Available on the CMS information server **CMS AN-18-062**

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2019/07/01
Archive Hash: 97d67b8-D
Archive Date: 2019/06/25

Search for SM tttt in the same-sign dilepton and multi-lepton final states at $\sqrt{s} = 13$ TeV with the full Run 2 dataset

L. Bauerdick[1], K. Burkett[1], O. Gutsche[1], S. Jindariani[1], H. Weber[1], M. Derdzinski[2], B. Hashemi[2], D. Gilbert[2], D. Klein[2], S. May[2], D. Olivito[2], I. Suarez[2], F. Würthwein[2], A. Yagil[2], G. Zevi Della Porta[2], N. Amin[3], C. Campagnari[3], B. Marsh[3], S. Wang[3], C. Fangmeier[4], and F. Golf[4]

[1] Fermi National Accelerator Laboratory, Batavia, IL, USA
[2] University of California, San Diego, CA, USA
[3] University of California, Santa Barbara, CA, USA
[4] University of Nebraska, Lincoln, NE, USA

Abstract

This is the AN supporting the Run 2 tttt analysis. It is based on the 2016 tttt AN (AN-17-115) and the 2016 same-sign AN (AN-16-386).

Contents

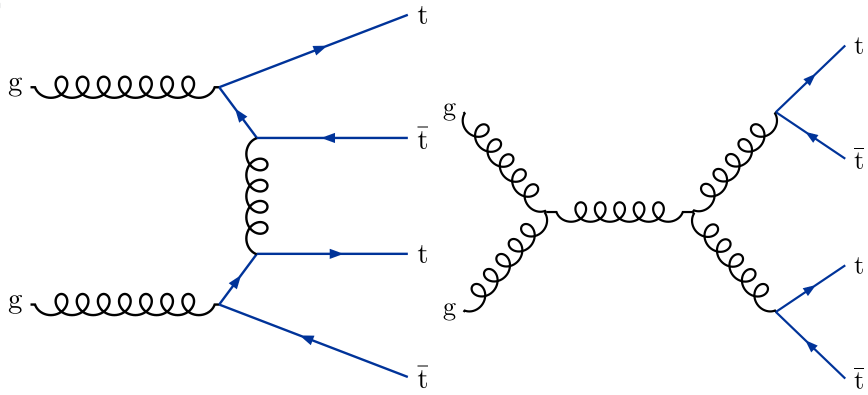
1 Introduction

The same-sign dilepton final state is often used in searches for new physics due to its ability to stifle all but some rare standard model backgrounds. One of these rare backgrounds is the production of four top quarks, $pp \rightarrow \mathrm{t\bar{t}t\bar{t}}$, primarily through the two processes depicted in Fig. 1. The lepton multiplicity composition in $\mathrm{t\bar{t}t\bar{t}}$ events is shown in Fig. 2. With four W bosons, the same-sign dilepton and tri-lepton final state, considering electrons and muons (and those from arising from leptonic tau decays), occurs in approximately 11.9% of $\mathrm{t\bar{t}t\bar{t}}$ events.

This Analysis Note describes a $\mathrm{t\bar{t}t\bar{t}}$ analysis using the full Run 2 data, which is based on two analyses of the 2016 data: the SUSY same-sign dilepton analysis and the corresponding $\mathrm{t\bar{t}t\bar{t}}$ analysis. The SUSY same-sign dilepton analysis used 35.9 fb^{-1} of 2016 collision data, is documented in [1], and published in [2]. The observed (expected) 95% confidence limit on cross section of four top quark production was found to be 42 (27^{+13}_{-8}) fb, and the cross section of the process calculated at NLO is $9.2^{+2.9}_{-2.4}$ fb [3].

The 2016 $\mathrm{t\bar{t}t\bar{t}}$ analysis, based on the same dataset as the SUSY search, was developed by modifying the selection of the SUSY analysis to take advantage of the jet and b-jet multiplicity associated with four top quark decays, as well as creating a separate category of events with 3 or more leptons. The resulting 2016 $\mathrm{t\bar{t}t\bar{t}}$ analysis is document in [4], and [5]. The observed (expected) 95% confidence limit on cross section of four top quark production was found to be 41.7 ($20.8^{+11.2}_{-6.9}$) fb. The expected significance, based on an NLO cross section for tttt of $9.2^{+2.9}_{-2.4}$ fb [3], was 1.05 standard deviations, corresponding to a measured signal strength parameter of $1.0^{+1.2}_{-0.94}$. The observed significance was 1.56 standard deviations (p-value of 6%), corresponding to a measured signal strength parameter of $1.83^{+1.42}_{-1.23}$. The results were also used to constrain the top quark Yukawa coupling of the Higgs boson, based on an LO*k-factor cross section of $12.2^{+5.0}_{-4.4}$ fb and its dependence on $|y_T/y_T^{SM}|$. The central (upper,lower) value of the theoretical cross section band, combined with our 95% observed limit on tttt production, resulted in a 95% CL limit $|y_T/y_T^{SM}| < 2.27$ (2.03,2.56). Additionally, an extension to the interpetation of heavy (pseudo)scalar bosons in the context of 2HDM in the manner of the SUSY analysis was studied, but it was not included in the tttt publication due to the very limited improvements with respect to the SUSY analysis, as discussed in Section 9 of Ref. [4].

All details of the 2016 analysis can be found in the SUSY AN [1] (Fake and Flip backgrounds) and the TOP AN [4] (everything else).



(a) a (b) b

Figure 1: Leading order diagrams for $\mathrm{t\bar{t}t\bar{t}}$ production.

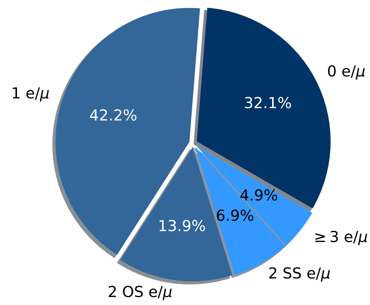


Figure 2: Lepton multiplicity for decay of 4 W bosons.

With respect of the 2016 tttt analysis, the Run2 analysis includes the following changes and improvements:

- Increased luminosity of $35.922 + 41.53 + 58.83 = 136.3\ \mathrm{fb}^{-1}$
- Update the cut-and-count signal regions from 8 to 18 bins in N(lep)/N(jets)/N(b-jets)
- Note that this was the case for ANv7 and previous versions Update the correction to ttW+bb and ttZ+bb based on the Data/MC discrepancy in the ttbb/ttjj ratio from 1.7 ± 0.6 (TOP-16-010) to 1.0 ± 0.35. This is based on the preliminary results of TOP-17-021, which finds a ttbb/ttjj ratio of 1.0 ± 0.35.
- This is the current prescription for ANv8 and on Maintain the correction to ttW+bb and ttZ+bb based on the Data/MC discrepancy in the ttbb/ttjj ratio of 1.7 ± 0.6 (TOP-16-010).
- Update $\mathrm{t\bar{t}t\bar{t}}$ cross section to the latest and most precise estimate, from 9.2 fb [3] to 11.97 fb [6].
- Reduced normalization uncertainty on $\mathrm{t\bar{t}H}$ from 50% to 25% motivated by HIG-17-035.
- Studied two BDTs, one with a long list and one with a short list of variables. See Section 6.2.
- Optimized hyperparameter choice for BDT. See Section 6.3.

Upon looking at the 2017 and 2018 datasets, several more changes were adopted in order to deal with data-taking issues:

- Since the dilepton+HT triggers were not active for the first period of 2017 data taking (RunB, corresponding to 12% of the data), we decided to use the dilepton triggers throughout the 2017 run. The difference resulting from this change is expected to be at the few % level, although we will only know the exact number after all Trigger Scale Factors will be computed.
- Since an increase in non-prompt leptons was observed in our control regions, we tightened the isolation requirements with respect to 2016, as described in Section 4.3.

Issues relating to the missing HEM15/16 sectors starting in Run 319077 of 2018 data-collection were found to impact the fake background only negligibly, and no action was taken to specifically deal with this as the estimation method is data-driven. Additionally, the L1 prefiring issue affecting the efficiency of EGamma objects at high $|\eta|$ was found to have a small effect (%-level) which can be accounted for by a scale factor applied to simulation. See appendix D for checks, and also see the 2LSS+multi-lepton SUSY analysis presentation at the SUSY Leptonic Meeting (slides 21-25 for HEM and slides 26-28 for L1 prefiring at `https://indico.cern.ch/event/770079/`).

Several changes were considered, but it was decided not to implement them, as motivated below:

- Many additional variables were considered for the BDT, including resolved and boosted top-tagging variables, but they were not found to bring significant ($> 2\%$) gain. See Section 6.2.
- Separate BDT trainings for different background categories were considered, but since all backgrounds were found to have a similar BDT shape, it was decided to continue with a single training based on the sum of all backgrounds (slide 14 from Nick at `https://indico.cern.ch/event/709496/`).
- The LeptonMVA selection (combining ID, isolation and nearby jet information) was studied, but ultimately not adopted. Due to the different kinematics, it was found to increase ttV backgrounds by more than the tttt signal, resulting in no gain in significance. See Nick's presentation at the Top Cross Section meeting (slides 4-8 at `https://indico.cern.ch/event/709496/`).
- Jet and b-jet p_{T} thresholds were studied in the range of 20-40 GeV. Based on expected significance, it was decided to stay with 40 GeV threshold for jets, and with 25 GeV threshold for b-jets. See Appendix C.
- Including events with tau leptons. See Appendix C.
- Events with a 3rd lepton, with $10 < p_{\mathrm{T}} < 20$ GeV and not forming a Z mass, were previously included in the 2-lepton signal regions. We considered removing these events, or moving them to a separate categories, but found that both treatments actually worsened (by a few %) the significance, as a non-negligible fraction of $\mathrm{t\bar{t}t\bar{t}}$ events falls into this category.

An extensive changelog with respect to ANv7 can be found in Appendix E.

Changelog with respect to ANv9

- Incorporation of lepton scale factors for 2017, 2018 (muons are incomplete for 2017, 2018, and electrons are incomplete for 2018)
- Switched to preliminary Autumn18V3 JECs applied to 2018 data, MC (with no residuals for data)
- Added Appendix F and Appendix G showing the unblinding of 2016 and 2017 datasets, respectively.
- For convenience, significance numbers will also be tracked in `https://docs.google.com/spreadsheets/d/140yqrUwEsJtOJ8OmDdOun-J8iUrV49aAdUkvfNjpn8k/edit#gid=0`

Changelog with respect to ANv10

- Updated to Autumn18V8 JECs applied to 2018 data, MC (includes residuals for 2018, split into eras A, B, C, and D)
- Updated to latest recommended b-tag SFs (V1) for 2018 MC
- Retrained the BDT with final samples and these updated corrections
- Added Appendix H which shows the results of unblinding 2018 and the full Run2 dataset, with a retrained BDT

- Updated documentation to reflect the full unblinding
- For convenience, significance numbers are tracked in `https://docs.google.com/spreadsheets/d/140yqrUwEsJtOJ8OmDdOun-J8iUrV49aAdUkvfNjpn8k/edit#gid=0`

Changelog with respect to ANv12

- Included final lepton SFs (electrons, muons), finalized 2018 muon scale factors
- Included 50ipb from new re-reco golden JSON in 2018
- Updated to latest luminosity recommendation for 2018 (59.6 fb^{-1}), with a total of 137.2 fb^{-1}
- Switched to final luminosity uncertainty of 2.5% for 2018
- Included latest 2018 samples for heavy (pseudo)scalar boson interpretation

2 Samples

2.1 Collision data

The full Run 2 analysis uses an amount of pp collision data corresponding to an integrated luminosity of 35.9 fb^{-1}, 41.5 fb^{-1}, and 59.6 fb^{-1}, for 2016, 2017, and 2018, totaling 137.2 fb^{-1}.

The following datasets are used:

- DoubleEG/EGamma: two-electron channel,
- DoubleMuon: two-muon channel,
- MuonEG: crossed channel targeting events with one muon and one electron.
- JetHT: used to recover efficiency for dilepton+H_{T} triggers in period 2016H (although no additional events are found in this sample).

The Reco campaigns used in this analysis are listed in table 1, together with the corresponding JSON file. Events present in multiple datasets are properly handled.

Table 1

Year	campaign	JSON file
2016	23Sep2016-v1	Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16
2017	31March2018-v1	Cert_294927-306462_13TeV_EOY2017ReReco_Collisions17
2018	17Sep2018-v1(2018A,B,C)-PromptReco(2018D)	Cert_314472-325175_13TeV_PromptReco_Collisions18_JSON.txt

In order to remove detector noise and unphysical events such beam-halo events, the recommended JetMET filters [7] have been applied.

2.2 Monte Carlo Simulation

The simulation samples used in this analysis have been produced in the Summer16 campaign for comparing to 2016 data, in the RunIIFall17MiniAOD-94X campaign for comparing to 2017 data, and in the RunIIAutumn18MiniAOD-102X campaign for 2018 data. The samples, almost exclusively at NLO, are listed in Table 2.

To improve on the MadGraph5_aMC@NLO modeling of the multiplicity of additional jets from initial-state radiation (ISR) and final-state radiation (FSR), $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ MC events are reweighted based on the number of ISR or FSR jets ($N_J^{\mathrm{ISR/FSR}}$). The reweighing is based on the Data/MC ratio in the light-flavor jet multiplicity in dilepton $\mathrm{t\bar{t}}$ events (using MADGRAPH5_aMC@NLO MC), as shown in Figure 3 and described in Ref. [8] for 2016 Data and MC. The method requires exactly two b-tagged jets, and assumes that all other jets are ISR or FSR. The reweighting factors vary between 0.86 and 0.77 for $N_J^{\mathrm{ISR/FSR}}$ between 1 and 4 (where dilepton ttV plus 4 ISR/FSR jets results in N_{jets} =8, corresponding to our highest SR bin). For 2017 MC, new weights were derived since the MC uses a new Pythia tune. The corresponding Data/MC ratios are shown in Figure 4, where the different plots are based on different number of partons simulated at the matrix element. Reweighting factors are applied to $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$, and $\mathrm{t\bar{t}}$ (from inclusive sample reweighting). Following the procedure used in the SUSY group, we take one half of the deviation from unity as the systematic uncertainty in these reweighting factors.

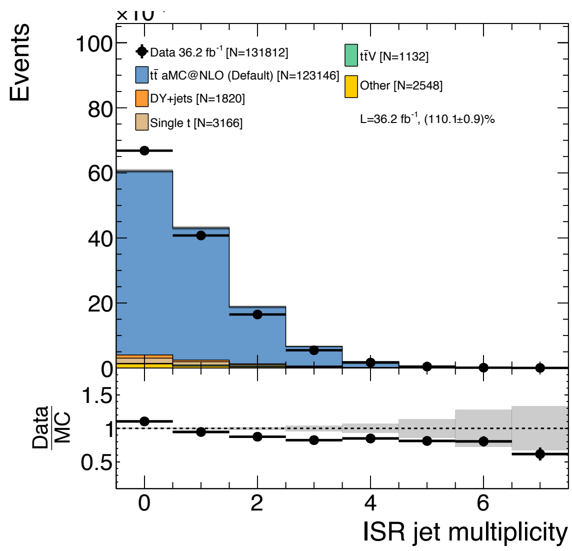


Figure 3: Distribution of the number of light jets (labeled ISR jets in the plot, but refering to ISR and FSR jets) in a dilepton $\mathrm{t\bar{t}}$ sample in 2016 data, compared to 2016 MC.

To improve the modeling of the flavor of additional jets, the $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$ and $\mathrm{t\bar{t}H}$ simulation is also corrected to account for the measured ratio of $\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$ cross sections reported in TOP-16-010. More details on these corrections and their uncertainties are provided in Section 9.

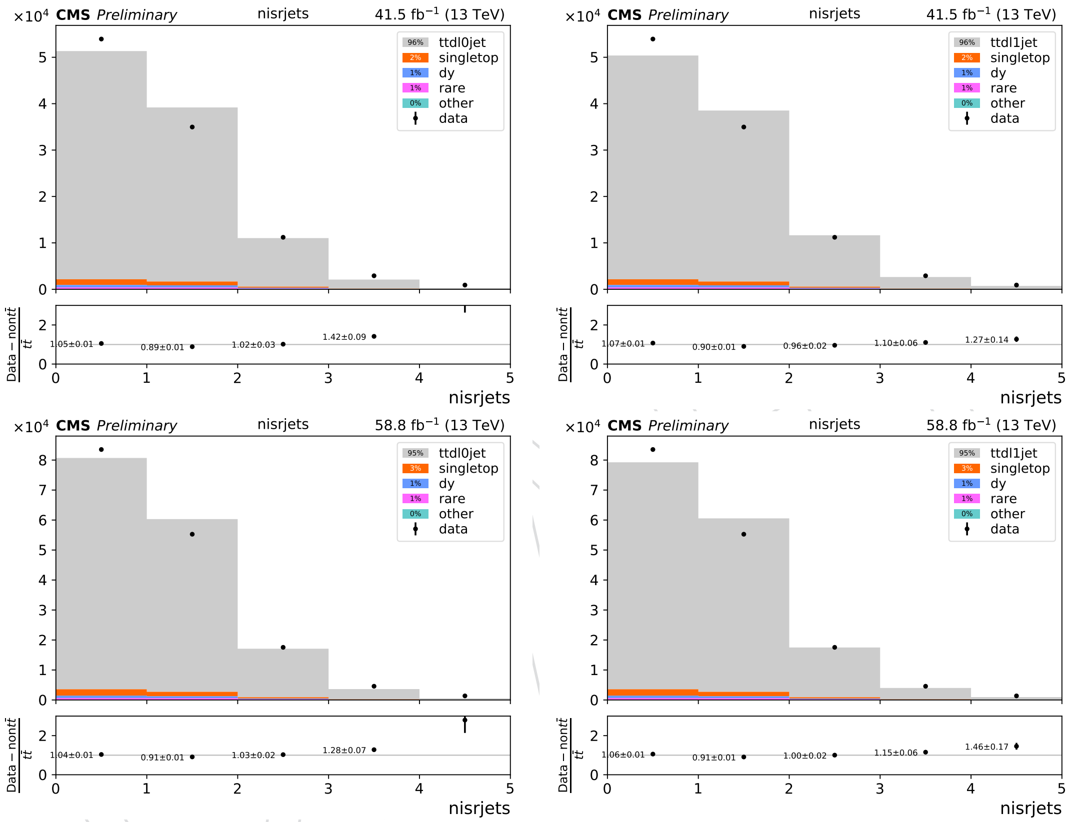


Figure 4: Distribution of the number of light jets in a dilepton $t\bar{t}$ sample with 0 additional partons (left) or 1 additional parton (right) for 2017 data (top) and 2018 data (bottom) compared to their respective MC samples. In the case of 0 additional partons, the reweighing factor in the last bin is taken instead from the previous bin.

Table 2: Signal and background samples from Summer16 and Fall17 campaigns. For each sample, the name has to be completed with the Summer16 string `RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6_v*` for 2016, and with the Fall17 string `RunIIFall17MiniAODv2-PU2017_12Apr2018_94X_mc2017_realist` for 2017. For 2018, the string is `RunIIAutumn18MiniAOD-102X_upgrade2018`. Starting in 2017, the new Pythia tune "TuneCP5" was used instead of the CUETP8M2T4 and CUETP8M1 tunes. The $t\bar{t}t\bar{t}$ signal samples are highlighted in red. The ISR/FSR variation samples for $t\bar{t}t\bar{t}$ are not necessary starting in 2017, since the Parton Shower variation weights are already included in the nominal sample. Remaining missing samples QCD samples are only used as a crosscheck, and the analysis does not depend on them.

sample name	σ (pb)	2016	2017	2018
/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6020.85	X	X	X
/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18610	X	X	X
/WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	61334.9	X	X	X
/TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	831.762	X	X	X
/TTJets_TuneCP5_13TeV-amcatnloFXFX-pythia8	831.762	X	X	X
/TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043	X	X	X
/TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529	X	X	X
/TTZToLL_M-1to10_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.0493	X	X	X
/ttHToNonbb_M125_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	0.2710	X	X	X
/tZq_ll_4f_13TeV-amcatnlo-pythia8_TuneCUETP8M1 (ext1)	0.0758	X	X	X
/TTTT_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	0.01197	-	X	X
/TTTT_TuneCUETP8M2T4_13TeV-amcatnlo-isrup-pythia8	0.01197	X	-	X
/TTTT_TuneCUETP8M2T4_13TeV-amcatnlo-fsrup-pythia8	0.01197	X	-	X
/TTTT_TuneCUETP8M2T4_13TeV-amcatnlo-isrdown-pythia8	0.01197	X	-	X
/TTTT_TuneCUETP8M2T4_13TeV-amcatnlo-fsrdown-pythia8	0.01197	X	-	X
/TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8	2.967	X	X	X
/TTGamma_SingleLeptFromT_TuneCUETP8M2T4_13TeV-amcatnlo-pythia8	0.77	X	X	X
/TTGamma_SingleLeptFromTbar_TuneCUETP8M2T4_13TeV-amcatnlo-pythia8	0.769	X	X	X
/TTGamma_Dilept_TuneCUETP8M2T4_13TeV-amcatnlo-pythia8	0.632	X	X	X
/WGToLNuG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	405.271	X	-	X
/WGToLNuG_TuneCP5_13TeV-madgraphMLM-pythia8	405.271	-	X	-
/ZGTo2LG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	123.9	X	X	X
/WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.4297	X	X	X
/ZZTo4L_13TeV_powheg_pythia8	1.256	X	X	X
/ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.01398	X	X	X
/WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.05565	X	X	X
/WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.1651	X	X	X
/WZG_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.04123	X	X	X
/WWG_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2147	X	X	X
/WWW_4F_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2086	X	X	X
/WWTo2L2Nu_DoubleScattering_13TeV-pythia8	0.16975	X	X	X
/ST_tWll_5f_LO_13TeV-MadGraph-pythia8	0.01123	X	X	X
/WpWpJJ_EWK-QCD_TuneCUETP8M1_13TeV-madgraph-pythia8	0.03711	X	X	X
/GluGluHToZZTo4L_M125_13TeV_powheg2_JHUgenV6_pythia8	0.01181	X	X	X
/VHToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8	0.9561	X	X	X
/TTHH_TuneCUETP8M2T4_13TeV-madgraph-pythia8	0.000757	X	X	X
/TTZH_TuneCUETP8M2T4_13TeV-madgraph-pythia8	0.001535	X	X	X
/TTZZ_TuneCUETP8M2T4_13TeV-madgraph-pythia8	0.001982	X	X	X
/TTWZ_TuneCUETP8M2T4_13TeV-madgraph-pythia8	0.003884	X	X	X
/TTTW_TuneCUETP8M2T4_13TeV-madgraph-pythia8	0.000788	X	X	X
/TTTJ_TuneCUETP8M2T4_13TeV-madgraph-pythia8	0.000474	X	X	X
/TTWH_TuneCUETP8M2T4_13TeV-madgraph-pythia8	0.001582	X	X	X
/TTWW_TuneCUETP8M2T4_13TeV-madgraph-pythia8	0.01150	X	X	X
/QCD_Pt-20toInf_MuEnrichedPt15_TuneCUETP8M1_13TeV_pythia8	720648000 × 0.00042	X	X	TODO
/QCD_Pt-15to20_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	1273190000 × 0.003	X	X	TODO
/QCD_Pt-20to30_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	558528000 × 0.0053	X	X	TODO
/QCD_Pt-30to50_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	139803000 × 0.01182	X	X	TODO
/QCD_Pt-50to80_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	19222500 × 0.02276	X	X	TODO
/QCD_Pt-80to120_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia	2758420 × 0.03844	X	X	TODO
/QCD_Pt-120to170_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	469797 × 0.05362	X	X	TODO
/QCD_Pt-170to300_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	117989 × 0.07335	X	X	TODO
/QCD_Pt-470to600_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	645.528 × 0.12242	X	X	TODO
/QCD_Pt-600to800_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	187.109 × 0.13412	X	X	TODO
/QCD_Pt-1000toInf_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	10.4305 × 0.15544	X	X	TODO
/QCD_Pt-20to30_EMEnriched_TuneCUETP8M1_13TeV_pythia8	557600000 × 0.0096	X	X	TODO
/QCD_Pt-30to50_EMEnriched_TuneCUETP8M1_13TeV_pythia8	136000000 × 0.073	X	X	TODO
/QCD_Pt-50to80_EMEnriched_TuneCUETP8M1_13TeV_pythia8	19800000 × 0.146	X	X	TODO
/QCD_Pt-80to120_EMEnriched_TuneCUETP8M1_13TeV_pythia8	2800000 × 0.125	X	X	TODO
/QCD_Pt-120to170_EMEnriched_TuneCUETP8M1_13TeV_pythia8	477000 × 0.132	X	X	TODO
/QCD_Pt-170to300_EMEnriched_TuneCUETP8M1_13TeV_pythia8	114000 × 0.165	X	X	TODO
/QCD_Pt-300toInf_EMEnriched_TuneCUETP8M1_13TeV_pythia8	9000 × 0.15	X	X	TODO
/QCD_Pt_20to30_bcToE_TuneCUETP8M1_13TeV_pythia8	557627000 × 0.00059	X	X	TODO
/QCD_Pt_30to80_bcToE_TuneCUETP8M1_13TeV_pythia8	159068000 × 0.00255	X	X	TODO
/QCD_Pt_80to170_bcToE_TuneCUETP8M1_13TeV_pythia8	3221000 × 0.01183	X	X	TODO
/QCD_Pt_170to250_bcToE_TuneCUETP8M1_13TeV_pythia8	105771 × 0.02492	X	X	TODO
/QCD_Pt_250toInf_bcToE_TuneCUETP8M1_13TeV_pythia8	21094.1 × 0.03375	X	X	TODO

3 Triggers

The analysis uses two slightly different trigger strategies for 2016 and 2017/2018 data. For 2016 data, the dilepton + H_{T} triggers are used, since they have slightly looser lepton requirements (they do not apply isolation), and are therefore slightly more efficient than the pure dilepton triggers. For 2017, these triggers were not available in RunB, which represent 12% of the 2017 dataset, and therefore the pure dilepton triggers are used for the entirety of 2017 and 2018. We considered also using the pure dilepton triggers for 2016, to further simplify this treatment, but we prefer to avoid unnecessary changes to the published result, wherever possible. The full set of triggers used for selecting events in the signal regions are listed in Table 3.

Table 3: Summary of the signal triggers

2016		
$H_{T,off}$	Channel	Trigger Name
> 300 GeV	$\mu\mu$	HLT_DoubleMu8_Mass8_PFHT300
	ee	HLT_DoubleEle8_CaloIdM_TrackIdM_Mass8_PFHT300
	$e\mu$	HLT_Mu8_Ele8_CaloIdM_TrackIdM_Mass8_PFHT300
2017		
$H_{T,off}$	Channel	Trigger Name
all	$\mu\mu$	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8
	ee	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_
	$e\mu$	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ
2018		
$H_{T,off}$	Channel	Trigger Name
all	$\mu\mu$	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8
	ee	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_
	$e\mu$	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ

An additional set of triggers is developed for a control sample selection for fake rate measurement. They require either a single lepton, or a single lepton and a jet in order to increase a purity of the collected data sample. The full list can be found in Table 4. Two sets of the auxiliary triggers are introduced in order to collect data with different online lepton IDs corresponding to the two sets of signal dilepton triggers.

3.1 2016 trigger efficiency

Computation of trigger efficiencies and scale factors for 2016 data is described in Section 4 of [1]. Summarizing these results, the dilepton+H_{T} trigger efficiency is found to be the product of lepton leg efficiencies (93-98% for electrons, 85-95% for muons), and the HT efficiency (98-99%). As a result the trigger efficiency for events passing the baseline selection in the 2016 data is greater than 95% for ee and eμ events, and about 92% for $\mu\mu$ events.

3.2 2017 and 2018 trigger efficiency

This section is identical to same-sign SUS analysis (AN-18-280), which provides trigger efficiency maps. These are currently being applied for all 3 years.

Table 4: Summary of the control triggers ordered by lepton flavor and p_T.

$H_{T,off}$	Channel	Trigger Name
2016		
> 300 GeV	μ	HLT_Mu8 HLT_Mu17
	e	HLT_Ele8_CaloIdM_TrackIdM_PFJet30 HLT_Ele17_CaloIdM_TrackIdM_PFJet30
2017		
all	μ	HLT_Mu8_TrkIsoVVL HLT_Mu17_TrkIsoVVL
	e	HLT_Ele8_CaloIdL_TrackIdL_IsoVL_PFJet30 HLT_Ele23_CaloIdL_TrackIdL_IsoVL_PFJet30
2018		
all	μ	HLT_Mu8_TrkIsoVVL HLT_Mu17_TrkIsoVVL
	e	HLT_Ele8_CaloIdL_TrackIdL_IsoVL_PFJet30 HLT_Ele23_CaloIdL_TrackIdL_IsoVL_PFJet30

4 Object Definition and Selection

The object selection is based on the one used in the SUSY 2016 analysis documented in [1]. In particular, the same lepton ID is used for same-sign and trilepton events.

The somewhat complicated definitions of isolation and lepton IDs were arrived through dedicated studies within the ”SUSY fake lepton working group” before data taking started in 2015. The goal was to have a fairly efficient lepton selection, with low fake rate, and with a reliable method to estimate nonprompt lepton backgrounds with minimal sample dependence starting from ”fakeable objects” (FO). The studies are documented in Ref. [9].

4.1 Electron identification

The electrons are reconstructed by associating ECAL-clusters and GSF tracks. Only electrons within the tracker and ECAL acceptance are considered: $|\eta| < 2.5$.

The electron identification is performed using a multivariate (MVA) discriminant built with shower-shape variables ($\sigma_{i\eta i\eta}$, $\sigma_{i\phi i\phi}$, the cluster circularity, widths along η and ϕ, R_9, H/E, E_{inES}/E_{raw}), track-cluster matching variables (E_{tot}/p_{in}, E_{Ele}/p_{out}, $\Delta\eta_{in}$, $\Delta\eta_{out}$, $\Delta\phi_{in}$, $1/E - 1/p$) and track quality variables (χ^2 of the KF and GSF tracks, the number of hits used by the KF/GSF filters, fbrem). A complete description of the discriminant and training used can be found in [10–12]. Three identification working points, summarized in the Table 5, are used in this analysis, following the prescriptions of the SUSY lepton scale factors group [13]. Signal events are selected using the tight working point while the loose working points are used in the estimate of background arising from mistakenly identified or non prompt leptons. Separate loose working points are derived for regions where isolated and non isolated dilepton triggers are used.

Different cuts are used in 2016,2017, and 2018 based on the Spring16_GeneralPurpose_V1, Fall17NoIso, and Fall17V2NoIso trainings, respectively. Note that the cuts for 2018 are with respect to the “raw” BDT output (i.e., the values are not forced to lie between -1 and 1 by squashing the output with a sigmoid function). Raw values can be obtained from squashed output via $\text{raw} = \frac{1}{2}\ln\left(\frac{1+\text{squashed}}{1-\text{squashed}}\right)$. The working points depend of the momentum and pseudorapidity of the electrons: for 2016 (resp. 2017/2018) in the region between 15 (resp. 10) and 25 GeV, the working point is given by a linear interpolation between the higher and lower value, following $\min[A, \max[B, A + C * (p_T - 15(10))]]$, where A is the cut at 15(10) GeV, B is the cut at 25 GeV, and $C = (B - A)/10(15)$. Electrons with $p_T < 10$ GeV are only not used to select signal events, so the Tight WP is not assigned for them.

As the electron charge can be determined with three different techniques [14], and as the mismeasurement of the lepton charge can lead to accept more background after the selection, the requirement to get a consistent charge measurement with the three methods is considered in some of the electron definitions used in the analysis. To reject electrons originating from photon conversion, two variables are considered: the number of missing pixel hits and a conversion veto based on the vertex fit probability. Finally, impact parameter variables are also considered: impact parameter in the transverse plane d_0, impact parameter along the z axis d_z, and the impact parameter significance in the detector space SIP_{3D}.

4.2 Muon identification

Two working points are considered for the muon identification. The loose working point follows the ”muon POG Loose ID” described in [15], while the tight working point is given by the list of requirements known as the ”muon Medium Id”, defined in [16]. Only muons within

Table 5: Lower cut on the electron MVA discriminant for the various electron ID used in this analysis.

2016

pseudorapidity region	momentum [GeV]	loose WP (non iso trigger)	loose WP (iso trigger)	tight WP
$0 < \|\eta\| < 0.8$	$5 < p_T < 10$	-0.30	-0.46	N/A
$0 < \|\eta\| < 0.8$	$10 < p_T < 15$	-0.86	-0.48	0.77
$0 < \|\eta\| < 0.8$	$p_T > 25$	-0.96	-0.85	0.52
$0.8 < \|\eta\| < 1.479$	$5 < p_T < 10$	-0.36	-0.03	N/A
$0.8 < \|\eta\| < 1.479$	$10 < p_T < 15$	-0.85	-0.67	0.56
$0.8 < \|\eta\| < 1.479$	$p_T > 25$	-0.96	-0.91	0.11
$1.479 < \|\eta\| < 2.5$	$5 < p_T < 10$	-0.63	0.06	N/A
$1.479 < \|\eta\| < 2.5$	$10 < p_T < 15$	-0.81	-0.49	0.48
$1.479 < \|\eta\| < 2.5$	$p_T > 25$	-0.95	-0.83	-0.01

2017

pseudorapidity region	momentum [GeV]	loose WP (non iso trigger)	loose WP (iso trigger)	tight WP
$0 < \|\eta\| < 0.8$	$5 < p_T < 10$	-0.135	0.488	N/A
$0 < \|\eta\| < 0.8$	$10 < p_T < 25$	$-0.930 + \frac{0.043}{15} \times (p_T\text{-}10)$	$-0.788 + \frac{0.148}{15} \times (p_T\text{-}10)$	$0.2 + 0.032 \times (p_T\text{-}10)$
$0 < \|\eta\| < 0.8$	$p_T > 25$	-0.887	-0.64	0.68
$0.8 < \|\eta\| < 1.479$	$5 < p_T < 10$	-0.417	-0.045	N/A
$0.8 < \|\eta\| < 1.479$	$10 < p_T < 25$	$-0.930 + \frac{0.04}{15} \times (p_T\text{-}10)$	$-0.850 + \frac{0.075}{15} \times (p_T\text{-}10)$	$0.1 + 0.025 \times (p_T\text{-}10)$
$0.8 < \|\eta\| < 1.479$	$p_T > 25$	-0.890	-0.775	0.475
$1.479 < \|\eta\| < 2.5$	$5 < p_T < 10$	-0.470	0.176	N/A
$1.479 < \|\eta\| < 2.5$	$10 < p_T < 25$	$-0.942 + \frac{0.032}{15} \times (p_T\text{-}10)$	$-0.810 + \frac{0.077}{15} \times (p_T\text{-}10)$	$-0.1 + 0.028 \times (p_T\text{-}10)$
$1.479 < \|\eta\| < 2.5$	$p_T > 25$	-0.910	-0.733	0.320

2018

pseudorapidity region	momentum [GeV]	loose WP (non iso trigger)	loose WP (iso trigger)	tight WP
$0 < \|\eta\| < 0.8$	$5 < p_T < 10$	0.053	1.320	N/A
$0 < \|\eta\| < 0.8$	$10 < p_T < 25$	$-0.106 + 0.062 \times (p_T\text{-}25)$	$1.204 + 0.066 \times (p_T\text{-}25)$	$4.277 + 0.112 \times (p_T\text{-}25)$
$0 < \|\eta\| < 0.8$	$p_T > 25$	-0.106	1.204	4.277
$0.8 < \|\eta\| < 1.479$	$5 < p_T < 10$	-0.434	0.192	N/A
$0.8 < \|\eta\| < 1.479$	$10 < p_T < 25$	$-0.769 + 0.038 \times (p_T\text{-}25)$	$0.084 + 0.033 \times (p_T\text{-}25)$	$3.152 + 0.060 \times (p_T\text{-}25)$
$0.8 < \|\eta\| < 1.479$	$p_T > 25$	-0.769	0.084	3.152
$1.479 < \|\eta\| < 2.5$	$5 < p_T < 10$	-0.956	0.362	N/A
$1.479 < \|\eta\| < 2.5$	$10 < p_T < 25$	$-1.461 + 0.042 \times (p_T\text{-}25)$	$-0.123 + 0.053 \times (p_T\text{-}25)$	$2.359 + 0.087 \times (p_T\text{-}25)$
$1.479 < \|\eta\| < 2.5$	$p_T > 25$	-1.461	-0.123	2.359

the muon system acceptance $|\eta| < 2.4$ are considered. Impact parameter selection is also applied on muons, and use the variables already defined in the previous section. Contrary to the electrons, only one charge can be reconstructed for a muon track. The quality of the charge reconstruction is then given for the muons by a quality criteria on the track reconstruction : $\delta p_T(\mu)/p_T(\mu) < 0.2$.

4.3 Lepton isolation

The lepton isolation is constructed using three different variables:

- the mini isolation I_{mini}. Requiring I_{mini} below a given threshold ensures that the lepton is locally isolated, even in boosted topologies. The impact of pileup is mitigated using the so-called effective area correction:

$$I_{\text{mini}} = \frac{\sum_R p_T(h^\pm) - \max(0, \sum_R p_T(h^0) + p_T(\gamma) - \rho \mathcal{A} \left(\frac{R}{0.3}\right)^2}{p_T(\ell)}. \quad (1)$$

where ρ is the pileup energy density, where $\sum_R p_T(h^\pm)$, $\sum_R p_T(h^0)$ and $\sum_R p_T(\gamma)$ refers to the sum of the transverse momentum of the charged hadrons, neutral hadrons and photons, respectively, within a cone R, dependent of the lepton p_T:

$$R = \frac{10}{\min(\max(p_T(\ell), 50), 200)} \tag{2}$$

The effective areas $\mathcal{A}$ used are listed in Table 6.

- The second variable is the ratio of the lepton p_T and of the jet matched to the lepton p_T: $p_T^{\text{ratio}} = \frac{p_T(\ell)}{p_T(\text{jet})}$. This jet is matched geometrically to the lepton, and in most of the case is the jet containing the lepton. If no jet is clustering the lepton, then the closest one is chosen. The use of p_T^{ratio} is a simple way to identify leptons in quite boosted topologies, without any jet reclustering. In order to avoid an over-correction on prompt leptons, the application of the jet energy correction is only applied on the hadronic part of the jet, using the following formula jet $= \ell +$ (jet-PU-ℓ) $* JEC - PU$, where ℓ is the lepton, PU the pileup energy clustered into the jet, and JEC the jet energy scale correction to be applied. This formula is applied at the Lorentz vector level. This approach is commonly used in the B2G group and is blessed by the JetMET POG.
- The last variable used is the p_T^{rel} variable:

$$p_T^{\text{rel}} = \frac{|(\vec{p}(\text{jet}) - \vec{p}(\ell)) \times \vec{p}(\ell)|}{|\vec{p}(\text{jet}) - \vec{p}(\ell)|}. \tag{3}$$

This variable allows to recover leptons from accidental with jets. Similarly to p_T^{ratio}, the jet energy scale correction are only applied on the hadronic part of the considered jet.

Using those three variables, a lepton is considered isolated if the following condition is respected:

$$I_{\text{mini}} < I_1 \wedge (p_T^{\text{ratio}} > I_2 \vee p_T^{\text{rel}} > I_3) \tag{4}$$

The $I_i, i = 1, 2, 3$ values depends of the flavor of the lepton: as the probability to misidentify a jet is higher for the electrons, tighter isolation values are used. The loose lepton isolation is significantly relaxed, as well as an extra definition (fakeable) used for the fake lepton background estimation. The different values are summarized in the Table 7. The logic beyond that isolation is a relaxing of the local isolation, compensated by the requirement that the lepton carries the major part of the energy of the corresponding jet, or if not, if the lepton is considered as accidentally overlapping with a jet.

More details and figures about the lepton isolation can be found in the previous version of the analysis described in [17].

4.4 Trigger emulation selection

Due to the differences between online and offline reconstruction and selection, some leptons can be selected offline but not online. Even though we simulate this effect by applying trigger selection in the simulation, we still choose to reduce online/offline differences by applying a set of offline requirements to ensure that electrons will pass the trigger. Effectively this increases the efficiency of our trigger selection, reducing the role of trigger Data/MC scale factors. The requirements are summarized in Table 8.

Table 6: Effective areas for muons and electrons.

2016

Muons		Electrons	
$\lvert\eta\rvert$ range	$\mathcal{A}(\mu)$ neutral	$\lvert\eta\rvert$ range	$\mathcal{A}$(e) neutral
0.0 − 0.8	0.0735	0.0 − 1.0	0.1752
0.8 − 1.3	0.0619	1.0 − 1.479	0.1862
1.3 − 2.0	0.0465	1.479 − 2.0	0.1411
2.0 − 2.2	0.0433	2.0 − 2.2	0.1534
2.2 − 2.5	0.0577	2.2 − 2.3	0.1903
		2.3 − 2.4	0.2243
		2.4 − 2.5	0.2687

2017/2018

Muons		Electrons	
$\lvert\eta\rvert$ range	$\mathcal{A}(\mu)$ neutral	$\lvert\eta\rvert$ range	$\mathcal{A}$(e) neutral
0.0 − 0.8	0.0566	0.0 − 1.0	0.1440
0.8 − 1.3	0.0562	1.0 − 1.479	0.1562
1.3 − 2.0	0.0363	1.479 − 2.0	0.1032
2.0 − 2.2	0.0119	2.0 − 2.2	0.0859
2.2 − 2.5	0.0064	2.2 − 2.3	0.1116
		2.3 − 2.4	0.1321
		2.4 − 2.5	0.1654

Table 7: Isolation working points

2016

isolation value	loose WP (e/μ)	μ (Medium) WP	e (Tight) WP
I_1	0.4	0.16	0.12
I_2	0	0.76	0.80
I_3	0	7.2	7.2

2017/2018

isolation value	loose WP (e/μ)	μ (Medium) WP	e (Tight) WP
I_1	0.4	0.11	0.07
I_2	0	0.74	0.78
I_3	0	6.8	8.0

4.5 Lepton definitions

The loose, tight and fakeable lepton definition, based the different identification and isolation definitions described in the previous sections, are summarized in the Table 9.

4.6 Jets

Jets are reconstructed from particle flow candidates, clustered with the anti-kt algorithm and with a cone size of $\Delta R < 0.4$. Only jets above a transverse momentum $p_T > 40$ GeV and within the tracker acceptance $|\eta| < 2.4$ are considered. To reject noise and mis-measured jets, the selected jets have to fulfill several identification criteria.

For the 2016 data analysis the following criteria, corresponding to the loose JetID selection, are applied:

- neutral hadronic energy fraction < 0.99

Table 8: Electron selection criteria used for the emulation of the high level trigger selection .

variable	barrel, $\|\eta\| < 1.4442$	endcaps, $1.566 < \|\eta\| < 2.4$
Identification criteria		
$\|\Delta\eta In\| <$	0.01	0.01
$\|\Delta\phi In\| <$	0.04	0.08
$\sigma I\eta I\eta <$	0.011	0.031
$H/E <$	0.08	0.08
$\|1/E - 1/p\| <$	0.01	0.01
Isolation criteria (Not used with Dilepton+H_T triggers, since those triggers do not apply isolation.)		
relative ecal PFCluster isolation (dR=0.3)	0.45	
relative hcal PFCluster isolation (dR=0.3)	0.25	
relative track p_T isolation	0.2	

variable	muons			electrons		
	loose	fakable	tight	loose	fakable	tight
identification	loose ID	medium ID	medium ID	loose WP	loose WP	tight WP
isolation	loose WP	loose WP	μ WP	loose WP	loose WP	e WP
HLT emulation	–	–	–	–	– /yes	– /yes
d_0 (cm)	0.05	0.05	0.05	0.05	0.05	0.05
d_z (cm)	0.1	0.1	0.1	0.1	0.1	0.1
SIP_{3D}	–	< 4	< 4	–	< 4	< 4
missing inner hits	–	–	–	≤ 1	$= 0$	$= 0$
conversion veto	–	–	–	yes	yes	yes
tight charge	-	yes	yes	-	yes	yes

Table 9: Summary of the lepton selection.

- neutral electromagnetic energy fraction < 0.99
- number of constituents > 1
- charged hadron fraction > 0
- charged multiplicity > 0
- charged EM fraction < 0.99

For the 2017 and 2018 data analysis we followed the JetMet recommendation to switch to the Tight JetID selection defined as the following set of conditions:

- neutral hadronic energy fraction < 0.9
- neutral electromagnetic energy fraction < 0.9
- number of constituents > 1
- charged hadron fraction > 0
- charged multiplicity > 0

The most recently recommended set of jet energy corrections (JEC) is applied:

- 2016: Summer16_23Sep2016 (corresponding Global Tag: 80X_dataRun2_2016SeptRepro_v7)
- 2017: Fall17_17Nov2017B_V32
- 2018: Autumn18_RunX_V8

To avoid double counting due to jets matched geometrically with a lepton, the jet the closest matched to a fakeable or tight lepton within $\Delta R < 0.4$ is not considered as a jet in the event.

From those selected jets, the key variable H_{T} is defined by

$$H_{\mathrm{T}} = \sum_{jets} p_{\mathrm{T}}.$$

4.7 Jets originating from b quarks

Jets defined in the previous section - with the only difference of a lower p_{T} threshold down to 25 GeV - can be promoted as jets originating from b quark. The "deepCSV" discriminator is used for all years. The discriminator value for each jet is obtained by summing the *"x:probb"* and *"x:probbb"* discriminators. In 2016, x = `deepFlavourJetTags`, in 2017, x = `pfDeepCSVJetTags`, and in 2018, x = `pfDeepCSVJetTags` [18]. The medium working point requires the discriminant value to be greater than 0.6324 in 2016, 0.4941 in 2017, and 0.4184 in 2018.

When the jet multiplicity and H_{T} are computed, b-jets with $p_{\mathrm{T}} > 40$ GeV are counted along with standard jets.

4.8 Missing transverse energy

The missing transverse energy $\not\!\!E_{\mathrm{T}}$ is defined as the negative sum of the transverse momentum of all particle flow particles reconstructed in an event. The jet energy corrections are propagated to the $\not\!\!E_{\mathrm{T}}$ following the procedure described in [19].

For the analysis of 2017 data we follow the recommendation from the MET group [20] and exclude from the MET calculation low p_T jets and unclustered PF candidates in the most forward region of the ECAL (MET "v2" recipe).

5 Signal Extraction Strategy: cut-based

5.1 Baseline selection

With respect to Ref. [1], the baseline region was modified for the 2016 analysis to take advantage of the four top kinematics. The baseline region is designed to reject the majority of background (while preserving signal) before splitting kinematic phase space to form the signal regions.

With respect to the 2016 data analysis, the definition of individual signal regions has been modified (5.3), while the control regions (5.2) and the overall baseline selection (below) have remained the same.

Selected events must contain at least two tight (as defined in Table 9) same-sign leptons. To reject low-mass resonances, events with a pair of same-sign electrons with an invariant mass $m_{\ell\ell} < 12\,\text{GeV}$ are rejected. Events with three leptons of the same charge are also rejected.

In total, the baseline selection consists of the following:

- tight same-sign dileptons,
- $N_{\text{jets}} \geq 2$,
- $N_{\text{b jets}} \geq 2$,
- $E\!\!\!/_{\text{T}} > 50\,\text{GeV}$,
- $H_{\text{T}} > 300\,\text{GeV}$,
- $p_{\text{T,lep1}} \geq 25\,\text{GeV}$, $p_{\text{T,lep2}} \geq 20\,\text{GeV}$.

Events containing a third loose (as defined in Table 9) lepton that makes a DY candidate (opposite charge, same flavor pair with $|m_{\ell\ell} - m_Z| < 15\,\text{GeV}$ or $m_{\ell\ell} < 12\,\text{GeV}$) with one of the two SS dileptons are rejected from the signal regions, and a subset of them is assigned to the $\text{t}\bar{\text{t}}\text{Z}$ control region mentioned below. The transverse momentum of the third lepton is required to be larger than 5(7) GeV for muons (electrons).

If a Z candidate is not found, a tight third lepton present in the event with $p_{\text{T}} > 20\,\text{GeV}$ contributes to the N_{leps} count.

5.2 Control regions selection

Two control regions have been introduced to simultaneously constrain two dominant SM backgrounds: $\text{t}\bar{\text{t}}\text{W}$ and $\text{t}\bar{\text{t}}\text{Z}$.

The control region for $\text{t}\bar{\text{t}}\text{Z}$ ("CRZ") consists of events passing the baseline selection with three leptons (where the third lepton has $p_{\text{T}} > 20\,\text{GeV}$ and passes the tight selection), with a Z boson candidate as described above.

The control region for $\text{t}\bar{\text{t}}\text{W}$ ("CRW") consists of events with a same-sign lepton pair passing the tight lepton requirement, fewer than 6 jets, and exactly 2 b-tagged jets.

5.3 Signal regions selection

Signal regions are formed by subdividing the baseline region by number of leptons (N_{leps}), number of b-jets ($N_{\text{b jets}}$), and number of jets (N_{jets}).

The 2016+2017+2018 analysis signal regions and CRW/CRZ are tabulated in Table 10, and shown graphically in Figure 5, while the old 2016 signal regions are shown in Table 11 for reference.

Table 10: Definition of the 14 SRs and two CRs, CRW and CRZ, for the full Run2 analysis.

N_{leps}	$N_{\text{b jets}}$	N_{jets}	Region
2	2	≤ 5	CRW
		6	SR1
		7	SR2
		≥ 8	SR3
	3	5	SR4
		6	SR5
		7	SR6
		≥ 8	SR7
	≥ 4	≥ 5	SR8
≥ 3	2	5	SR9
		6	SR10
		≥ 7	SR11
	≥ 3	4	SR12
		5	SR13
		≥ 6	SR14
inverted Z-veto			CRZ

Table 11: SR definitions and CRW definition for the 2016 analysis.

