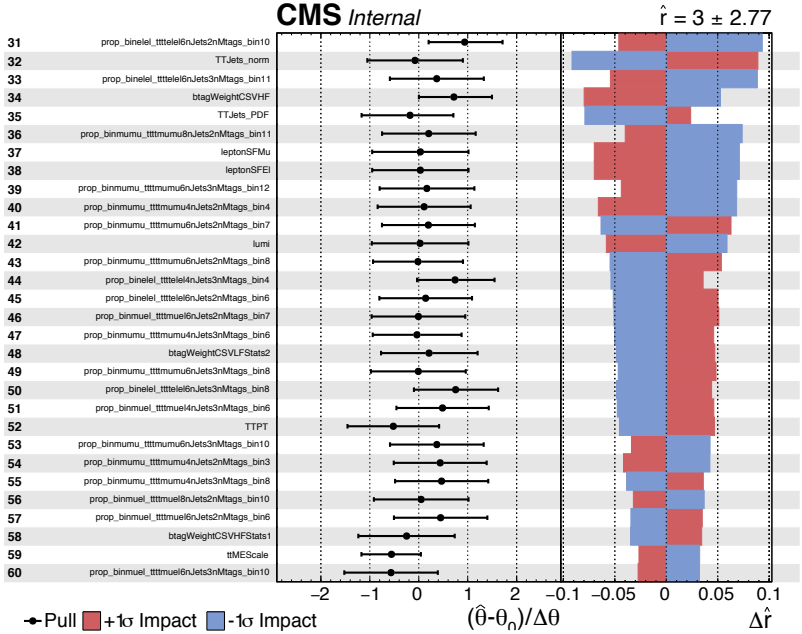
Summary

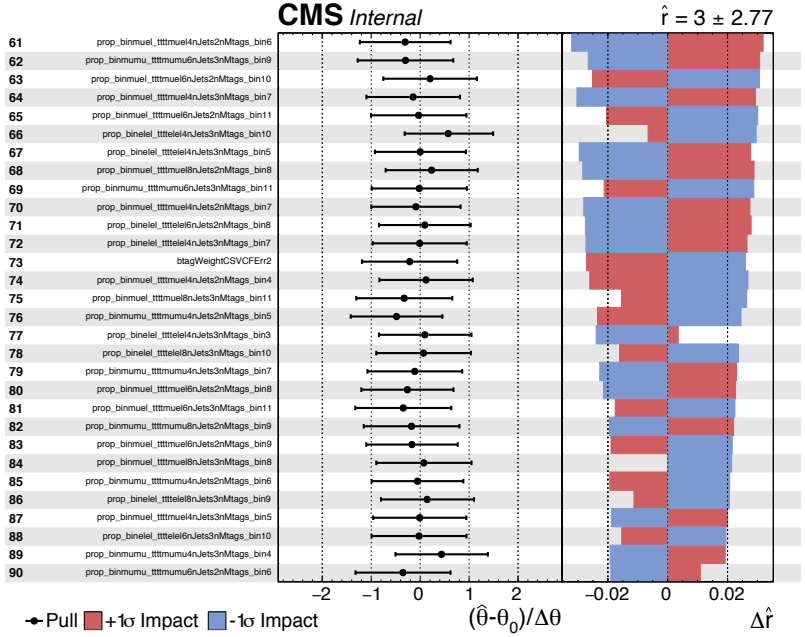
- There are two options to get 10× MC stat extension:
  - (Preferred) Two samples with different gen level cuts
    (Total of 18M events passing the filters):
    - OS **9M events**: $HT > 500, nJets + nLep \geq 7, nLep = 2$
    - SL **9M events**: $HT > 500, nJets + nLep \geq 9, nLep = 1$
  - Combined sample
    (Total of 26M events passing the combined filter):
    - SL+OS: $HT > 500, nJets + nLep \geq 8, nLep \geq 1$

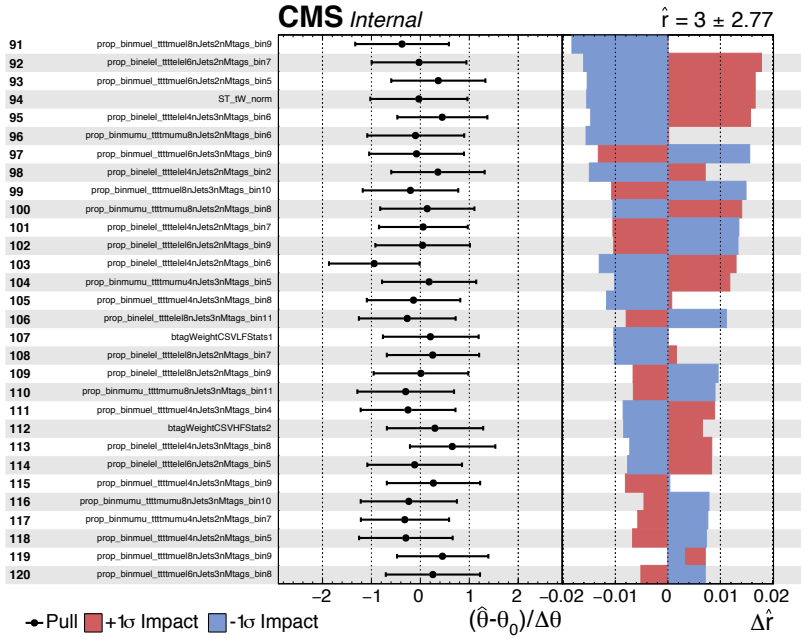
Definitions

- Variable definitions:
  - $HT = \sum_{j:p_T>30,|\eta|<2.4} p_T(j)$
  - nJets= number of gen jets with $p_T > 30$
    No $\eta(j)$ cut due to different CMSSW filter logic.
  - Jets in HT and multiplicity definition may include jets from isolated leptons
- Lepton cuts are applied using LHE filter, similar to NoFullyHadronicDecays filter applied to single-top MC samples

impacts of nuisance parameters







impacts of nuisance parameters