N_{leps}	$N_{\text{b jets}}$	N_{jets}	Region
2	2	≤ 5	CRW
		6	SR1
		7	SR2
		≥ 8	SR3
	3	5, 6	SR4
		≥ 7	SR5
	≥ 4	≥ 5	SR6
≥ 3	2	≥ 5	SR7
	≥ 3	≥ 4	SR8
inverted Z-veto			CRZ

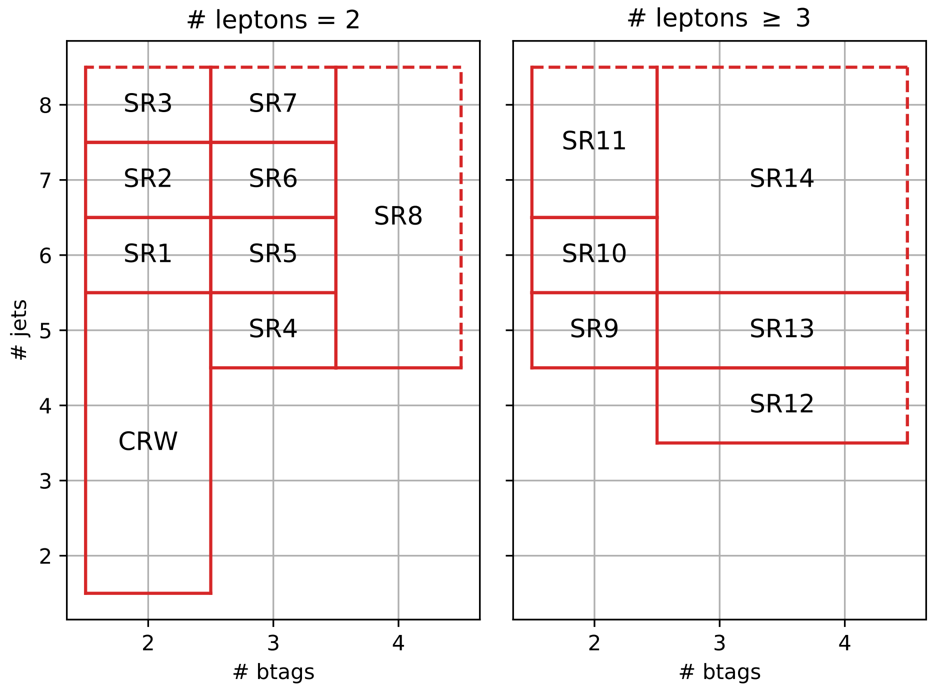


Figure 5: Signal regions and CRW (15 regions) with CRZ not displayed. Dotted edges indicate no explicit boundary.

6 Signal Extraction Strategy: BDT

6.1 Intro

A BDT approach was studied for the 2016 analysis, indicating that a 10-20% improvement in the expected significance could be reached with respect to the cut-based approach, in an optimistic scenario without additional systematics on the BDT input variables. For the 2016 analysis, the cut-based approach was chosen, resulting in a faster publication and a slightly lower performance. For the full Run 2 analysis, we have revisited the BDT, with an extended study of the input variables and hyperparameters, as discussed below. Some of the studies below are based on 75 fb^{-1} of integrated luminosity, as they were performed with a 2016+2017 analysis in mind, but their conclusions also apply to the full Run 2 analysis.

A BDT was explored using a well-known python package *xgboost* for a gradient boosted decision tree. Both ROOT-based TMVA, and python-based *xgboost* were explored, with the latter being used in the end with appropriate binning since it shows slightly higher performance. The sections will be structured as follows. A list of variables used as inputs to the BDT will be motivated and described, followed by the hyperparameters used in the training of the decision tree. The end-result comparison with the nominal cut-based analysis will be presented, and studies contributing to this result (e.g., training sets/parameters, etc) will follow.

6.2 Selection and variables

BDT training is susceptible to low statistics, so in order to increase statistics, a relaxed baseline selection for training was created, consisting of

- Lepton 1 $p_T > 15$
- Lepton 2 $p_T > 15$
- $H_T > 250$
- MET > 30
- Njets ≥ 2
- Nbtags ≥ 1

Signal was defined to be from four top MC and background was taken as all of the SM backgrounds previously described and taken in this analysis. Variations of this are discussed in the last section of miscellaneous studies.

Over 40 variables were considered. Ranked in approximate descending order of discriminative power, as reported by TMVA, they are

- Njets
- Nbtags
- Nleps
- $m_T(\ell_2)$
- $m(\ell_1, \ell_2)$
- MET
- $m_T(\ell_1)$
- H_T^b: H_T made from b jets
- $m(j_1, j_2)$: invariant mass of leading two jets
- $m(\ell_1, j_2)$

- $\Delta\phi(j_1, j_2)$
- $\Delta\phi(\ell_1, j_1)$
- $p_T(j_i)$ for $i = 1 - 8$
- q_1: sign of the same-sign lepton pair
- $\Delta\eta(\ell_1, \ell_2)$
- H_T
- H_T^{ratio}: Ratio of H_T of first four leading jets to rest
- $m(\ell_1, j_1)$
- Nlooseb: number of btags passing loose threshold
- Ntightb: number of btags passing tight threshold
- max($m(j)/p_T$): ratio of jet mass to momentum to discriminate merged jets
- Wcands: number of jet pairs with invariant mass within 30GeV of the W mass

Other more general variables like $m(\ell_{\text{i}}, j_{\text{j}})$ were also considered but showed little to no discrimination power, where i and j encompass the first two leading objects. The same generalization applied to other indexed variables. Note that this ranking takes into account the correlation between variables, explaining why H_T is not as highly ranked, since discriminative power is first taken from Njets. Past approximately 22 variables, no extra AUC (area under curve) was gained, so it was determined to proceed with this smaller set of variables for simplicity. The AUC metric is the area under the signal versus background efficiency Receiver Operating Characteristic (ROC) curve, bounded by 0 and 1, where 1 is equivalent to perfect discrimination between signal and background, and 0.5 represents discrimination no better than random guessing.

The 19 approximately most performant variables were then selected to continue optimization. They are

- (a) Nbtags
- (b) Njets
- (c) Nlooseb
- (d) MET
- (e) Ntightb
- (f) $p_T(\ell_2)$
- (g) $m(\ell_1, j_1)$
- (h) $p_T(j_1)$
- (i) $p_T(j_7)$
- (j) $\Delta\phi(\ell_1, \ell_2)$
- (k) $p_T(j_6)$
- (l) max($m(j)/p_T(j)$)
- (m) Nleps
- (n) $p_T(\ell_1)$
- (o) $\Delta\eta(\ell_1, \ell_2)$
- (p) $p_T(j_8)$
- (q) H_T^b

- (r) $p_T(\ell_3)$
- (s) q_1

Kinematic distributions for these input variables are shown in Figure 6.

6.3 Hyperparameters

Hyper-parameter tuning was performed in order to maximize discrimination ($s/\sqrt{s+b}$ and AUC) while avoiding over-training from limited statistics. The selected set of TMVA hyper-parameters is given by the strings

- NTrees=500
- nEventsMin=150
- MaxDepth=5
- BoostType=AdaBoost
- AdaBoostBeta=0.25
- SeparationType=GiniIndex
- nCuts=20
- PruneMethod=NoPruning

In total, the complete set of combinations considered in the hyperparameter scan were

- NTrees=200,500,1000
- nEventsMin=50,150,300
- MaxDepth=4,5,6
- BoostType=AdaBoost,Bagging,Grad
- AdaBoostBeta=0.1,0.25,0.8
- SeparationType=GiniIndex,CrossEntropy
- nCuts=5,20,100

A similar hyperparameter scan for xgboost yielded the parameters

- n_estimators = 500
- eta = 0.07
- max_depth = 5
- subsample = 0.6
- alpha = 8.0
- gamma = 2.0
- lambda = 1.0
- min_child_weight = 1.0
- colsample_bytree = 1.0

Note that "n_estimators" represents the number of trees, and "eta" is the learning rate. In particular, for a given learning rate, the number of trees and the depth of each tree are the most impactful hyper-parameters.

Using these optimal hyperparameters, Figure 7 compares ROC curves and $s/\sqrt{s+b}$ curves for TMVA and xgboost, showing that xgboost yields a significance performance than TMVA.

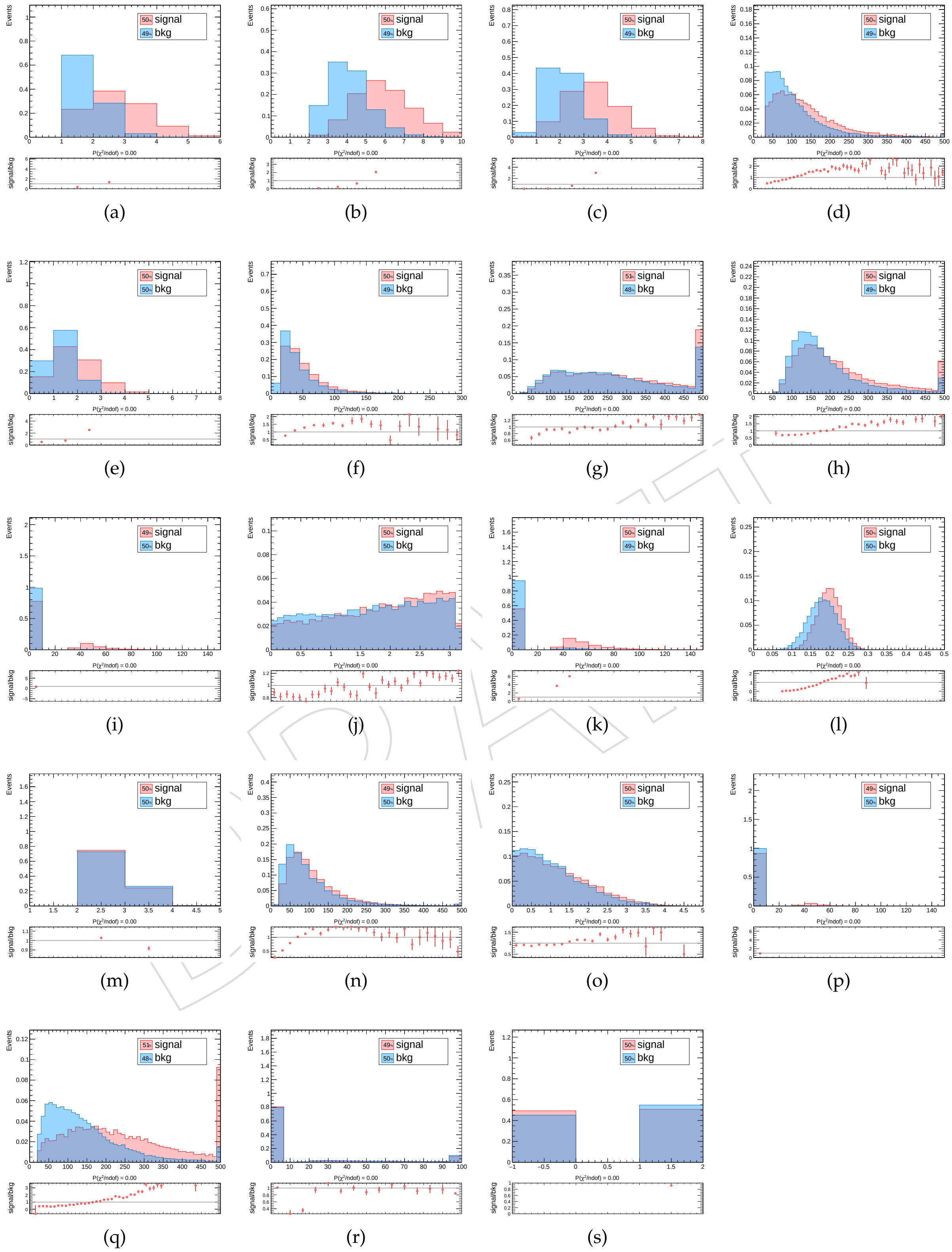
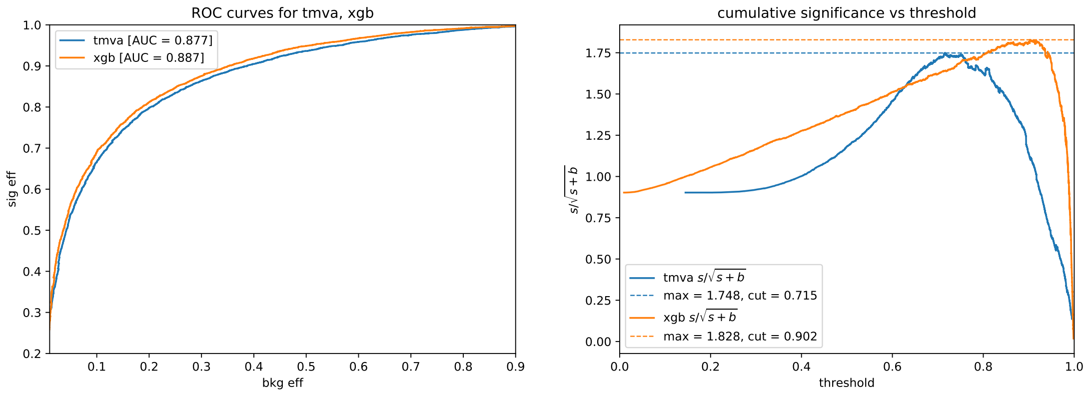


Figure 6: Distributions of kinematic inputs to the BDT. Integrals have been normalized so only shapes are relevant here.

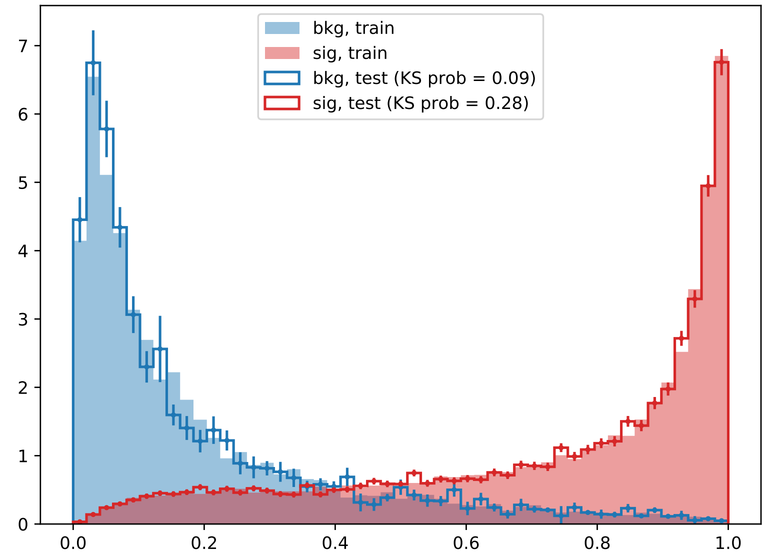
Figure 8 shows that xgboost discriminator shapes for train and test sets separately for background and signal. There appears to be no signs of overtraining, given the above hyperparameters.

The raw xgboost output discriminant using these 19 variables is shown in Figure 9, lumping rare MC processes into ”Rares”, and fakes/flips into ”Others”.



(a) (b)

Figure 7: Left: ROC curves for TMVA and xgboost with AUC metric displayed in the legend. Right: Cumulative significance vs threshold value on the discriminant for TMVA and xgboost. The maximum of this curve is marked by a dotted line.

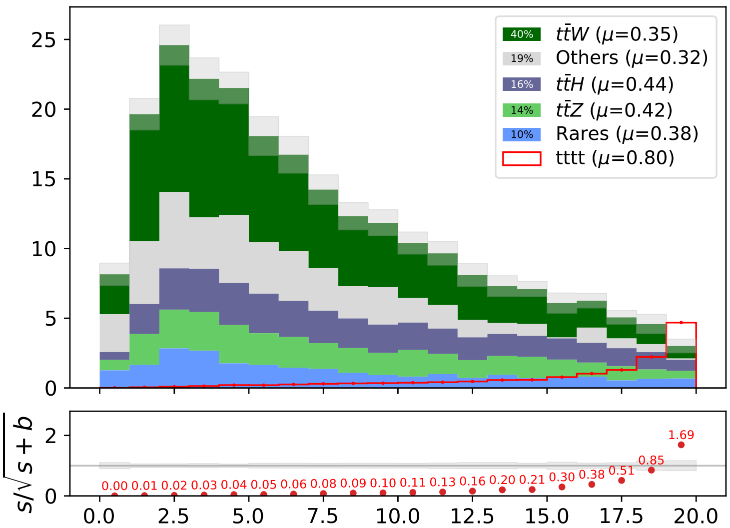


(a)

Figure 8: Raw xgboost discriminator shape for train and test sets separately for background and signal.

6.4 Comparison with cut-based

For 18 bins in total, there are 17 bins along the discriminator output from 0 to 1 corresponding to the bin edges of `0.0000, 0.0362, 0.0659, 0.1055, 0.1573, 0.2190, 0.2905, 0.3704, 0.4741, 0.6054, 0.7260, 0.8357, 0.9076, 0.9506, 0.9749, 0.9884,`



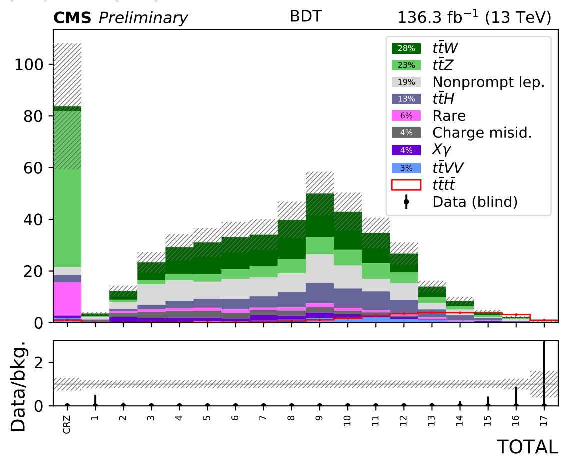
(a)

Figure 9: Raw xgboost BDT discriminator shape with a panel showing the cumulative $s/\sqrt{s+b}$ with MC scaled to 75 fb^{-1}.

0.9956, 1.0000 . The initial bin contains events matching the CRZ selection to help constrain the $t\bar{t}Z$ background. Increasing the binning beyond this does not result in a gain in significance, and also gives issues with sample statistics.

Note that the binning was chosen to create a shape different from the raw xgboost discriminator output. The resulting shape is similar to the output of an AdaBoost algorithm from TMVA, and has the effect of distributing signal events normally contained within 1-2 bins into a few more, allowing for a better constraint in the fit.

In order to compare properly with the cut-based, the region in which the BDT discriminant is calculated and applied corresponds to the nominal baseline selection. The 18-bin BDT shown in Figure 10 yields an expected significance of 2.68σ, with the 16-bin cut-based result yielding an expected significance of 2.46σ, which is a gain slightly larger than 8%.



(a)

Figure 10: BDT discriminant in discrete signal bins, including CRZ in the first bin.

6.5 Misc studies, all numbers and text is from 2016

6.5.1 Definition of background

Training is performed on the full mix of backgrounds. Backgrounds modeled with simulation are normalized to a common luminosity, i.e. weighted events are used. For comparison, we also trained separate BDTs for each of the ttV backgrounds, but saw almost no gain in expected sensitivity. Note that signal-background have similar separation for each of the individual background processes, as shown in Figure 11.

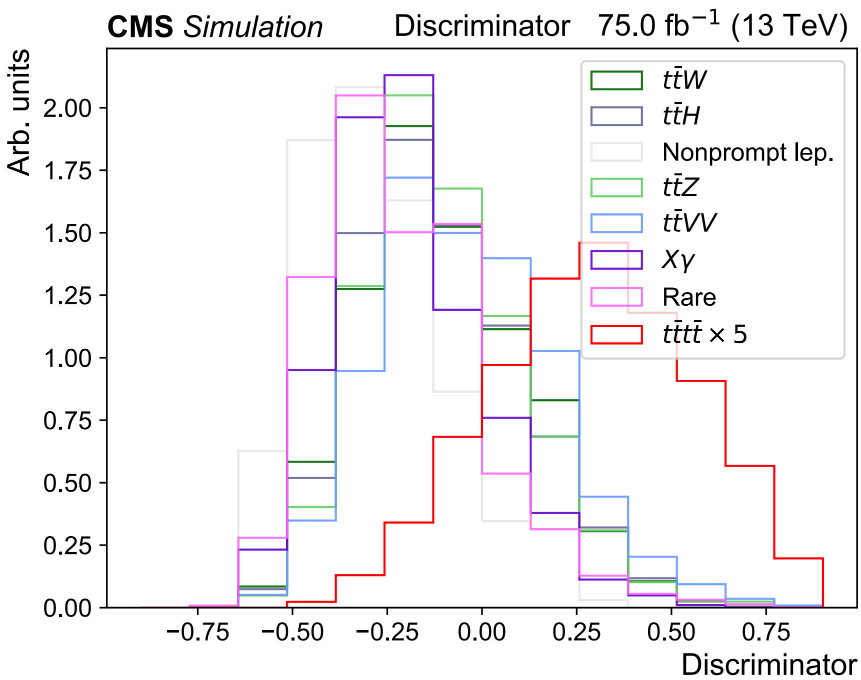


Figure 11: Unstacked BDT discriminator shape for signal and background process

7 Validation of kinematic variables

In this section we describe the control-regions to validate with data the kinematic variables we use in defining our cut-based signal regions: H_{T}, $\not{E}_{\mathrm{T}}$, N_{jets}, $N_{\mathrm{b\,jets}}$. In addition, the variables used for the BDT are also studied: H_T^b, Nlooseb, Ntightb, $\Delta\phi(\ell_1, \ell_2)$, $\Delta\eta(\ell_1, \ell_2)$, $p_T(\ell_1)$, $p_T(\ell_2)$, $p_T(\ell_3)$, q_1, $m(\ell_1, j_1)$, $\max(m(j)/p_T(j))$ and the p_{T} for jets 1,6,7, and 8.

The control regions considered are opposite-sign dilepton (Section 7.1), tight/loose same-sign dilepton (Section 7.2). CRW and CRZ, which are closest to the signal region, will be presented in Section 10.

7.1 Opposite-sign dilepton events

In this control region, the same requirements on H_{T}, $\not{E}_{\mathrm{T}}$ and N_{jets} as in our inclusive *baseline* selection are applied (see Section 5), but we require two opposite sign tight leptons and we remove the Z/γ^* veto. This control region coincides with the application region we use for the data-driven method to estimate the *Charge flips* background (see Section 8).

Distributions are shown in Figs. 12,13,14 (2017, 2018, 2016+2017+2018) for the main variables, and Figs. 15,16,17 for additional variables used in the BDT. The overall agreement is very good, but there are some discrepancies in the $N_{\mathrm{b\,jets}}$ and BDT distributions that we will monitor as we include additional corrections to the MC. The ISR/FSR corrections described in Section 2, corresponding to Figure 4, is applied to the $\mathrm{t\bar{t}}$ sample here. Note that quantities like the p_{T} for jet 7,8 correspond to high (5+) ISR/FSR jet multiplicities in OS $\mathrm{t\bar{t}}$. These are anyway not directly probed in the signal regions, consisting mainly of $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ for which the matrix element provides more partons.

7.2 Same-sign tight+fail dilepton events

In this control region, the same requirements on H_{T}, $\not{E}_{\mathrm{T}}$ and N_{jets} as in our inclusive *baseline* selection are applied (see Section 5), but we require one tight lepton and one same-sign lepton failing the tight requirement. The control region is enriched in events with one fake lepton. It corresponds to the application region for the *fake-rate* method, the data-driven estimate of backgrounds with fake-leptons (see Section 8). The only difference is that in the fake-rate application we also allow events where the two leptons fail the tight requirements.

Distributions are shown in Figs. 18,19,20 (2017, 2018, 2016+2017+2018) for the main variables, and Figs. 21,22,23 for additional variables used in the BDT. There is an overall underestimate, consistent with what was seen in 2016. The underestimate is flavor-independent, and does not show large trends in the main kinematics. Such an underestimate in loose-not-tight leptons, if paired with a data FakeRate equal or larger to the MC one, can indicate that the Nonprompt lepton background is underestimated by a pure MC prediction. Since this background is predicted from data, that is not an issue with the analysis strategy.

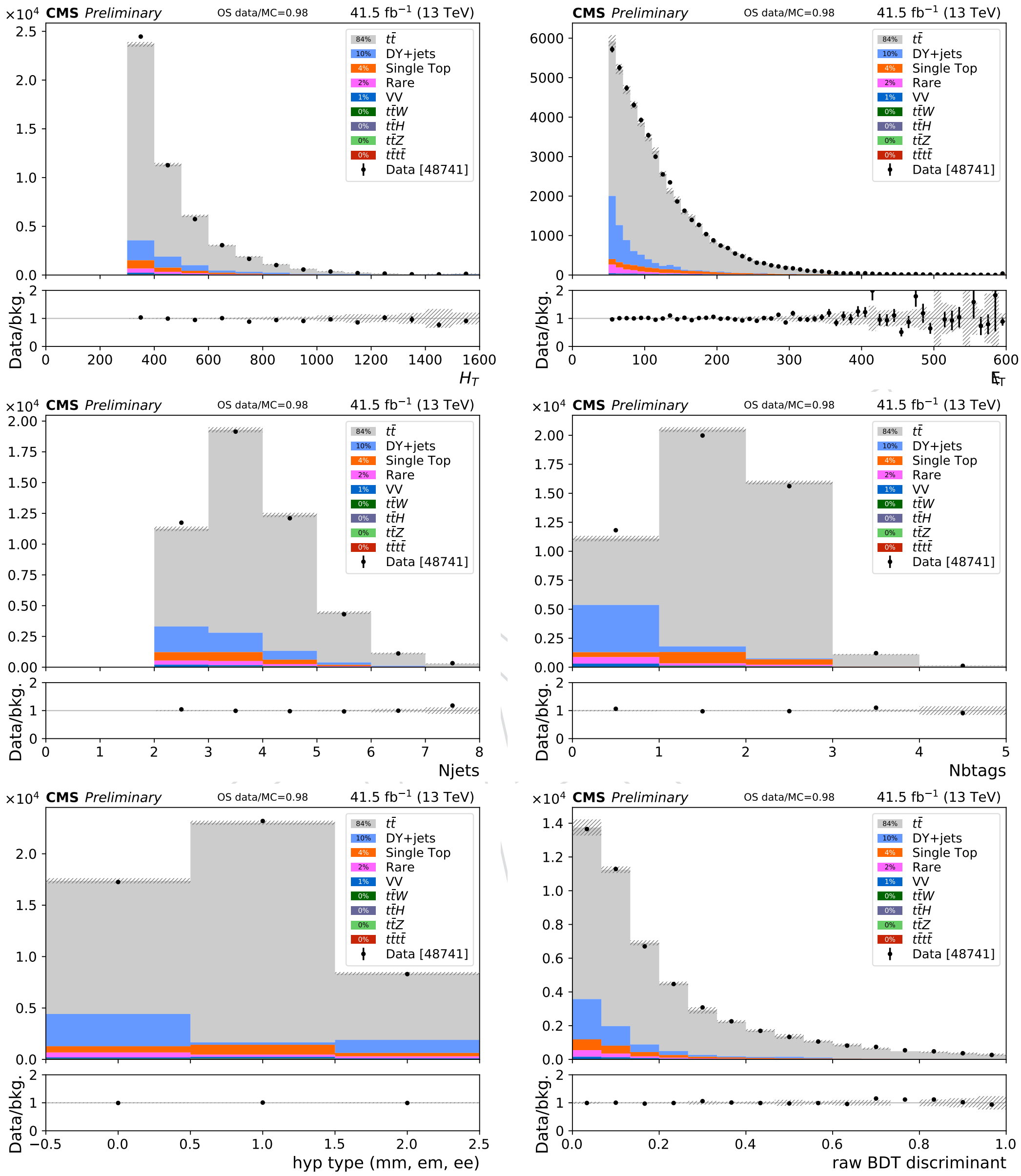


Figure 12: Data to simulation comparisons for 2017. From left top to right bottom the H_{T}, $\not\!\!E_{\mathrm{T}}$, N_{jets}, $N_{\mathrm{b\,jets}}$, lepton flavor and raw BDT discriminant distributions are shown for the opposite-sign dilepton baseline region. Shaded band shows MC stat uncertainty, except in the case of the discriminator distribution, where it has been added in quadrature with scale, btag, JEC, and JER variations

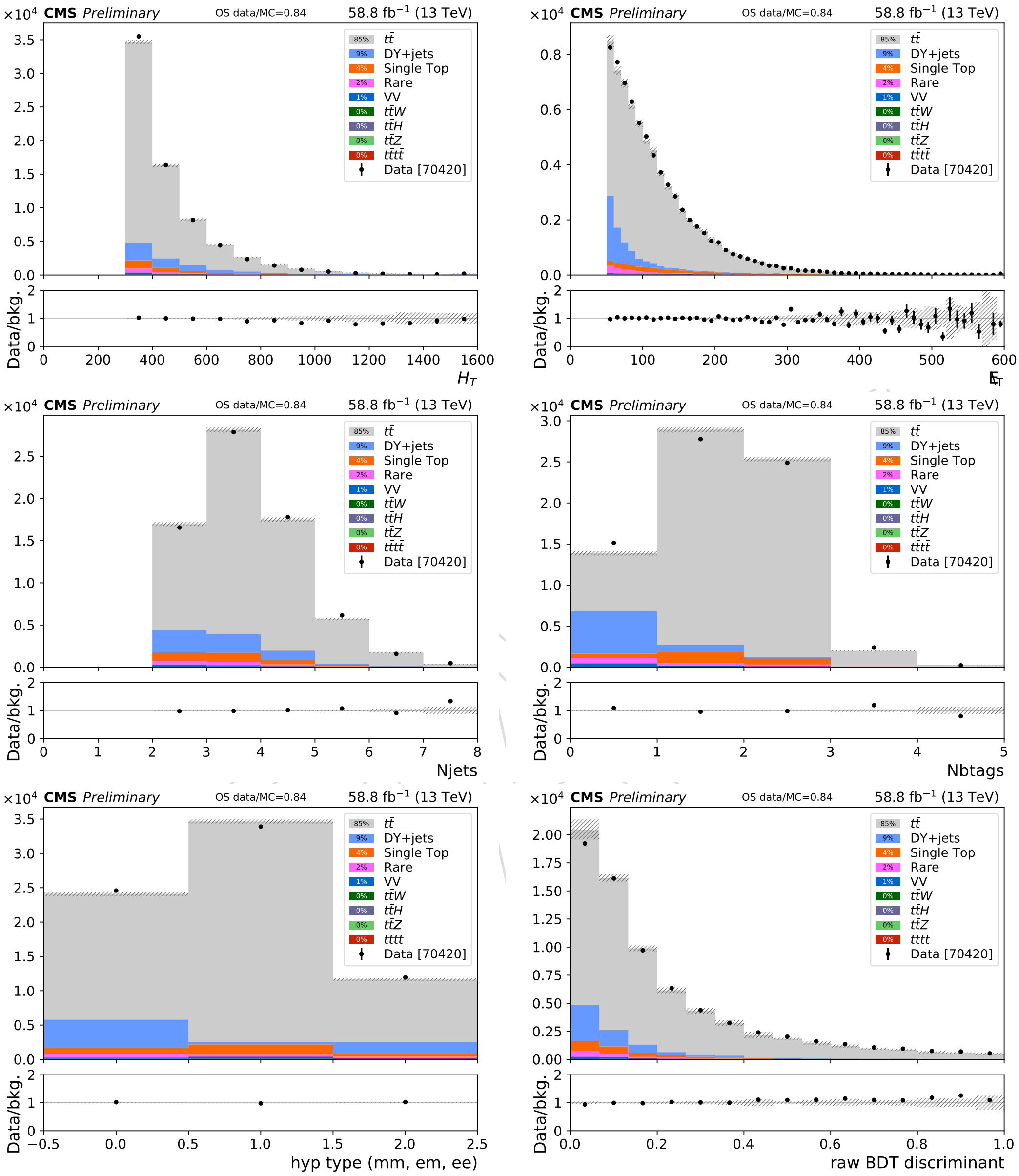


Figure 13: Data to simulation comparisons for 2018. From left top to right bottom the H_T, $\not\!\!E_T$, N_{jets}, $N_{b\ jets}$, lepton flavor and raw BDT discriminant distributions are shown for the opposite-sign dilepton baseline region. Shaded band shows MC stat uncertainty, except in the case of the discriminator distribution, where it has been added in quadrature with scale, btag, JEC, and JER variations.

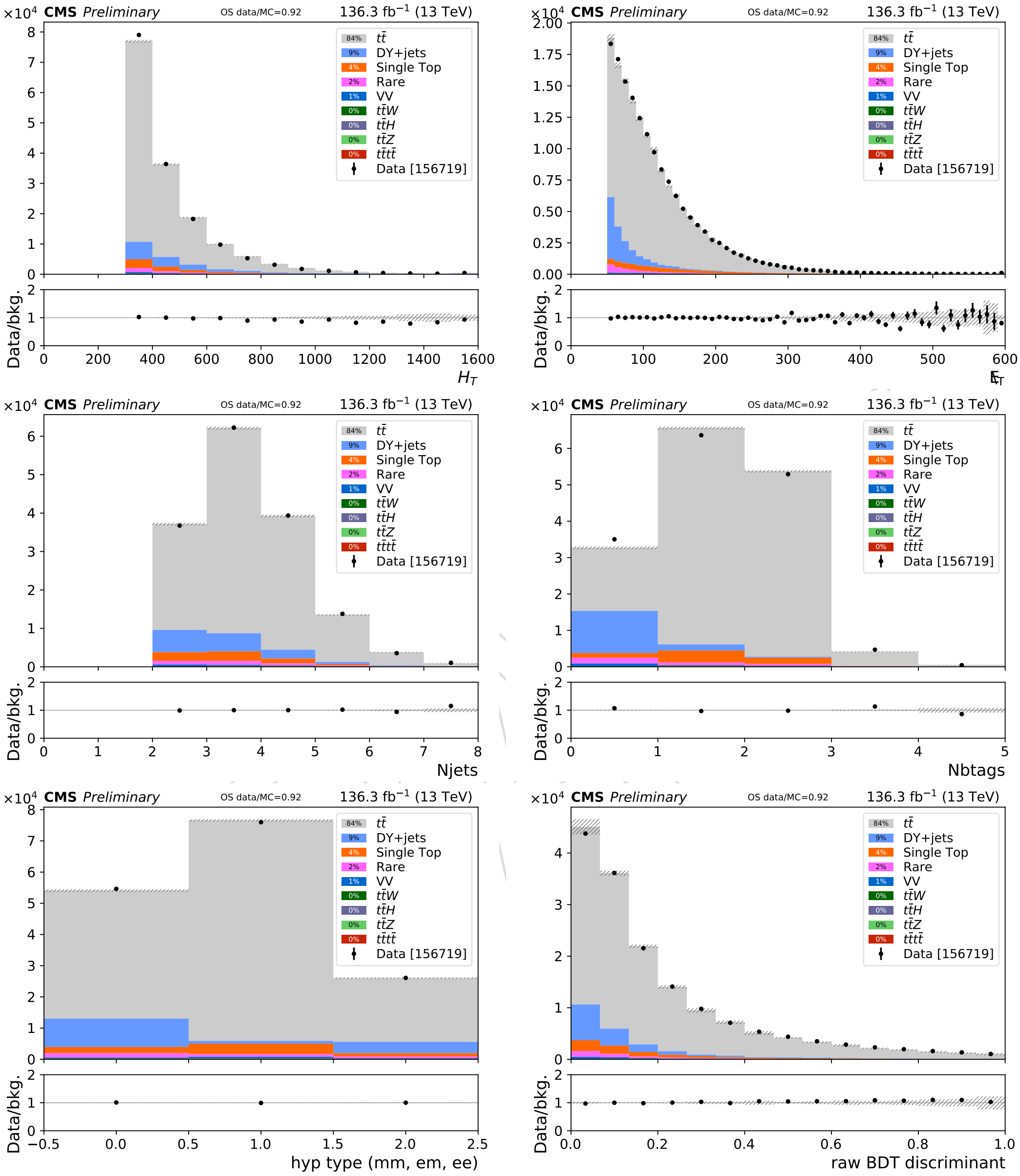


Figure 14: Data to simulation comparisons for 2016+2017+2018. From left top to right bottom the H_T, $\not\!\!E_T$, N_{jets}, $N_{b\,jets}$, lepton flavor and raw BDT discriminant distributions are shown for the opposite-sign dilepton baseline region. Shaded band shows MC stat uncertainty, except in the case of the discriminator distribution, where it has been added in quadrature with scale, btag, JEC, and JER variations

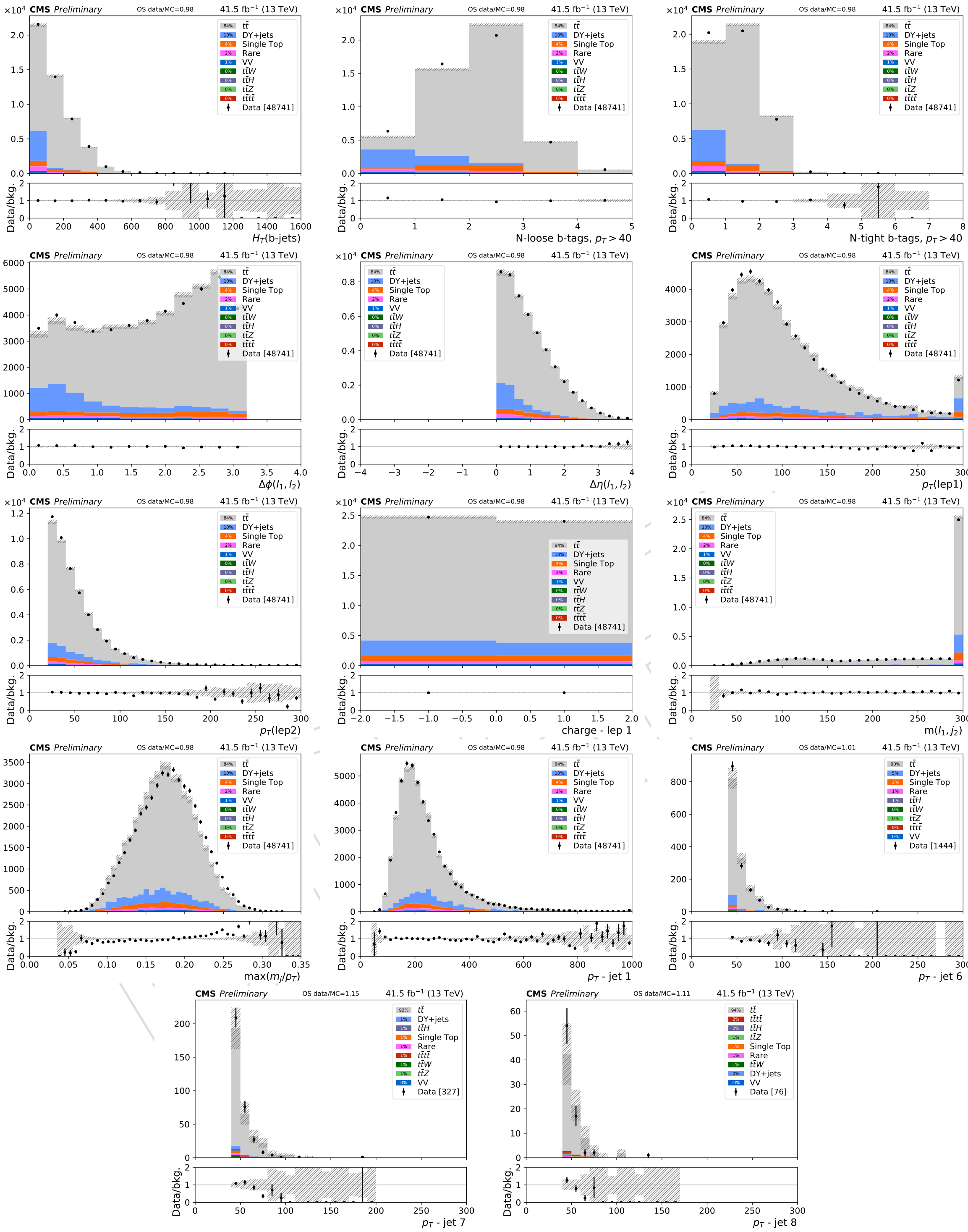


Figure 15: Data to simulation comparisons for 2017, for the additional variables used by the BDT. From left top to right bottom, H_T^b, Nlooseb, Ntightb, $\Delta\phi(\ell_1, \ell_2)$, $\Delta\eta(\ell_1, \ell_2)$, $p_T(\ell_1)$, $p_T(\ell_2)$, $p_T(\ell_3)$, q_1, $m(\ell_1, j_1)$, $\max(m(j)/p_T(j))$ and the p_{T} for jets 1,6,7, and 8; shown for the opposite-sign dilepton baseline region.

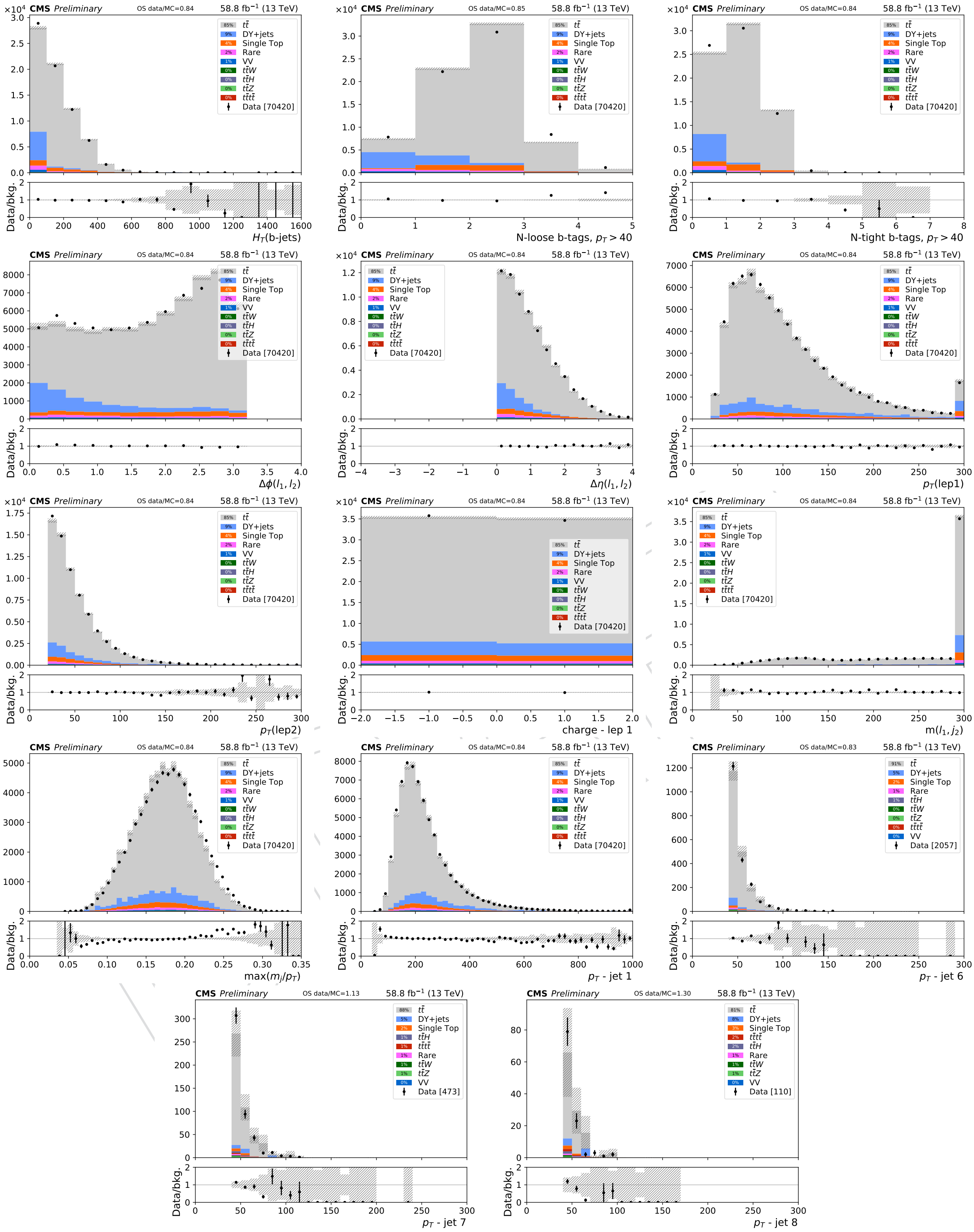


Figure 16: Data to simulation comparisons for 2018, for the additional variables used by the BDT. From left top to right bottom, H_T^b, Nlooseb, Ntightb, $\Delta\phi(\ell_1, \ell_2)$, $\Delta\eta(\ell_1, \ell_2)$, $p_T(\ell_1)$, $p_T(\ell_2)$, $p_T(\ell_3)$, q_1, $m(\ell_1, j_1)$, $\max(m(j)/p_T(j))$ and the p_T for jets 1,6,7, and 8; shown for the opposite-sign dilepton baseline region.

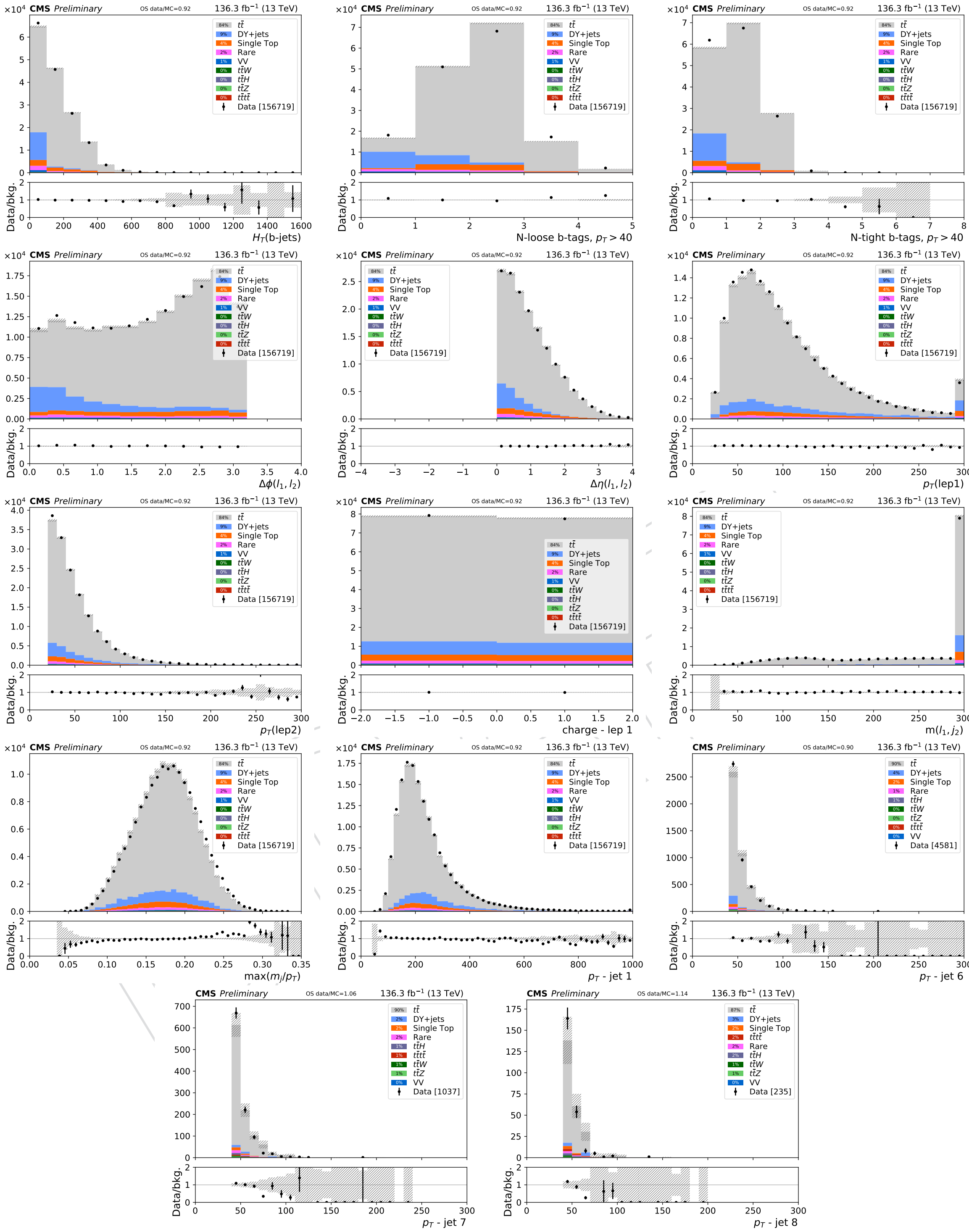


Figure 17: Data to simulation comparisons for 2016+2017+2018, for the additional variables used by the BDT. From left top to right bottom, H_T^b, Nlooseb, Ntightb, $\Delta\phi(\ell_1, \ell_2)$, $\Delta\eta(\ell_1, \ell_2)$, $p_T(\ell_1)$, $p_T(\ell_2)$, $p_T(\ell_3)$, q_1, $m(\ell_1, j_1)$, $\max(m(j)/p_T(j))$ and the p_T for jets 1,6,7, and 8; shown for the opposite-sign dilepton baseline region.

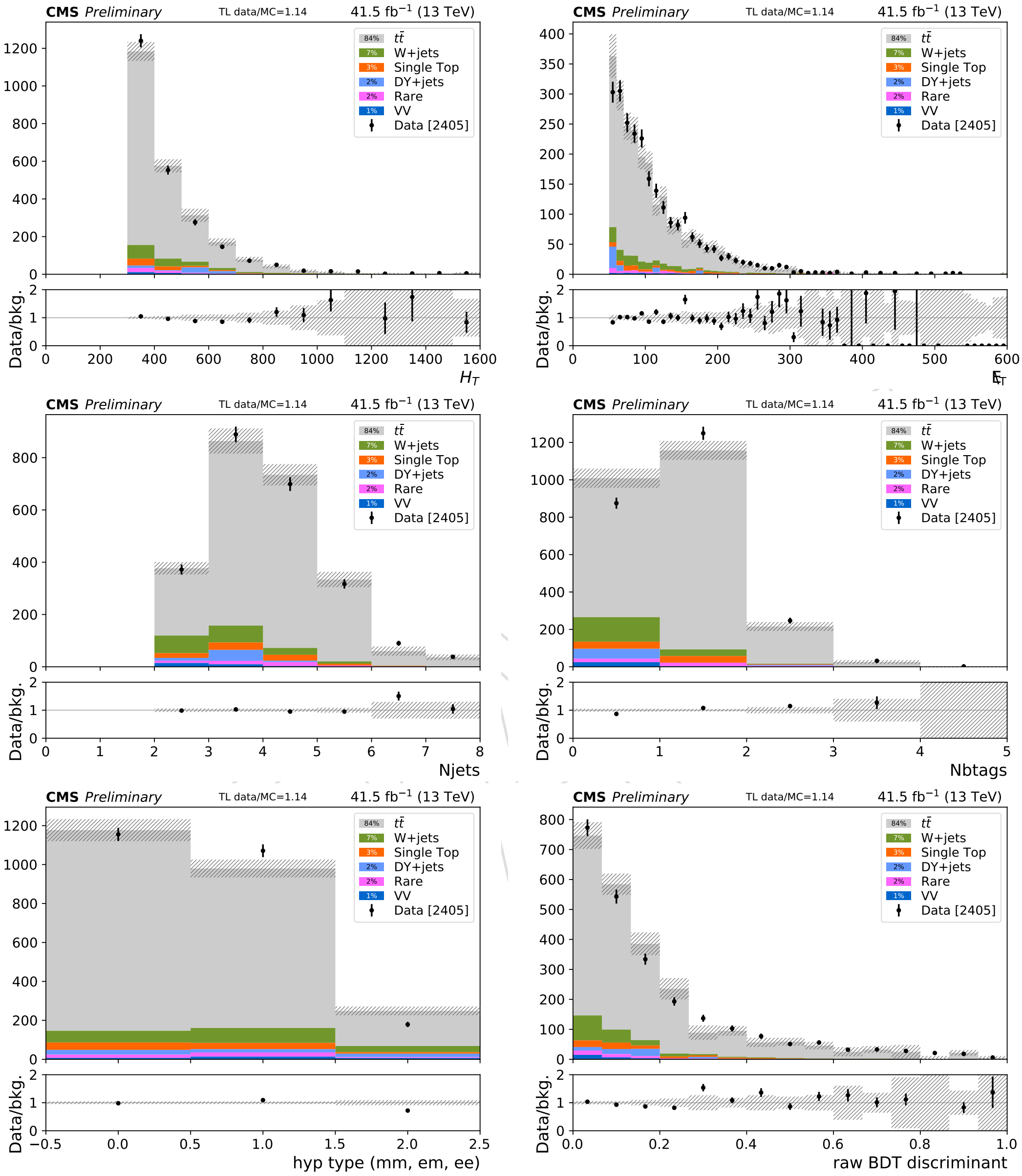


Figure 18: Data to simulation comparisons for 2017. From left top to right bottom the H_T, $\not{E}_T$, N_{jets}, $N_{b\ jets}$, lepton flavor and raw BDT discriminant distributions are shown for the same-sign tight+fail dilepton baseline region. Shaded band shows MC stat uncertainty, except in the case of the discriminator distribution, where it has been added in quadrature with scale, btag, JEC, and JER variations

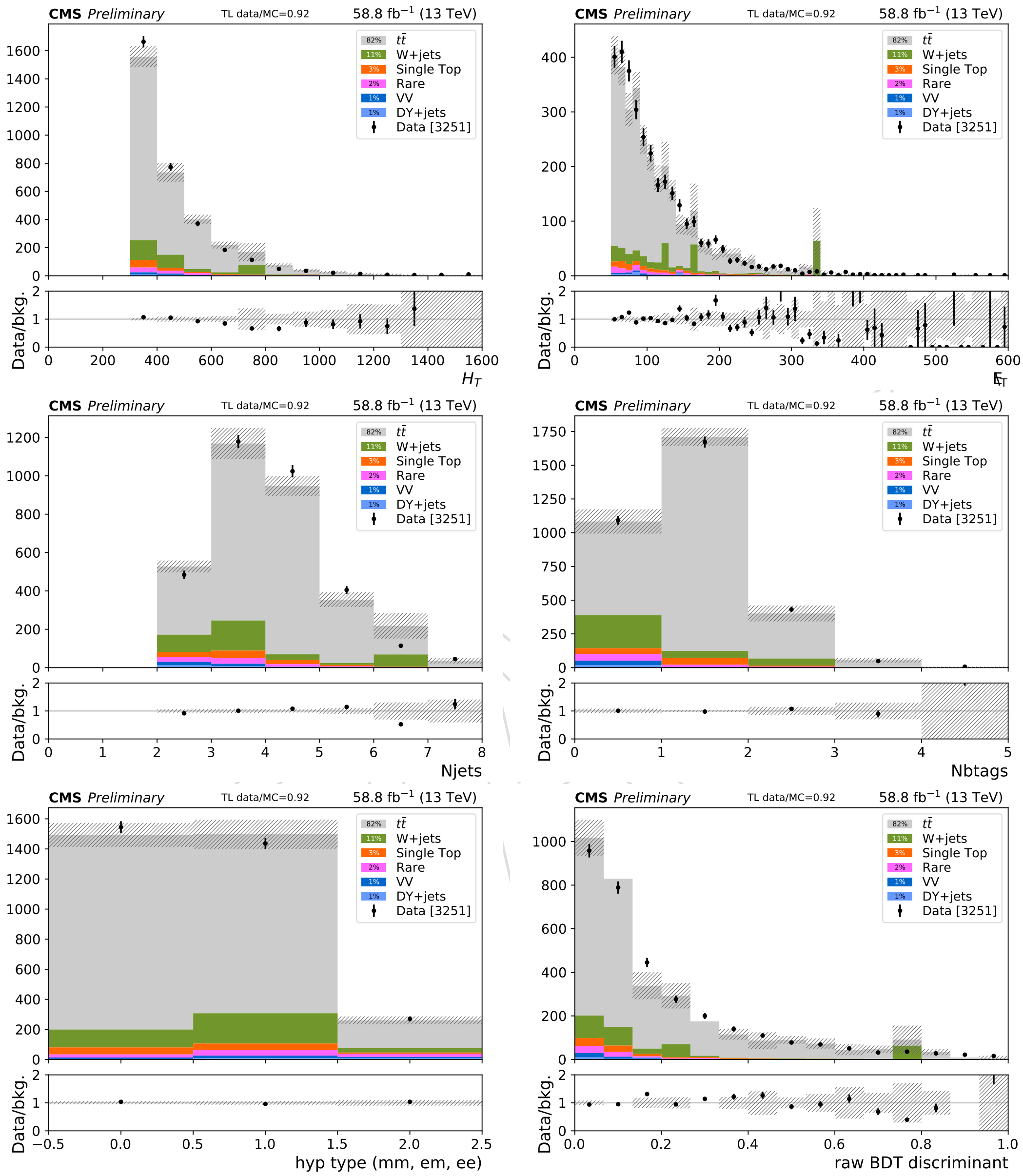


Figure 19: Data to simulation comparisons for 2018. From left top to right bottom the H_{T}, $\not{E}_{\mathrm{T}}$, N_{jets}, $N_{\mathrm{b\ jets}}$, lepton flavor and raw BDT discriminant distributions are shown for the same-sign tight+fail dilepton baseline region. Shaded band shows MC stat uncertainty, except in the case of the discriminator distribution, where it has been added in quadrature with scale, btag, JEC, and JER variations

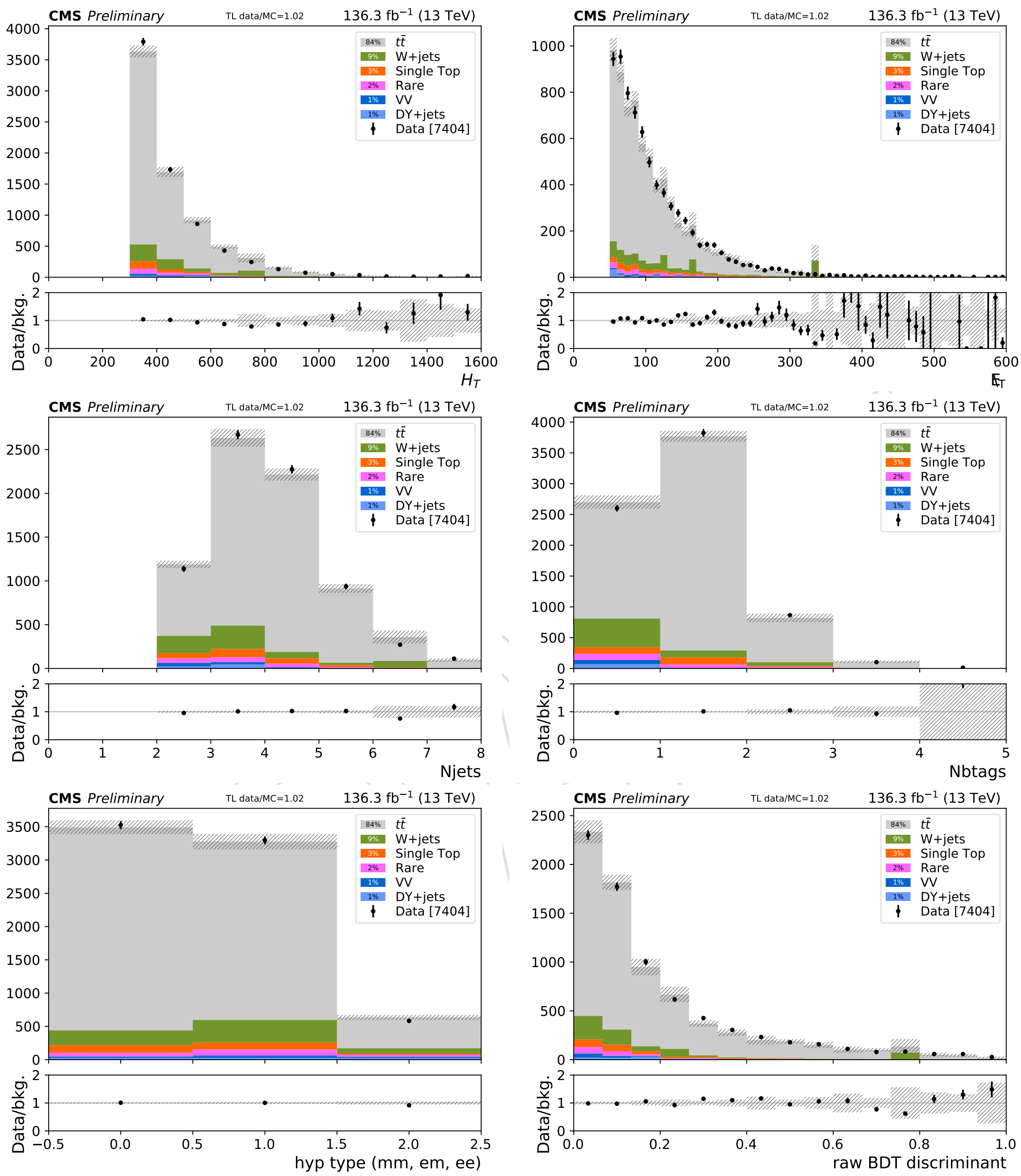


Figure 20: Data to simulation comparisons for 2016+2017+2018. From left top to right bottom the H_T, $\not{E}_T$, N_{jets}, $N_{\text{b jets}}$, lepton flavor and raw BDT discriminant distributions are shown for the same-sign tight+fail dilepton baseline region. Shaded band shows MC stat uncertainty, except in the case of the discriminator distribution, where it has been added in quadrature with scale, btag, JEC, and JER variations

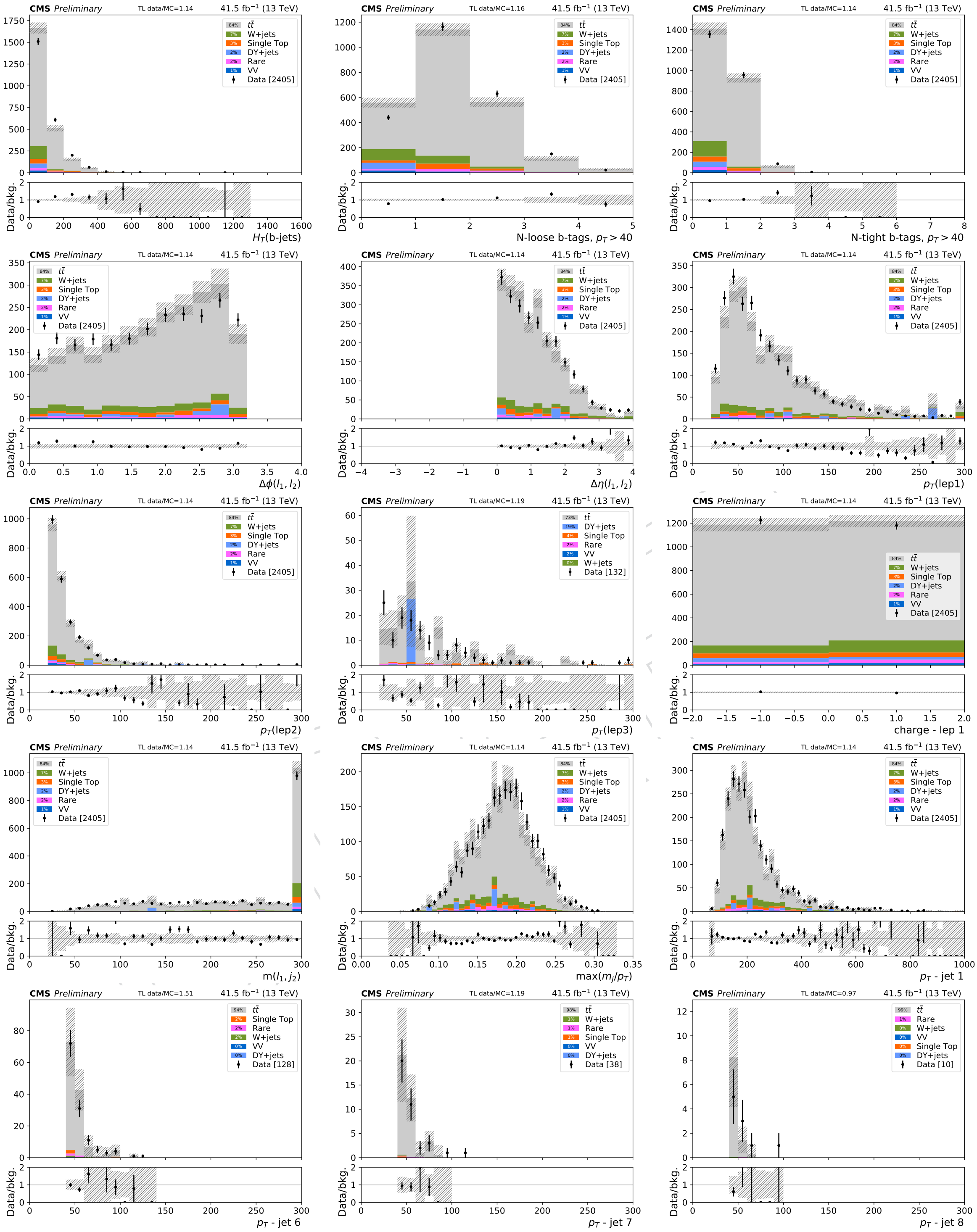


Figure 21: Data to simulation comparisons for 2017, for the additional variables used by the BDT. From left top to right bottom, H_T^b, Nlooseb, Ntightb, $\Delta\phi(\ell_1, \ell_2)$, $\Delta\eta(\ell_1, \ell_2)$, $p_T(\ell_1)$, $p_T(\ell_2)$, $p_T(\ell_3)$, q_1, $m(\ell_1, j_1)$, $\max(m(j)/p_T(j))$ and the p_{T} for jets 1,6,7, and 8; shown for the same-sign tight+fail dilepton baseline region.

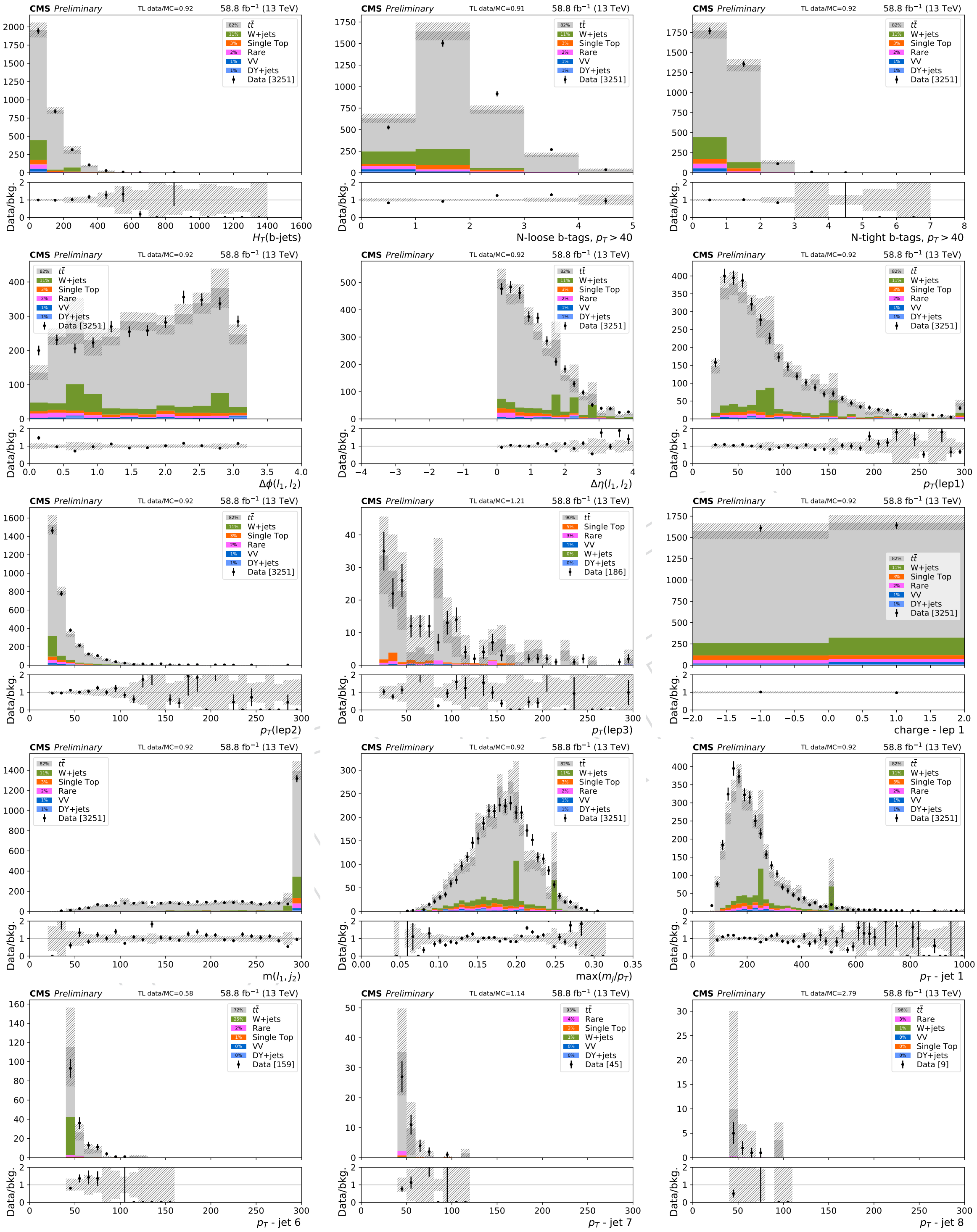


Figure 22: Data to simulation comparisons for 2018, for the additional variables used by the BDT. From left top to right bottom, H_T^b, Nlooseb, Ntightb, $\Delta\phi(\ell_1, \ell_2)$, $\Delta\eta(\ell_1, \ell_2)$, $p_T(\ell_1)$, $p_T(\ell_2)$, $p_T(\ell_3)$, q_1, $m(\ell_1, j_1)$, $\max(m(j)/p_T(j))$ and the p_T for jets 1,6,7, and 8; shown for the same-sign tight+fail dilepton baseline region.

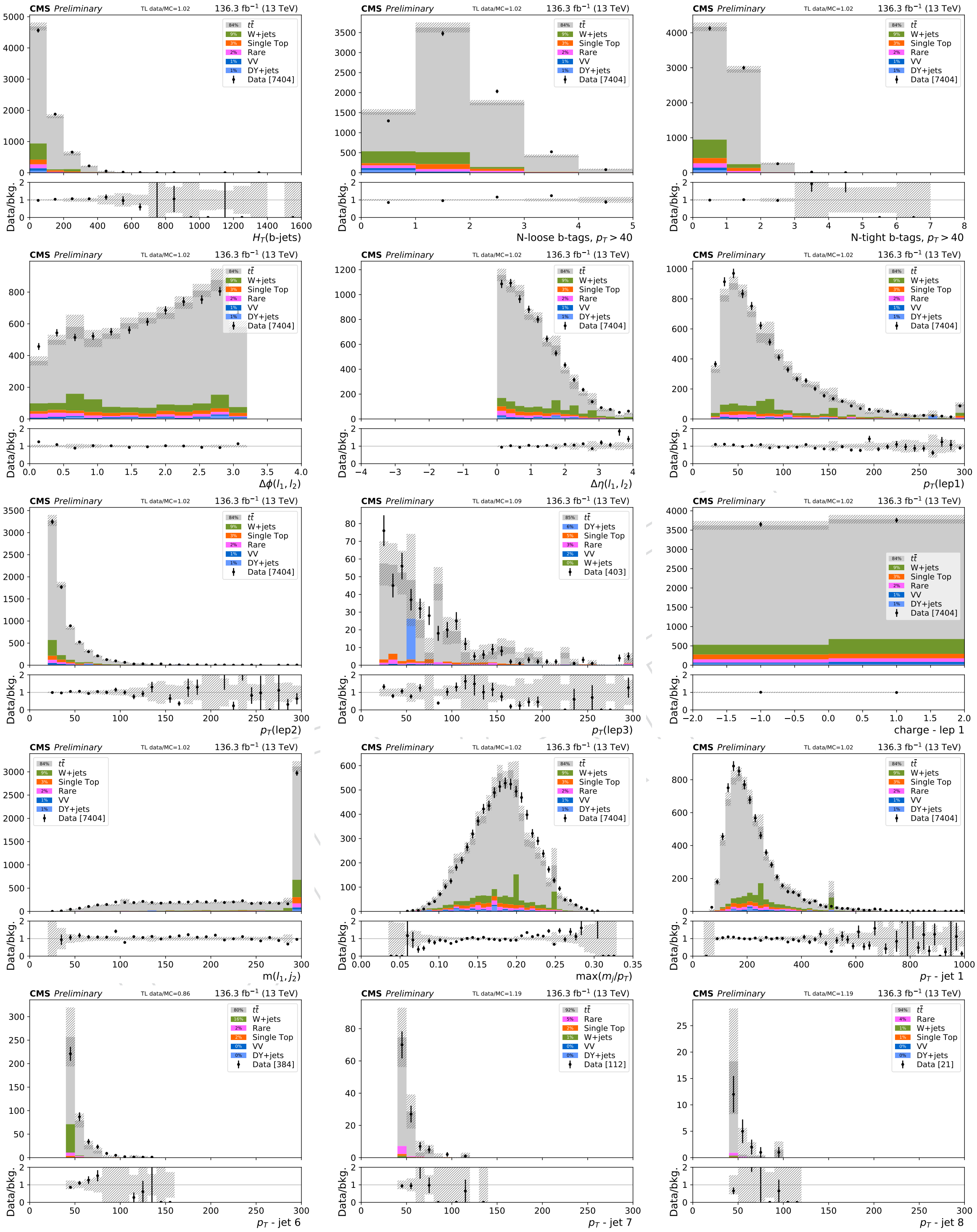


Figure 23: Data to simulation comparisons for 2016+2017+2018, for the additional variables used by the BDT. From left top to right bottom, H_T^b, Nlooseb, Ntightb, $\Delta\phi(\ell_1, \ell_2)$, $\Delta\eta(\ell_1, \ell_2)$, $p_T(\ell_1)$, $p_T(\ell_2)$, $p_T(\ell_3)$, q_1, $m(\ell_1, j_1)$, $\max(m(j)/p_T(j))$ and the p_T for jets 1,6,7, and 8; shown for the same-sign tight+fail dilepton baseline region.

7.3 Fake-enriched validation region in data

In this control region, the same requirements on H_{T}, $\not\!\!E_{\mathrm{T}}$ and N_{jets} as in our inclusive *baseline* selection are applied (see Section 5), except we relax the H_{T} requirement and require $N_{\mathrm{b\ jets}} = 1$. This region has a significant non-prompt component and allows us to check the overall closure of the method in data. In the plots, the fake background is data-driven. Distributions are shown in Figs. 24. The overall data/MC normalization factor in this region is 1.06. If fakes are entirely responsible for this discrepancy, and given that fakes constitute half of the background, this represents a 12% normalization increase of fakes, well within the 30% normalization uncertainty taken on this process.

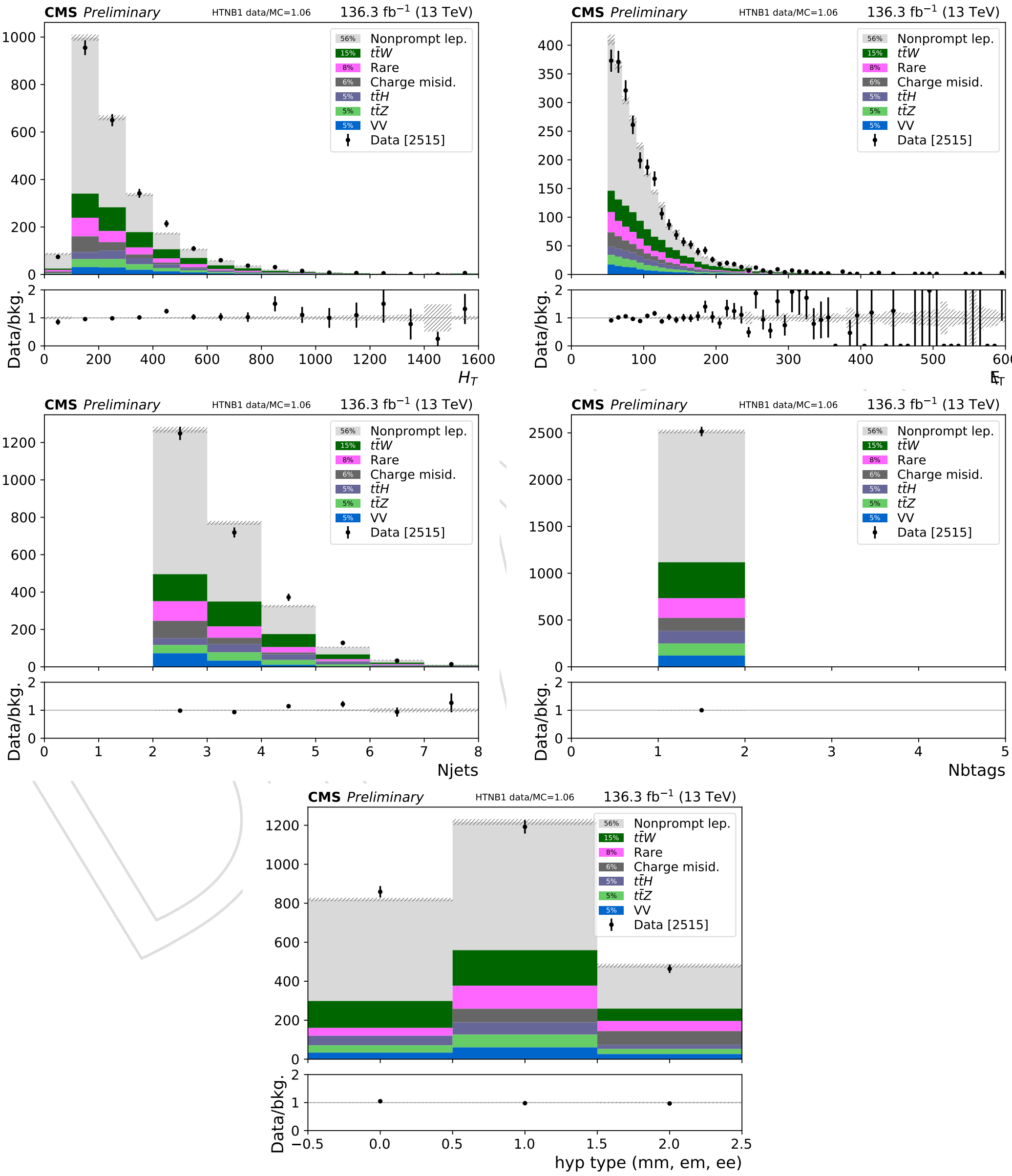


Figure 24: Data to simulation comparisons for 2016+2017+2018. From left top to right bottom the H_{T}, $\not\!\!E_{\mathrm{T}}$, N_{jets}, $N_{\mathrm{b\ jets}}$, lepton flavor and raw BDT discriminant distributions are shown for the same-sign dilepton region with $N_{\mathrm{b\ jets}} = 1$, $N_{\mathrm{jets}} \geq 2$, $E_{\mathrm{T}}^{\mathrm{miss}} \geq 50$. Shaded band shows MC stat uncertainty.

8 Background Estimations

Backgrounds for the same-sign dilepton final state can be divided in three categories:

- **Fake leptons:** "Non-Prompt" or "Fake" leptons are leptons from heavy-flavour decays, misidentified hadrons, muons from light-meson decays in flight, or electrons from unidentified photon conversions. Depending on the signal regions, this background is dominated by $t\bar{t}$ or W+jets processes.
- **Charge flips:** Charge misidentification, i.e. events with opposite-sign isolated leptons where the charge of one of the leptons is misidentified because of severe bremsstrahlung in the tracker material. This background, is relevant only for electrons and is negligible for muons.
- **Rare SM processes:** Rare SM processes yield same-sign leptons, mostly from ttW and ttZ. We also include the contribution from the SM Higgs boson produced in association with a vector boson or a pair of top quarks in this category of background. With the exception of $t\bar{t}Z$ and $t\bar{t}W$, rares are estimated from simulation.

8.1 Fake leptons

The fake lepton prediction is determined as described in the SUSY AN [1] using the SUSY same-sign analysis baseline selection. The description is also included below, for convenience.

The uncertainties related to this estimate are discussed in Section 9.3. Note that in the case of 0 events in the application region as an input to the extrapolation into the signal regions, the prompt-nonprompt events in simulation are multiplied by the data fake rate to obtain a prediction with associated error. This error is then set as the statistical upper limit on the data-driven prediction of 0.

8.1.1 The fake rate method

Background from fake leptons is estimated with the "fake rate" method. The number of events in the sample with at least one lepton that passes a loose selection but fails the full set of tight identification and isolation requirements (application region) is weighted using the "tight-to-loose" ratio, i.e. the probability that a loosely identified non-prompt lepton also passes the full set of requirements. This probability is measured as a function of lepton p_T and η, as well as event kinematics, in a control sample of QCD multijet events that are enriched in non-prompt leptons (measurement region). Such region is triggered by the auxiliary triggers defined in Table 4 and requires only one denominator lepton in the event, one recoiling jet with $\Delta R(jet, lep) > 1.0$ and low MET and MT to suppress the contribution from W and Z. The main systematic effects are the non-universality of the "tight-to-loose" ratio, particularly due to the dependency from the mother parton p_T and the flavor composition of the sample, and the prompt contribution in the measurement region.

In the SUSY same-sign 13 TeV analyses, developments were deployed [9][17] in the fake rate estimation in order to reduce the dependency to the mother parton p_T by using a new proxy of the lepton p_T:

$$\begin{aligned} if\ p_T^{rel} > I_3 : \quad & p_T \rightarrow \quad p_T \cdot (1 + max(0, I_m - I_1)) \\ else : \quad & p_T \rightarrow \quad max(p_T, p_T(jet) \cdot I_2) \end{aligned} \tag{5}$$

In addition to the corrected lepton p_T definition, another development has been introduced since [17]: for the electrons, the flavor dependency of the tight-to-loose ratio is reduced by extrapolating on both isolation and lepton MVA ID. These improvement are included in the 2016 and in the current (2016+2017) $t\bar{t}t\bar{t}$ analyses.

8.1.2 Fake Rate measurement

We derive different versions of the fake rate for muons and electrons, collected with triggers with an isolation requirement. The numerator and denominator selections are defined as in section 4.5.

Events for fake rate measurement need to pass the following requirements:

- pass a specific auxiliary trigger, described in Table 4
- only one denominator lepton (FO)
- at least one jet with ΔR(jet,FO)>1
- $E_{\mathrm{T}}^{\mathrm{miss}}$ <20 GeV, M_T <20 GeV.

The requirements on $E_{\mathrm{T}}^{\mathrm{miss}}$ and M_T are intended to suppress the contribution from prompt leptons in the measurement region. Such contribution is subtracted from the fake rate using DY,WJets, and $t\bar{t}$ Monte Carlo samples. For the 2016 data, these samples were normalized in the control region defined with $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV, 70< M_T <120 GeV and a tight lepton. For the 2017 data, due to the increased PU contribution, an improved normalization technique is used, based on template fits of the full M_T distribution. The M_T distribution in data with $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV and lepton $p_{\mathrm{T}} > 20$ GeV is fitted with the sum of two templates derived from MC, one for QCD and one for the sum of the electroweak processes (DY, WJets and tt). The normalization of the electroweak processes is extracted from the fit and half the difference between the normalization and unity is taken as an uncertainty. As the shape of the non-prompt component may not we well-modeled, we repeat the fit replacing the QCD MC template with a data-driven template extracted from events failing the isolation cut. The difference between the two fits is taken as an additional uncertainty in the normalization, added in quadrature with the uncertainty obtained from the fit above.

The nominal and alternative fits for electrons and muons are shown in Figure 25 (2017) and Figure 26 (2018), and the resulting normalization corrections for the electroweak samples are shown in Table 12. Statistical uncertainties on the measurement are assumed to be negligible. Results with 35.9 fb^{-1} of 2016 data are shown in [1], while the 2017 and 2018 results with 41.5 fb^{-1} and 59.6 fb^{-1} of data are shown in Figs. 27-29,28-30 below.

		isolated		non-isolated	
	template	e	μ	e	μ
2017	MC	1.215	1.222	1.208	1.202
	data	1.277	1.195	1.298	1.178
2018	MC	1.200	1.283	1.202	1.288
	data	1.268	1.252	1.297	1.250

Table 12: Normalization scale factors for electroweak samples derived with two different M_T templates for QCD: MC and data (the data template refers to the inverted isolation region).

8.1.3 Fake Rate closure in MC: $t\bar{t}$ and W+jets: 2017/2018 samples

Using these definitions, we tested the closure of the method in the baseline and signal regions and, inclusively in lepton p_{T}, for the most relevant kinematic distributions in 31 and 32. the level of closure obtained is typically at 30% or better, similar to the 2016 one. The closure in the electron channel showed a potential trend at high p_{T}, with deviations up to 60%.

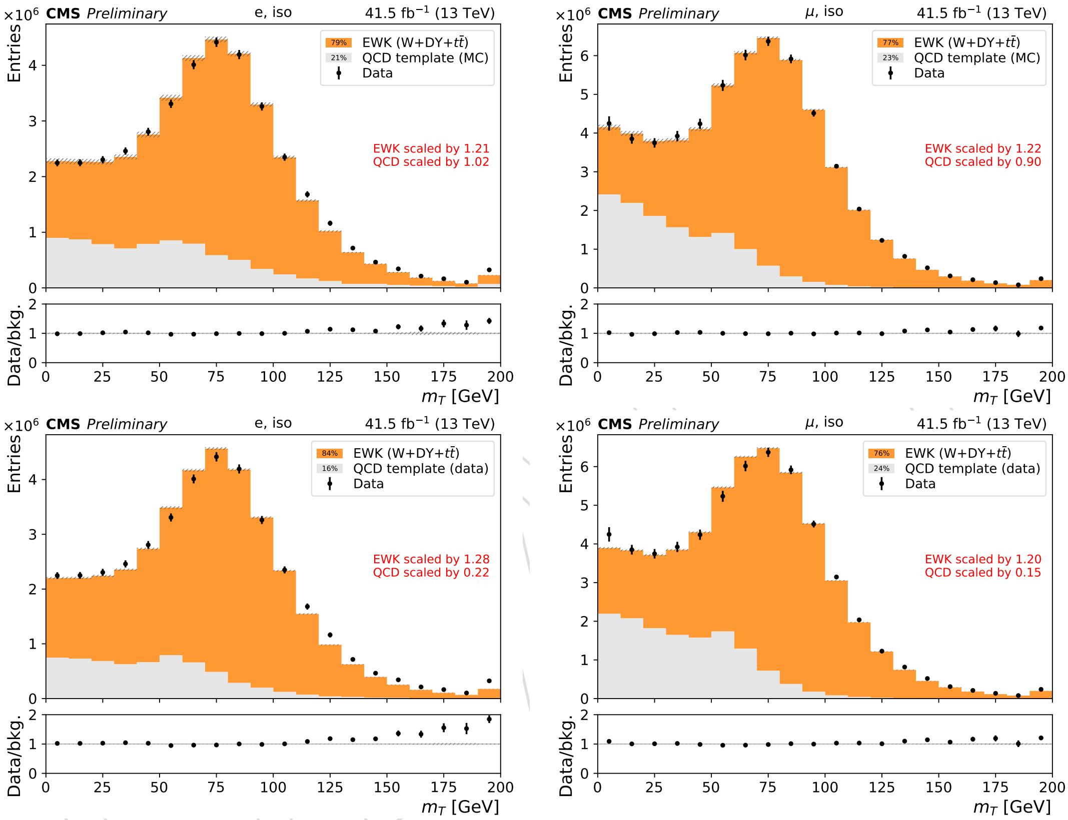


Figure 25: Isolated triggers, 2017 Data and MC: Fitted M_T distribution used to derive the normalization of electroweak samples (DY, WJets, tt) in the fake rate measurement region. Electrons are shown on the left, muons on the right. From top to bottom, the results from the nominal selection ($E_T^{miss} > 30$ GeV and lepton $p_T > 20$ GeV) with the QCD MC template and alternative data QCD template are shown.

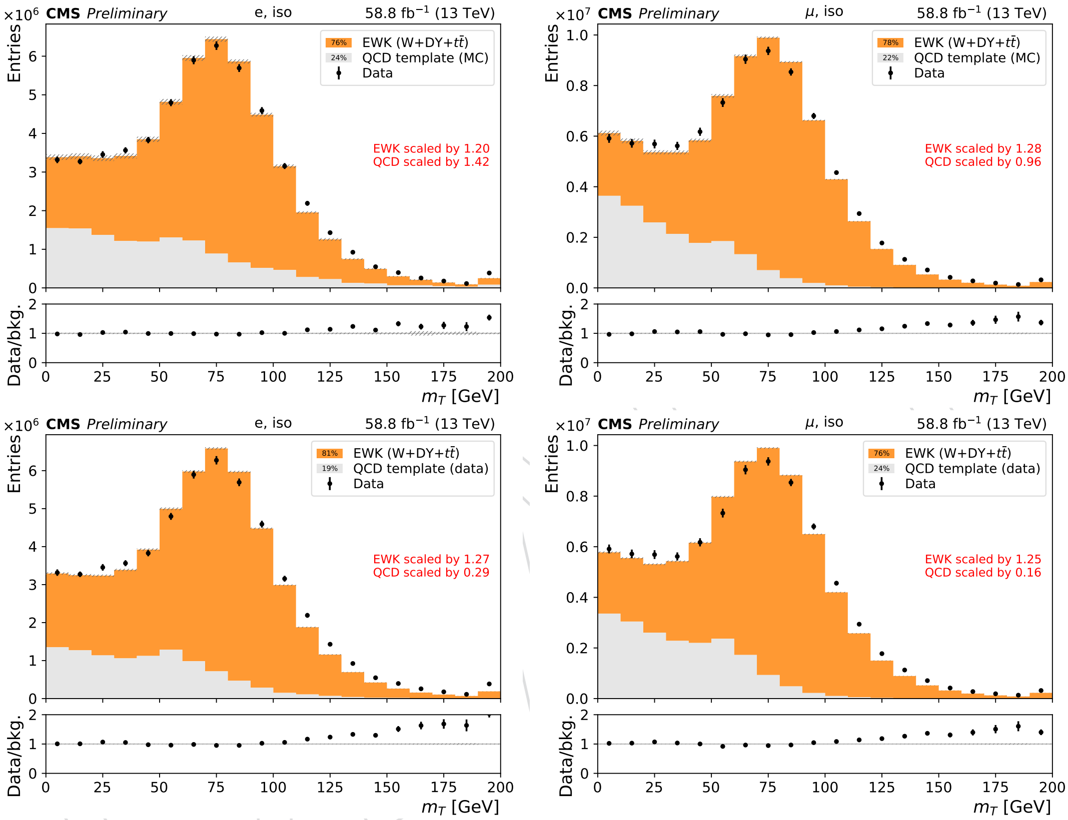


Figure 26: Isolated triggers, 2018 Data and MC : Fitted M_T distribution used to derive the normalization of electroweak samples (DY, WJets, tt) in the fake rate measurement region. Electrons are shown on the left, muons on the right. From top to bottom, the results from the nominal selection ($E_T^{miss} > 30$ GeV and lepton $p_T > 20$ GeV) with the QCD MC template and alternative data QCD template are shown.

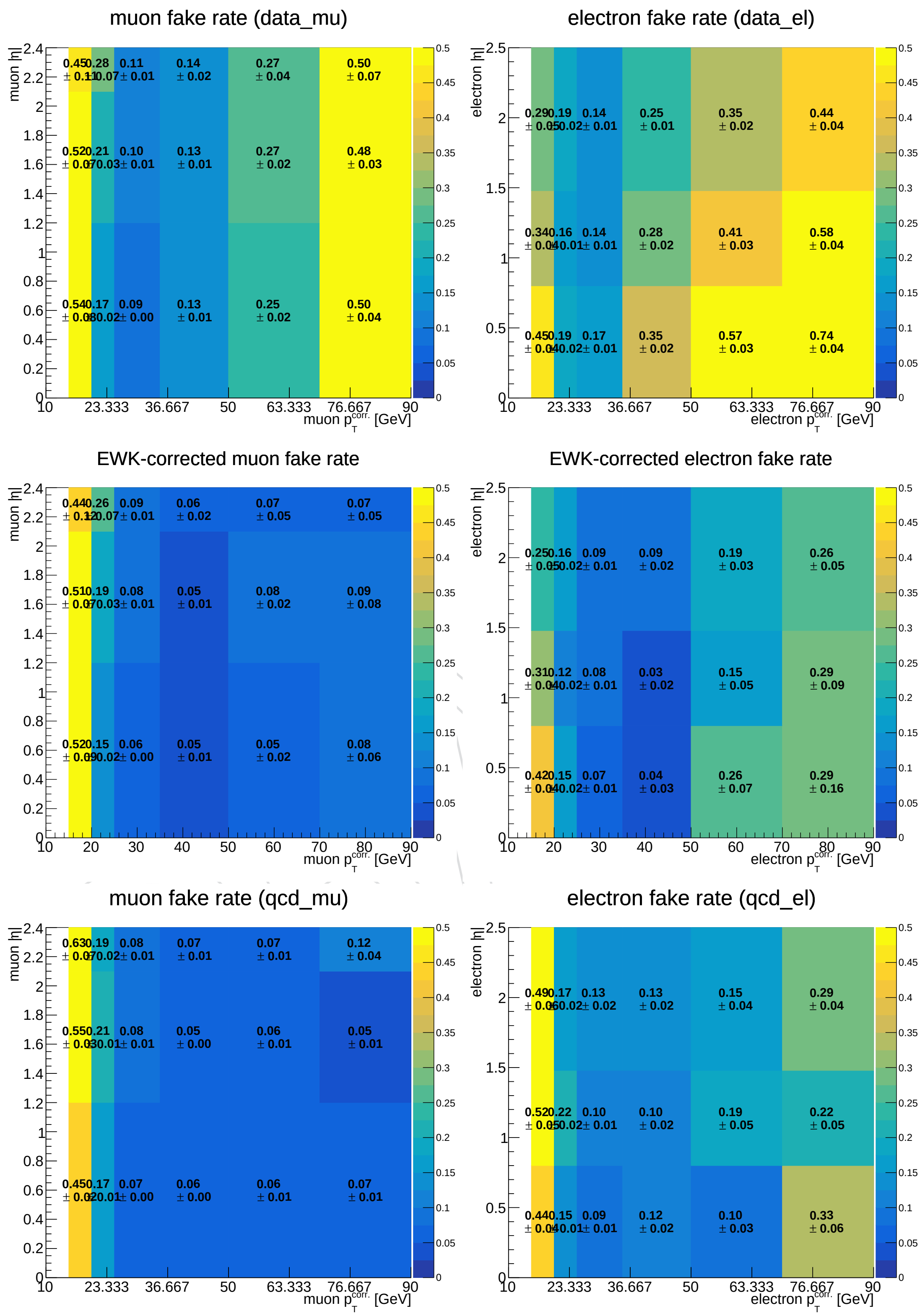


Figure 27: Isolated triggers: Fake rate for muons (left) and electrons (right) for: 2017 data uncorrected (top), 2017 data corrected for EWK contribution (middle) and 2017 QCD MC (bottom). Uncertainties are only statistical.

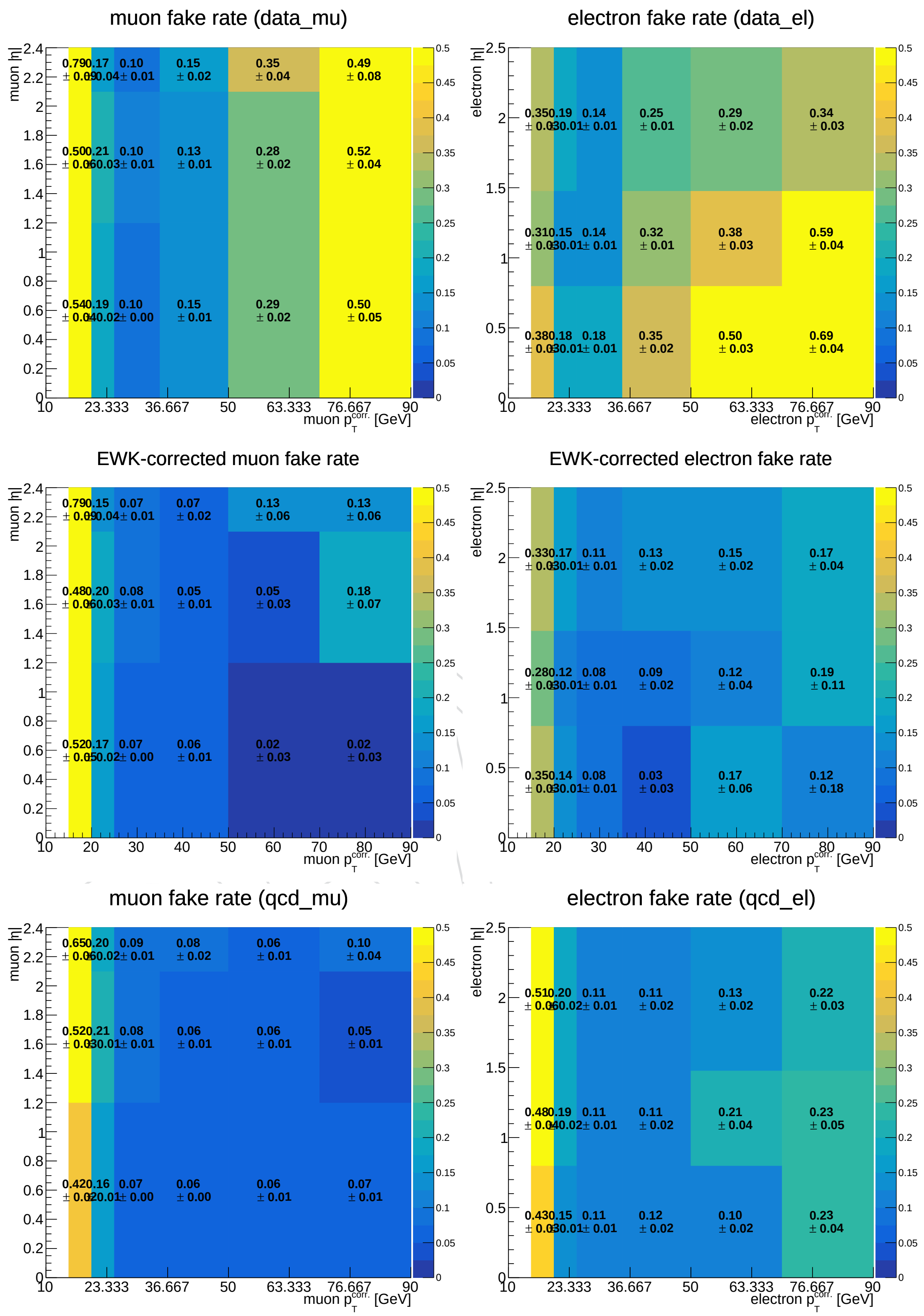


Figure 28: Isolated triggers: Fake rate for muons (left) and electrons (right) for: 2018 data uncorrected (top), 2018 data corrected for EWK contribution (middle) and 2017 QCD MC (bottom). Uncertainties are only statistical.

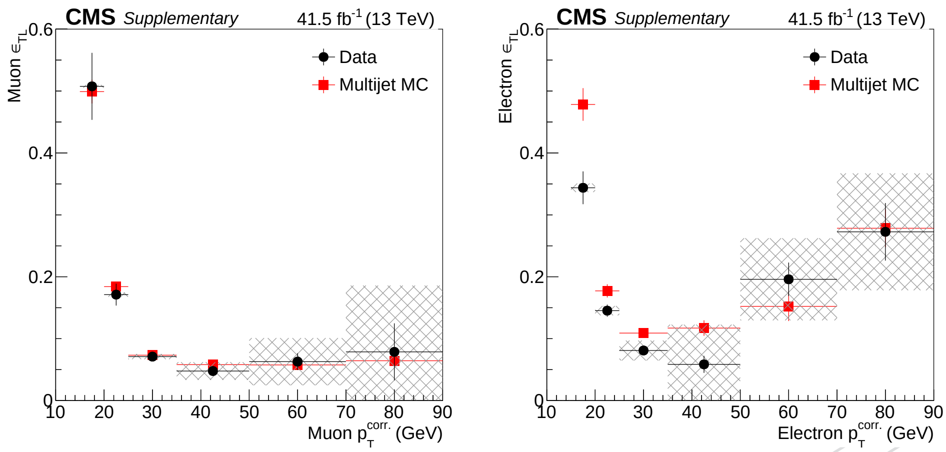


Figure 29: EWK-corrected data fake rate projected vs p_T for 2017 data (black) and 2017 QCD MC (red), for muons (left) and electrons (right). The shaded band in the projection is the systematic uncertainty related to the EWK contamination.

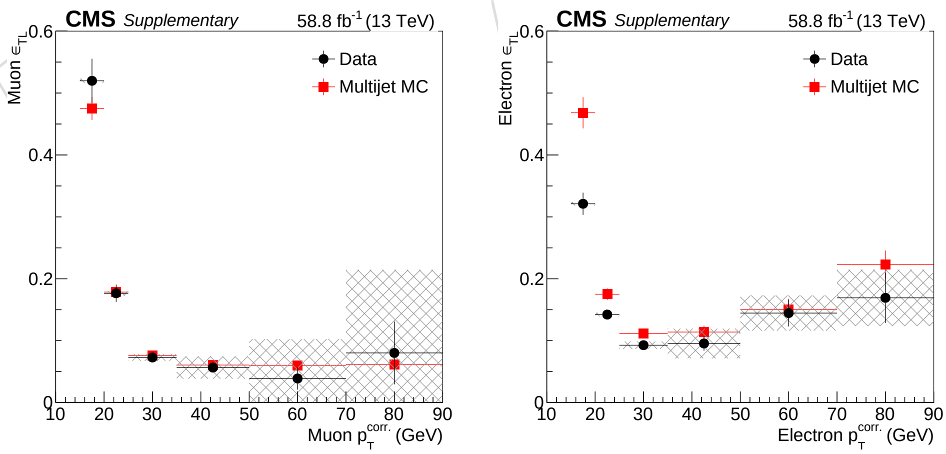


Figure 30: EWK-corrected data fake rate projected vs p_T for 2018 data (black) and 2017 QCD MC (red), for muons (left) and electrons (right). The shaded band in the projection is the systematic uncertainty related to the EWK contamination.

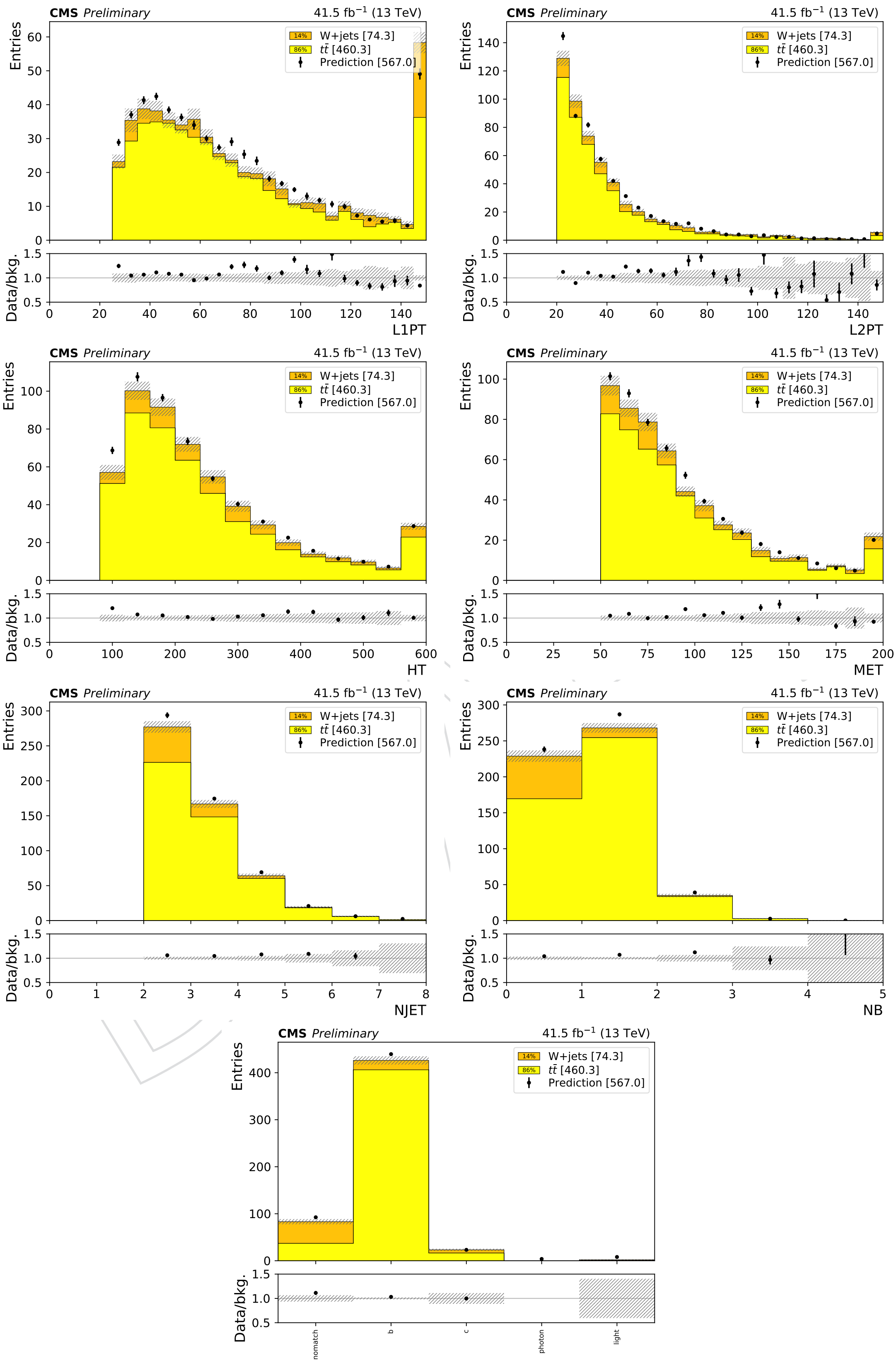


Figure 31: 2017 MC: Electron+muon fake rate closure for QCD measurement in MC soup.

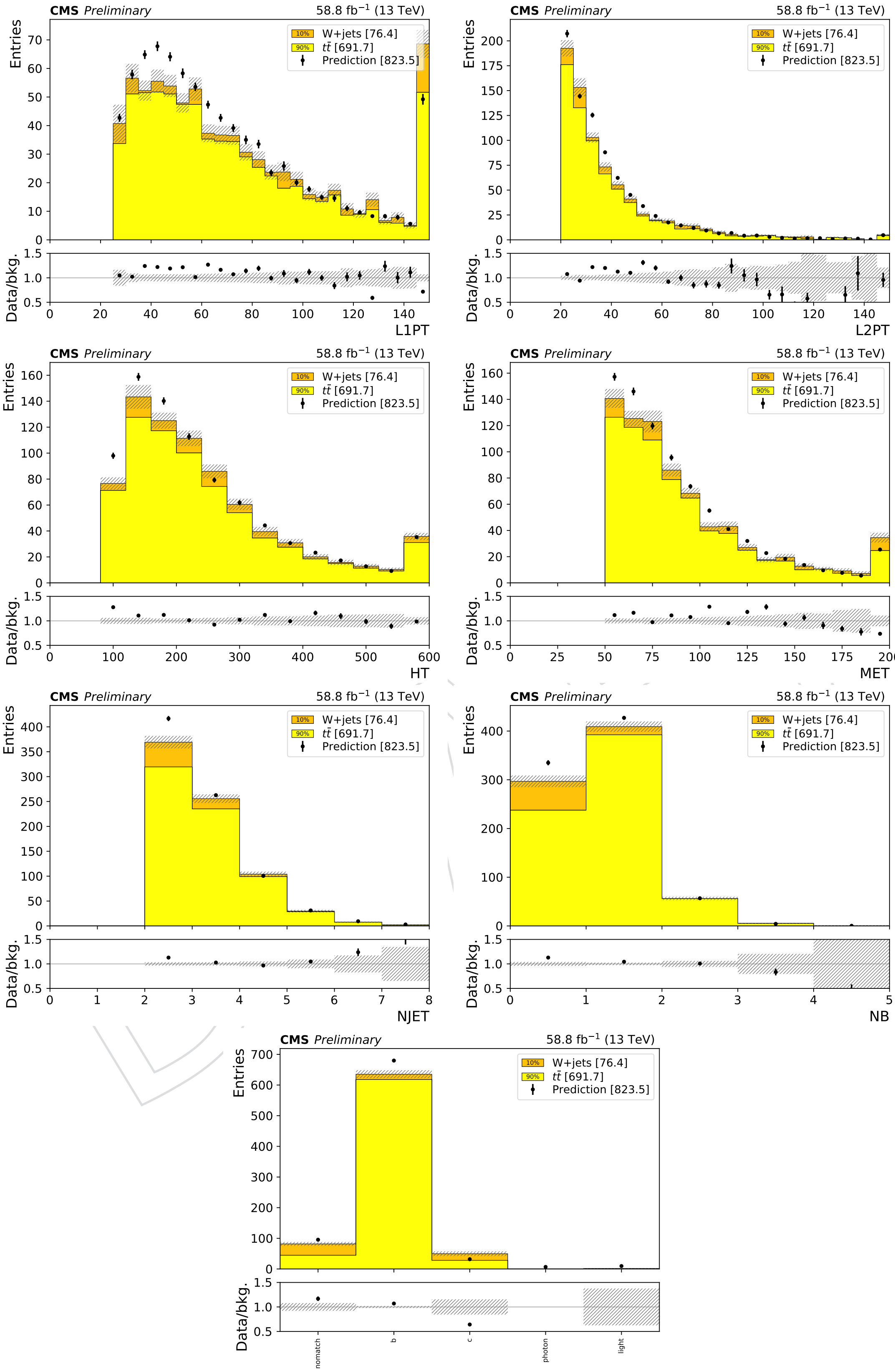


Figure 32: 2018 MC: Electron+muon fake rate closure for QCD measurement in MC soup. QCD Fake rate to be updated. 2018 bcToE QCD MC is not available yet

8.2 Charge misidentification

The charge flip prediction is determined as described in the SUSY AN [1] using the SUSY same-sign analysis baseline selection. The description is also included below, for convenience. The uncertainties related to this estimate are discussed in Section 9.3.

The background due to charge flips is estimated by selecting opposite-sign ee or eμ events passing the full kinematic selection and then weighting them by the p_{T} and η-dependent probability of electron charge mismeasurement

This probability, shown in Fig. 33, is obtained from a soup of TTbar and DY simulation and is then validated with a control data sample of same-sign Z$\rightarrow$ee events, using a $E_{\mathrm{T}}^{\mathrm{miss}} < 50\,\mathrm{GeV}$ requirement to be orthogonal to the signal region. The level of agreement in this control region is used to gauge the associated systematic uncertainty and to derive a correction to the MC-based rate estimation. In the 2016 data, we find good agreement between prediction and data in the control region [1]. In the 2017 and 2018 data, the MC Flip Rate is significantly lower than the 2016 one due to the upgraded pixel detector. However, the prediction in the same-sign Z$\rightarrow$ee region is found to only be about 30% lower than the observed number of events in this region, as shown in Figure 35. Consequently, the 2017 and 2018 charge misidentification prediction are scaled by nearly 50%, as seen in Figure 34. The yearly scaling is given in Table 13. Since we do not find significant trends in the lepton kinematics, so we do not consider p_{T} and η-dependent corrections. In addition to the statistical uncertainties, we apply a 20% systematic uncertainty on this background prediction for all years. In MC, the flip rate for muons is found to be O(10^{-6}) and is therefore neglected.

year	obs/pred
2016	1.01
2017	1.44
2018	1.41

Table 13: Ratio of observed flip rate in data to the flip rate in simulation. These are the multiplicative correction to the MC-based charge flip probabilities.

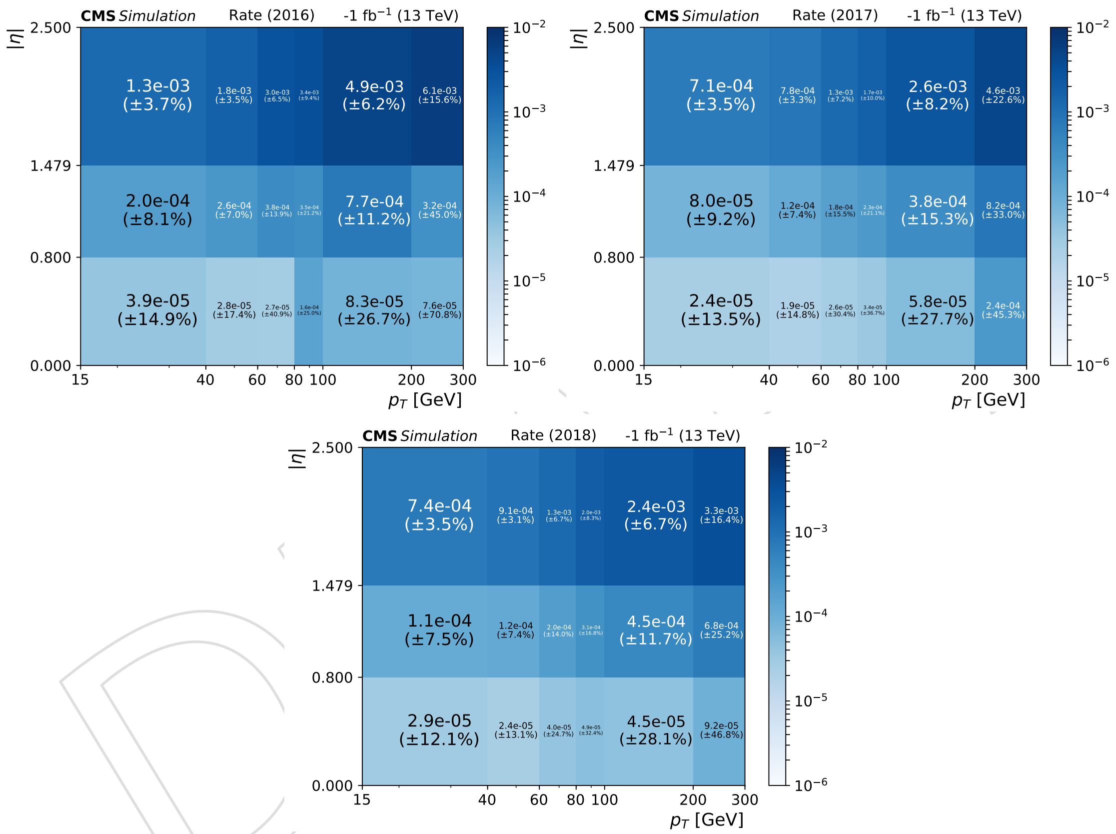


Figure 33: Electron charge flip rate for 2016, 2017, and 2018.

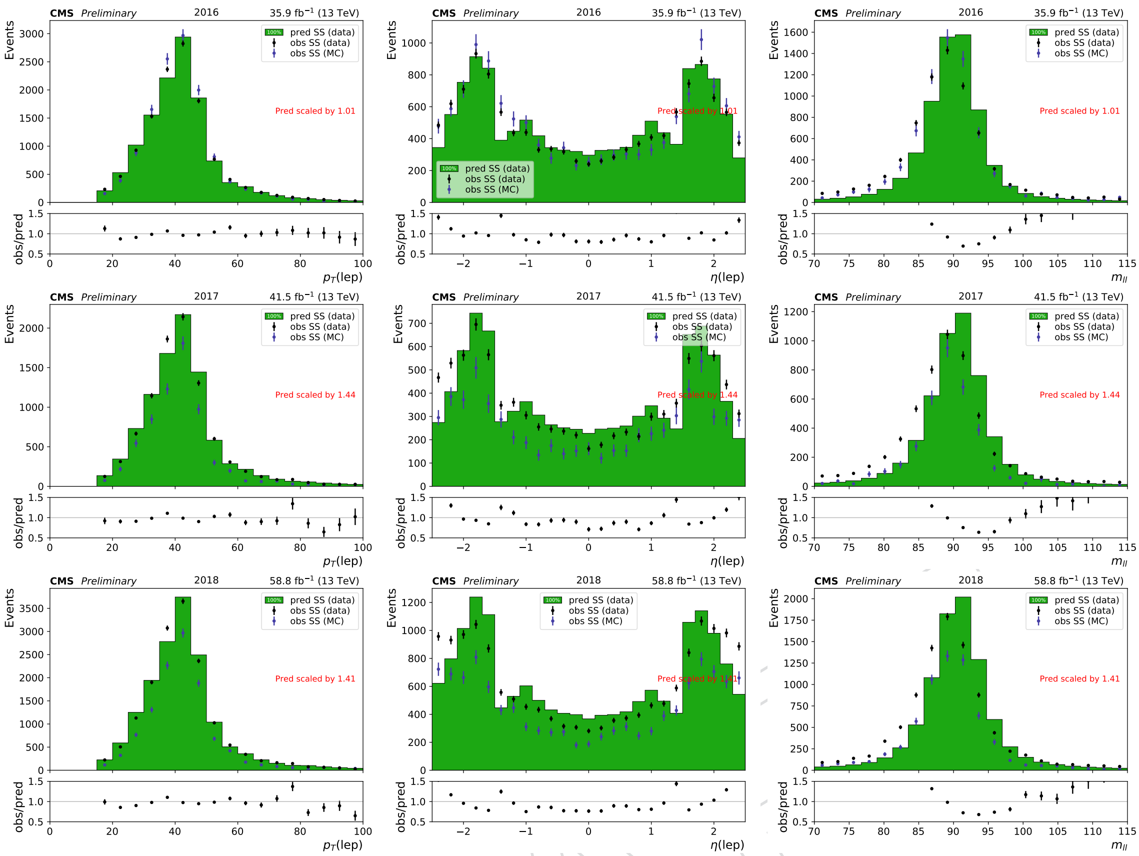


Figure 34: Predicted and observed lepton p_T (left) and η (middle) and $m_{\ell\ell}$ (right) in a same-sign Z→ee peak for years 2016, 2017, and 2018 from top to bottom. The prediction is normalized to the observed data.

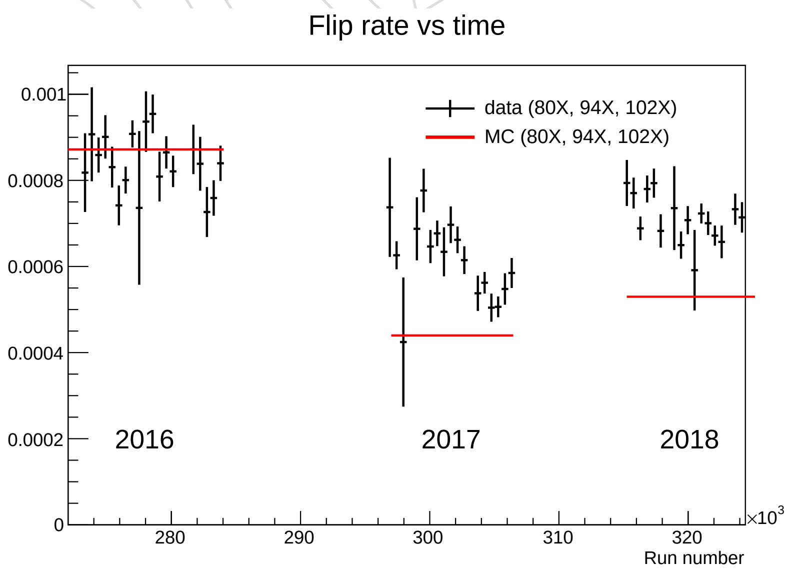
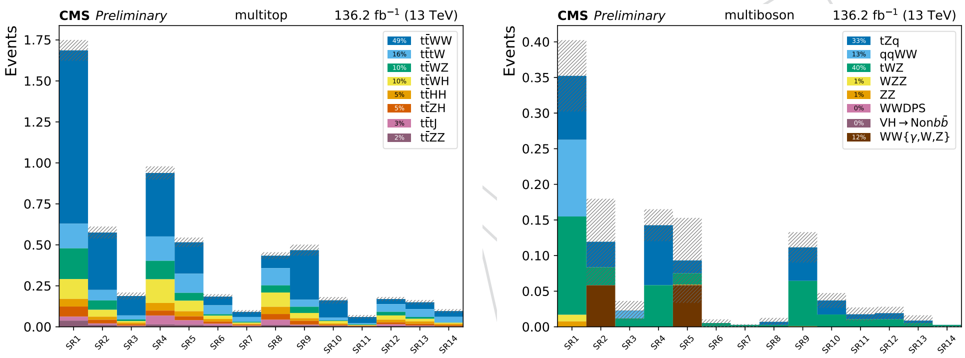


Figure 35: Electron charge flip rate in simulation and data as a function of time

8.3 Rare SM processes

Rare SM processes that result in the production of same-sign leptons are estimated using MC samples. The processes with the largest contributions, ttW, ttZ and ttH, are treated individually and assigned separate uncertainties. Processes with smaller contributions, including diboson (WZ, ZZ) and triboson (WWW, WWZ, WZZ, ZZZ) , Higgs (HZZ, VH), same-sign WW from both single (qqWW) and double-parton scattering (DPS WW), rare top channels (tZq), and ttt+W,ttt+j are grouped in a single category termed "Rare SM" in all plots. Similarly, processes where one of the leptons is an electron from an unidentified photon conversion are grouped in a category called "X+γ" in all plots: these include $W\gamma$, $Z\gamma$, $t\bar{t}\gamma$ and $t\gamma$. Both the "Rare SM" and "X+γ" categories are assigned a large ($\pm$50%) theoretical normalization uncertainty. The breakdown of the individual processes of the "Rare SM" category is shown in Figure 36 The theory uncertainties considered are described in more detail in Sec. 8.4.

In addition to the theoretical uncertainties, all samples are assigned uncertainties based on reconstruction, as summarized in Section 9, due to JES, b-tagging, lepton and trigger scale factors uncertainties and luminosity. Finally, the statistical uncertainty of the MC samples is also taken into account.



(a) (b)

Figure 36: Relative composition of multi-top (left) and multi-boson (right) rare backgrounds in the signal regions for all MC.

8.4 Theoretical uncertainties on TTW, TTZ and TTH

For the largest SM backgrounds, we carefully assign uncertainties on the overall normalization and on the shape across signal regions.

8.4.1 Normalization

For $t\bar{t}W$ and $t\bar{t}Z$, a 40% normalization uncertainty is applied, but the results are not sensitive to this initial choice, because the respective control regions included in the maximum likelihood fit, discussed in more detail in Sec. 11, are used to further constrain the normalization of these backgrounds.

For $t\bar{t}H$, the normalization uncertainty used in the 2016 analysis was 50%, to cover the 1.5 signal strength observed in the results of HIG-17-004. For the full Run2 analysis, we rely on the updated measurement of HIG-17-035, which finds a signal strength of $1.26^{+0.31}_{-0.26}$. We therefore apply a 25% normalization uncertainty.

8.4.2 Shape

To evaluate the theory uncertainties on the shape, we explore scale and PDF variations, as well as differences between LO and NLO samples, where both are available. These studies were performed with 2016 MC.

The PDF shape uncertainties for these samples are generally smaller than the MC statistical uncertainties on the background yield in each region, so no additional shape uncertainty is assigned to account for them.

The NLO scale uncertainties for these samples are fairly stable at around 15% for ttW, ttZ, and ttH, as shown in Figures 38, 37, 39. With the exception of ttH, these figures also show the LO scale uncertainty, which covers well the difference between LO and NLO. The LO agreement with the NLO shapes, within LO scale uncertainty, gives confidence in the NLO scale uncertainties as they are evaluated using the same variations of renormalization and factorization scales.

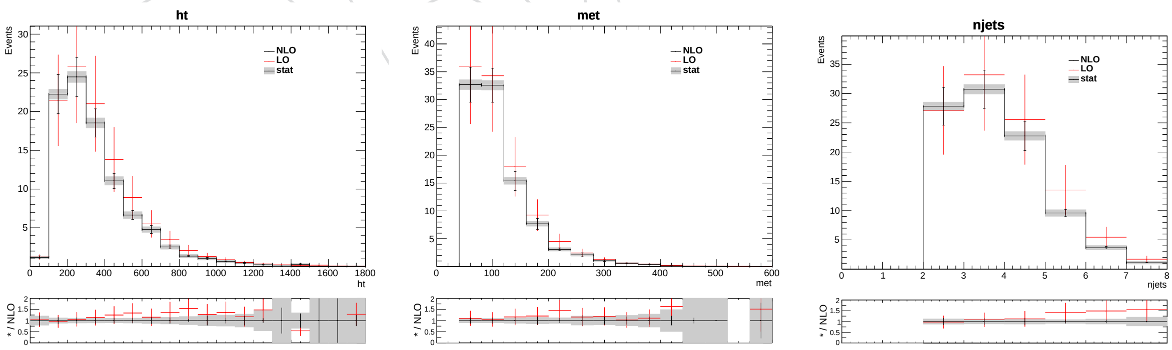


Figure 37: Comparison of LO and NLO kinematics for ttZ with scale uncertainties in the baseline region and statistical uncertainties on NLO. The error bars on the NLO (black) are the NLO uncertainties, while the error bars on the LO (red) are the LO uncertainties. The gray band is the stat uncertainty on the NLO

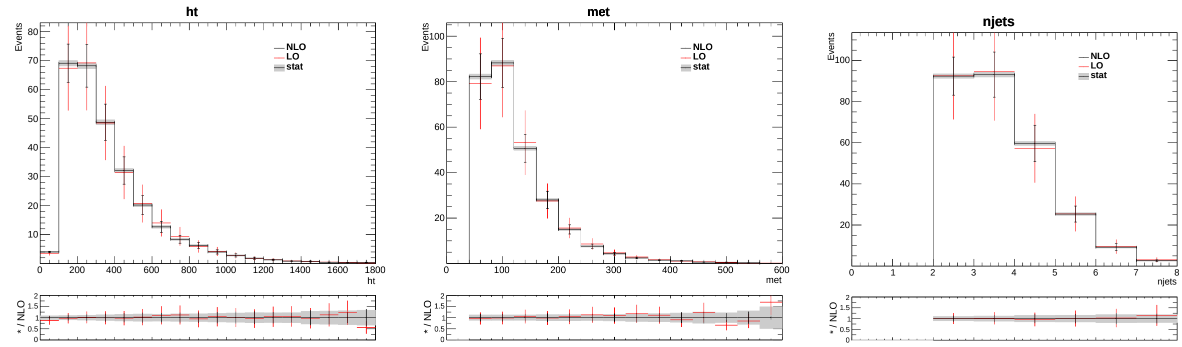


Figure 38: Comparison of LO and NLO kinematics for ttW with scale uncertainties in the baseline region and statistical uncertainties on NLO. The error bars on the NLO (black) are the NLO uncertainties, while the error bars on the LO (red) are the LO uncertainties. The gray band is the stat uncertainty on the NLO

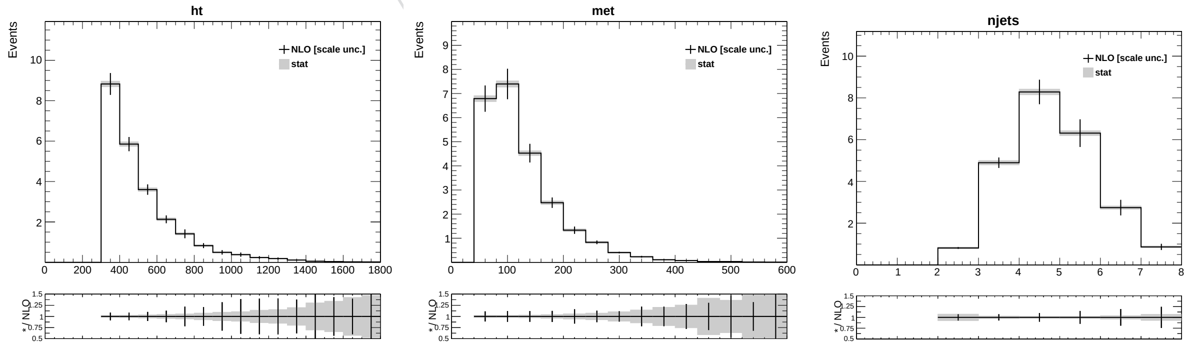


Figure 39: ttH NLO with NLO scale uncertainties in the baseline region. The gray band is the stat uncertainty on the NLO

9 Systematic Uncertainties

With respect to the 2016 SUSY same-sign analysis [1], the 2016 tttt analysis only changed uncertainties on $t\bar{t}Z$, $t\bar{t}H$, and $t\bar{t}W$, including also uncertainty for $t\bar{t}W$+bb,$t\bar{t}Z$+bb, JER, and adding ISR/FSR variations from dedicated samples. With respect to the 2016 tttt analysis [4], the cross-section for $t\bar{t}H$ has a 25% uncertainty taken as a systematic. Other uncertainties, which have changed as a result of new measurements, are marked in blue if the change is also applied to 2016 data. All uncertainties marked in black have remained unchanged since 2016. These changes are reflected in Table 14.

9.1 Correlation model for 2016+2017+2018

In order to combine the 2016+2017+2018 datasets, we have to make assumptions about the correlation of uncertainties in the different years. Since the analysis is statistically dominated, these correlations are expected to have a very small influence. The current model we have chosen makes the following assumptions,

- Statistical uncertainties (FO CR stat, OS CR stat, MC stat): uncorrelated
- Luminosity, JEC, JER, b-tag, PU, lepton eff., HLT, eff.: uncorrelated (derived from

source	magnitude	effect on yield
luminosity	2.5, 2.3, 2.5%	2.5, 2.3, 2.5% (2016-2018)
JES	$1-8\%$	$1-15\%$
JER	$1-10\%$	$1-10\%$
ttW, ttZ ISR/FSR	$< 15\%$	$1-8\%$
ttW/Z+bb	$< 30\%$	$< 15\%$
b-tag efficiency	$\sim 1-20\%$	$1-15\%$
pileup	5%	0-5%
lepton efficiencies	$2-5\%$ (leg)	$2-10\%$
HLT efficiencies	$1-5\%$ (leg)	$2-7\%$
FO CR stat.	$1-100\%$	$1-100\%$ (fake bkg. only)
FR extrapolation	30-60%	30-60% (fake bkg. only)
EWK subtraction in FR	25-50% (FR)	$1-30\%$ (fake bkg. only)
OS CR stat.	$4-100\%$	$4-100\%$ (charge misId. only)
charge misId.	20%	20% (charge misId. only)
ttW norm.	40%	40 % ($t\bar{t}W$ only)
ttZ norm.	40%	40 % ($t\bar{t}Z$ only)
ttH norm.	25%	25% ($t\bar{t}H$ only)
MonteCarlo stat.	$1-25\%$	$1-25\%$
QCD scales and PDFs	$\times 0.5/\times 2$	$10-20\%$ ($t\bar{t}W$,$t\bar{t}Z$,$t\bar{t}H$)
other bkgs.	50%	50% (Rare and $X+\gamma$)

Table 14: Summary of the sources of uncertainties, their magnitude and their effects. The second column indicates the magnitude of the yield variation. Reported uncertainties are representative for the most relevant signal regions.

different datasets)
- ISR/FSR: uncorrelated (derived from different datasets and with respect to different MC samples)
- ttW/Z+bb, ttH normalization: correlated (based on theory predictions and CMS measurements)
- ttW and ttZ normalizations: correlated (based on theory predictions and consistent Control Regions)
- Rare and X+γ normalizations: correlated (based on theory predictions)
- QCD scale and PDF uncertainties: correlated (based on theory predictions)
- EWK subtraction in FR: uncorrelated (based on differently prescaled triggers in data)
- FR extrapolation: correlated (method is the same, closure in MC shows the same features)
- Charge misid.: uncorrelated (agreement in validation region is quite different, probably due to different Pixel detector conditions)

We ran limits under three correlation assumptions with 18 cut-based bins. With respect to the nominal correlation model, fully uncorrelating (correlating) nuisances between the 3 years decreases (increases) the expected significance by 2%.

9.2 Experimental sources of uncertainties

One of the main experimental sources of uncertainty is the knowledge of the jet energy scale (JES), affecting all the simulated backgrounds and considered signals. The 13 TeV uncertainties vary the jet energy scale by 1–8%, depending of the transverse momentum and pseudorapidity of the jet. The impact of these uncertainties is assessed by shifting the jet energy correction factors for each jet up and down by $\pm 1\sigma$ before the calculation of all kinematic quantities. The variations are correlated among the different signal regions as bin-by-bin migration is allowed. Variations are asymmetric in nature and are used as asymmetric nuisance parameters in the result interpretation. The JES uncertainties are propagated to the missing transverse energy and all jet-related variables (number of jets, H_{T}, number of b-jets) used in this analysis. As some of the simulation samples are statistically limited, the size of the JES uncertainties can reach high values in several regions not well populated by several processes. As usually those variations impact mostly background samples that do not contribute much to those regions, we consider large variations as they are obtained out of the box. Most populated signal regions shows yield variations of 8% when the jet energy scale is varied by one standard deviation.

The uncertainty related to the knowledge of the jet energy resolution (JER) can be considered as well for the simulated backgrounds and the signals. The effect of smearing jets on the acceptance of events is assessed in the same manner as the JES uncertainty.

A similar approach is used for the uncertainties associated to the corrections for the b-tagging efficiencies for light and bottom flavor jets, which are parametrized as a function of p_{T}, η. The variation of the scale factor is at maximum of the order of 1–20% per jet, and lead to an overall effect on yield included between 1 and 15% depending on the signal region and on the topology of the events included in those signal regions. If considering only highly populated signal regions to get an overview of the main effects on the background yields, the bulk of the $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}H}$ yield varies by $\sim$8% and the $\mathrm{t\bar{t}Z}$ yield by $\sim$6%.

When applying b-tagging efficiency scale factors, it is possible to apply period-dependent scale factors, rather than an average scale factor for the entire 2016 dataset. This is relevant for signal regions where several mistags from light jets are expected, since the mistag SFs for light jets are time-dependent. When checking the flavor composition of our main background $t\bar{t}W$, as in Figure 40, we find that none of our $N_{\mathrm{b\,jets}}$ bins are dominated by events with multiple light jets, so the period-dependent scale factors are not necessarily warranted. As expected, when we do apply them, we find negligible differences in our background predictions with respect to the nominal scale factors. For 2017, we apply period-dependent b-tagging SFs.

Events from processes such ttW and ttZ, which have two b quarks from the decay of the two top quarks, primarly enter regions requiring three or more b-tagged jets by either the mis-tag of a charm quark from the hadronic decay of a W boson or from the production of additional heavy flavor, primarily from an ISR or FSR gluon which splits to a $b\bar{b}$ pair.

In 2016, events in simulation containing additional HF quarks not from top or W decays were scaled up to account for $\mathrm{SF}_{\sigma(t\bar{t}b\bar{b})/\sigma(t\bar{t}j\bar{j})} \sim 1.7 \pm 0.6$ as measured by TOP-16-010. In the high Nbtag signal region bins where this effect is dominant, this resulted in a systematic uncertainty up to 20%. In 2017 and 2018, we apply the same correction as 2016 onto $t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$, and $t\bar{t}H$, as it is observed to bring agreement in the high statistics opposite-sign $t\bar{t}$ control region.

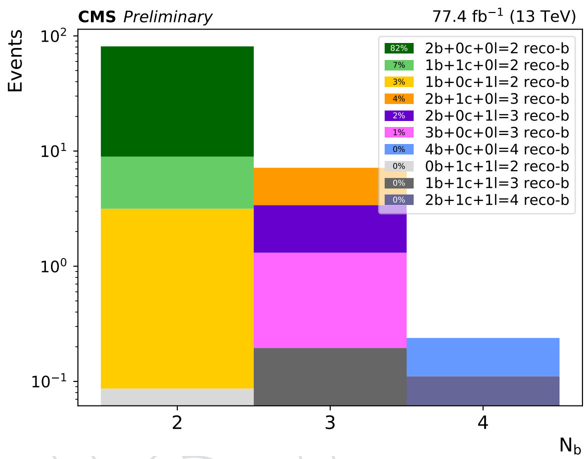


Figure 40: Flavor composition breakdown of the $t\bar{t}W$ sample, in bins of number of b-tagged jets for all MC

Trigger efficiency scale factors obtained from data are applied to correct the simulation.

Lepton efficiency scale factors [13], accounting for differences between the Data and MC for reconstruction and identification of electrons and muons, are applied to all MC events. They result in uncertainties of approximately 2-3% for muons and 3-5% for electrons for 2016.

The simulation is reweighted to match the expected data distribution in the number of collision per events; the uncertainty on the minimum bias cross-section is propagated to the final yields with an effect at the level of 4% or less across the three years.

The $t\bar{t}W$ and $t\bar{t}Z$ simulation is reweighted to match the number of additional ISR/FSR jets observed in data, as described in Section 2. An uncertainty equivalent to 50% of the difference between 1 and the reweighting factor is applied to this procedure. In 2016, since the reweighing factor can be as large as 0.77, this results in an uncertainty of 15% on the highest N_{jets} signal regions. In 2017 and 2018, the reweighing factor is as large as 1.4, which results in a similar uncertainty to 2016 on the highest N_{jets} signal regions.

9.3 Uncertainties from data-driven background estimations

The following uncertainties are defined for the data-driven background estimations, based on the statistics of the control region and the extrapolation from control to signal region.

For the nonprompt lepton and charge misidentified lepton backgrounds, the statistical uncertainty from the control sample varies greatly depending on the signal region considered. Both the flip rate and the fake rate are smaller than 1, so the statistical uncertainty is always smaller (or much smaller, for charge misidentification) than the size of the predicted background. In addition to the statistical uncertainty, the nonprompt lepton background is assigned an overall normalization uncertainty of 30%, based on a comparison of the non-closure of two alternative methods (nominal fake-rate and in-situ fake-rate). This uncertainty is increased to 60% for electrons with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$, to account for the trends observed at high p_{T} in the electron closure for both methods. On top of the uncertainties mentioned above, the nonprompt lepton background prediction also includes an uncertainty related to the electroweak contamination subtraction in the region where the fake rate is computed. To evaluate this uncertainty the fake-rate is computed based on varying the size of the electroweak contamination, and the effect is propagated through the whole analysis. The overall effect on the nonprompt lepton background yield lies between 1 and 50% depending of the signal region considered. The charge misidentified lepton background is assigned a systematic uncertainty of 20%.

For $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$, the respective control regions are used only to set the overall normalization, while the shape is taken from simulation with the experimental uncertainties defined above and the theoretical uncertainties described in section 8.4. The two control regions are included in the fit with a 40% normalization nuisance parameter, and the fit constrains this parameter to approximately 20% given the statistics of the control region, where the corresponding post-fit constraint was 30% in the 2016 analysis.

9.4 Systematic uncertainties from statistical sources

The statistical precision given by the Monte Carlo samples has to be taken into account for all rare SM processes. Within the framework of the HiggsCombine tool, we use the Barlow-Beeston [21] method of handling both MC statistics and data-driven background statistics (for the nonprompt and misidentified-charge backgrounds) via the "autoMCStats" parameter in the datacard.

9.5 Signal uncertainties

The $\mathrm{t\bar{t}t\bar{t}}$ signal is assigned uncertainties based on all the effects described above. The uncertainty values are summarized in Table 15. For the experimental uncertainties we report the range of their effect across SRs, while for the theoretical uncertainties we separate the correlated (acceptance*efficiency, $\mathcal{A}\epsilon$) and uncorrelated (shape) effects. The $\mathcal{A}\epsilon$ of our SR selection, including branching ratio, is $1.5^{+0.02}_{-0.03}\%$, where the uncertainties represent the QCD scale variations. The effect of QCD scale variations on the shape of $\mathrm{t\bar{t}t\bar{t}}$ within the set of SRs is shown in Figure 41: is as large as 10%. PDF uncertainties on acceptance and shape are both negligible as they are smaller than 1%

Uncertainties for ISR and FSR variations during sample generation for $\mathrm{t\bar{t}t\bar{t}}$ are assessed by considering 4 additional samples and are shown in Figure 42. The variation in acceptance (shape) is 1% (<8%) for ISR and 6% (<10%) for FSR.

For the SM $\mathrm{t\bar{t}t\bar{t}}$ measurement, as well as the BSM interpretations related to off-shell effects (low mass particles, yukawa coupling, $\hat{H}$), no uncertainty is assigned to the cross-section itself, since we are measuring this process. For the BSM interpretations related to on-shell effects (heavy scalar and pseudoscalar in 2HDM and dark matter models), the SM $t\bar{t}t\bar{t}$ process is treated as a background and assigned its SM cross section and uncertainty $12.0^{+2.2}_{-2.5}$ fb [6].

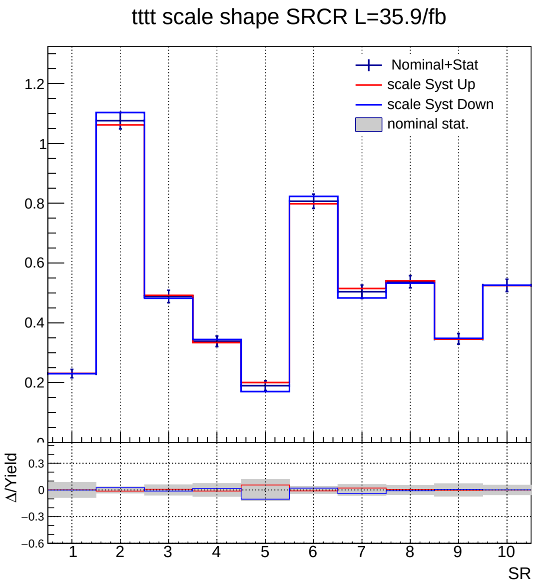


Figure 41: QCD scale uncertainties on the shape of the $t\bar{t}t\bar{t}$ signal.

source	magnitude	effect on yield
jet ES	$1-8\%$	$1-15\%$
JER	$1-10\%$	$1-10\%$
b-tag efficiency	$\sim 1-20\%$	$1-15\%$
pileup	5%	$1-4\%$
ISR	-	<8%
FSR	-	<10%
lepton efficiencies	$2-5\%$ (leg)	$2-10\%$
HLT efficiencies	$1-5\%$ (leg)	$2-7\%$
MonteCarlo stat.	$1-100\%$	$< 8\%$
QCD scales (acceptance)	$\times 0.5/\times 2$	2%
QCD scales (shape)	$\times 0.5/\times 2$	1-10%
PDFs (acceptance)	Envelope	$\leq 1\%$
PDFs (shape)		$\leq 1\%$
α_S (acceptance)		$\leq 6\%$
α_S (shape)		$\leq 1\%$

Table 15: Summary of the sources of uncertainties for $t\bar{t}t\bar{t}$ signal.

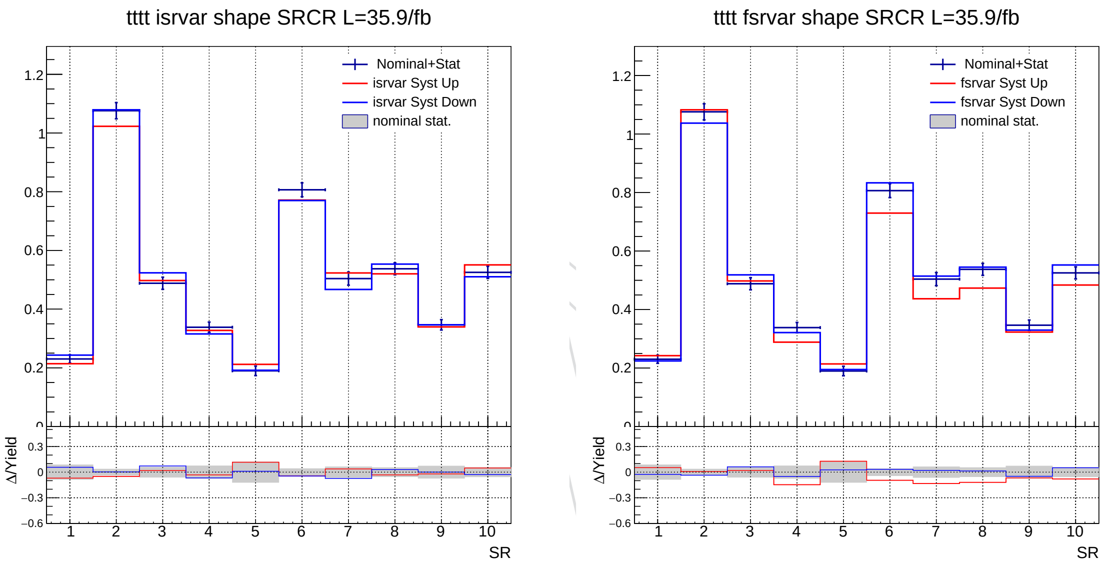


Figure 42: Effect of ISR (left) and FSR (right) variations on the $t\bar{t}t\bar{t}$ signal.

9.6 Summary of theoretical uncertainties on MC-based processes

Table 16 shows a summary of the cross-section, scale, and pdf uncertainties on the MC-based processes, including signal, indicating whether the nuisance contains an acceptance and/or shape component. Associated with each row, there are 3 nuisances, parameterized as one log-normal, and two shape nuisances. Some notes about individual processes follows.

- Signal $\mathrm{t\bar{t}t\bar{t}}$ has no log-normal nuisance associated with cross-section uncertainty because it is being measured.
- Both $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ do not have an acceptance component in the scale/pdf variations because the 40% log-normal is taken to generally cover cross-section$\times$acceptance as they are constrained by dedicated control regions.
- $\mathrm{t\bar{t}H}$ has a 25% cross-section log-normal to reflect the result of HIG-17-035 as mentioned in Section 8.4.
- For X+Gamma and ttVV, the largest cross-section uncertainty on any constituent subprocess is taken on the whole process. Thus, x+gamma obtains an 11% normalization uncertainty driven by $t\bar{t}+\gamma$ (+9.9%, -11.2%), and ttVV obtains a 11% normalization uncertainty from ttWW (+8.6% -11.3%).
- For rares, we obtain uncertainties directly from MC samples by calculating the summed yield of rares for nominal variation, scale up, scale down, pdf up, and pdf down. Adding scale and pdf variations in quadrature yields a 20% variation for 2016 and 2017 samples. We proceed with a log-normal of 20% on the sum of these processes.

	lnN	scale		pdf	
		acceptance	shape	acceptance	shape
$\mathrm{t\bar{t}t\bar{t}}$		x	x	x	x
$\mathrm{t\bar{t}W}$	1.40		x		x
$\mathrm{t\bar{t}Z}$	1.40		x		x
$\mathrm{t\bar{t}H}$	1.25	x	x	x	x
X+γ	1.11	x	x	x	x
$\mathrm{t\bar{t}}VV$	1.11	x	x	x	x
Rares	1.20	x	x	x	x

Table 16: Summary of the theory uncertainties

10 Kinematics

Prefit kinematic distributions of the 2017 and 2018 data events passing the baseline selection and falling into the $t\bar{t}W$ control region are shown in Figs. 43-45. Similarly, distributions of events falling into the $t\bar{t}Z$ control region are shown in Figs. 47-45. The background prediction for the sum of signal regions is shown in Figs. 51-53.

process	SF (2016)	SF (2017)	SF (2018)	SF (Run2)
ttz	1.58	1.005	1.207	1.258
ttw	1.347	1.35	1.156	1.299
tth	1.087	1.089	1.045	1.088
tttt	1.175	0.845	1.451	1.053
fakes	1.064	1.163	1.081	1.125
xg	1.06	1.035	1.015	1.014
rares	1.055	1.017	1.023	1.017
ttvv	1.028	1.018	1.02	1.011
flips	1.016	1.007	0.999	1.001

Table 17: Postfit/prefit scale factors with the BDT analysis. These are the ratio of the normalizations before and after the fit, and they are impacted by all nuisance parameters affecting each background. For this reason, they are not necessarily expected to be consistent between years, as are the individual nuisance parameters themselves (shown in Appendix A).

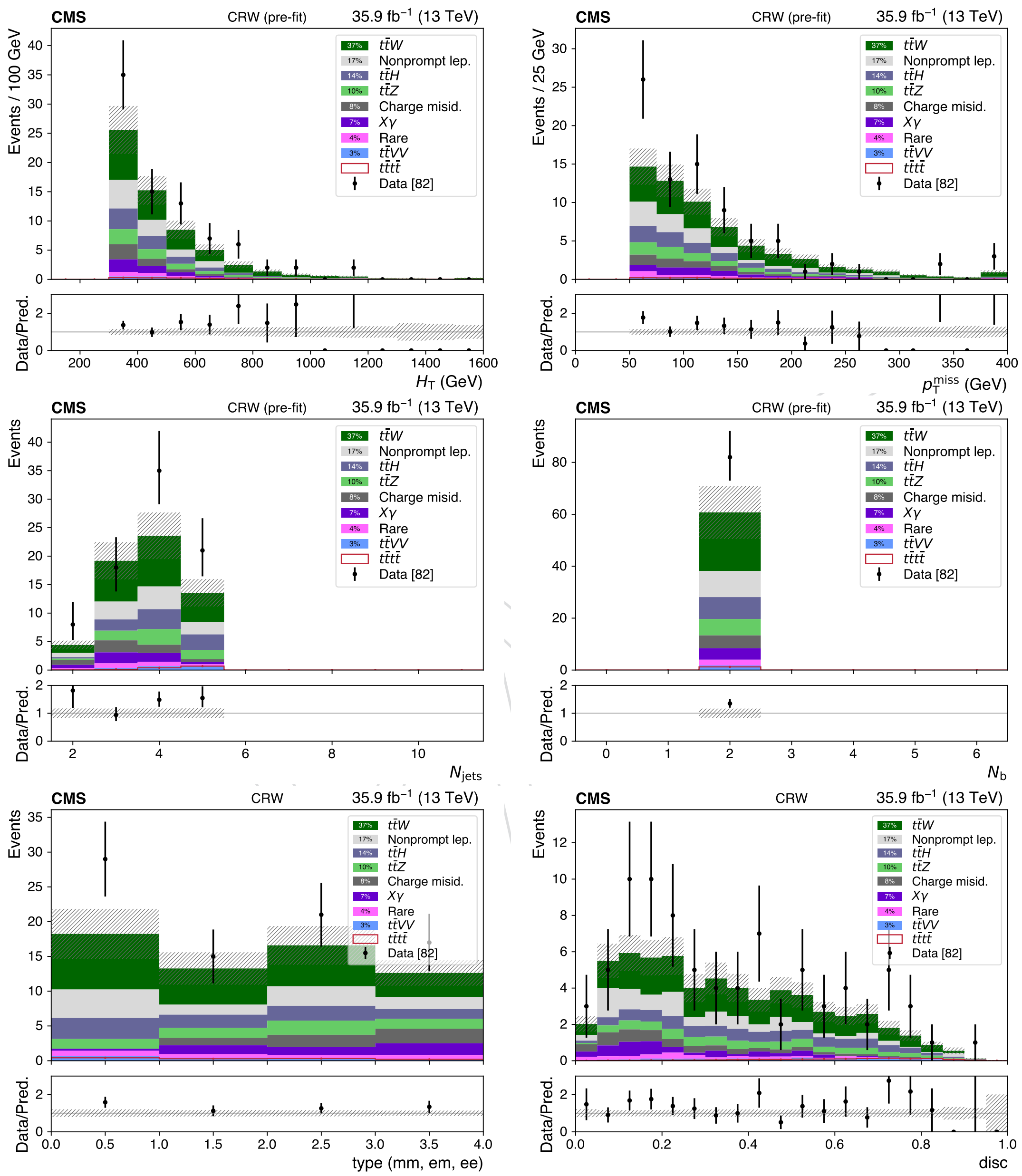


Figure 43: 2016 data and predictions: Prefit distributions of the main analysis variables in the $t\bar{t}W$ control region: H_T, E_T^{miss}, N_{jets}, $N_{b\,jets}$, lepton flavor, and raw BDT discriminant where the last bin includes the overflow.

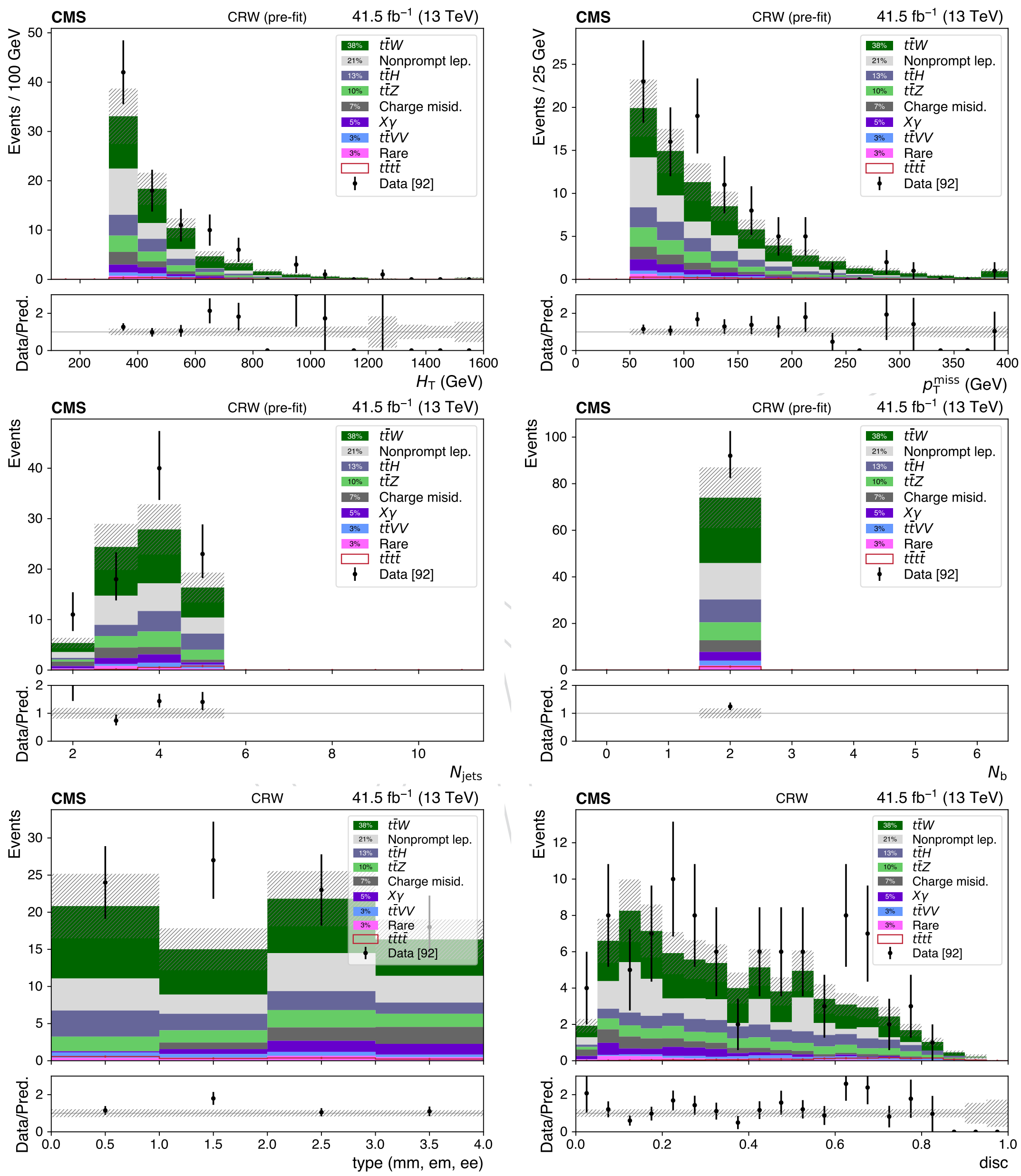


Figure 44: 2017 data and predictions: Prefit distributions of the main analysis variables in the $t\bar{t}W$ control region: H_T, E_T^{miss}, N_{jets}, $N_{b\ jets}$, lepton flavor, and raw BDT discriminant where the last bin includes the overflow.

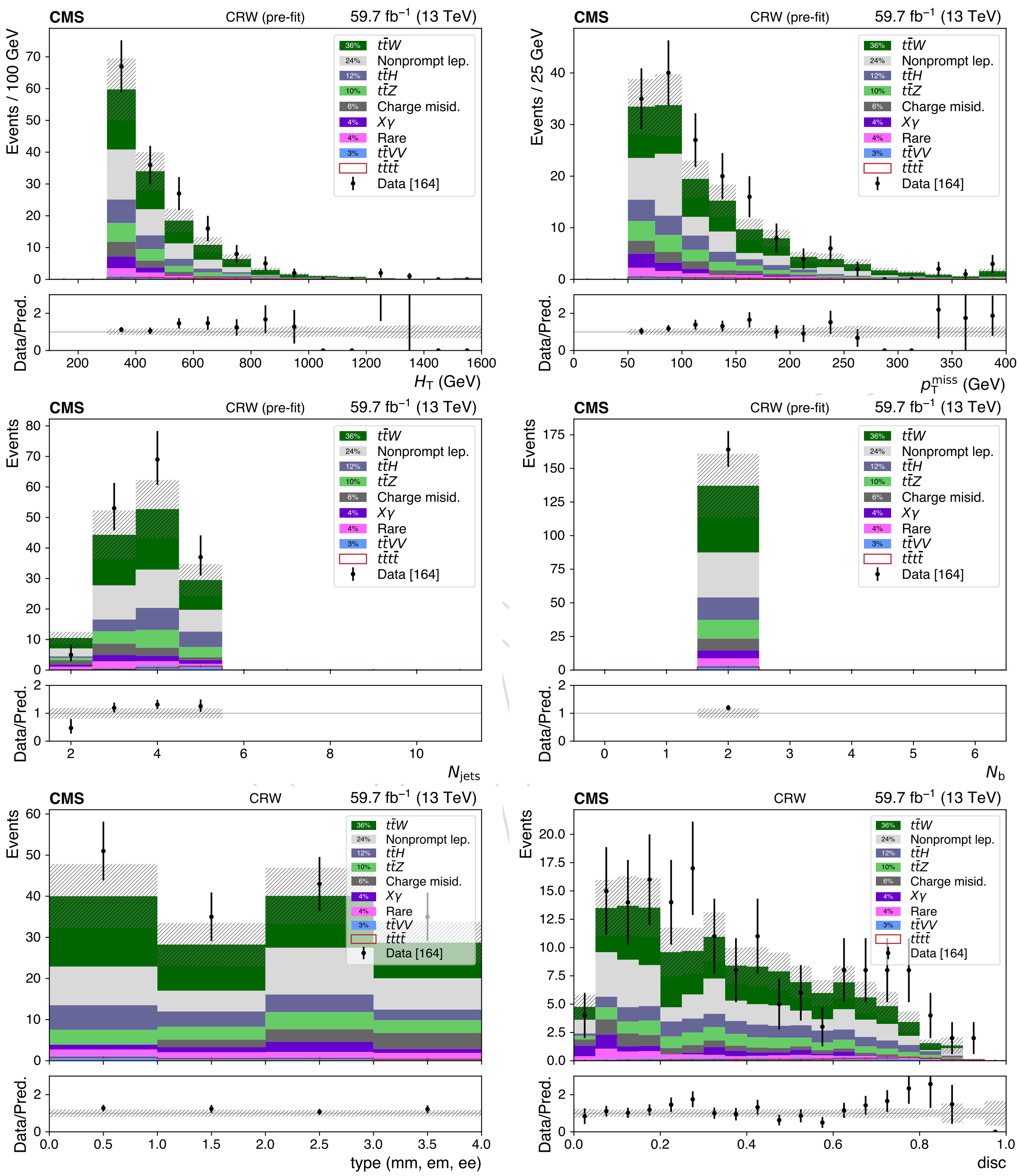


Figure 45: 2018 data and predictions: Prefit distributions of the main analysis variables in the $t\bar{t}W$ control region: H_T, E_T^{miss}, N_{jets}, $N_{b\,jets}$, lepton flavor, and raw BDT discriminant where the last bin includes the overflow.

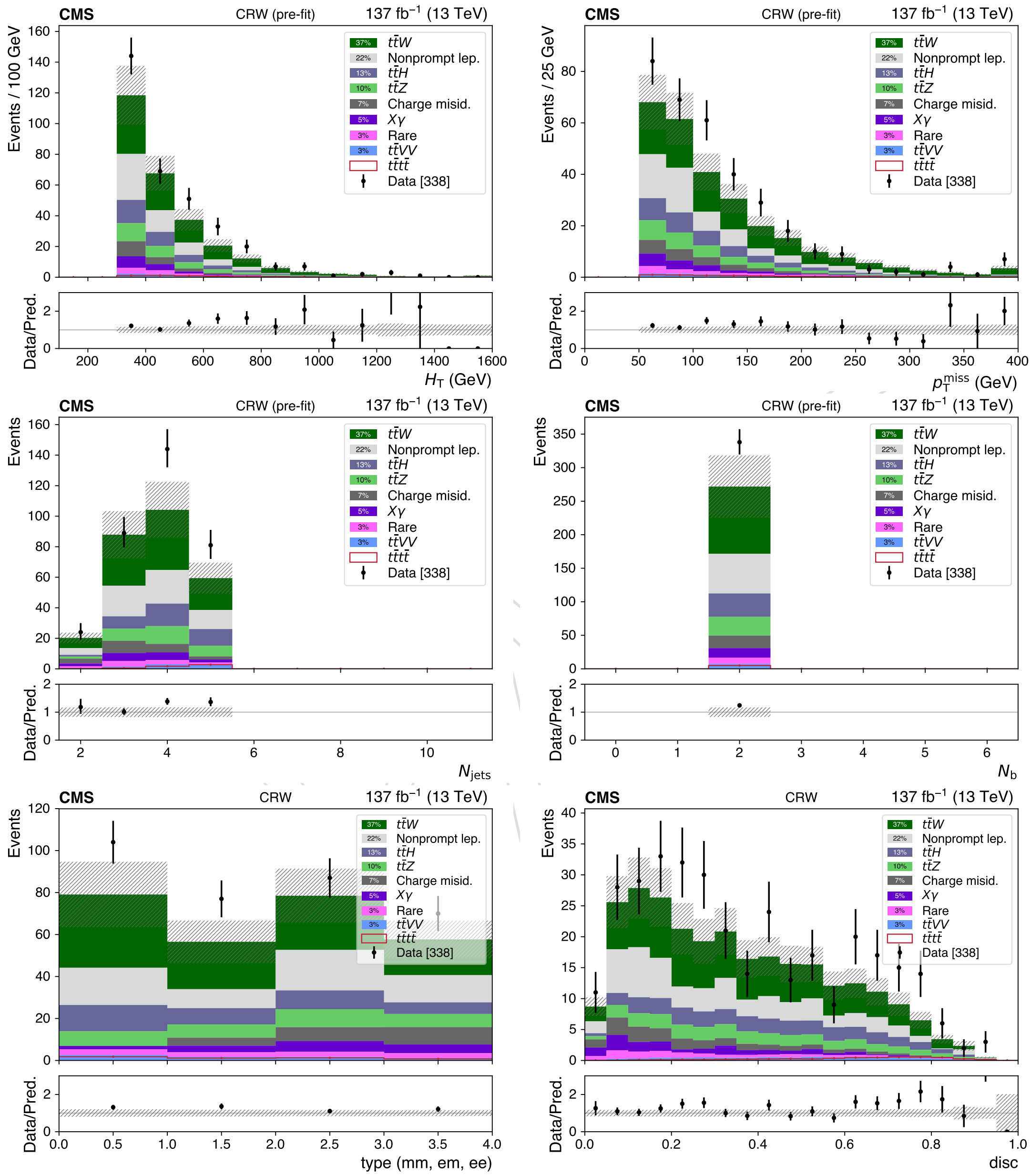


Figure 46: Run2 data and predictions: Prefit distributions of the main analysis variables in the $t\bar{t}W$ control region: H_T, E_T^{miss}, N_{jets}, $N_{b\ jets}$, lepton flavor, and raw BDT discriminant where the last bin includes the overflow.

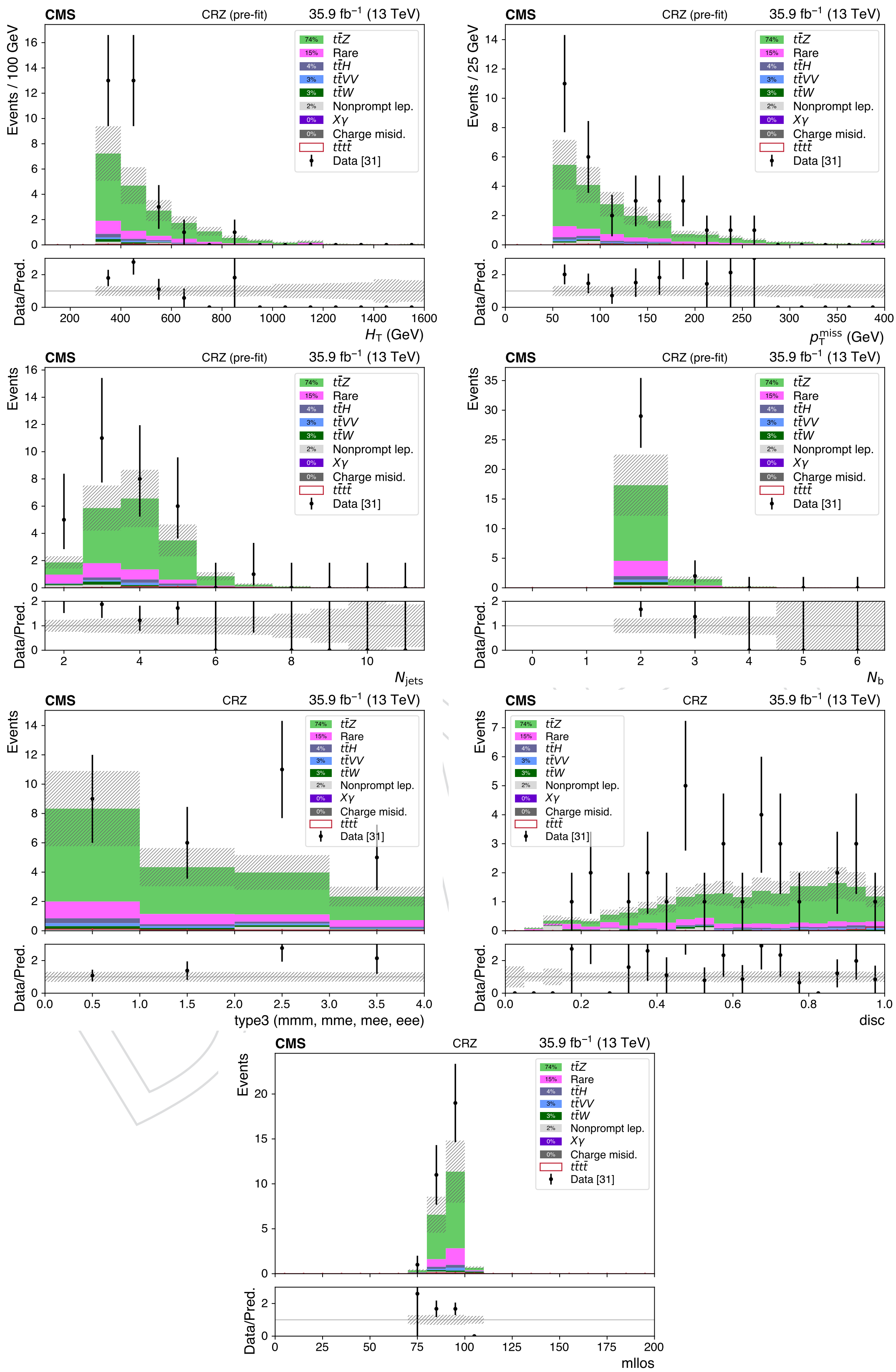


Figure 47: 2016 data and predictions: Prefit distributions of the main analysis variables in the $t\bar{t}Z$ control region: H_T, E_T^{miss}, N_{jets}, $N_{b\,jets}$, lepton flavor, raw BDT discriminant, and $m_{\ell\ell}$ of the OSSF pair, where the last bin includes the overflow.

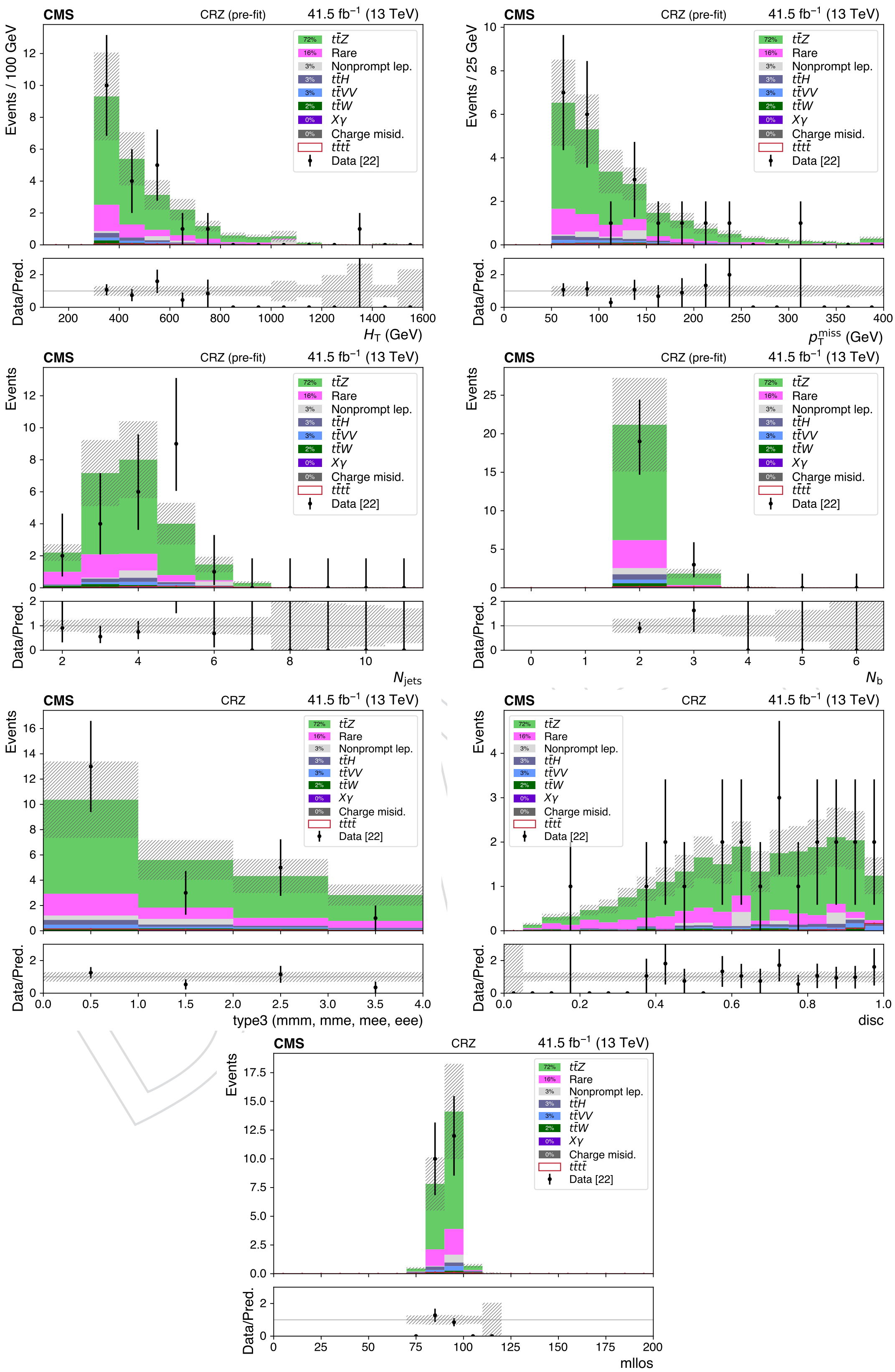


Figure 48: 2017 data and predictions: Prefit distributions of the main analysis variables in the $t\bar{t}Z$ control region: H_T, E_T^{miss}, N_{jets}, $N_{b\ jets}$, lepton flavor, raw BDT discriminant, and $m_{\ell\ell}$ of the OSSF pair, where the last bin includes the overflow.

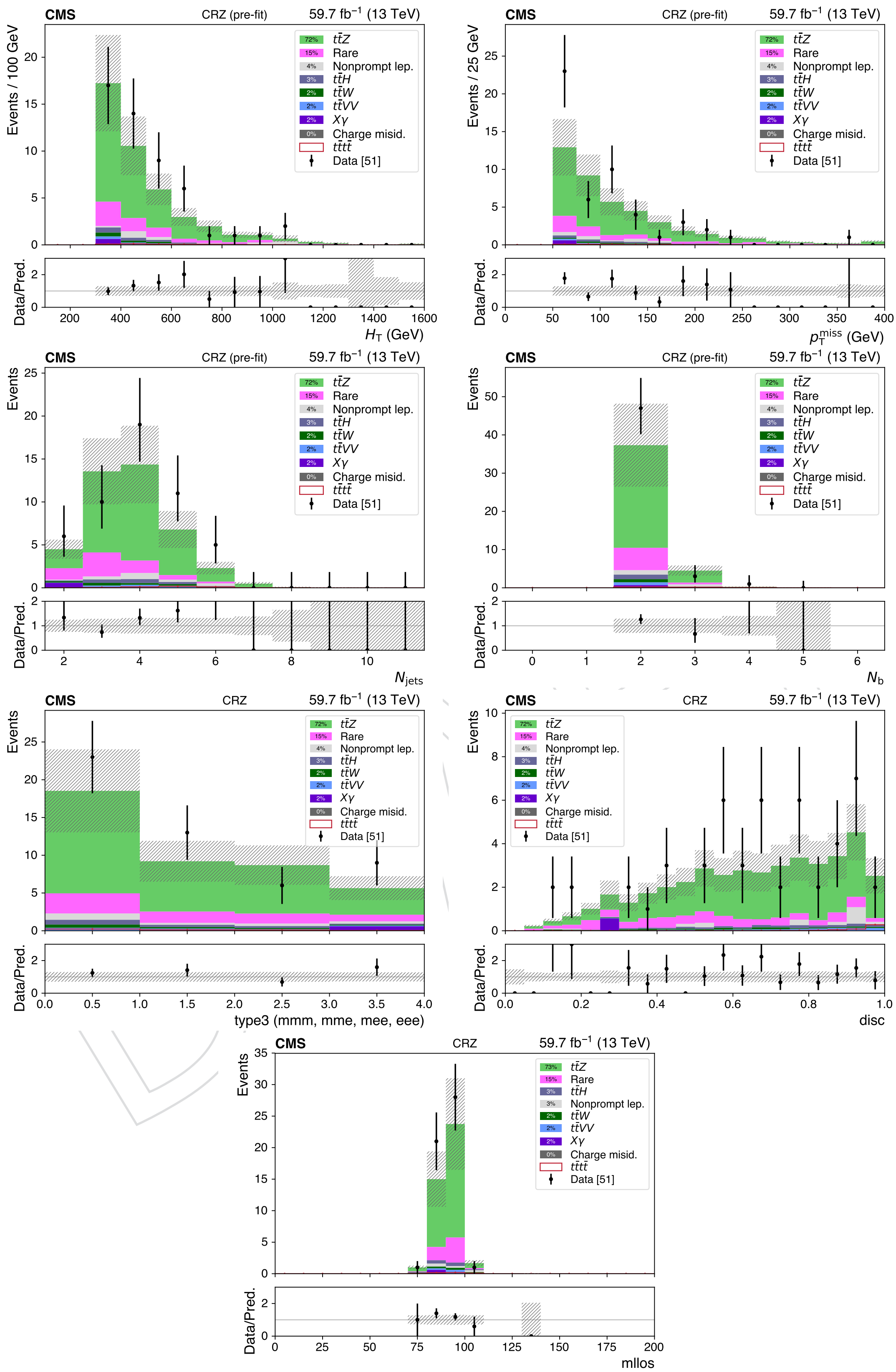


Figure 49: 2018 data and predictions: Prefit distributions of the main analysis variables in the $t\bar{t}Z$ control region: H_T, E_T^{miss}, N_{jets}, $N_{b\ jets}$, lepton flavor, raw BDT discriminant, and $m_{\ell\ell}$ of the OSSF pair, where the last bin includes the overflow.

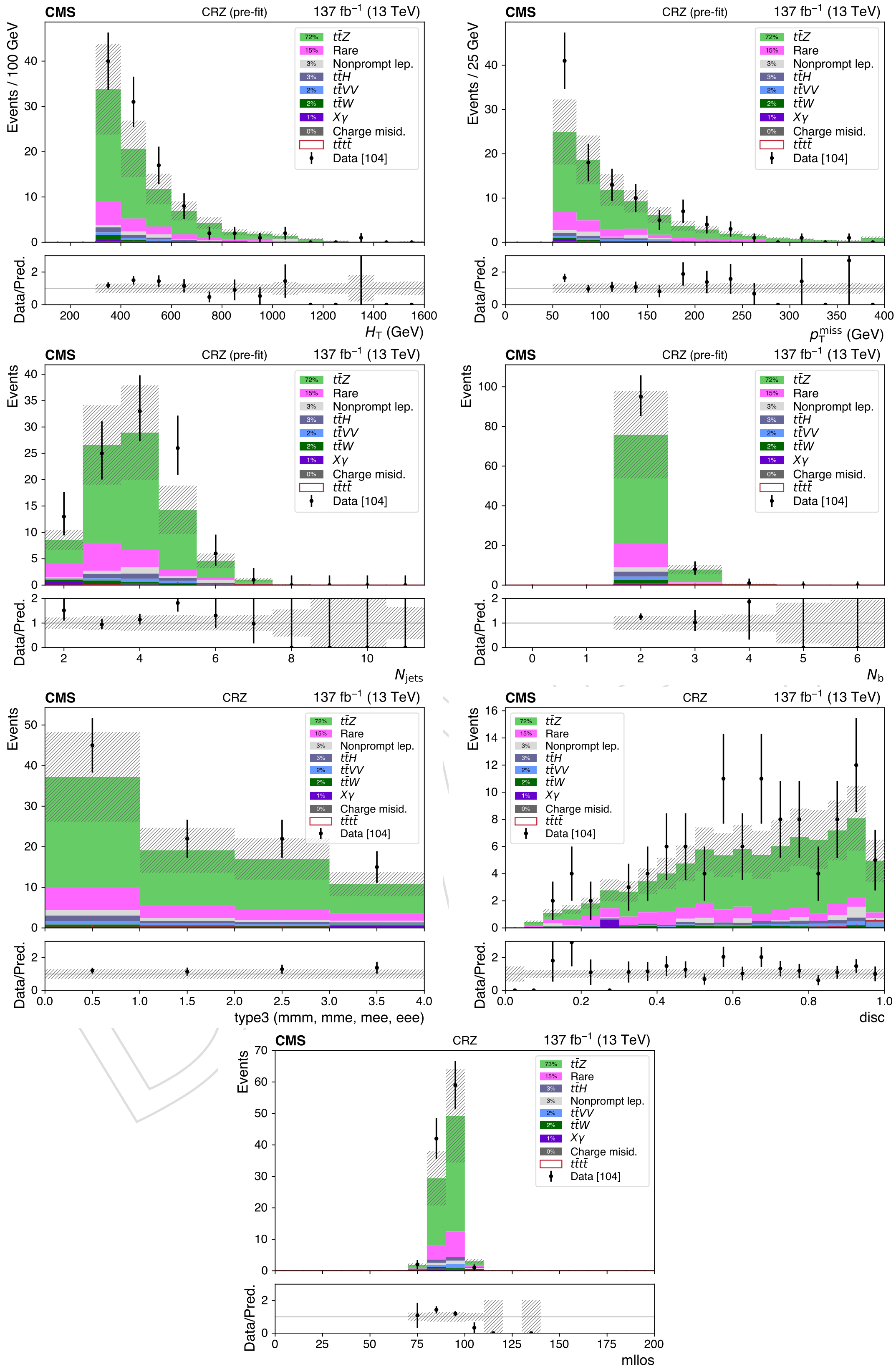


Figure 50: Run2 data and predictions: Prefit distributions of the main analysis variables in the $t\bar{t}Z$ control region: H_T, E_T^{miss}, N_{jets}, $N_{b\,jets}$, lepton flavor, raw BDT discriminant, and $m_{\ell\ell}$ of the OSSF pair, where the last bin includes the overflow.

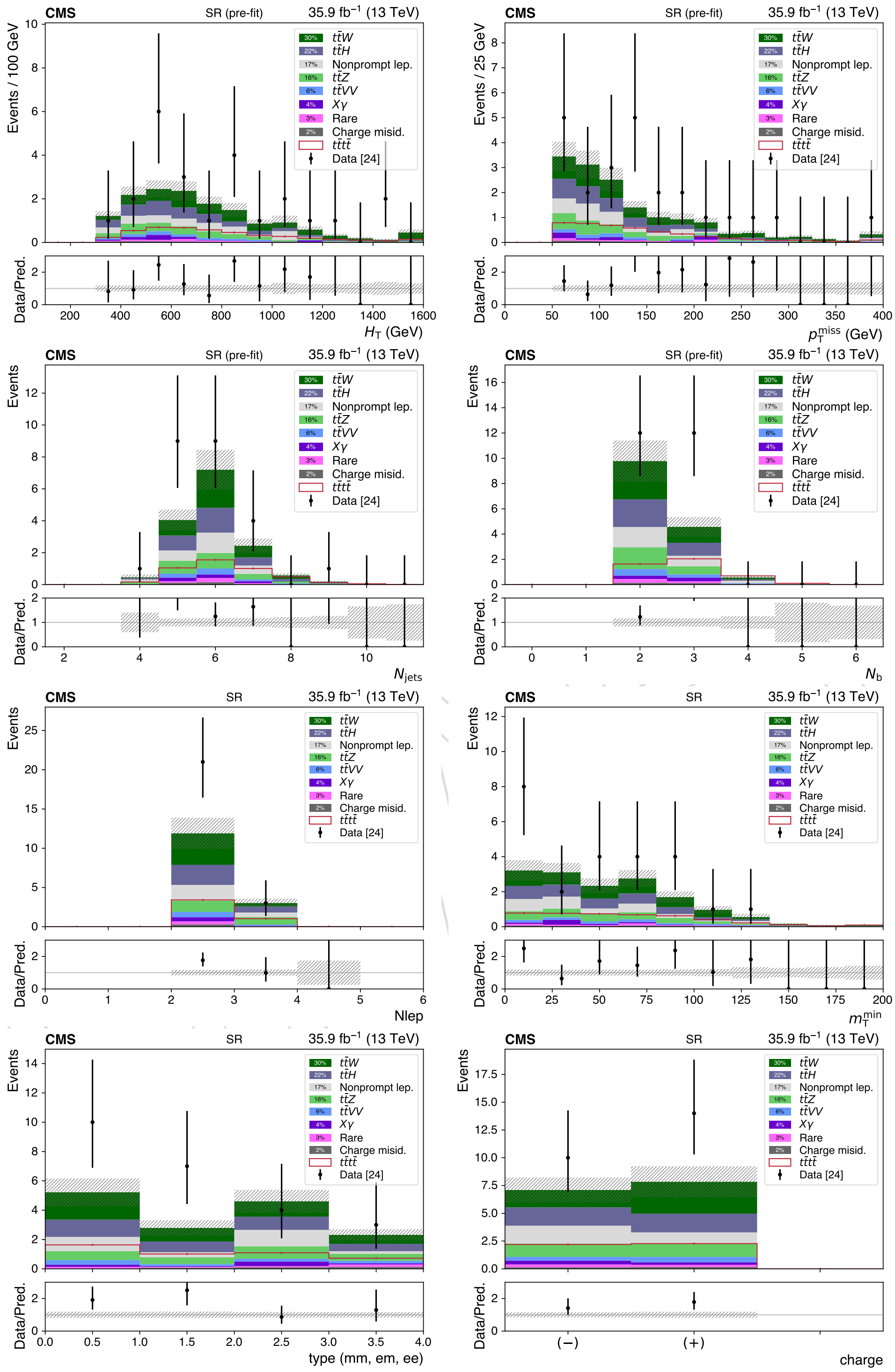


Figure 51: 2016 predictions: prefit distributions of the main analysis variables in the sum of signal regions: H_T, E_T^{miss}, M_T^{min}, N_{jets}, $N_{b\,jets}$, lepton flavor, and $m_{\ell\ell}$ of the OSSF pair, where the last bin includes the overflow.

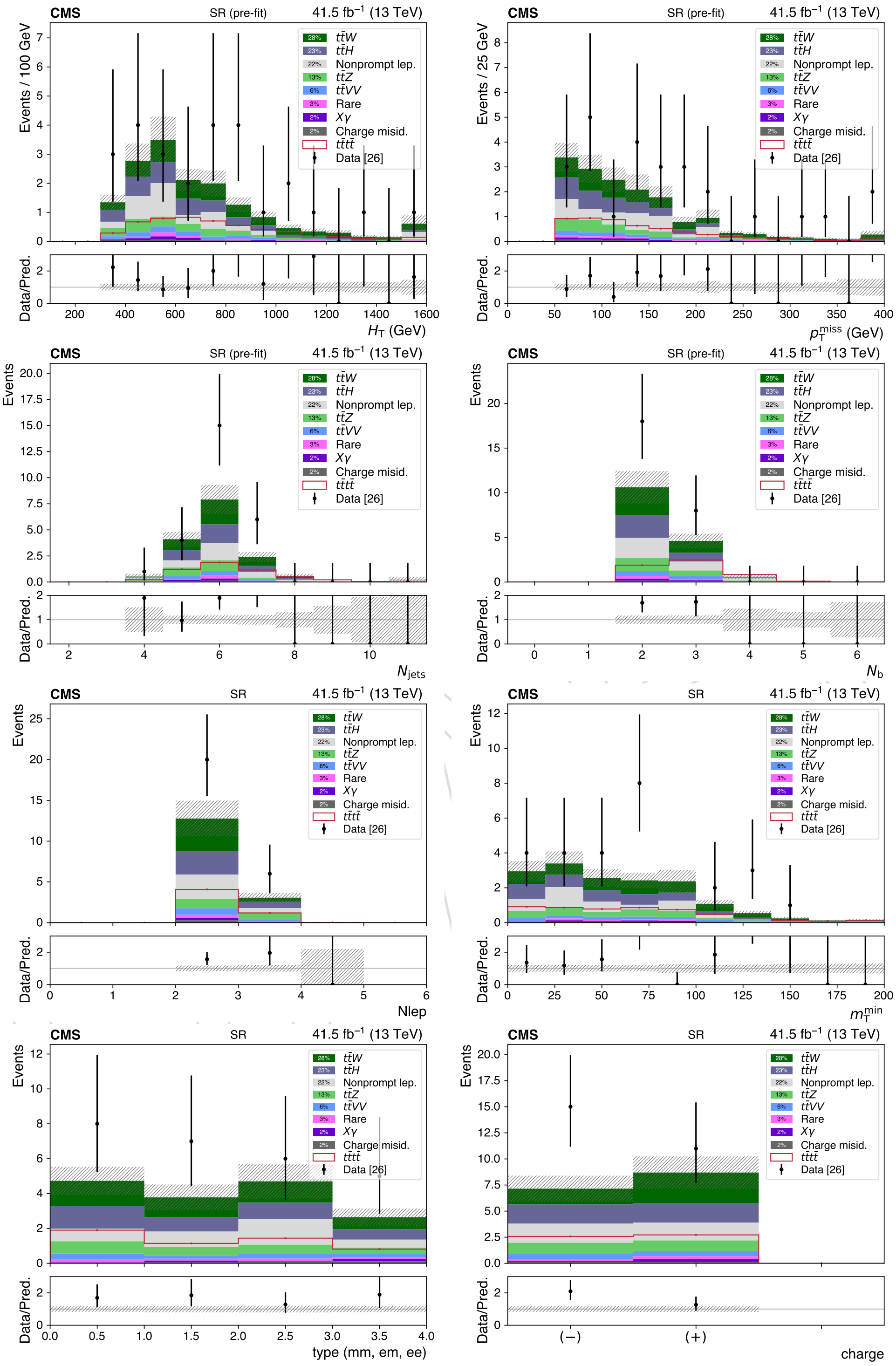


Figure 52: 2017 predictions: prefit distributions of the main analysis variables in the sum of signal regions: H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, $M_{\mathrm{T}}^{\mathrm{min}}$, N_{jets}, $N_{\mathrm{b\,jets}}$, lepton flavor, and $m_{\ell\ell}$ of the OSSF pair, where the last bin includes the overflow.

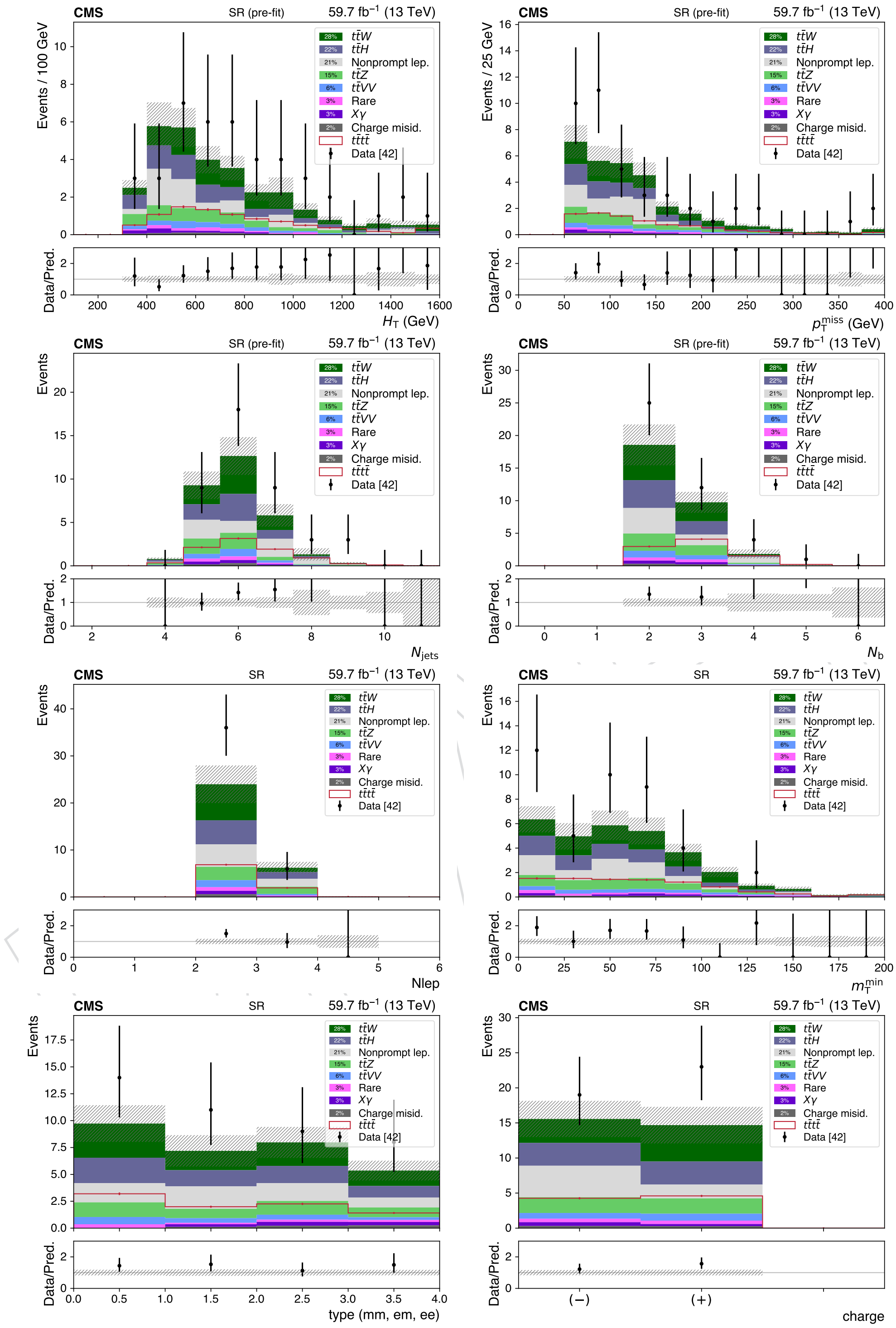


Figure 53: 2018 predictions: prefit distributions of the main analysis variables in the sum of signal regions: H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, $M_{\mathrm{T}}^{\mathrm{min}}$, N_{jets}, $N_{\mathrm{b\,jets}}$, lepton flavor, and $m_{\ell\ell}$ of the OSSF pair, where the last bin includes the overflow.

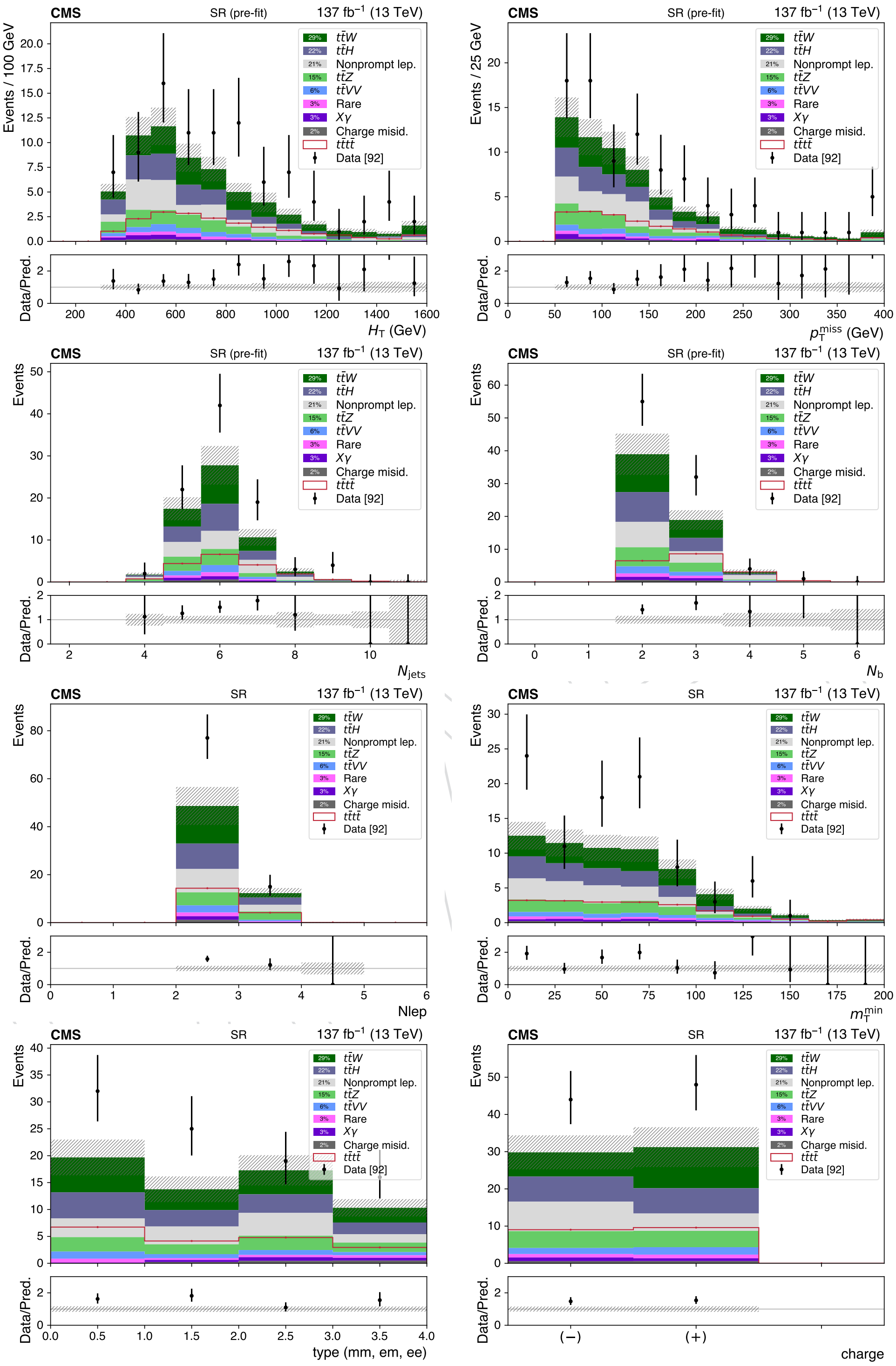


Figure 54: run2 predictions: prefit distributions of the main analysis variables in the sum of signal regions: H_T, E_T^{miss}, M_T^{min}, N_{jets}, $N_{b\ jets}$, lepton flavor, and $m_{\ell\ell}$ of the OSSF pair, where the last bin includes the overflow.

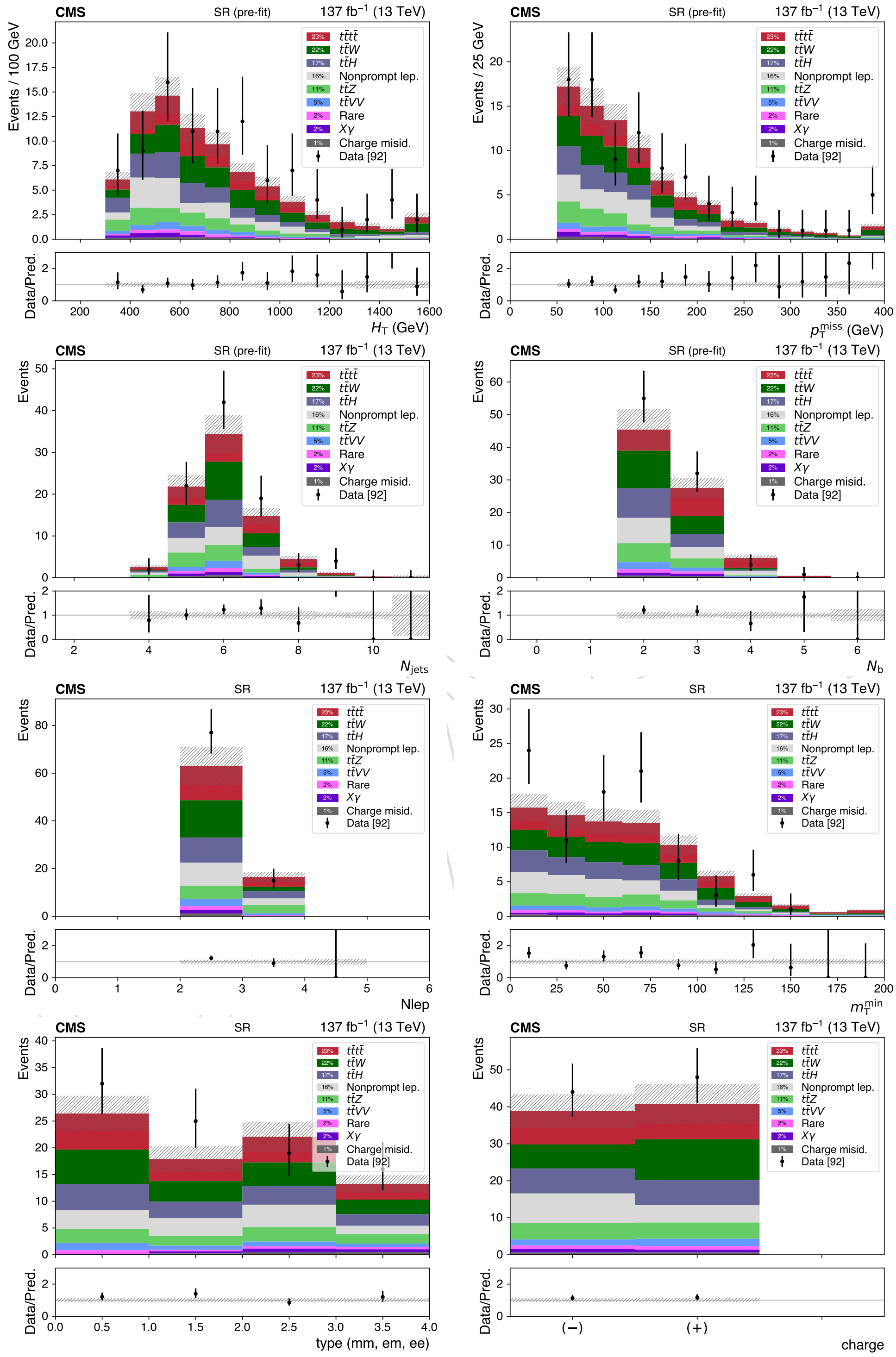


Figure 55: Same as Figure 54 but with stacked signal, assuming signal strength of 1.

11 Results

Results for the 2016+2017+2018 ("Run 2") data are obtained with the data in "golden JSON" given in Sec.2 , for a total luminosity of 77.4 fb^{-1}. The expected yields in the signal regions as well as the $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ control regions, and BDT regions are shown in Figure 56 for prefit for 2016/2017/2018/Run2. The postfits plot for the two analyses are shown in Figure 57. The 2016,2017,2018,Run 2 cut-based numerical yields can be found in Table 18,19,20,21 for prefit, and analogous tables for the BDT analysis can be found in Table 22,23,24,25. Postfit tables for Run 2 cut-based and BDT can be found in Table 26 and Table 27, respectively. . Note that with respect to 2016, also the theoretical cross section assumed for the $\mathrm{t\bar{t}t\bar{t}}$ process has changed, from 9.2 to 11.96 fb.

A binned likelihood fit is performed over signal regions using the Higgs Combine tool [22]; exclusion limits at 95% CL are calculated with the Asymptotic CLs method.

For reference, the commands used to extract upper limits, significances, measurement values, and obtain the NLL vs μ, respectively, are:

- `combine -M Asymptotic card.txt --noFitAsimov`
- `combine -M Significance card.txt -t -1 --expectSignal=1 --significance`
- `combine -M FitDiagnostics card.txt -t -1 --expectSignal=1 --robustFit=1 --saveShapes --saveOverallShapes --saveWithUncertainties -n name`
- `combine -M MultiDimFit card.txt -t -1 --expectSignal=1 --algo grid --centeredRange=2.0 --saveFitResult --redefineSignalPOI r --robustFit=1 --saveNLL`

`-t -1 --expectSignal=1` and `--noFitAsimov` are omitted for observed results.

Using the Run2 data, the *cut-based* analysis sets an observed (expected) upper limit on the production cross section of 20.04 fb ($9.35^{+4.29}_{-2.88}$ fb), assuming the signal process does not exist. The observed (expected) significance is 1.712 (2.478) standard deviations, corresponding to a measured observed (expected) signal strength parameter of $0.784^{+0.514}_{-0.469}$ ($1.000^{+0.480}_{-0.433}$)

The *BDT* analysis sets an observed (expected) upper limit on the production cross section of 22.51 fb ($8.46^{+3.91}_{-2.57}$ fb), assuming the signal process does not exist. The observed (expected) significance is 2.561 (2.699) standard deviations, corresponding to a measured observed (expected) signal strength parameter of $1.052^{+0.483}_{-0.437}$ ($1.000^{+0.442}_{-0.401}$)

The likelihood scan for Run2 is shown in Fig. 58.

Several interpretations of these results are presented in the following sections

- Section 12.1: 2HDM with a heavy scalar or pseudoscalar boson decaying to on-shell $\mathrm{t\bar{t}}$
- Section 12.2: top yukawa coupling constant
- Section 12.3: off-shell top-philic scalar or pseudoscalar boson
- Section 12.4: EFT oblique Higgs parameter $\hat{H}$
- Section 12.5: Simplified dark matter with scalar or pseudoscalar mediator

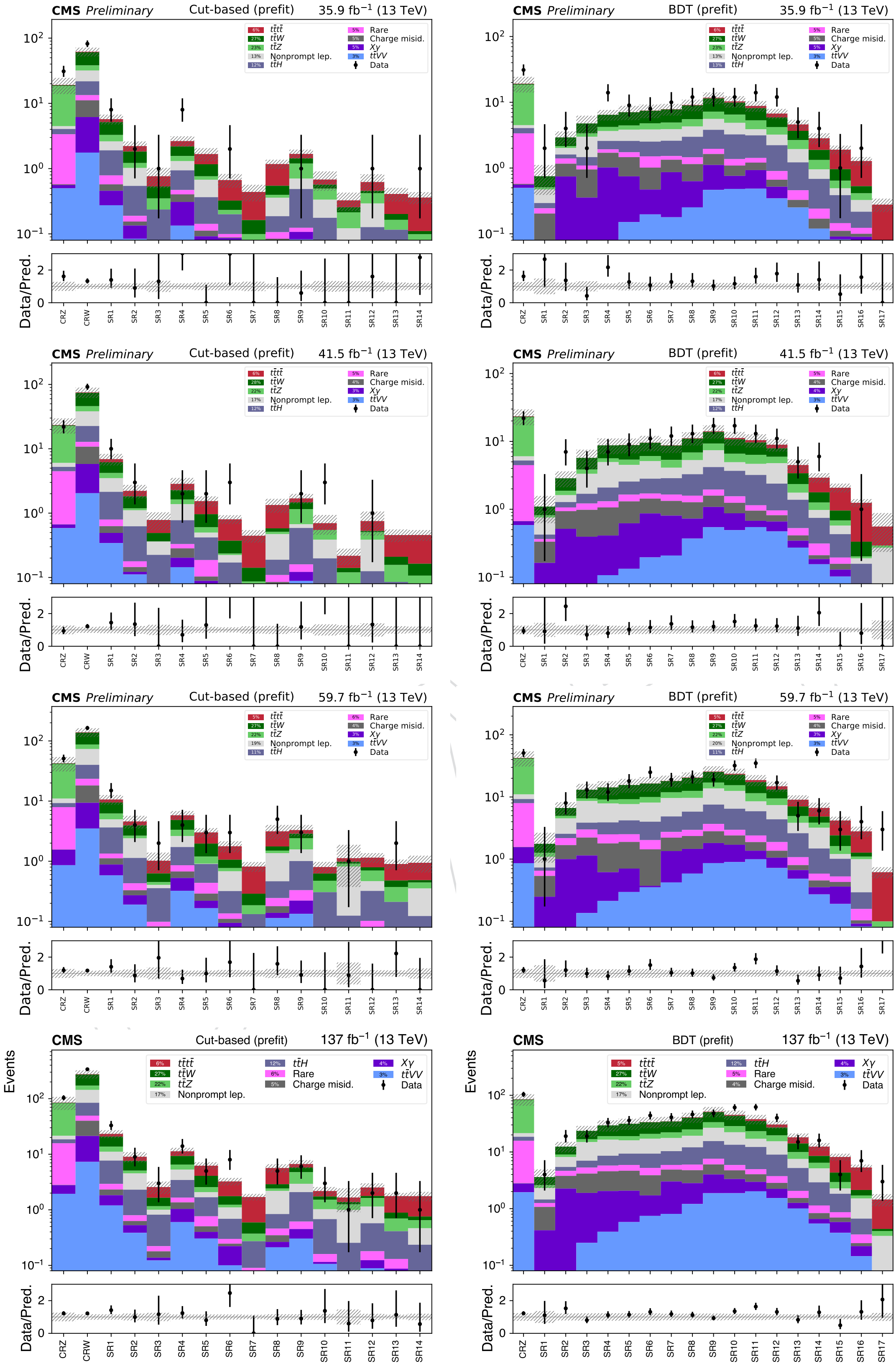


Figure 56: Data yields compared to prefit for cut-based (left) and BDT (right) analyses separately for data periods 2016, 2017, 2018, Run2 from top to bottom.

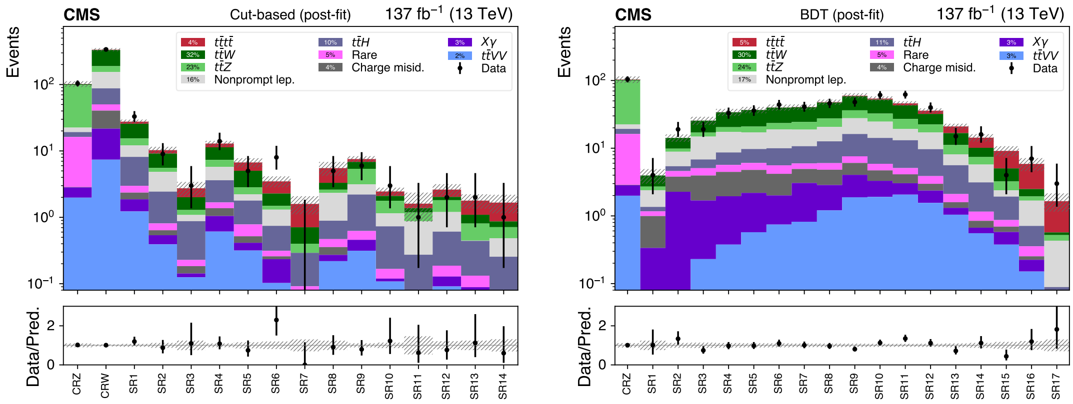


Figure 57: Data yields compared to postfit for cut-based (left) and BDT (right) analyses.

Table 18: Prefit event yields in SR+CR regions for 2016.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.50± 0.18	14.01± 5.20	0.69± 0.18	0.50± 0.06	0.06± 0.01	2.79± 0.58	0.02± 0.00	0.42± 0.17	18.98± 5.34	31	0.25± 0.04
CRW	22.56± 8.01	6.31± 2.31	8.44± 2.14	1.76± 0.21	4.41± 0.71	2.24± 0.49	4.95± 0.95	10.05± 3.82	60.72± 9.61	82	1.23± 0.10
SR1	1.87± 0.70	0.68± 0.26	1.12± 0.31	0.27± 0.04	0.18± 0.07	0.21± 0.06	0.12± 0.02	0.70± 0.32	5.16± 0.97	8	0.58± 0.07
SR2	0.61± 0.25	0.20± 0.10	0.35± 0.10	0.08± 0.02	0.05± 0.01	0.03± 0.01	0.02± 0.00	0.48± 0.23	1.83± 0.39	2	0.39± 0.02
SR3	0.18± 0.12	0.11± 0.06	0.17± 0.07	0.05± 0.01	0.00± 0.01	0.01± 0.01	0.01± 0.00	0.00± 0.11	0.53± 0.24	1	0.23± 0.05
SR4	0.63± 0.25	0.26± 0.14	0.47± 0.13	0.13± 0.03	0.17± 0.12	0.07± 0.01	0.09± 0.02	0.36± 0.21	2.19± 0.48	8	0.46± 0.05
SR5	0.40± 0.19	0.09± 0.03	0.22± 0.08	0.06± 0.02	0.03± 0.01	0.03± 0.01	0.02± 0.00	0.31± 0.29	1.17± 0.39	0	0.51± 0.05
SR6	0.09± 0.06	0.03± 0.05	0.10± 0.04	0.01± 0.01	0.07± 0.05	0.01± 0.00	0.00± 0.00	0.00± 0.04	0.32± 0.15	2	0.34± 0.04
SR7	0.06± 0.04	0.04± 0.02	0.05± 0.02	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.05	0.16± 0.10	0	0.28± 0.08
SR8	0.15± 0.08	0.04± 0.03	0.08± 0.03	0.05± 0.01	0.00± 0.00	0.02± 0.01	0.03± 0.01	0.17± 0.11	0.55± 0.17	0	0.63± 0.06
SR9	0.25± 0.10	0.49± 0.18	0.36± 0.10	0.08± 0.02	0.02± 0.00	0.02± 0.01	0.00± 0.00	0.22± 0.11	1.44± 0.28	1	0.24± 0.06
SR10	0.08± 0.03	0.16± 0.07	0.14± 0.04	0.02± 0.01	0.00± 0.00	0.01± 0.01	0.00± 0.00	0.16± 0.19	0.57± 0.22	0	0.11± 0.03
SR11	0.04± 0.02	0.09± 0.05	0.05± 0.02	0.01± 0.00	0.00± 0.00	0.01± 0.01	0.00± 0.00	0.05± 0.04	0.26± 0.10	0	0.07± 0.02
SR12	0.04± 0.02	0.13± 0.05	0.09± 0.03	0.02± 0.00	0.00± 0.00	0.02± 0.00	0.00± 0.00	0.17± 0.18	0.46± 0.19	1	0.17± 0.02
SR13	0.04± 0.02	0.05± 0.03	0.08± 0.02	0.01± 0.00	0.02± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.06	0.20± 0.09	0	0.21± 0.02
SR14	0.01± 0.02	0.05± 0.03	0.04± 0.02	0.01± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.11± 0.05	1	0.25± 0.05

Table 19: Prefit event yields in SR+CR regions for 2017.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.53± 0.20	16.57± 6.03	0.77± 0.22	0.59± 0.08	0.08± 0.04	3.79± 0.81	0.00± 0.00	0.79± 0.42	23.12± 6.24	22	0.32± 0.02
CRW	28.09± 9.95	7.70± 2.83	9.81± 2.67	2.04± 0.25	3.77± 0.57	1.99± 0.42	4.98± 0.95	15.62± 8.52	73.99±13.76	92	1.48± 0.07
SR1	1.97± 0.72	0.70± 0.27	1.35± 0.38	0.34± 0.05	0.15± 0.06	0.17± 0.06	0.14± 0.03	1.35± 1.05	6.17± 1.42	10	0.75± 0.04
SR2	0.64± 0.28	0.08± 0.07	0.45± 0.14	0.11± 0.02	0.01± 0.01	0.03± 0.01	0.02± 0.00	0.48± 0.37	1.83± 0.52	3	0.39± 0.04
SR3	0.13± 0.09	0.00± 0.02	0.17± 0.06	0.03± 0.01	0.00± 0.00	0.02± 0.00	0.01± 0.00	0.14± 0.16	0.49± 0.24	0	0.29± 0.05
SR4	0.65± 0.25	0.26± 0.14	0.45± 0.14	0.14± 0.02	0.06± 0.04	0.05± 0.02	0.08± 0.02	0.60± 0.49	2.28± 0.67	2	0.57± 0.05
SR5	0.36± 0.20	0.10± 0.04	0.23± 0.08	0.08± 0.01	0.01± 0.01	0.08± 0.02	0.01± 0.00	0.08± 0.09	0.96± 0.27	2	0.58± 0.04
SR6	0.14± 0.06	0.01± 0.04	0.11± 0.04	0.03± 0.00	0.01± 0.00	0.02± 0.01	0.01± 0.00	0.05± 0.06	0.37± 0.12	3	0.43± 0.03
SR7	0.05± 0.03	0.01± 0.01	0.04± 0.02	0.01± 0.00	0.01± 0.01	0.01± 0.00	0.01± 0.00	0.00± 0.06	0.14± 0.08	0	0.30± 0.04
SR8	0.07± 0.05	0.04± 0.03	0.08± 0.04	0.05± 0.01	0.02± 0.01	0.02± 0.01	0.01± 0.00	0.28± 0.28	0.58± 0.30	0	0.76± 0.08
SR9	0.32± 0.16	0.48± 0.18	0.43± 0.12	0.09± 0.01	0.03± 0.01	0.04± 0.01	0.00± 0.00	0.08± 0.07	1.47± 0.30	2	0.23± 0.01
SR10	0.00± 0.03	0.16± 0.07	0.15± 0.05	0.03± 0.01	0.01± 0.01	0.01± 0.00	0.00± 0.00	0.19± 0.21	0.55± 0.23	3	0.15± 0.02
SR11	0.02± 0.01	0.06± 0.03	0.05± 0.02	0.01± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.02	0.14± 0.06	0	0.08± 0.02
SR12	0.02± 0.01	0.12± 0.05	0.08± 0.03	0.03± 0.00	0.01± 0.00	0.01± 0.00	0.00± 0.00	0.26± 0.30	0.53± 0.30	1	0.23± 0.02
SR13	0.05± 0.03	0.07± 0.03	0.06± 0.02	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.06	0.21± 0.09	0	0.25± 0.02
SR14	0.06± 0.03	0.05± 0.03	0.04± 0.02	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.04	0.16± 0.08	0	0.29± 0.03

Table 20: Prefit event yields in SR+CR regions for 2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.88± 0.33	30.20±11.07	1.34± 0.36	0.86± 0.12	0.70± 0.40	6.36± 1.28	0.01± 0.00	1.72± 0.60	42.07±11.49	51	0.49± 0.03
CRW	49.69±18.72	13.94± 4.93	16.64± 4.19	3.50± 0.43	5.88± 1.15	5.25± 1.10	8.80± 1.75	33.41± 9.80	137.09±23.04	164	2.24± 0.09
SR1	3.27± 1.39	1.25± 0.47	2.31± 0.63	0.58± 0.09	0.30± 0.09	0.22± 0.05	0.24± 0.05	1.31± 0.62	9.48± 1.96	15	1.18± 0.07
SR2	1.29± 0.53	0.26± 0.16	0.71± 0.24	0.19± 0.04	0.08± 0.04	0.10± 0.03	0.06± 0.01	1.27± 0.79	3.95± 1.09	4	0.65± 0.05
SR3	0.18± 0.17	0.04± 0.08	0.26± 0.09	0.05± 0.01	0.01± 0.02	0.02± 0.00	0.03± 0.01	0.05± 0.04	0.63± 0.31	2	0.39± 0.07
SR4	1.51± 0.63	0.72± 0.32	0.87± 0.27	0.32± 0.05	0.20± 0.06	0.18± 0.05	0.15± 0.03	0.89± 0.42	4.85± 1.01	4	0.97± 0.05
SR5	0.77± 0.32	0.28± 0.18	0.46± 0.15	0.17± 0.03	0.05± 0.03	0.15± 0.05	0.07± 0.01	0.01± 0.05	1.97± 0.54	3	1.05± 0.06
SR6	0.29± 0.15	0.09± 0.04	0.18± 0.07	0.05± 0.01	0.04± 0.01	0.02± 0.00	0.02± 0.00	0.36± 0.24	1.06± 0.33	3	0.71± 0.08
SR7	0.10± 0.05	0.05± 0.03	0.09± 0.04	0.02± 0.00	0.00± 0.00	0.02± 0.01	0.00± 0.00	0.00± 0.04	0.29± 0.11	0	0.53± 0.07
SR8	0.26± 0.16	0.14± 0.09	0.22± 0.09	0.11± 0.02	0.03± 0.01	0.07± 0.01	0.04± 0.01	0.90± 0.75	1.76± 0.83	5	1.39± 0.13
SR9	0.47± 0.23	0.84± 0.36	0.68± 0.19	0.13± 0.02	0.09± 0.02	0.11± 0.09	0.00± 0.00	0.56± 0.34	2.87± 0.71	3	0.44± 0.04
SR10	0.15± 0.09	0.17± 0.08	0.23± 0.07	0.05± 0.01	0.00± 0.01	0.02± 0.01	0.00± 0.00	0.00± 0.04	0.63± 0.17	0	0.18± 0.04
SR11	0.07± 0.05	0.10± 0.06	0.09± 0.03	0.02± 0.00	0.01± 0.01	0.01± 0.01	0.00± 0.00	0.69± 0.75	0.98± 0.75	1	0.14± 0.02
SR12	0.09± 0.06	0.23± 0.09	0.22± 0.06	0.05± 0.01	0.03± 0.01	0.03± 0.01	0.00± 0.00	0.14± 0.11	0.79± 0.20	0	0.36± 0.03
SR13	0.11± 0.06	0.16± 0.08	0.15± 0.05	0.04± 0.01	0.00± 0.00	0.02± 0.01	0.00± 0.00	0.00± 0.05	0.48± 0.16	2	0.42± 0.04
SR14	0.04± 0.02	0.09± 0.06	0.09± 0.04	0.02± 0.01	0.00± 0.00	0.02± 0.00	0.00± 0.00	0.23± 0.25	0.49± 0.27	0	0.46± 0.06

Table 21: Prefit event yields in SR+CR regions for 2016+2017+2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	1.92± 0.70	60.77±21.92	2.80± 0.72	1.94± 0.22	0.84± 0.40	12.94± 2.74	0.03± 0.00	2.94± 1.10	84.18±22.57	104	1.05± 0.06
CRW	100.33±36.18	27.95±10.15	34.89± 8.58	7.30± 0.79	14.06± 1.83	9.48± 2.08	18.73± 2.14	59.08±21.64	271.81±46.38	338	4.94± 0.17
SR1	7.11± 2.70	2.63± 0.96	4.78± 1.29	1.20± 0.15	0.63± 0.14	0.60± 0.15	0.49± 0.06	3.36± 1.41	20.80± 3.77	33	2.50± 0.12
SR2	2.53± 0.98	0.53± 0.28	1.51± 0.46	0.39± 0.06	0.14± 0.05	0.17± 0.04	0.10± 0.01	2.23± 1.24	7.60± 1.88	9	1.43± 0.07
SR3	0.49± 0.33	0.16± 0.11	0.60± 0.22	0.12± 0.02	0.02± 0.02	0.04± 0.01	0.04± 0.01	0.19± 0.18	1.65± 0.61	3	0.92± 0.12
SR4	2.78± 1.02	1.24± 0.53	1.78± 0.48	0.60± 0.07	0.44± 0.14	0.29± 0.07	0.32± 0.04	1.85± 0.87	9.31± 1.75	14	2.00± 0.09
SR5	1.54± 0.64	0.47± 0.21	0.92± 0.29	0.31± 0.04	0.10± 0.04	0.26± 0.07	0.10± 0.01	0.40± 0.28	4.10± 0.97	5	2.13± 0.09
SR6	0.52± 0.20	0.14± 0.10	0.39± 0.14	0.10± 0.02	0.12± 0.05	0.05± 0.01	0.02± 0.00	0.41± 0.27	1.76± 0.48	8	1.49± 0.09
SR7	0.22± 0.09	0.10± 0.04	0.18± 0.07	0.04± 0.01	0.01± 0.01	0.03± 0.01	0.01± 0.00	0.00± 0.08	0.59± 0.19	0	1.11± 0.10
SR8	0.48± 0.24	0.22± 0.12	0.38± 0.14	0.21± 0.03	0.05± 0.02	0.11± 0.02	0.08± 0.01	1.36± 0.81	2.89± 0.93	5	2.77± 0.16
SR9	1.03± 0.44	1.80± 0.66	1.47± 0.39	0.30± 0.04	0.15± 0.02	0.16± 0.09	0.00± 0.00	0.86± 0.45	5.78± 1.16	6	0.90± 0.08
SR10	0.23± 0.10	0.49± 0.17	0.52± 0.16	0.11± 0.02	0.01± 0.01	0.04± 0.01	0.00± 0.00	0.35± 0.24	1.74± 0.39	3	0.44± 0.06
SR11	0.13± 0.07	0.25± 0.11	0.19± 0.06	0.04± 0.01	0.01± 0.01	0.02± 0.01	0.00± 0.00	0.74± 0.74	1.38± 0.77	1	0.29± 0.03
SR12	0.15± 0.08	0.48± 0.17	0.39± 0.10	0.09± 0.01	0.04± 0.01	0.06± 0.02	0.00± 0.00	0.57± 0.37	1.78± 0.46	2	0.75± 0.04
SR13	0.20± 0.10	0.28± 0.11	0.28± 0.09	0.07± 0.01	0.02± 0.01	0.04± 0.01	0.00± 0.00	0.00± 0.09	0.89± 0.24	2	0.88± 0.05
SR14	0.11± 0.05	0.18± 0.09	0.17± 0.07	0.04± 0.01	0.00± 0.00	0.03± 0.01	0.00± 0.00	0.23± 0.26	0.76± 0.32	1	1.00± 0.09

Table 22: Prefit event yields in BDT regions for 2016.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.50± 0.20	14.01± 5.02	0.69± 0.15	0.50± 0.06	0.06± 0.01	2.79± 0.59	0.02± 0.00	0.42± 0.19	18.98± 5.15	31	0.25± 0.04
SR1	0.24± 0.11	0.04± 0.03	0.05± 0.04	0.00± 0.00	0.00± 0.15	0.04± 0.11	0.20± 0.04	0.18± 0.12	0.75± 0.36	2	0.00± 0.00
SR2	0.77± 0.41	0.13± 0.08	0.21± 0.09	0.02± 0.01	0.72± 0.29	0.23± 0.10	0.42± 0.08	0.41± 0.22	2.92± 0.68	4	0.00± 0.00
SR3	1.38± 0.79	0.38± 0.15	0.37± 0.14	0.05± 0.02	0.30± 0.36	0.14± 0.09	0.60± 0.11	1.53± 0.72	4.76± 1.48	2	0.01± 0.00
SR4	2.23± 0.96	0.57± 0.26	0.66± 0.20	0.07± 0.04	0.95± 0.45	0.22± 0.08	0.69± 0.13	1.06± 0.42	6.44± 1.17	14	0.02± 0.02
SR5	2.61± 1.19	0.62± 0.29	0.84± 0.24	0.15± 0.03	0.60± 0.30	0.27± 0.13	0.72± 0.14	1.24± 0.51	7.05± 1.72	9	0.04± 0.01
SR6	2.65± 1.23	0.92± 0.38	1.02± 0.24	0.20± 0.04	0.27± 0.27	0.49± 0.14	0.54± 0.10	1.33± 0.67	7.43± 1.72	8	0.05± 0.04
SR7	2.87± 1.17	1.00± 0.48	1.08± 0.24	0.18± 0.04	0.69± 0.17	0.18± 0.11	0.47± 0.09	1.30± 0.56	7.78± 1.57	10	0.08± 0.04
SR8	3.44± 1.42	1.23± 0.60	1.53± 0.32	0.25± 0.05	0.38± 0.13	0.30± 0.11	0.52± 0.10	1.29± 0.61	8.96± 1.82	12	0.17± 0.03
SR9	4.24± 1.77	1.63± 0.73	1.85± 0.41	0.47± 0.06	0.65± 0.18	0.32± 0.13	0.52± 0.10	1.76± 0.72	11.44± 2.37	12	0.32± 0.04
SR10	3.40± 1.43	1.61± 0.77	1.88± 0.40	0.48± 0.06	0.28± 0.16	0.28± 0.06	0.33± 0.06	1.62± 0.77	9.88± 2.05	12	0.45± 0.04
SR11	2.83± 1.21	1.43± 0.64	1.72± 0.39	0.49± 0.05	0.45± 0.19	0.21± 0.05	0.28± 0.05	0.74± 0.47	8.14± 1.61	14	0.67± 0.07
SR12	1.68± 0.73	0.98± 0.47	1.33± 0.31	0.35± 0.06	0.17± 0.07	0.24± 0.06	0.14± 0.03	0.99± 0.51	5.88± 1.18	12	0.86± 0.04
SR13	1.03± 0.44	0.72± 0.31	0.76± 0.18	0.25± 0.03	0.00± 0.02	0.13± 0.03	0.08± 0.01	0.72± 0.38	3.68± 0.80	5	0.92± 0.04
SR14	0.56± 0.33	0.37± 0.14	0.44± 0.14	0.12± 0.02	0.00± 0.06	0.07± 0.01	0.05± 0.01	0.27± 0.18	1.89± 0.54	4	0.95± 0.05
SR15	0.32± 0.16	0.15± 0.09	0.24± 0.07	0.08± 0.01	0.01± 0.01	0.03± 0.01	0.02± 0.00	0.10± 0.07	0.96± 0.27	1	0.94± 0.06
SR16	0.12± 0.06	0.09± 0.05	0.09± 0.04	0.03± 0.01	0.06± 0.06	0.02± 0.00	0.01± 0.00	0.11± 0.09	0.53± 0.18	2	0.75± 0.07
SR17	0.03± 0.03	0.00± 0.01	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.05± 0.04	0	0.23± 0.05

Table 23: Prefit event yields in BDT regions for 2017.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.53± 0.22	16.57± 5.93	0.77± 0.18	0.59± 0.07	0.08± 0.04	3.79± 0.86	0.00± 0.00	0.79± 0.46	23.12± 6.17	22	0.32± 0.02
SR1	0.28± 0.22	0.06± 0.04	0.06± 0.04	0.01± 0.01	0.16± 0.12	0.03± 0.02	0.17± 0.03	0.33± 0.39	1.10± 0.47	1	0.00± 0.00
SR2	0.99± 0.48	0.26± 0.14	0.21± 0.08	0.02± 0.01	0.50± 0.22	0.13± 0.06	0.42± 0.08	0.34± 0.31	2.86± 0.75	7	0.00± 0.00
SR3	1.87± 1.00	0.53± 0.25	0.48± 0.15	0.06± 0.01	0.33± 0.20	0.23± 0.18	0.59± 0.11	1.63± 1.11	5.72± 1.81	4	0.01± 0.01
SR4	3.13± 1.31	0.67± 0.28	0.78± 0.22	0.11± 0.03	0.30± 0.34	0.25± 0.13	0.65± 0.12	2.86± 1.58	8.75± 2.28	7	0.02± 0.01
SR5	3.42± 1.49	0.88± 0.39	0.92± 0.23	0.13± 0.02	0.49± 0.18	0.33± 0.12	0.69± 0.13	1.94± 1.20	8.80± 2.18	9	0.05± 0.01
SR6	3.48± 1.51	0.94± 0.49	1.12± 0.28	0.20± 0.03	0.67± 0.33	0.27± 0.09	0.57± 0.11	2.31± 1.44	9.54± 2.45	11	0.07± 0.02
SR7	3.34± 1.47	1.23± 0.57	1.22± 0.29	0.21± 0.03	0.60± 0.14	0.32± 0.07	0.52± 0.10	1.23± 0.86	8.68± 1.99	12	0.08± 0.02
SR8	4.11± 1.76	1.55± 0.83	1.70± 0.38	0.37± 0.05	0.36± 0.13	0.32± 0.14	0.54± 0.10	2.01± 1.25	10.96± 2.47	13	0.21± 0.02
SR9	4.48± 1.96	1.89± 0.91	2.24± 0.52	0.54± 0.07	0.54± 0.07	0.36± 0.11	0.51± 0.10	3.20± 1.82	13.76± 3.12	17	0.35± 0.03
SR10	4.15± 1.66	1.63± 0.82	2.13± 0.48	0.49± 0.07	0.38± 0.20	0.30± 0.10	0.37± 0.07	1.25± 1.43	10.72± 2.54	17	0.52± 0.04
SR11	3.29± 1.44	1.69± 0.69	2.06± 0.47	0.55± 0.07	0.13± 0.13	0.19± 0.06	0.28± 0.05	1.46± 1.21	9.64± 2.27	13	0.79± 0.05
SR12	1.98± 0.90	1.17± 0.47	1.42± 0.33	0.47± 0.06	0.19± 0.03	0.14± 0.10	0.16± 0.03	2.38± 1.37	7.92± 1.87	11	1.03± 0.05
SR13	1.04± 0.51	0.60± 0.34	0.86± 0.22	0.27± 0.03	0.07± 0.03	0.11± 0.02	0.07± 0.01	0.37± 0.42	3.39± 0.94	5	1.09± 0.05
SR14	0.52± 0.26	0.28± 0.13	0.44± 0.12	0.16± 0.02	0.03± 0.02	0.10± 0.02	0.03± 0.00	0.23± 0.26	1.78± 0.47	6	1.14± 0.05
SR15	0.28± 0.16	0.11± 0.06	0.24± 0.08	0.10± 0.02	0.03± 0.01	0.04± 0.01	0.03± 0.01	0.14± 0.16	0.96± 0.29	0	1.12± 0.06
SR16	0.13± 0.09	0.01± 0.02	0.09± 0.03	0.04± 0.00	0.00± 0.00	0.03± 0.00	0.01± 0.00	0.03± 0.06	0.33± 0.14	1	0.91± 0.06
SR17	0.00± 0.00	0.02± 0.02	0.01± 0.00	0.00± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.25± 0.32	0.29± 0.32	0	0.26± 0.03

Table 24: Prefit event yields in BDT regions for 2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.88± 0.31	30.20±11.24	1.34± 0.35	0.86± 0.13	0.70± 0.42	6.36± 1.55	0.01± 0.00	1.72± 0.66	42.07±11.64	51	0.49± 0.03
SR1	0.50± 0.38	0.18± 0.10	0.08± 0.06	0.01± 0.01	0.24± 0.20	0.11± 0.17	0.29± 0.05	0.36± 0.36	1.76± 0.89	1	0.00± 0.00
SR2	1.68± 0.87	0.44± 0.28	0.38± 0.20	0.03± 0.02	0.97± 0.49	0.50± 0.19	0.68± 0.13	1.97± 1.04	6.65± 1.59	8	0.00± 0.01
SR3	3.58± 1.52	0.91± 0.49	0.87± 0.31	0.14± 0.04	1.01± 0.30	0.77± 0.24	1.08± 0.20	4.74± 2.28	13.09± 3.14	13	0.02± 0.01
SR4	5.17± 2.33	1.24± 0.71	1.39± 0.42	0.21± 0.04	0.40± 0.49	0.72± 0.30	1.17± 0.22	3.95± 1.41	14.25± 3.64	12	0.04± 0.01
SR5	6.35± 2.41	1.54± 1.04	1.69± 0.50	0.30± 0.06	0.41± 0.29	0.62± 0.35	1.19± 0.22	3.47± 1.57	15.56± 3.92	18	0.06± 0.01
SR6	6.47± 2.64	1.85± 0.89	1.98± 0.54	0.35± 0.05	0.02± 0.65	0.59± 0.21	1.06± 0.20	4.12± 1.99	16.44± 4.36	25	0.10± 0.03
SR7	6.14± 2.33	2.26± 1.26	2.13± 0.58	0.42± 0.08	0.93± 0.20	0.52± 0.13	0.87± 0.16	4.73± 2.16	18.00± 3.90	19	0.15± 0.03
SR8	7.76± 2.87	2.88± 1.72	2.86± 0.77	0.59± 0.07	0.87± 0.24	0.59± 0.26	0.89± 0.17	3.86± 1.57	20.31± 4.19	21	0.30± 0.04
SR9	8.29± 3.19	3.30± 2.43	3.86± 0.98	0.86± 0.11	0.95± 0.24	0.87± 0.27	0.99± 0.19	6.14± 2.60	25.28± 5.48	19	0.46± 0.07
SR10	7.39± 2.78	2.96± 1.84	3.52± 0.95	0.90± 0.10	0.80± 0.34	0.56± 0.28	0.70± 0.13	5.98± 2.66	22.81± 4.82	32	0.83± 0.04
SR11	5.86± 2.36	2.99± 1.60	3.50± 0.93	0.99± 0.12	0.37± 0.07	0.52± 0.15	0.52± 0.10	2.59± 1.35	17.34± 3.78	35	1.38± 0.06
SR12	3.77± 1.37	2.03± 1.05	2.49± 0.68	0.72± 0.09	0.34± 0.21	0.52± 0.15	0.31± 0.06	3.04± 1.69	13.20± 2.82	17	1.65± 0.08
SR13	2.14± 0.88	1.08± 0.68	1.54± 0.42	0.48± 0.07	0.18± 0.04	0.27± 0.09	0.13± 0.03	1.22± 0.51	7.05± 1.55	5	2.00± 0.09
SR14	0.98± 0.45	0.66± 0.34	0.90± 0.26	0.27± 0.04	0.08± 0.02	0.13± 0.03	0.10± 0.02	1.71± 1.10	4.84± 1.39	6	1.83± 0.10
SR15	0.79± 0.37	0.42± 0.20	0.49± 0.14	0.19± 0.02	0.17± 0.12	0.10± 0.02	0.07± 0.01	0.19± 0.10	2.41± 0.60	3	1.78± 0.14
SR16	0.18± 0.10	0.10± 0.09	0.14± 0.06	0.08± 0.01	0.01± 0.00	0.05± 0.01	0.01± 0.00	0.70± 0.47	1.27± 0.51	4	1.53± 0.11
SR17	0.00± 0.02	0.05± 0.05	0.02± 0.01	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.03	0.10± 0.08	3	0.52± 0.07

Table 25: Prefit event yields in BDT regions for 2016+2017+2018.

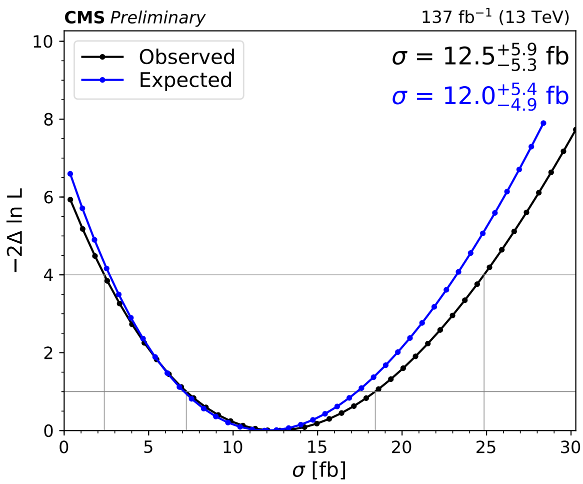
	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	1.92± 0.76	60.77±19.35	2.80± 0.61	1.94± 0.25	0.84± 0.35	12.94± 2.43	0.03± 0.00	2.94± 0.97	84.18±19.84	104	1.05± 0.05
SR1	1.02± 0.66	0.28± 0.15	0.19± 0.13	0.02± 0.01	0.40± 0.27	0.18± 0.24	0.66± 0.07	0.86± 0.55	3.61± 1.31	4	0.00± 0.00
SR2	3.45± 1.75	0.83± 0.45	0.80± 0.35	0.07± 0.03	2.19± 0.62	0.85± 0.27	1.52± 0.17	2.71± 1.28	12.43± 2.62	19	0.01± 0.01
SR3	6.83± 3.29	1.82± 0.80	1.72± 0.57	0.25± 0.05	1.65± 0.44	1.14± 0.38	2.26± 0.26	7.90± 2.79	23.58± 5.19	19	0.04± 0.02
SR4	10.53± 4.72	2.47± 1.13	2.83± 0.80	0.39± 0.08	1.66± 0.65	1.18± 0.45	2.51± 0.29	7.86± 2.55	29.44± 6.65	33	0.09± 0.03
SR5	12.38± 5.22	3.04± 1.56	3.45± 0.96	0.58± 0.10	1.49± 0.58	1.23± 0.54	2.59± 0.29	6.64± 2.23	31.41± 7.00	36	0.15± 0.03
SR6	12.60± 5.48	3.71± 1.54	4.12± 1.03	0.74± 0.09	0.97± 0.87	1.35± 0.37	2.17± 0.25	7.76± 2.77	33.41± 7.42	44	0.22± 0.06
SR7	12.35± 5.14	4.50± 2.09	4.43± 1.11	0.81± 0.14	2.21± 0.36	1.02± 0.26	1.86± 0.21	7.27± 2.42	34.45± 6.85	41	0.31± 0.07
SR8	15.32± 6.19	5.66± 2.87	6.10± 1.45	1.21± 0.15	1.61± 0.37	1.21± 0.43	1.95± 0.22	7.16± 2.61	40.22± 8.27	46	0.68± 0.05
SR9	17.01± 7.14	6.82± 3.69	7.95± 1.84	1.87± 0.25	2.14± 0.35	1.56± 0.41	2.02± 0.23	11.11± 4.16	50.48±10.65	48	1.12± 0.08
SR10	14.93± 6.06	6.20± 3.08	7.53± 1.71	1.87± 0.24	1.46± 0.46	1.14± 0.36	1.40± 0.16	8.86± 3.31	43.40± 8.48	61	1.81± 0.07
SR11	11.98± 5.15	6.11± 2.67	7.27± 1.67	2.03± 0.26	0.95± 0.19	0.92± 0.22	1.07± 0.12	4.79± 2.15	35.12± 7.15	62	2.83± 0.14
SR12	7.43± 3.07	4.18± 1.74	5.24± 1.22	1.54± 0.19	0.70± 0.22	0.89± 0.26	0.61± 0.07	6.41± 2.60	27.00± 5.09	40	3.54± 0.12
SR13	4.20± 1.87	2.40± 1.18	3.15± 0.77	1.01± 0.14	0.26± 0.06	0.52± 0.14	0.28± 0.03	2.31± 0.88	14.13± 3.00	15	4.01± 0.14
SR14	2.06± 1.01	1.32± 0.51	1.78± 0.48	0.55± 0.07	0.11± 0.08	0.30± 0.05	0.18± 0.02	2.21± 1.13	8.50± 1.93	16	3.93± 0.15
SR15	1.38± 0.73	0.68± 0.29	0.96± 0.28	0.37± 0.05	0.21± 0.11	0.17± 0.03	0.12± 0.01	0.44± 0.24	4.33± 1.13	4	3.84± 0.20
SR16	0.43± 0.25	0.20± 0.12	0.32± 0.13	0.15± 0.02	0.07± 0.04	0.10± 0.02	0.03± 0.00	0.84± 0.48	2.13± 0.66	7	3.19± 0.17
SR17	0.03± 0.03	0.07± 0.05	0.03± 0.02	0.02± 0.00	0.00± 0.00	0.02± 0.00	0.00± 0.00	0.25± 0.30	0.44± 0.33	3	1.01± 0.10

Table 26: Postfit event yields in SR+CR regions for 2016+2017+2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	2.72± 0.56	75.97±10.55	3.00± 0.67	1.98± 0.24	0.85± 0.37	13.37± 2.24	0.03± 0.00	3.24± 1.03	101.17±10.12	104	0.83± 0.49
CRW	142.07±27.77	34.87± 4.83	37.34± 8.06	7.39± 0.83	14.31± 1.79	9.82± 1.74	18.78± 2.53	66.66±19.89	331.24±18.64	338	3.88± 2.28
SR1	10.22± 2.17	3.29± 0.47	5.16± 1.17	1.23± 0.15	0.64± 0.13	0.62± 0.12	0.49± 0.07	3.99± 1.61	25.64± 2.09	33	1.98± 1.18
SR2	3.61± 0.86	0.69± 0.23	1.63± 0.41	0.39± 0.06	0.15± 0.05	0.17± 0.04	0.10± 0.01	2.40± 1.04	9.15± 1.26	9	1.13± 0.65
SR3	0.72± 0.34	0.20± 0.10	0.65± 0.20	0.13± 0.02	0.02± 0.02	0.04± 0.01	0.04± 0.01	0.21± 0.17	2.01± 0.58	3	0.73± 0.42
SR4	4.02± 0.93	1.58± 0.34	1.94± 0.46	0.61± 0.07	0.44± 0.13	0.31± 0.06	0.32± 0.04	2.13± 0.84	11.34± 1.25	14	1.58± 0.90
SR5	2.20± 0.57	0.61± 0.13	1.01± 0.27	0.32± 0.04	0.10± 0.04	0.26± 0.05	0.10± 0.02	0.44± 0.27	5.03± 0.77	5	1.68± 0.95
SR6	0.80± 0.21	0.18± 0.09	0.43± 0.13	0.10± 0.02	0.13± 0.05	0.05± 0.01	0.02± 0.00	0.56± 0.29	2.29± 0.40	8	1.20± 0.67
SR7	0.31± 0.11	0.12± 0.04	0.20± 0.06	0.04± 0.01	0.01± 0.01	0.03± 0.01	0.01± 0.00	0.00± 0.08	0.71± 0.20	0	0.88± 0.48
SR8	0.70± 0.28	0.28± 0.12	0.42± 0.15	0.22± 0.03	0.05± 0.02	0.11± 0.02	0.08± 0.01	1.44± 0.81	3.31± 0.95	5	2.20± 1.27
SR9	1.46± 0.41	2.24± 0.34	1.58± 0.36	0.31± 0.05	0.14± 0.02	0.16± 0.09	0.00± 0.00	0.94± 0.46	6.84± 0.80	6	0.71± 0.39
SR10	0.33± 0.11	0.63± 0.14	0.56± 0.14	0.11± 0.02	0.01± 0.01	0.05± 0.01	0.00± 0.00	0.42± 0.26	2.10± 0.31	3	0.35± 0.22
SR11	0.18± 0.07	0.32± 0.08	0.20± 0.05	0.04± 0.01	0.01± 0.01	0.02± 0.01	0.00± 0.00	0.60± 0.72	1.38± 0.75	1	0.23± 0.14
SR12	0.22± 0.08	0.61± 0.12	0.42± 0.10	0.09± 0.01	0.04± 0.01	0.06± 0.01	0.00± 0.00	0.59± 0.39	2.03± 0.48	2	0.59± 0.34
SR13	0.29± 0.11	0.36± 0.12	0.31± 0.09	0.07± 0.01	0.02± 0.01	0.04± 0.01	0.00± 0.00	0.00± 0.11	1.09± 0.28	2	0.69± 0.39
SR14	0.16± 0.05	0.23± 0.07	0.18± 0.06	0.04± 0.01	0.00± 0.00	0.03± 0.01	0.00± 0.00	0.23± 0.27	0.87± 0.30	1	0.80± 0.45

Table 27: Postfit event yields in BDT regions for 2016+2017+2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	2.50± 0.54	77.26±11.71	3.09± 0.64	1.99± 0.21	0.85± 0.35	13.33± 2.56	0.03± 0.00	3.23± 1.18	102.28±11.58	104	1.11± 0.43
SR1	1.25± 0.49	0.31± 0.10	0.19± 0.10	0.02± 0.01	0.32± 0.25	0.17± 0.19	0.66± 0.07	1.04± 0.64	3.95± 0.96	4	0.00± 0.00
SR2	4.35± 1.28	0.97± 0.32	0.83± 0.29	0.08± 0.02	2.20± 0.54	0.79± 0.23	1.53± 0.17	3.46± 1.44	14.19± 1.76	19	0.01± 0.01
SR3	8.53± 2.40	2.17± 0.50	1.75± 0.53	0.23± 0.05	1.46± 0.47	1.12± 0.38	2.26± 0.26	8.01± 2.77	25.53± 3.53	19	0.04± 0.03
SR4	13.52± 3.37	3.04± 0.68	3.00± 0.76	0.38± 0.08	1.57± 0.51	1.19± 0.40	2.52± 0.29	8.74± 3.07	33.96± 4.01	33	0.08± 0.05
SR5	15.80± 3.69	3.73± 0.93	3.68± 0.92	0.57± 0.08	1.61± 0.49	1.27± 0.46	2.59± 0.29	7.41± 2.61	36.67± 3.96	36	0.15± 0.07
SR6	16.34± 3.87	4.50± 0.99	4.41± 1.03	0.75± 0.08	1.26± 0.69	1.40± 0.33	2.18± 0.25	8.99± 3.45	39.81± 4.16	44	0.23± 0.12
SR7	15.96± 3.61	5.53± 1.27	4.75± 1.10	0.82± 0.12	2.25± 0.31	1.03± 0.26	1.86± 0.21	8.12± 3.05	40.32± 3.73	41	0.31± 0.16
SR8	19.81± 4.39	6.88± 1.74	6.61± 1.47	1.21± 0.14	1.65± 0.32	1.21± 0.34	1.95± 0.22	7.97± 2.80	47.28± 4.33	46	0.72± 0.28
SR9	22.22± 4.94	8.57± 2.22	8.63± 1.89	1.88± 0.21	2.15± 0.30	1.53± 0.42	2.02± 0.24	11.50± 3.92	58.51± 5.22	48	1.18± 0.46
SR10	19.68± 4.28	7.79± 1.84	8.31± 1.78	1.91± 0.21	1.39± 0.37	1.18± 0.29	1.40± 0.16	10.48± 3.93	52.15± 4.28	61	1.91± 0.74
SR11	15.96± 3.65	7.82± 1.52	8.03± 1.73	2.06± 0.23	0.98± 0.17	0.97± 0.22	1.07± 0.12	6.10± 2.62	43.00± 3.52	62	2.98± 1.19
SR12	9.66± 2.22	5.30± 1.00	5.81± 1.26	1.56± 0.18	0.81± 0.21	0.89± 0.24	0.61± 0.07	7.46± 3.03	32.10± 3.04	40	3.74± 1.41
SR13	5.57± 1.34	3.01± 0.66	3.50± 0.77	1.04± 0.12	0.27± 0.05	0.51± 0.12	0.28± 0.03	2.52± 0.93	16.71± 1.62	15	4.25± 1.63
SR14	2.74± 0.74	1.67± 0.28	1.97± 0.48	0.56± 0.07	0.11± 0.07	0.31± 0.06	0.18± 0.02	2.61± 1.23	10.14± 1.24	16	4.17± 1.59
SR15	1.81± 0.53	0.84± 0.23	1.06± 0.26	0.38± 0.04	0.20± 0.10	0.17± 0.03	0.12± 0.01	0.45± 0.22	5.03± 0.77	4	4.09± 1.55
SR16	0.55± 0.21	0.26± 0.10	0.36± 0.12	0.15± 0.02	0.07± 0.04	0.10± 0.02	0.03± 0.00	0.97± 0.54	2.49± 0.61	7	3.37± 1.25
SR17	0.05± 0.03	0.09± 0.05	0.04± 0.02	0.02± 0.00	0.00± 0.00	0.02± 0.00	0.00± 0.00	0.34± 0.35	0.57± 0.36	3	1.08± 0.42



(a)

Figure 58: Observed/expected likelihood scans for Run 2 with the BDT analysis

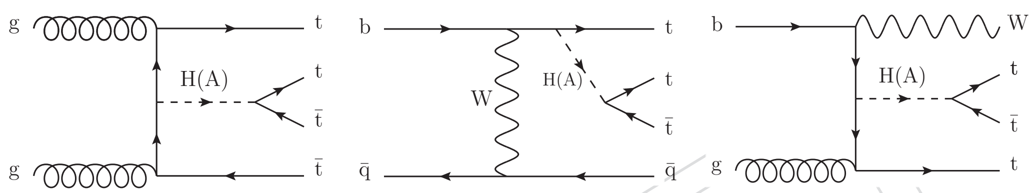
12 Results: interpretations

12.1 Type-II 2HDM

12.1.1 Introduction

Final states with 3 or more top quarks occur in many scenarios of new physics.

In in two Higgs doublet models (2HDM) [23, 24] of type-II, the couplings of the CP-even scalar h become exactly SM-like in the so-called ”alignment limit”, $\sin(\beta - \alpha) \rightarrow 1$. In such models, the couplings of the heavy scalar and pseudoscalar to the SM vector bosons are suppressed, vanishing as $\cos(\beta - \alpha) \rightarrow 0$. In such a limit, production is predominantly via gluon-fusion and then associated production with either $b\bar{b}$ or $t\bar{t}$. The sensitivity of the direct search for resonant $t\bar{t}$ production is significantly reduced due to interference with SM QCD production [25, 26]. As a result, at low $\tan\beta$, three and four top final states as seen in the diagrams in Figure 59 provide a promising window to probe this scenario [27].



(a) $t\bar{t}$H (b) tHq (c) tHW

Figure 59: Diagrams for scalar (pseudoscalar) production in association with top quarks. In some scenarios, these heavy bosons will preferentially decay to a pair of top quarks, generating final states with three or four top quarks.

12.1.2 Simulation

We interpret the results of the SM $t\bar{t}t\bar{t}$ search as limits on the top-associated production of a scalar or pseudoscalar with subsequent decay to a pair of top quarks.

We generate a one-dimensional grid of points for scalar masses between 350 and 550 GeV in 20 GeV steps for the three processes shown in Figure 59 at LO with MadGraph in the 5-flavor scheme using the NN23LO pdf set. The $t\bar{t}H$ and tWH processes were generated with up to one additional parton, while the tHq process was generated with no additional partons due to the presence of a light flavor quark in the lowest order ME.

For the purpose of interpretation, we use LO cross sections for the production of a heavy scalar or pseudoscalar boson using the 2HDMtII_NLO MadGraph model with the `NNPDF30_lo_as_0130` PDF. Processes involving H and A mediators are generated separately and charged higgses ($H^{\pm}$) are decoupled by setting their mass to 10TeV. We fix $\sin(\beta - \alpha) = 1$ for the alignment condition, and use $\tan\beta$ as an additional parameter. For consistency with previous results and after verifying consistency with the results of [27], $\tan\beta = 1$ is used for one-dimensional exclusions as a function of mediator mass. To construct a two-dimensional exclusion plane, we calculate cross-sections for the same mediator mass grid with $\tan\beta$ values ranging from 0.5 to 3. Note that with the type-II 2HDM, the branching ratio to up-type quarks (e.g., top quark) is proportional to $\frac{1}{\tan\beta}$.

For the pseudoscalar case, we use the same events as the scalar scenario but with cross sections determined for the case of a CP-odd scalar with the same mass.

12.1.3 Kinematic comparison of scalar and pseudoscalar diagrams

Figure 60 shows analysis-level quantities for $t\bar{t}H$ and $t\bar{t}A$, with the mediator mass set to 450 GeV at $\tan\beta = 1$. The events were simulated with the `2HDMtII_NLO` MG5 model and reconstructed with the 2017 `MINIAODSIM` workflow. In general, acceptance and kinematics are similar between scalar and pseudoscalar processes, especially the final distribution of yields with the BDT-based regions, where a KS statistical test yields a p-value of 0.8. For this reason, scalar samples can be used for both processes as the production cross-section is the primary difference.

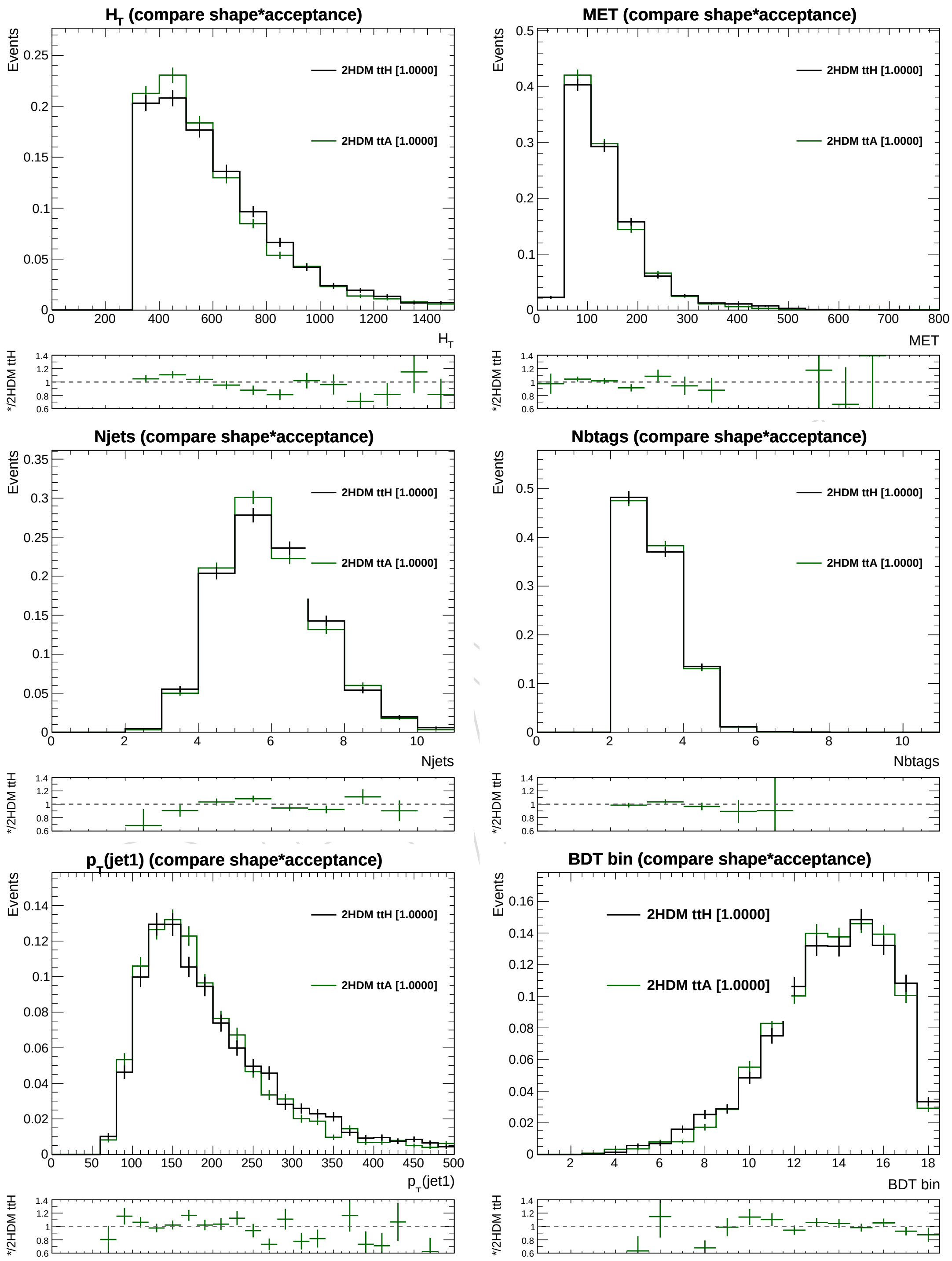


Figure 60: Distributions comparing H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, N_{jets}, $N_{\mathrm{b\ jets}}$, p_{T} of the leading jet, and the BDT signal region yields for scalar and pseudoscalar processes. The cross-section has been normalized away, so both shapes and acceptance are relevant here.

12.1.4 Cross-sections vs $\tan\beta$

Cross sections for the scalar and pseudoscalar processes are plotted in Figure 61 and summarized in Table 28, separated by the value of $\tan\beta$. Comparing these limits with the Type-2 2HDM cross sections with $\tan\beta = 1$ in the alignment limit, $\sin(\beta - \alpha) = 1$, we exclude scalar (pseudoscalar) masses up to 470 (550) GeV. Alternatively, we consider the simplified model of dark matter defined in Ref. [28], which includes a Dirac fermion dark matter candidate, χ, in addition to H/A, and where the couplings of H/A to SM fermions and χ are determined by parameters g_{SM} and g_{DM}, respectively. In this model, exclusions similar to the 2HDM ones are reached by assuming $g_{\mathrm{SM}} = 1$ and $g_{\mathrm{DM}} = 1$, and taking $m_{\mathrm{H/A}} > 2m_{\chi}$

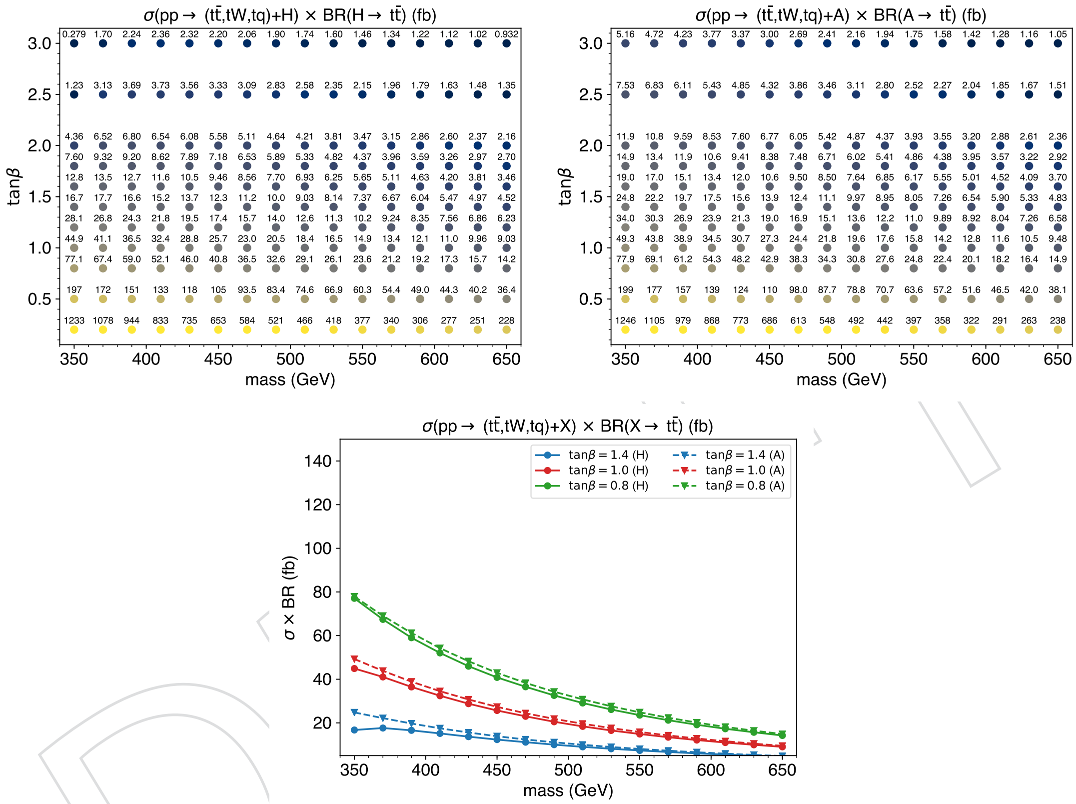


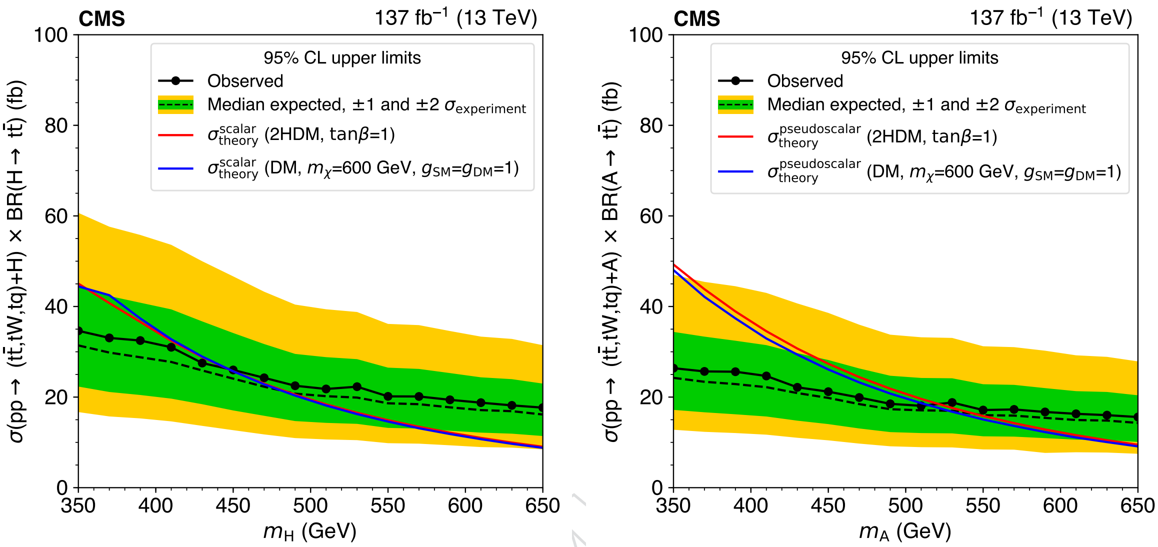
Figure 61: Cross-sections in units of fb in the plane of $\tan\beta$ vs mediator mass for a heavy scalar boson (top left) and heavy pseudoscalar boson (top right) The bottom plot cotains 1-D projections for both mediators for a few values of $\tan\beta$.

Table 28: Cross sections for the case of scalar (H) and pseudoproduction (A), assuming a heavy higgs boson with SM-like top quark couplings. All cross-sections are reported in units of fb, and masses in units of GeV. Based on Figure 5 of arXiv:1605.08744, the LO scale uncertainties are close to 40% for ttH–¿tttt, mH 400. This is consistent with the LO uncertainty on SM tttt. Single top processes, tHX–¿tttX have a smaller uncertainty (15%) as they have one less power of α_S.

$\tan\beta$	mass	σ_{ttH}	σ_{tWH}	σ_{tqH}	σ_{ttA}	σ_{tWA}	σ_{tqA}	$\tan\beta$	mass	σ_{ttH}	σ_{tWH}	σ_{tqH}	σ_{ttA}	σ_{tWA}	σ_{tqA}
	350	81.4	38.0	77.9	117	36.3	45.7		350	31.8	14.9	30.4	45.9	14.2	17.8
	370	71.4	34.3	66.7	103	33.0	40.4		370	27.9	13.4	26.1	40.4	12.9	15.8
	390	62.5	31.0	57.5	91.1	29.9	35.7		390	24.4	12.1	22.5	35.6	11.7	13.9
	410	55.2	28.1	49.9	80.1	27.1	31.7		410	21.6	11.0	19.5	31.3	10.6	12.4
	430	48.8	25.4	43.4	70.8	24.7	28.2		430	19.1	9.94	17.0	27.6	9.63	11.0
	450	43.6	23.1	37.9	62.4	22.3	25.1		450	17.0	9.02	14.8	24.4	8.73	9.83
	470	39.1	21.0	33.4	55.4	20.3	22.3		470	15.3	8.22	13.0	21.6	7.94	8.74
0.5	490	35.0	19.1	29.3	49.2	18.5	20.0	0.8	490	13.7	7.47	11.4	19.2	7.23	7.82
	510	31.5	17.4	25.7	43.9	16.9	18.0		510	12.3	6.78	10.0	17.2	6.62	7.02
	530	28.3	15.8	22.7	39.1	15.4	16.1		530	11.1	6.17	8.89	15.3	6.03	6.29
	550	25.6	14.4	20.2	35.0	14.1	14.5		550	10.0	5.63	7.90	13.7	5.53	5.65
	570	23.2	13.1	18.0	31.3	12.9	13.0		570	9.07	5.13	7.03	12.2	5.05	5.07
	590	21.0	12.0	16.0	28.1	11.8	11.7		590	8.21	4.70	6.25	11.0	4.61	4.57
	610	19.1	11.0	14.2	25.2	10.8	10.6		610	7.46	4.29	5.56	9.83	4.21	4.13
	630	17.3	10.1	12.8	22.6	9.89	9.54		630	6.77	3.95	4.99	8.83	3.86	3.72
	650	15.8	9.24	11.4	20.4	9.09	8.64		650	6.16	3.61	4.46	7.96	3.55	3.37
	350	18.5	8.66	17.7	29.0	8.98	11.3		350	11.6	5.40	11.1	20.0	6.20	7.80
	370	17.0	8.19	15.9	25.6	8.18	10.0		370	11.1	5.34	10.4	17.7	5.67	6.92
	390	15.1	7.49	13.9	22.6	7.43	8.85		390	10.0	4.98	9.25	15.7	5.15	6.14
	410	13.4	6.85	12.2	19.9	6.76	7.87		410	9.01	4.60	8.18	13.8	4.68	5.45
	430	11.9	6.24	10.6	17.6	6.13	7.00		430	8.07	4.21	7.18	12.2	4.25	4.86
	450	10.7	5.68	9.32	15.5	5.56	6.26		450	7.25	3.86	6.33	10.8	3.86	4.34
	470	9.61	5.18	8.22	13.8	5.06	5.57		470	6.54	3.53	5.60	9.55	3.51	3.86
1.0	490	8.63	4.70	7.19	12.2	4.61	4.98	1.2	490	5.88	3.21	4.92	8.49	3.20	3.46
	510	7.77	4.28	6.35	10.9	4.20	4.47		510	5.31	2.94	4.34	7.58	2.92	3.10
	530	7.00	3.91	5.62	9.74	3.85	4.01		530	4.79	2.67	3.85	6.75	2.66	2.78
	550	6.34	3.56	5.00	8.71	3.51	3.60		550	4.35	2.44	3.43	6.04	2.45	2.50
	570	5.75	3.25	4.44	7.80	3.21	3.23		570	3.94	2.23	3.07	5.41	2.23	2.24
	590	5.20	2.98	3.96	6.99	2.94	2.92		590	3.57	2.05	2.73	4.85	2.05	2.02
	610	4.73	2.72	3.54	6.27	2.69	2.63		610	3.25	1.87	2.43	4.35	1.86	1.83
	630	4.29	2.50	3.16	5.63	2.47	2.38		630	2.95	1.72	2.18	3.91	1.71	1.65
	650	3.91	2.28	2.83	5.08	2.25	2.15		650	2.69	1.58	1.95	3.52	1.57	1.49
	350	6.89	3.23	6.62	14.6	4.52	5.69		350	5.28	2.48	5.08	11.2	3.47	4.36
	370	7.29	3.52	6.84	13.0	4.15	5.07		370	5.58	2.70	5.24	9.92	3.18	3.89
	390	6.85	3.42	6.32	11.5	3.77	4.49		390	5.24	2.62	4.85	8.77	2.89	3.44
	410	6.27	3.20	5.70	10.1	3.43	4.00		410	4.80	2.45	4.37	7.72	2.63	3.07
	430	5.67	2.97	5.05	8.92	3.12	3.56		430	4.34	2.27	3.88	6.83	2.39	2.73
	450	5.13	2.73	4.48	7.88	2.83	3.17		450	3.93	2.09	3.44	6.03	2.17	2.44
	470	4.66	2.52	3.99	6.99	2.57	2.83		470	3.57	1.93	3.06	5.35	1.97	2.17
1.4	490	4.21	2.30	3.52	6.22	2.34	2.53	1.6	490	3.22	1.77	2.71	4.76	1.80	1.94
	510	3.81	2.10	3.12	5.55	2.15	2.28		510	2.91	1.62	2.39	4.25	1.64	1.75
	530	3.44	1.92	2.78	4.95	1.96	2.04		530	2.64	1.48	2.13	3.79	1.50	1.56
	550	3.13	1.76	2.48	4.43	1.79	1.83		550	2.40	1.35	1.90	3.39	1.37	1.41
	570	2.85	1.61	2.21	3.97	1.64	1.65		570	2.18	1.24	1.70	3.04	1.25	1.26
	590	2.58	1.48	1.97	3.55	1.50	1.49		590	1.98	1.14	1.51	2.72	1.15	1.14
	610	2.35	1.36	1.76	3.19	1.37	1.34		610	1.80	1.04	1.35	2.44	1.05	1.03
	630	2.14	1.25	1.58	2.86	1.26	1.21		630	1.64	0.956	1.21	2.19	0.964	0.929
	650	1.95	1.15	1.42	2.58	1.15	1.10		650	1.49	0.877	1.09	1.98	0.884	0.840
	350	3.12	1.47	3.01	8.73	2.72	3.42		350	1.79	0.845	1.73	7.00	2.19	2.75
	370	3.84	1.86	3.62	7.79	2.51	3.06		370	2.68	1.31	2.53	6.26	2.03	2.47
	390	3.79	1.90	3.51	6.89	2.28	2.72		390	2.80	1.41	2.60	5.55	1.84	2.19
	410	3.55	1.83	3.24	6.07	2.07	2.42		410	2.69	1.39	2.46	4.89	1.68	1.96
	430	3.26	1.72	2.92	5.37	1.88	2.15		430	2.50	1.32	2.25	4.33	1.53	1.74
	450	2.97	1.59	2.61	4.75	1.71	1.92		450	2.31	1.24	2.03	3.83	1.39	1.56
	470	2.72	1.48	2.34	4.21	1.56	1.71		470	2.12	1.16	1.83	3.40	1.26	1.39
1.8	490	2.46	1.35	2.08	3.75	1.42	1.54	2.0	490	1.93	1.07	1.64	3.02	1.15	1.24
	510	2.24	1.25	1.84	3.35	1.30	1.38		510	1.76	0.987	1.46	2.70	1.05	1.12
	530	2.03	1.14	1.64	2.98	1.19	1.23		530	1.60	0.904	1.31	2.41	0.963	1.00
	550	1.85	1.05	1.47	2.67	1.08	1.11		550	1.47	0.833	1.17	2.16	0.878	0.900
	570	1.69	0.961	1.31	2.39	0.992	0.998		570	1.34	0.766	1.05	1.93	0.807	0.807
	590	1.53	0.884	1.18	2.14	0.910	0.900		590	1.22	0.706	0.940	1.73	0.736	0.730
	610	1.40	0.811	1.05	1.92	0.830	0.812		610	1.11	0.648	0.841	1.55	0.673	0.658
	630	1.27	0.746	0.945	1.73	0.762	0.734		630	1.02	0.597	0.756	1.40	0.618	0.594
	650	1.16	0.687	0.849	1.56	0.696	0.664		650	0.928	0.551	0.681	1.26	0.567	0.538
	350	0.498	0.240	0.488	4.38	1.40	1.75		350	0.112	0.056	0.112	2.97	0.978	1.22
	370	1.28	0.633	1.22	3.95	1.30	1.58		370	0.682	0.349	0.668	2.69	0.915	1.11
	390	1.50	0.768	1.41	3.51	1.19	1.41		390	0.899	0.476	0.866	2.40	0.836	0.992
	410	1.52	0.800	1.41	3.10	1.08	1.25		410	0.948	0.514	0.902	2.12	0.764	0.886
	430	1.46	0.784	1.32	2.74	0.985	1.12		430	0.933	0.514	0.868	1.88	0.695	0.791
	450	1.37	0.748	1.22	2.43	0.895	1.00		450	0.889	0.502	0.810	1.66	0.634	0.706
	470	1.27	0.706	1.11	2.16	0.816	0.893		470	0.835	0.477	0.746	1.48	0.578	0.630
2.5	490	1.17	0.657	1.00	1.92	0.743	0.800	3.0	490	0.774	0.449	0.678	1.32	0.526	0.565
	510	1.07	0.610	0.898	1.71	0.681	0.718		510	0.714	0.419	0.612	1.18	0.480	0.507
	530	0.982	0.564	0.808	1.53	0.622	0.646		530	0.656	0.389	0.553	1.05	0.440	0.456
	550	0.901	0.520	0.727	1.37	0.568	0.580		550	0.604	0.360	0.499	0.940	0.402	0.409
	570	0.825	0.480	0.654	1.23	0.521	0.521		570	0.554	0.332	0.451	0.842	0.368	0.368
	590	0.753	0.445	0.589	1.10	0.474	0.471		590	0.507	0.308	0.406	0.754	0.337	0.332
	610	0.690	0.409	0.529	0.986	0.435	0.425		610	0.466	0.284	0.365	0.677	0.307	0.300
	630	0.631	0.377	0.476	0.887	0.398	0.384		630	0.426	0.262	0.329	0.609	0.281	0.272
	650	0.578	0.347	0.429	0.800	0.365	0.348		650	0.391	0.243	0.298	0.549	0.258	0.245

12.1.5 Exclusions

With the Run2 BDT analysis and the previous model assumptions, heavy scalar (pseudoscalar) bosons are excluded as shown in Figure 62, as a function of mediator mass for $\tan\beta = 1$, and in Figure 63, as a function of mediator mass and $\tan\beta$, considering one new particle at a time and also including a scenario with $m_{\mathrm{H}} = m_{\mathrm{A}}$ inspired by a special case of Type-2 2HDM, the hMSSM [**?**]. For this interpretation the SM $\mathrm{t\bar{t}t\bar{t}}$ process is treated as a background and assigned its SM cross section and uncertainty $12.0^{+2.2}_{-2.5}$ fb [6].



(a) (b)

Figure 62: Observed and expected limits on heavy scalar (left) and pseudoscalar (right) processes as a function of the (pseudo)scalar mass using FullSim samples. The parameter $\tan\beta$ is assumed to be 1 here. The DM model is described in Section 12.5.

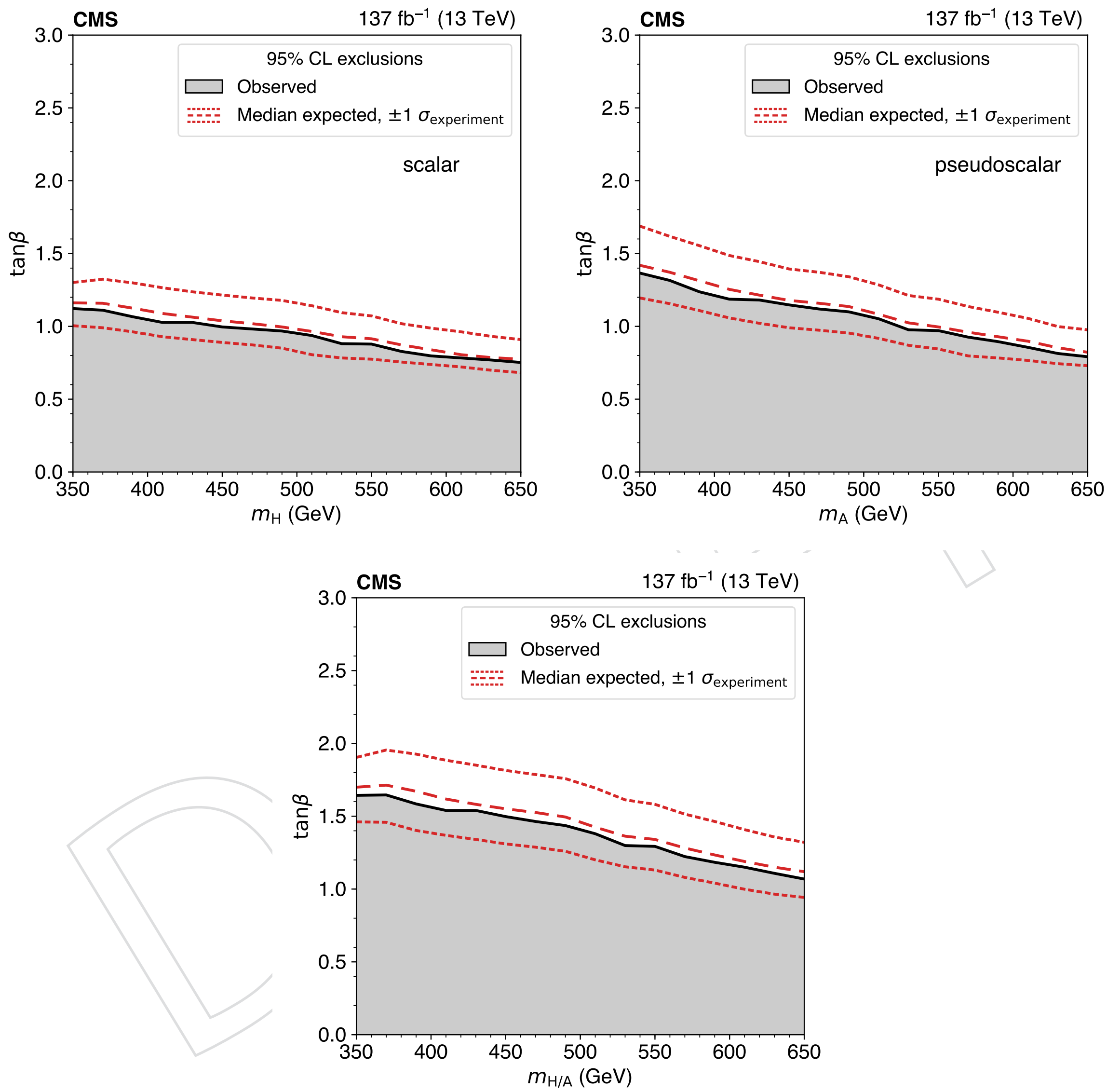


Figure 63: Two-dimensional observed and expected exclusions of heavy scalar only (top left), pseudoscalar only (top right), and both scalar+pseudoscalar (both) as a function of the mediator mass and $\tan\beta$. In each case, other 2HDM particles (except SM higgs) are decoupled.

12.2 Top Yukawa coupling

12.2.1 Introduction

In the SM there are contributions to $pp \to t\bar{t}t\bar{t}$ from diagrams with virtual Higgs bosons, see for example Figure 64. The amplitude corresponding to these diagrams is proportional to the square of the top Yukawa coupling.

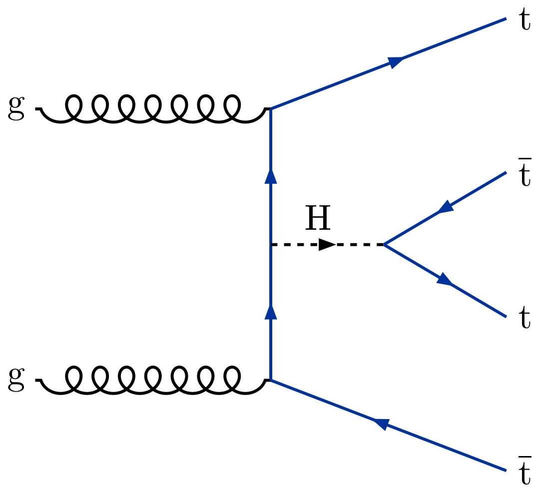


Figure 64: One of the Feynman diagrams for $t\bar{t}t\bar{t}$ including a virtual Higgs.

Using the notation of Reference [29] the $t\bar{t}t\bar{t}$ cross-section can be written as

$$\sigma(t\bar{t}t\bar{t}) = \sigma^{SM}(t\bar{t}t\bar{t})_{g+Z/\gamma} + k_t^4 \sigma^{SM}(t\bar{t}t\bar{t})_H + k_t^2 \sigma_{\text{int}}^{SM} \quad (6)$$

where $k_t \equiv y_t / y_t^{SM}$, y_t is the top Yukawa coupling, and y_t^{SM} is its value in the SM. In equation 6 the first term on the right hand side corresponds to the SM contribution to the cross section from diagrams with virtual gluons or Z/γ, the second term is the contribution from diagrams with virtual H bosons, and the third term is the interference between the two. Therefore, given a theoretical calculation and a measurement of $\sigma(t\bar{t}t\bar{t})$, one can put constraints on $|y_t/y_t^{SM}|$.

The authors of Reference [29] have calculated the cross-section terms at LO, and have provided us privately with the uncertainties under variations of the factorization and renormalizations scales. These are given in Table 29.

	$[\mu/2, \mu, 2\mu]$
$\sigma^{SM}(t\bar{t}t\bar{t})_{g+Z/\gamma}$	[14.104, 9.997, 6.378] fb
$\sigma^{SM}(t\bar{t}t\bar{t})_H$	[1.625, 1.167, 0.7655] fb
σ_{int}^{SM}	[-2.152, -1.547, -0.999] fb

Table 29: LO calculation of the terms in equation 6 from Reference [29]. The scale variations are private communications from the authors.

We have investigated the possibility of a full NLO treatment of the interpretation, and discussed with the authors of the NLO calculation [6], but Madgraph could not provide what we need yet. So, we have decided to continue using the LO calculation for this interpretation. The authors of Reference [29] have also argued that it is appropriate to apply the overall NLO/LO k-factor of 1.27 calculated at 14 TeV for the **total** $\sigma(t\bar{t}t\bar{t})$ cross-section [30] to the individual components in equation 6. This would then result in an NLO cross-section of $12.2^{+5.0}_{-4.4}$ fb. This is in agreement with the NLO calculation of $11.97^{+2.15}_{-2.51}$ fb [6].[1] The authors of Reference [29] then go on to extract a limit on $|y_t/y_t^{SM}|$ based on their calculation and the then-available experimental limit on $\sigma(t\bar{t}t\bar{t})$. Following this procedure in the 2016 result, in Fig. 65 (left) we show the measuerement of the cross section and its upper limit, as well as its SM prediction as a function of the absolute value of the ratio of the top quark Yukawa to its SM value ($|y_t/y_t^{SM}|$). The central (upper,lower) value of the theoretical cross section band resulted in a 95% CL limit $|y_t/y_t^{SM}| < 2.27$ (2.03,2.56) (in the 2016 result).

Noting that ttH is a non-negligible background, showing a flat line for the observed cross section of $t\bar{t}t\bar{t}$ (and observed upper limit) as a function of κ_t is not completely correct, as it neglects the κ_t^2 scaling of ttH, since $\sigma(t\bar{t}H) \propto (y_t/y_t^{SM})^2/(\Gamma_H/\Gamma_H^{SM})$, where Γ_H is the total width of the Higgs. The limit plot therefore is also shown, on the right, for the case of the $t\bar{t}H$ background is scaled by $(y_t/y_t^{SM})^2$ for each value of $|y_t/y_t^{SM}|$. As the $t\bar{t}H$ background becomes larger, fewer $t\bar{t}t\bar{t}$ events are allowed by the fit, and the observed $t\bar{t}t\bar{t}$ upper limit and the measured $t\bar{t}t\bar{t}$ cross sections becomes smaller. With this prescription, the central (upper,lower) value of the theoretical cross section band results in a 95% CL limit $|y_t/y_t^{SM}| < 2.10^{+0.22}_{-0.27}$ (in the 2016 result).

For the full Run2 BDT analysis, this procedure resulted in a 95% CL limit of the central, upper, and lower values of the theoretical cross section provide respective 95% CL limits for $|y_t/y_t^{SM}|$ < 1.7, < 1.4, and < 2.0, as shown in Figure 66.

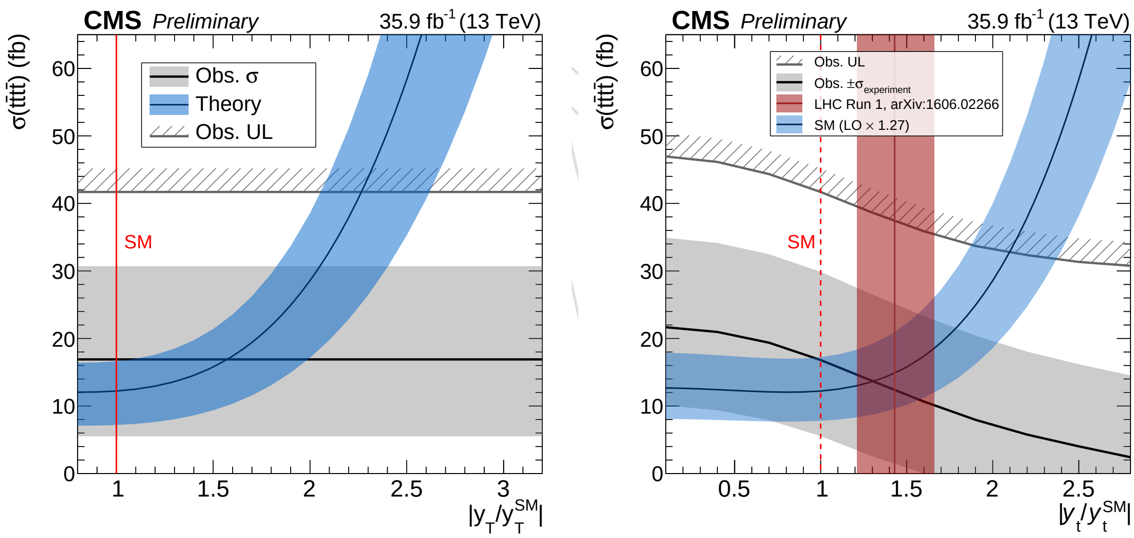


Figure 65: (2016-only results, the next plot shows the full run 2 results).The expected $t\bar{t}t\bar{t}$ cross-section, $\sigma(t\bar{t}t\bar{t})$, as a function of $|y_T/y_T^{SM}|$, the absolute value of the ratio of the top quark Yukawa coupling and its SM value (diagonal band), compared with the measured $\sigma(t\bar{t}t\bar{t})$ (horizontal band), and its 95% CL upper limit (horizontal line). The right plot includes the scaling of the ttH background as a function of y_T as discussed in the text.

12.2.2 Alternative statistical treatment

Note that this section is from 2016, but is left here as a reference

An alternative analysis, still based on the LO cross-section (and its uncertainty) from Reference [29], scaled up by the k-factor of 1.27, consists of interpreting the experimental likelihood

[1] Note that the uncertainties in the full NLO calculation are smaller, but this is to be expected.

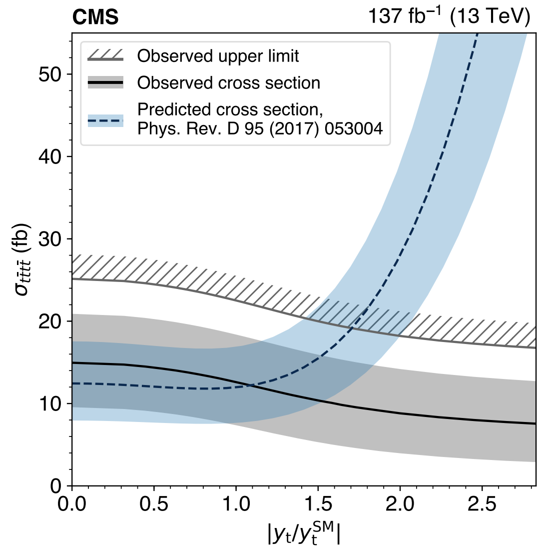


Figure 66: The expected $t\bar{t}t\bar{t}$ cross-section, $\sigma(t\bar{t}t\bar{t})$, as a function of $|y_T/y_T^{SM}|$, the absolute value of the ratio of the top quark Yukawa coupling and its SM value (diagonal band), compared with the measured $\sigma(t\bar{t}t\bar{t})$ (horizontal band), and its 95% CL upper limit (horizontal line)

as the posterior pdf for the cross-section.[2] We then construct (correlated) pdfs for the three terms on the right side of equation 6, and extract a pdf for $|y_T/y_T^{SM}|$ by propagating uncertainties using equation 6 and a toy MC method. We do this in two ways, by either taking the correlated pdfs as bifurcated gaussians (truncated to not allow a change of sign), or as "flat" within the limits of Table 29. The 95% CL limit would then be $|y_T/y_T^{SM}| < 2.45$ (biforcated gaussians pdfs) or $|y_T/y_T^{SM}| < 2.25$ (flat pdfs). These results are to be compared with the "constant ttH" interpretation of Figure 65 (left), which resulted in a 95% CL limits of $|y_T/y_T^{SM}| < 2.27$ (2.03,2.56).

We prefer to present the results according to Figure 65. This is in the same spirit as what is customary in the SUSY group, where the experimental limits are clearly separated from the theoretical uncertainties in the interpretation (but maybe the TOP group does it differently). In addition, we are reluctant to give too much weight to an error analysis that includes the systematic uncertainty from Reference [29], which is somewhat of a hack (a LO calculation with the pieces of the LO calculation scaled by the same k-factor as obtained by another group for the full calculation at a slightly different energy...). Incidentally, the limit on $|y_T/y_T^{SM}|$ quoted in Reference [29] ignores the uncertainty on the theoretical calculation. In any case, we are open to suggestions.

12.2.3 Kinematic dependence on top yukawa coupling

Since MC is taken with a nominal top yukawa coupling and only the cross-section is scaled in the scan above, it is worth verifying that there is no significant kinematic dependence on the top yukawa coupling value. This is checked with LO MG5 (MG5_aMC_v2_6_3_2) with default parameters, including a dynamical scale choice) using the following proc card.

```
set default_unset_couplings 99
```

[2] This would correspond to using a Baysean method with a flat prior in the cross-section. It is not entirely kosher since the likelihood has been profiled with respect to the nuisances, while Baysean approaches usually require marginalization.

```
set group_subprocesses Auto
set ignore_six_quark_processes False
set loop_optimized_output True
set loop_color_flows False
set gauge unitary
set complex_mass_scheme False
set max_npoint_for_channel 0
set nb_core 4


import model sm
define p = g u c d s u~ c~ d~ s~
define p = p b b~
generate p p > t t~ t t~ QED=99


output ftlo_ytscan
launch

# >>> import numpy as np
# >>> x = np.arange(0.4,2.2,0.1)
# >>> print ",".join(map(lambda y: str(round(y,1)),x))
# paste output in scan brackets below

set param_card yukawa 6 scan:[69.2,...,363.3]
```

The LHE for $|y_T/y_T^{SM}|$ between 0.4 and 2.2 in steps of 0.1 was carried through the 2016 MINIAOD-SIM workflow with a slightly modified Pythia fragment which imposes a dilepton (or more) filter for better statistical uncertainties. The relevant part of the fragment is

```
'JetMatching:setMad = off',
'JetMatching:scheme = 1',
'JetMatching:merge = on',
'JetMatching:jetAlgorithm = 2',
'JetMatching:etaJetMax = 5.',
'JetMatching:coneRadius = 1.',
'JetMatching:slowJetPower = 1',
'JetMatching:qCut = 59.',
'JetMatching:nQmatch = 5', #4 corresponds to 4-flavour scheme (no matching of b-
'JetMatching:nJetMax = 0', #number of partons in born matrix element for highest
'JetMatching:doShowerKt = off', #off for MLM matching, turn on for shower-kT mat
'6:m0 = 172.5',
'24:mMin = 0.1',
'23:mMin = 0.1',
'ResonanceDecayFilter:filter = on',
'ResonanceDecayFilter:exclusive = off', #off: require at least the specified num
'ResonanceDecayFilter:eMuAsEquivalent = off', #on: treat electrons and muons as
'ResonanceDecayFilter:eMuTauAsEquivalent = on', #on: treat electrons, muons , an
'ResonanceDecayFilter:allNuAsEquivalent = on', #on: treat all three neutrino fla
'ResonanceDecayFilter:daughters = 11,11',
```

```
'Check:abortIfVeto = on',
```

To validate the MadGraph setup, cross-sections for points along the scan are compared with Equation 6 (using values from Table 29) and they are found to agree within 2%, as shown in Figure 67

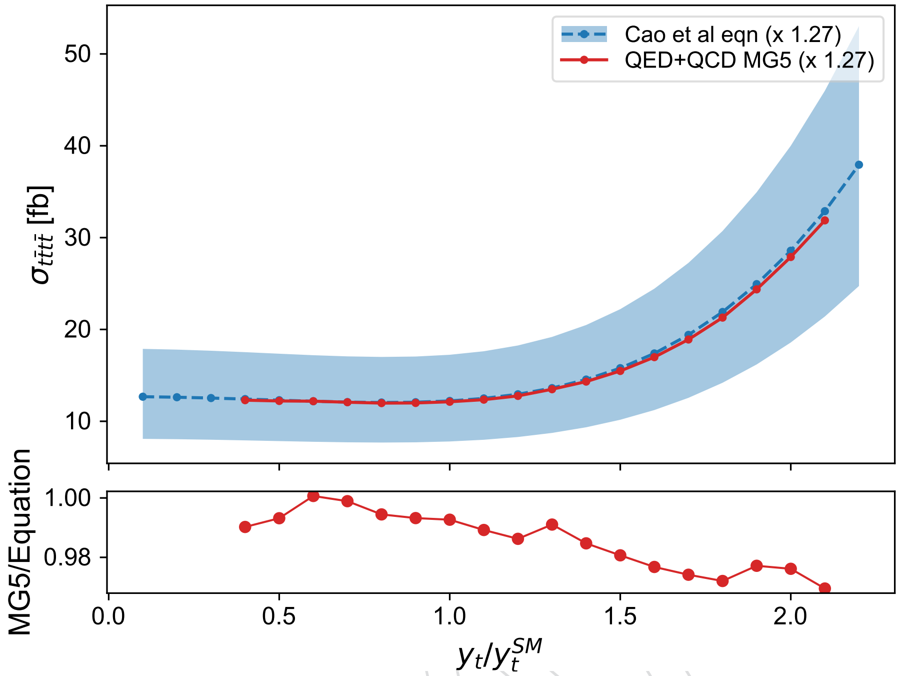


Figure 67: Calculated cross-section from Equation 6 and from MG5 for varying $|y_T/y_T^{SM}|$ values

Finally, plots of various kinematic quantities are shown in Figure 68 for the analysis baseline selection, also allowing opposite-sign events in addition to same-sign events for augmented statistics. There is no visible significant difference in the kinematic plots for different values of $|y_T/y_T^{SM}|$.

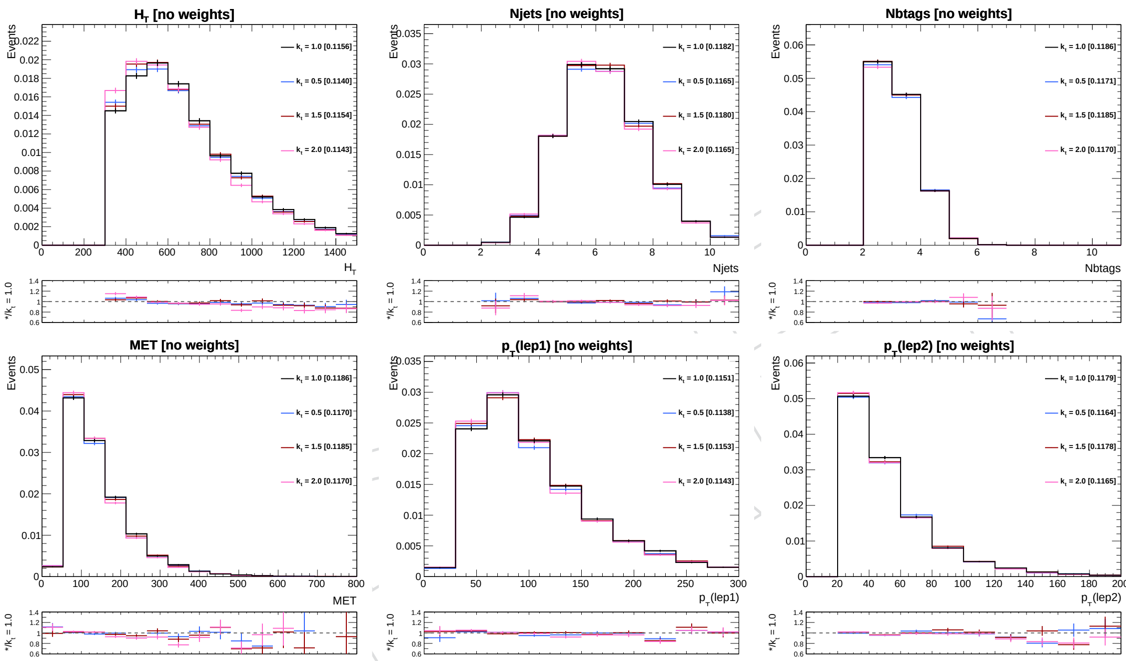


Figure 68: Kinematic quantities (left to right, top to bottom: H_T, N_{jets}, N_{btags}, MET, p_T(lep1), p_T(lep2)) for the baseline selection with same-sign and opposite-sign dileptons for $|y_T/y_T^{SM}|$ of 0.5, 1.0, 1.5, 2.0. All histograms are normalized to the same cross-section.

12.3 Off-shell mediators decaying to top quark pairs

12.3.1 Introduction

The production of $t\bar{t}t\bar{t}$ may also be influenced by a neutral scalar mediator (ϕ) or neutral vector mediator (Z') which couple to top quarks and have masses less than twice the mass of the top quark, distinguishing them from from similar processes within the 2HDM framework, for example. The off-shell contributions to the SM $t\bar{t}t\bar{t}$ production can be large, as shown in Ref. [31]. For a large range of masses, the authors have shown that kinematics are identical when considering these additional processes, so that the total $t\bar{t}t\bar{t}$ cross-section is subject to a simple rescaling. We consider coupling terms in the lagrangian of the form

$$\mathcal{L}_{Z'} = -g_{tZ'}\bar{t}_R \not{Z}' t_R \qquad \mathcal{L}_{\phi} = -g_{t\phi}\bar{t}_L \phi t_R$$

and calculate leading order cross-sections for the process

```
generate p p > t~ t t~ t QED=2
```

with MadGraph UFO models provided by the authors of Ref. [31], which have been copied to `https://github.com/aminnj/FTInterpretations/tree/master/models`. We also considered single top processes, similar to those for the 2HDM exclusions, but found their cross-section to be small compared to $t\bar{t}$-associated production, so they are not included.

Due to the approximate independence of kinematics on the coupling strength and mediator mass, we are able to use the upper limit result from the nominal analysis to place constraints on couplings $g_{tZ'}$ and $g_{t\phi}$ as a function of masses $m_{Z'}$ and m_{ϕ}, respectively, without the use of dedicated signal samples. The nominal analysis uses a NLO $t\bar{t}t\bar{t}$ sample, so to justify this procedure, we first show that the nominal NLO $t\bar{t}t\bar{t}$ sample indeed has good shape agreement with a LO $t\bar{t}t\bar{t}$ generated with these models (setting coupling strengths to 0) in Figure 69. Next, we show various couplings and mass points near the exclusion boundary compared to the LO SM $t\bar{t}t\bar{t}$ sample in Figure 70 for the Z' mediator, and Figure 71 for the ϕ mediator. Based on the level of agreement for both mediator types, we include a 10% additional normalization uncertainty on the SM $t\bar{t}t\bar{t}$ signal to conservatively cover minor acceptance differences.

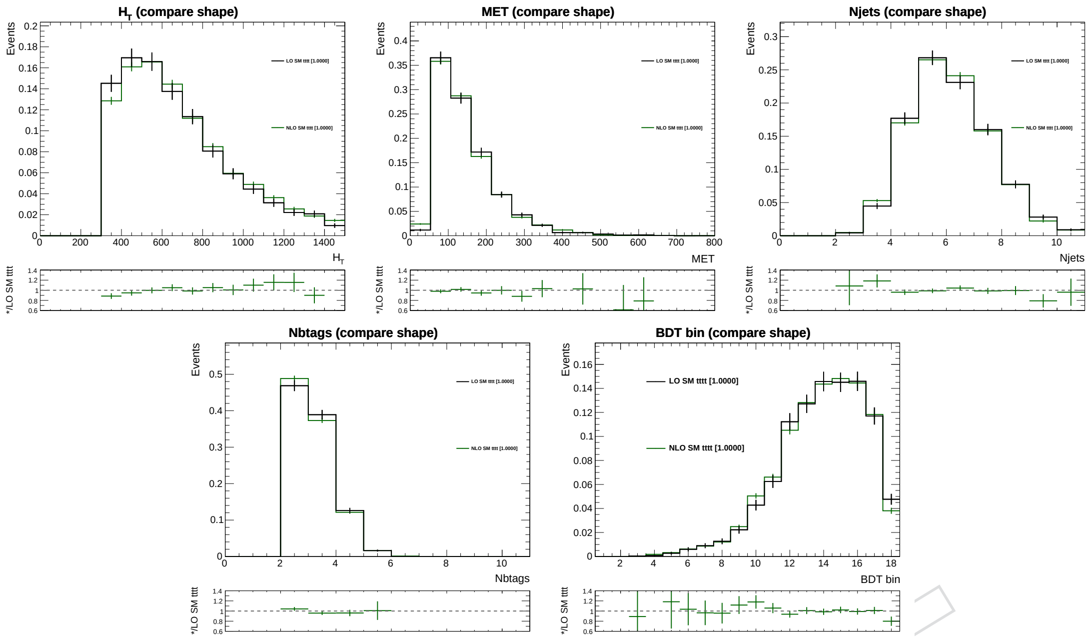


Figure 69: Distributions comparing H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, N_{jets}, $N_{\mathrm{b\ jets}}$, and the BDT signal region yields for the nominal NLO and LO $\mathrm{t\bar{t}t\bar{t}}$ samples. Only shapes are relevant here.

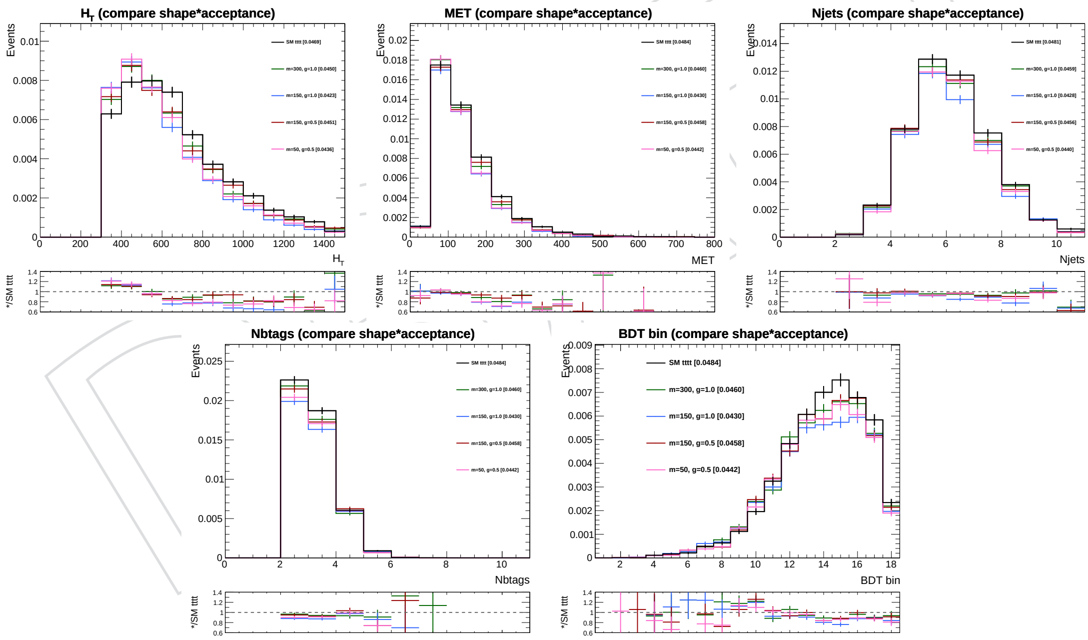


Figure 70: Distributions comparing H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, N_{jets}, $N_{\mathrm{b\ jets}}$, and the BDT signal region yields for the LO $\mathrm{t\bar{t}t\bar{t}}$ samples and various vector mediator mass points. Both shape and acceptance are relevant here.

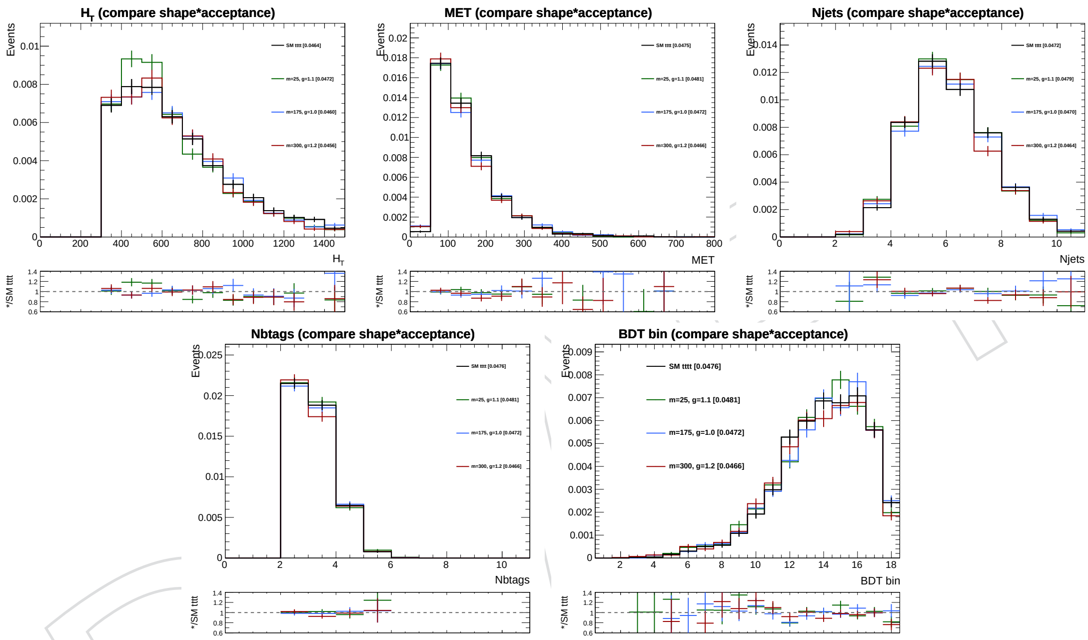


Figure 71: Distributions comparing H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$, N_{jets}, $N_{\mathrm{b\ jets}}$, and the BDT signal region yields for the LO $\mathrm{t\bar{t}t\bar{t}}$ samples and various scalar mediator mass points. Both shape and acceptance are relevant here.

The BDT observed upper limit on $t\bar{t}t\bar{t}$ production, including the extra 10% uncertainty previously motivated, is 23 fb^{-1}. Taking the ratio with the SM cross-section gives approximately 1.9. This is represented as the horizontal dashed line in Figure 72, which includes curves for cross-sections (normalized to SM) for various mediator masses as a function of coupling strengths. Intersections between the cross-section curves and the horizontal dashed line are marked with vertical dashed lines. To obtain smoother values, a quadratic interpolation between generated points is used. These intersections form the exclusion boundary as shown in Figure 73.

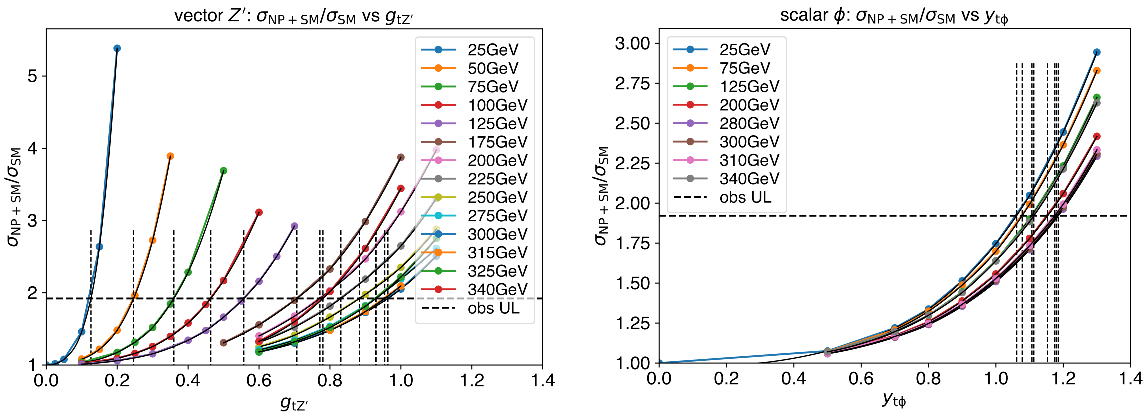


Figure 72: Cross-section (normalized to SM) as a function of coupling values with varying mediator masses for vector (left) and scalar (right) mediators. The observed analysis upper limit is shown as a horizontal dashed line.

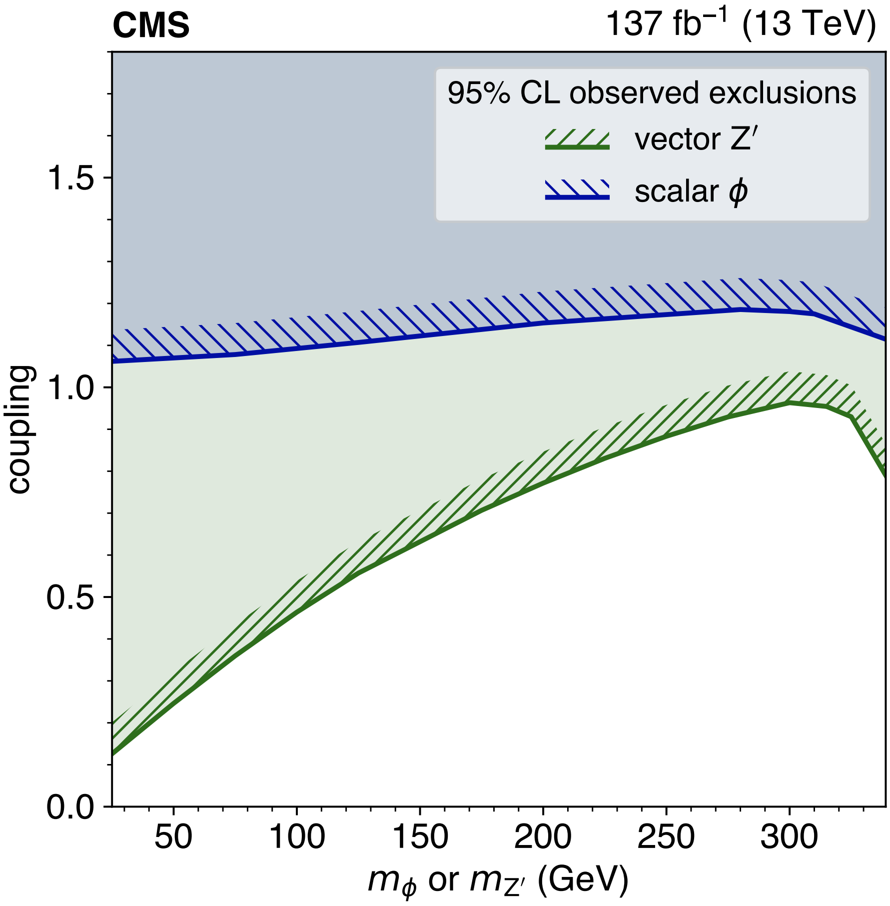


Figure 73: Observed 95% CL exclusions in the plane of coupling-mass for off-shell vector and scalar mediators.

12.4 Oblique Higgs parameter

12.4.1 Cross-section calculation

In a universal effective field theory framework, the Higgs oblique parameter $\hat{H}$, defined as the Wilson coefficient of the dimension-6 operator modifying the Higgs boson propagator, can result in deviations of the SM $t\bar{t}t\bar{t}$ cross-section, as shown in Ref. [32]. These (off-shell) deviations can be constrained to a level which is competitive with constraints from on-shell processes.

The two main characteristic effects of this oblique parameter are an additional contact term in the SM Higgs boson propagator

$$P_h(p^2) \approx \frac{i}{p^2 - m_h^2} - \frac{i\hat{H}}{m_h^2},$$

and a rescaling of the fermionic higgs couplings

$$\kappa_f = 1 - \hat{H}.$$

Using the latest combined fits of ATLAS for the (on-shell) fermionic couplings, with 80 fb^{-1} of 13TeV data, the authors of Ref. [32] find a constraint on the oblique parameter of $\hat{H} < 0.16$ at 95% CL.

The authors also calculate that the cross-section of (off-shell) $t\bar{t}t\bar{t}$ is subject to a fractional modification (with respect to the SM cross-section) at 14 TeV, given by,

$$\frac{\sigma_{\hat{H}+\text{SM}}}{\sigma_{\text{SM}}} = 1 + 0.03\left(\frac{\hat{H}}{0.04}\right) + 0.15\left(\frac{\hat{H}}{0.04}\right)^2.$$

For an oblique parameter value of 0.1, the formula predicts a doubling of the SM cross-section of $t\bar{t}t\bar{t}$ with MG 2.6.1 and 2.6.5

The SM model within MadGraph was modified to take into account the extra term in the propagator, as well as the rescaling of the top-yukawa coupling, by changing only the numerator of the scalar propagator affecting the Higgs boson:

$$\frac{i}{p^2 - m^2 + im\Gamma} \to \frac{(1-\hat{H})^2\left(1 - \frac{\hat{H}}{m^2}(p^2 - m^2)\right)}{p^2 - m^2 + im\Gamma}.$$

Note that the complex i in the numerator of the default propagator is removed due to an internal inconsistency found in the latest versions of MadGraph, as verified by the authors. Explicitly, we took the SM model UFO file and modified the numerator of `S` in file `propagators.py` to be

```
"(1-hhat)*(1-hhat)*(1-(hhat/(Mass(id)*Mass(id)))" +
"*(P('mu', id) * P('mu', id) - Mass(id) * Mass(id)))"
```

instead of `"i"`, and added the line

```
propagator = Prop.S,
```

to the SM Higgs particle definition in file `particles.py`. Finally, for convenience of scanning the parameter $\hat{H}$, we added

```
hhat = Parameter(
    name = 'hhat',
    nature = 'external',
    type = 'real',
    value = 0.,
    texname = '\\text{hhat}',
    lhablock = 'PROP',
    lhacode = [ 1 ],
)
```

to file `parameters.py`. The final model can be found at `https://github.com/aminnj/FTInterpretations/tree/master/models/Oblique_UFO`.

12.4.2 Private generation details

We privately generated five different values of $\hat{H}$ with the above modifications to the SM model in MadGraph 2.6.5 with the `nn23lo1` PDF and default (dynamic) scale choices at leading order. Each parameter point consists of three datasets of 50k events each with LHE→MINIAODSIM configurations matching the `RunIISummer16MiniAODv3`, `RunIIFall17MiniAODv2`, and `RunIIAutumn18MiniAOD` campaigns. To increase the baseline selection efficiency, we used the dilepton Pythia filter described in Section 12.2, which requires events to have at least 2 generator leptons (of any flavor) with a filter efficiency of 0.3944. The cross-sections (times filter efficincies) for these parameter points are tabulated below.

$\hat{H}$	$\sigma \times \epsilon_{\text{filter}}$ (fb)
0.0	4.721
0.04	5.411
0.08	6.752
0.12	8.456
0.16	10.235

12.4.3 Comparison of kinematics/acceptance

Figure 74 shows analysis-level quantities for five different values of $\hat{H}$, one of which is analogous to SM $\mathrm{t\bar{t}t\bar{t}}$ ($\hat{H} = 0$). In the last BDT bin, the acceptance for the two values of non-zero $\hat{H}$ are within approximately 10% of the SM. In general, acceptance for those non-zero points is 10-20% higher than the SM in the higher BDT bins, with visible trends in jet/lepton momentum, and H_{T}. The nominal analysis upper limit cannot be directly used to place a constraint on $\hat{H}$ as acceptance increases for higher values of $\hat{H}$. However, the next section outlines the method of exclusion while directly using the nominal analysis result (assuming SM acceptance/kinematics). The subsequent section then takes into account acceptance/kinematic differences by showing the result with dedicated samples for each $\hat{H}$ point.

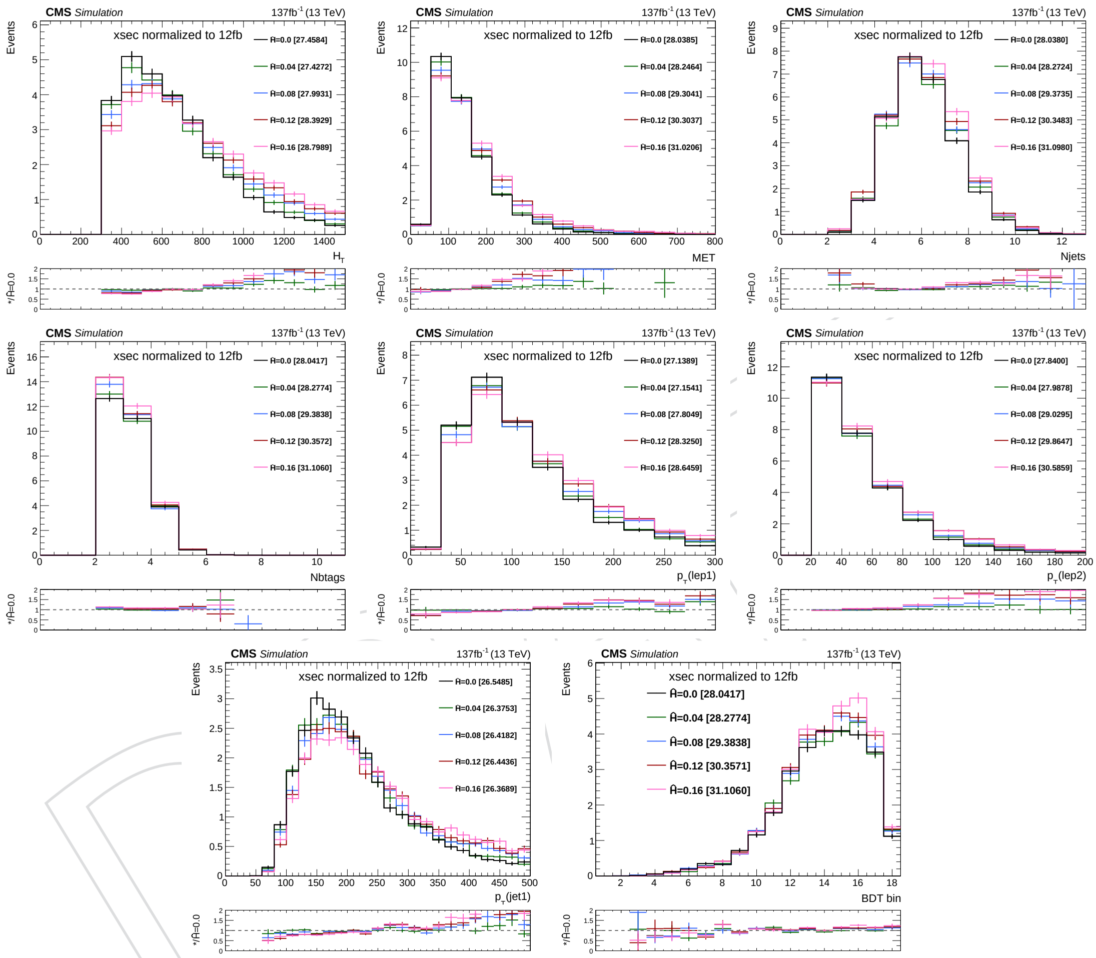


Figure 74: Distributions comparing H_T, E_T^{miss}, N_{jets}, $N_{b\,jets}$, p_T of the leading lepton, p_T of the subleading lepton, p_T of the leading jet, and the BDT signal region yields for three values of the oblique parameter $\hat{H}$ Both shapes and acceptance are relevant here, as cross-section has been normalized to $12\,\mathrm{fb}^{-1}$.

12.4.4 Exclusion, assuming SM kinematics

Through correspondence with the authors, we recalculated the 14TeV formula and find cross-sections to match closely with the predicted formula above, for values $\hat{H} < 0.04$, as the reference expands only to second order due to their tighter region of interest (HL-LHC projections). Figure 75 shows a recalculation of the 14TeV formula instead at 13TeV, as well as a cubic fit to the values in order to extract an upper limit on the oblique parameter using the nominal analysis BDT upper limit. We find $\hat{H} < 0.13$ at 95% CL.

As is the case for the top Yukawa interpretation from Section 12.2, the $t\bar{t}H$ background is also subject to modifications due to non-zero $\hat{H}$. The $t\bar{t}H$ background is affected by a cross-section scaling due to the reduction of the top yukawa constant ($y_t \to y_t - \hat{H}$). Thus, we can utilize the observed upper limits calculated for the $y_t < 1$ section of the curve in Figure 66. An increase in $\hat{H}$ corresponds to slight increase of the observed upper limit, which translates into a weaking of the $\hat{H}$ constraint by 0.005 (4% relative).

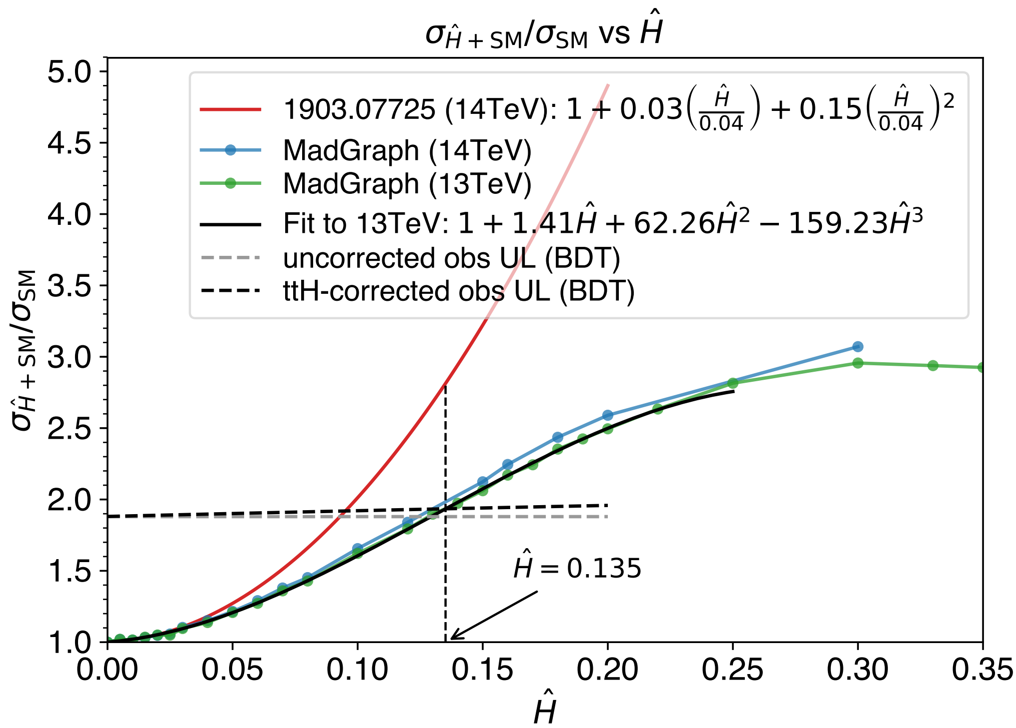


Figure 75: Cross-section (normalized to SM) as a function of oblique parameter $\hat{H}$. The red curve represents equation 5.4 from Ref. [32] at 14TeV. The blue and green curves are private calculations from MadGraph at 14TeV and 13TeV, respectively. The solid black curve is a cubic fit to the 13TeV private calculation, which is used to find the intersection with the (nearly) horizontal black dotted line (nominal analysis upper limit, corrected to account for deviations from $t\bar{t}H$, as described in the text. For the sake of comparison, the (completely) horizontal gray dotted line does not take this effect into account The intersection of the black dotted lines provides the upper limit on $\hat{H}$.

12.4.5 Exclusion using dedicated samples

Repeating the previous subsection with privately-generated dedicated samples yields Figure 76 and excludes $\hat{H} > 0.12$. Explicitly, we substitute the privately-generated $\hat{H}$-modified $t\bar{t}t\bar{t}$ sample for the $\mathrm{t\bar{t}t\bar{t}}$ signal, scale down the $\mathrm{t\bar{t}H}$ background normalization by $y_{\mathrm{t}}^2 = (1-\hat{H})^2$ (≈ 0.7 at $\hat{H} = 0.16$), and run the nominal limit-setting procedure.

Note that the upper limit becomes more stringent for higher values of $\hat{H}$, consistent with the increase in acceptance shown in Figure 74.

We have produced official gridpacks and asked the TOP MC contacts to generate official CMS samples using the configurations for 2016, 2017 and 2018 respectively. These samples are expected to be identical to the ones we generated privately, since the same process will be used.

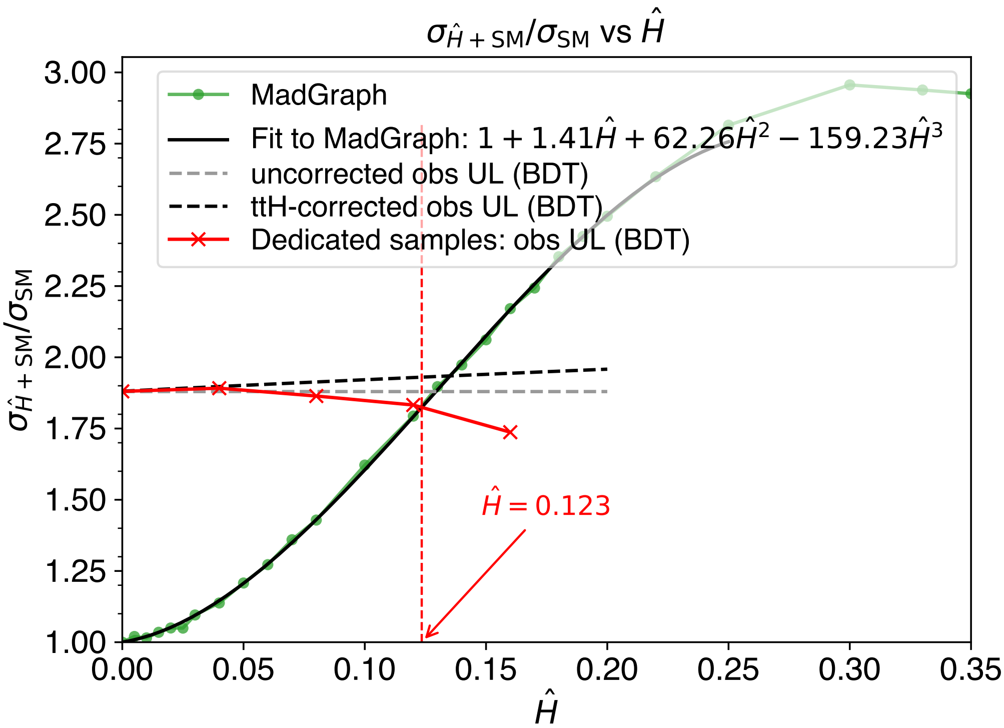


Figure 76: See caption for Figure 75. An additional red curve was added connecting the five upper limits calculated with dedicated samples. The intersection of a linear interpolation between these points (in red) with the black curve defines the 95% CL upper limit on $\hat{H}$.

12.5 Dark matter

12.5.1 Introduction

Upper limits on additional $t\bar{t}t\bar{t}$ production can be translated into exclusions/upper limits on simplified dark matter models which have a scalar or pseudoscalar mediator decaying into a pair of dark matter or standard model particles. The production of the mediator (decaying into invisible dark matter) in association with a pair of top quarks was performed by CMS with the 2016 dataset in Ref. [33]. Reference [34] found that production of the mediator in association with a single top quark contributes meaningfully to the total cross-section. Thus, a search that also included these single top associated production modes was carried out with the same dataset in Ref. [28], which we aim to complement in this analysis with an orthogonal final state. A brief overview is included below, and further theoretical details and motivation can be found in the Analysis Note for EXO-18-010 ([35]).

In the framework of a simplified dark matter model, where the scalar (ϕ) or pseudoscalar (a) mediator couples dark matter and SM particles, the relevant terms of the interaction lagrangian are of the form

$$\mathcal{L}_{\phi} = g_{\chi}\phi\bar{\chi}\chi + \frac{g_{\mathrm{q}}\phi}{\sqrt{2}}\sum_{\mathrm{f}} y_{\mathrm{f}}\bar{\mathrm{f}}\mathrm{f} \qquad \mathcal{L}_{a} = ig_{\chi}a\bar{\chi}\gamma^{5}\chi + \frac{ig_{\mathrm{q}}\mathrm{a}}{\sqrt{2}}\sum_{\mathrm{f}} y_{\mathrm{f}}\bar{\mathrm{f}}\gamma^{5}\mathrm{f}$$

where y_{f} are the fermionic yukawa couplings. The coupling constants g_{χ} and g_{q} give the relative strengths of the mediator coupling to dark matter and SM particles, and are used interchangeably with g_{DM} and g_{SM}, respectively. The model has four free parameters (g_{χ}, g_{q}, m_{χ}, and m_{a}) which is reduced to two with the assumption of $g_{\chi} = g_{\mathrm{q}} = 1$.

The relevant production diagrams are shown in Figure 77. When the mediator mass is above $2m_{\mathrm{top}} \approx 350\mathrm{GeV}$, on-shell decay to $t\bar{t}$ becomes kinematically accessible, resulting in 3 or 4 top quark final states, so we instead consider a version of the diagrams with a decay of the mediator into $t\bar{t}$ rather than a pair of (invisible) dark matter particles. Consequently, the production diagrams and kinematics are identical to that of the 2HDM interpretation which will allow us to use existing simulation samples for this interpretation, and we only need to calculate branching ratios into $t\bar{t}$.

In this way, our final state is complementary to the $t\bar{t}$ +MET final state used in Ref. [28] when the DM mass becomes large.

Using MadGraph 2.6.5, and the models from the official gridpacks used for the EXO-18-010 result we first verified consistency with the corresponding AN cross-sections before modifying the $\phi/\mathrm{a} \to \chi\bar{\chi}$ decay to be $\phi/\mathrm{a} \to t\bar{t}$. We use MadGraph to calculate cross-section (times branching ratio) for the $t\bar{t}$ associated process and two single-top processes for mediator masses matching the available 2HDM simulation samples at varying dark-matter masses. We use the `NNPDF30_lo_as_0130` PDF.

While we use MadGraph-calculated values in the end to account for all effects, we independently verified that relative branching ratios for scalar and pseudoscalar, respectively, are given by

$$\Gamma_{\mathrm{ratio}} \equiv \frac{\Gamma(\phi \to t\bar{t})}{\Gamma(\phi \to \chi\bar{\chi})} = \frac{3g_{\mathrm{SM}}^{2}y_{\mathrm{t}}^{2}}{2g_{\mathrm{DM}}^{2}}\left(\frac{M^{2}-4m_{\mathrm{t}}^{2}}{M^{2}-4m_{\chi}^{2}}\right)^{3/2} \tag{7}$$

$$\Gamma_{\mathrm{ratio}} \equiv \frac{\Gamma(\phi \to t\bar{t})}{\Gamma(\phi \to \chi\bar{\chi})} = \frac{3g_{\mathrm{SM}}^{2}y_{\mathrm{t}}^{2}}{2g_{\mathrm{DM}}^{2}}\left(\frac{M^{2}-4m_{\mathrm{t}}^{2}}{M^{2}-4m_{\chi}^{2}}\right)^{1/2} \tag{8}$$

where y_t is the top-quark yukawa coupling, M is the mediator mass, and ϕ represents both the scalar and pseudoscalar. Note that the only difference is in the exponent of the mass term. In practice, these formulae would be used after calculating the production cross-section with BR($\phi \rightarrow t\bar{t}$)=1 with MadGraph for each mediator mass value (e.g., $M = 450$GeV with $m_\chi >$ 450GeV so that the DM decay is kinematically supressed.) Then that production cross-section is multiplied by $\mathrm{BR}(t\bar{t}) = \frac{\Gamma(t\bar{t})}{\Gamma(t\bar{t})+\Gamma(\mathrm{DM})} = \left(1 - \Gamma_{\mathrm{ratio}}^{-1}\right)^{-1}$ in order to get the $\sigma{\cdot}$BR at an arbitrary m_χ. The formulae assume that there are only two accessible decay modes of the mediator.

12.5.2 Exclusions

As noted previously, the processes are identical to those of the 2HDM interpretation with a different labeling of the mediator, so we can calculate a single cross-section upper limit for a given mediator mass using pre-existing samples. For this interpretation the SM $t\bar{t}t\bar{t}$ process is treated as a background and assigned its SM cross section and uncertainty $12.0^{+2.2}_{-2.5}$ fb [6]. The single value is used to exclude along the vertical axis (m_{DM}) since kinematics remain identical and only the branching ratio to $t\bar{t}$ changes. This can be seen in Figure 78, which overlays exclusion contours on cross-sections calculated at discrete points. The diagonal line corresponds to the kinematic boundary between on- and off-shell mediator decays into dark matter, $m_{\mathrm{mediator}} = 2m_{\mathrm{DM}}$, above which $t\bar{t}$ is the dominant decay. For couplings set to unity, the region above the diagonal is excluded for both mediator types between approximately 350 and 500 GeV. When decreasing the relative branching ratio to DM ($g_{\mathrm{DM}} = 0.5$), DM masses down to 1 GeV are excluded up to a pseudoscalar mediator mass of 470GeV. Using Eqs. 7-8, Figure 79 shows that the gain is larger in the pseudoscalar case when compared to the scalar mediator due to a steeper decrease in $t\bar{t}$ branching ratio for the scalar mediator.

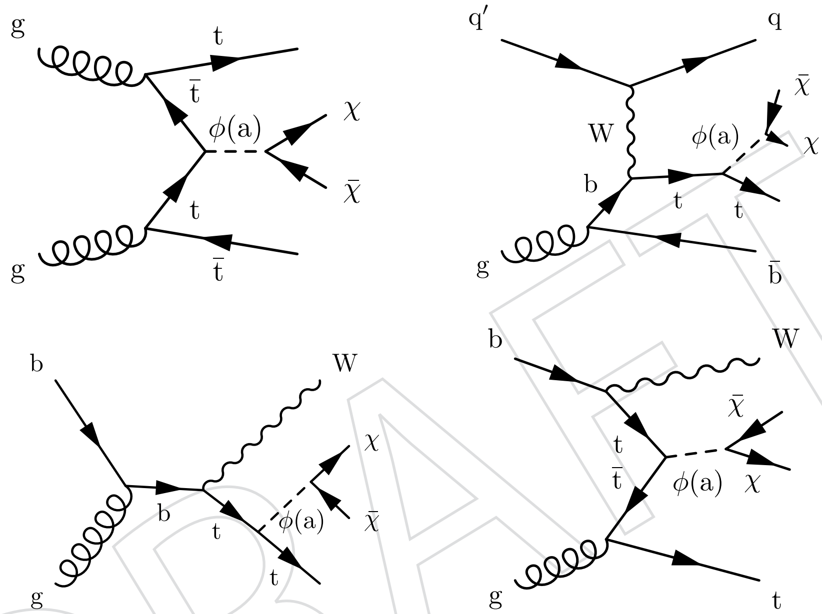


Figure 77: Diagrams for scalar (pseudoscalar) mediator production in association with a $t\bar{t}$ pair (top left), associated t-channel single top (top right), associated tW (bottom row). In these diagrams the mediator decays into a pair of invisible particles, while for the $t\bar{t}t\bar{t}$ analysis we are probing the equivalent diagrams where the mediator decays back to $t\bar{t}$. As the s-channel single top production cross-section is relatively negligible (by at least an order of magnitude), it is not included in this result.

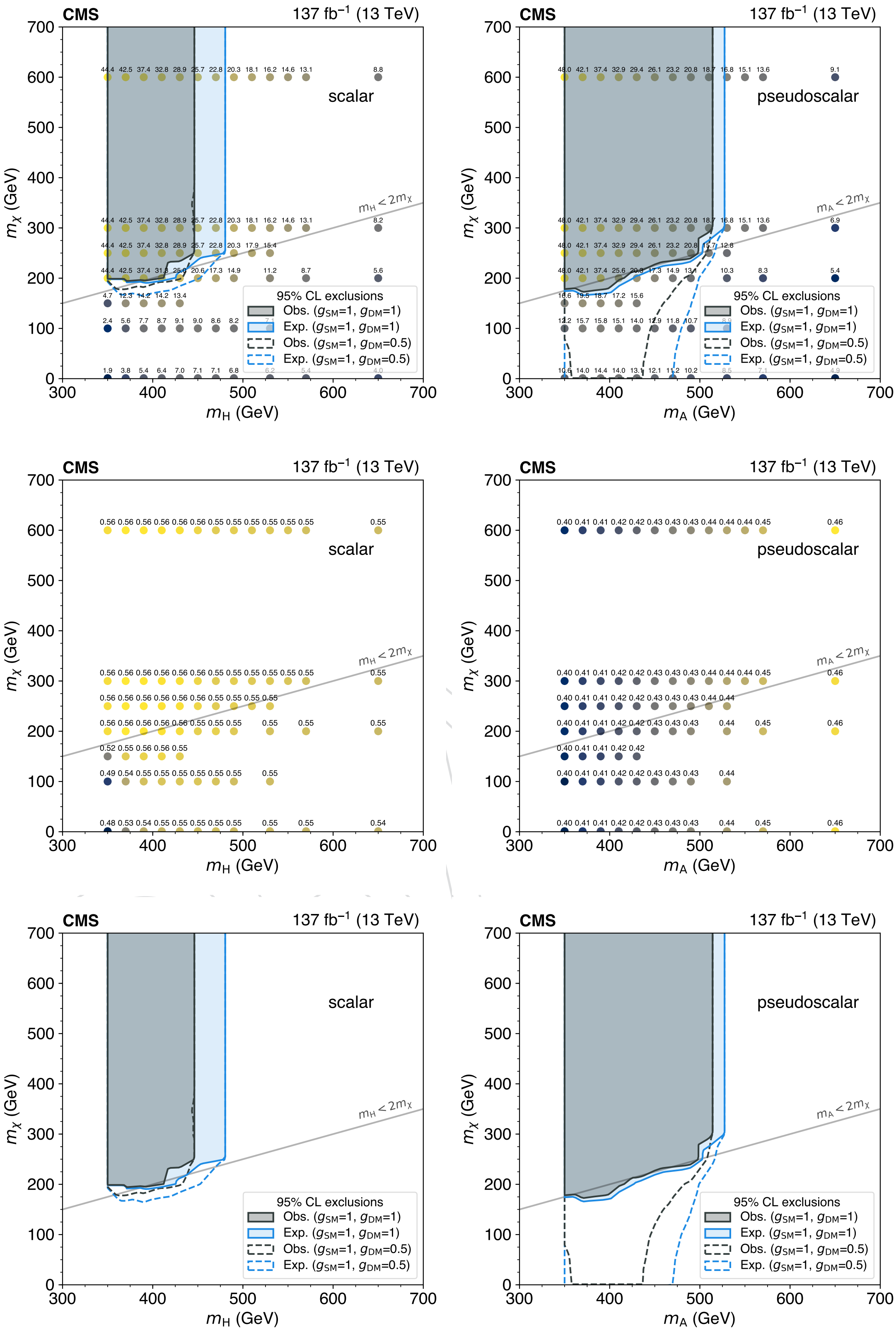


Figure 78: Expected and observed 95% CL exclusions in the plane of m_{DM}-$m_{mediator}$ for scalar (left) and pseudoscalar (right) mediators. The top row shows two exclusions: the nominal assumption ($g_{DM} = g_{SM} = 1$), and an alternate assumption ($g_{DM} = 0.5$, $g_{SM} = 1$). Calculated cross-section times branching ratio values for the first assumption are shown as markers with accompanying values in units of fb. The middle row shows the fraction of the total production cross-section attributed to single top processes. The the bottom row is identical to the top row but without cross-section value/markers for presentational purposes.

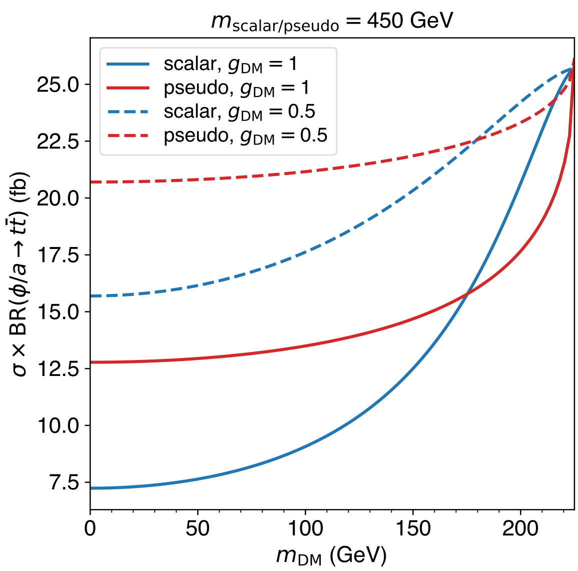


Figure 79: Cross-section times branching ratio to $t\bar{t}$ for scalar and pseudoscalar mediators with mass 450 GeV, shown at DM coupling values of $g_{\text{DM}} = 1$ and $g_{\text{DM}} = 0.5$, calculated with Eqs. 7-8.

References

[1] RA5 group, "Search for SUSY in same-sign dilepton events at 13 TeV", *CMS Physics Analysis Note* **AN-2016/228** (2016).

[2] CMS Collaboration, "Search for physics beyond the standard model in events with two leptons of same sign, missing transverse momentum, and jets in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", arXiv:1704.07323.

[3] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[4] tttt same-sign and multilepton group, "Search for SM tttt in the same-sign dilepton and multi-lepton final states at sqrt(s) = 13 TeV", *CMS Physics Analysis Note* **AN-2017/115** (2016).

[5] CMS Collaboration, "Search for standard model production of four top quarks with same-sign and multilepton final states in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$", *Eur. Phys. J.* **C78** (2018), no. 2, 140, doi:10.1140/epjc/s10052-018-5607-5, arXiv:1710.10614.

[6] R. Frederix, D. Pagani, and M. Zaro, "Large NLO corrections in $t\bar{t}W^{\pm}$ and $t\bar{t}t\bar{t}$ hadroproduction from supposedly subleading EW contributions", *JHEP* **02** (2018) 031, doi:10.1007/JHEP02(2018)031, arXiv:1711.02116.

[7] https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFilters.

[8] M. Franco Sevilla and A. Ovcharova, "Isr reweighting recommendations for moriond 2017", December, 2016. https://indico.cern.ch/event/592621/contributions/2398559/attachments/1383909/2105089/16-12-05_ana_manuelf_isr.pdf.

[9] CMS SUSY fake-leptons working group, "Studies of methods to estimate the non-prompt lepton background to searches for new physics", *CMS Physics Analysis Note* **AN-2014/261** (2015).

[10] P. Pigard, "Multivariate electron id in 8x", June, 2016. https://indico.cern.ch/event/482674/contributions/2206032/attachments/1292177/1931287/20160621_EGM_cms_week_v5.pdf.

[11] G. Zevi Della Porta, "Lepton id for the full 2016 dataset", November, 2016. https://indico.cern.ch/event/590228/contributions/2380031/attachments/1375541/2088587/EGMSUS_newIDs_17Nov16.pdf.

[12] V. H. et al., "Electron mva id for susy", April, 2017. https://indico.cern.ch/event/719317/contributions/2963816/attachments/1630110/2598062/MVAidSUSY_10Apr18_SUSYMeeting.pdf.

[13] S. L. S. F. team, "Lepton scale factors", July, 2016. https://twiki.cern.ch/twiki/bin/view/CMS/SUSLeptonSF.

[14] Baffioni, S. and others, "Electron Charge Identification using 8 TeV data", *CMS Physics Analysis Note* **AN-2014/164** (2015).

[15] G. Abbiendi et al., "Baseline muon selections", October, 2014. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId#Tight_Muon.

[16] G. Petrucciani and C. Botta, "Two step prompt muon identification", Jaunary, 2015. https://indico.cern.ch/event/368007/contribution/2/material/slides/0.pdf.

[17] RA5 group, "Search for SUSY in same-sign dilepton events at 13 TeV", *CMS Physics Analysis Note* **AN-2015/031** (2016).

[18] D. Ferencek et al., "b-tagging offline guide", November, 2014. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideBTagging.

[19] CMS Collaboration, "Performance of the missing transverse energy reconstruction by the CMS experiment in $\sqrt{s} = 8$ TeV pp data", (2014). arXiv:1411.0511. Submitted to *JINST*.

[20] https://twiki.cern.ch/twiki/bin/view/CMS/MissingETUncertaintyPrescription#Instructions_for_9_4_X_X_9_for_2.

[21] R. Barlow and C. Beeston, "Fitting using finite monte carlo samples", *Computer Physics Communications* **77** (1993), no. 2, 219 – 228, doi:https://doi.org/10.1016/0010-4655(93)90005-W.

[22] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit.

[23] K. J. F. Gaemers and F. Hoogeveen, "Higgs Production and Decay Into Heavy Flavors With the Gluon Fusion Mechanism", *Phys. Lett.* **B146** (1984) 347–349, doi:10.1016/0370-2693(84)91711-8.

[24] G. C. Branco et al., "Theory and phenomenology of two-Higgs-doublet models", *Phys. Rept.* **516** (2012) 1–102, doi:10.1016/j.physrep.2012.02.002, arXiv:1106.0034.

[25] D. Dicus, A. Stange, and S. Willenbrock, "Higgs decay to top quarks at hadron colliders", *Phys. Lett.* **B333** (1994) 126–131, doi:10.1016/0370-2693(94)91017-0, arXiv:hep-ph/9404359.

[26] N. Craig et al., "The Hunt for the Rest of the Higgs Bosons", *JHEP* **06** (2015) 137, doi:10.1007/JHEP06(2015)137, arXiv:1504.04630.

[27] N. Craig et al., "Heavy Higgs Bosons at Low $\tan\beta$: from the LHC to 100 TeV", arXiv:1605.08744.

[28] CMS Collaboration, "Search for dark matter produced in association with a single top quark or a top quark pair in proton-proton collisions at $\sqrt{s}$=13TeV", *Journal of High Energy Physics* **2019** (Mar, 2019) 141, doi:10.1007/JHEP03(2019)141.

[29] Q.-H. Cao, S.-L. Chen, and Y. Liu, "Probing higgs width and top quark yukawa coupling from $t\bar{t}h$ and $t\bar{t}t\bar{t}$ productions", *Phys. Rev. D* **95** (Mar, 2017) 053004, doi:10.1103/PhysRevD.95.053004.

[30] G. Bevilacqua and M. Worek, "Constraining BSM Physics at the LHC: Four top final states with NLO accuracy in perturbative QCD", *JHEP* **07** (2012) 111, doi:10.1007/JHEP07(2012)111, arXiv:1206.3064.

[31] E. Alvarez et al., "Four tops for lhc", *Nuclear Physics B* **915** (2017) 19 – 43, doi:https://doi.org/10.1016/j.nuclphysb.2016.11.024.

[32] C. Englert, G. F. Giudice, A. Greljo, and M. Mccullough, "The $\hat{H}$-Parameter: An Oblique Higgs View", arXiv:1903.07725.

[33] CMS Collaboration, "Search for Dark Matter Particles Produced in Association with a Top Quark Pair at $\sqrt{s}$=13TeV", *Phys. Rev. Lett.* **122** (Jan, 2019) 011803, doi:10.1103/PhysRevLett.122.011803.

[34] D. Pinna, A. Zucchetta, M. R. Buckley, and F. Canelli, "Single top quarks and dark matter", *Phys. Rev. D* **96** (Aug, 2017) 035031, doi:10.1103/PhysRevD.96.035031.

[35] e. a. T. Bose, "Search for dark matter in single top + MET final states", *CMS Physics Analysis Note* **AN-2017/327** (2017).

[36] J. Steggeman, "Tauid for 13 tev run: recommendation from the tau pog (version 43)", March, 2018. https://twiki.cern.ch/twiki/bin/view/CMS/TauIDRecommendation13TeV?rev=43#New_Tau_Isolation_Discriminators.

A Statistical checks

A.1 Impacts

The leading 30 nuisance impacts for two sets of impacts, expected and observed for cut-based and BDT analyses, are shown in Figure 117 (expected cut-based analysis), Figure 118 (observed cut-based analysis), Figure 119 (expected BDT analysis), and Figure 120 (observed BDT analysis). The leading expected nuisance in both cases corresponds to the $\sigma(\mathrm{ttbb})/\sigma(\mathrm{ttjj})$ scaling. Note that the "prop binSS" nuisances for MC statistics include (and are dominated by) tight-loose sideband statistics.

The obseved pulls show the most constrained/pulled nuisances correspond to normalization parameters for ttW and ttZ, as we would expect from the control regions. "TTWSF" is moved by approximately 1σ (0.8σ) with respect to the input nuisance sizes for the cut-based (BDT) analysis. "TTZSF" is moved up by approximately 0.6σ (0.7σ) for the cut-based (BDT) analysis.

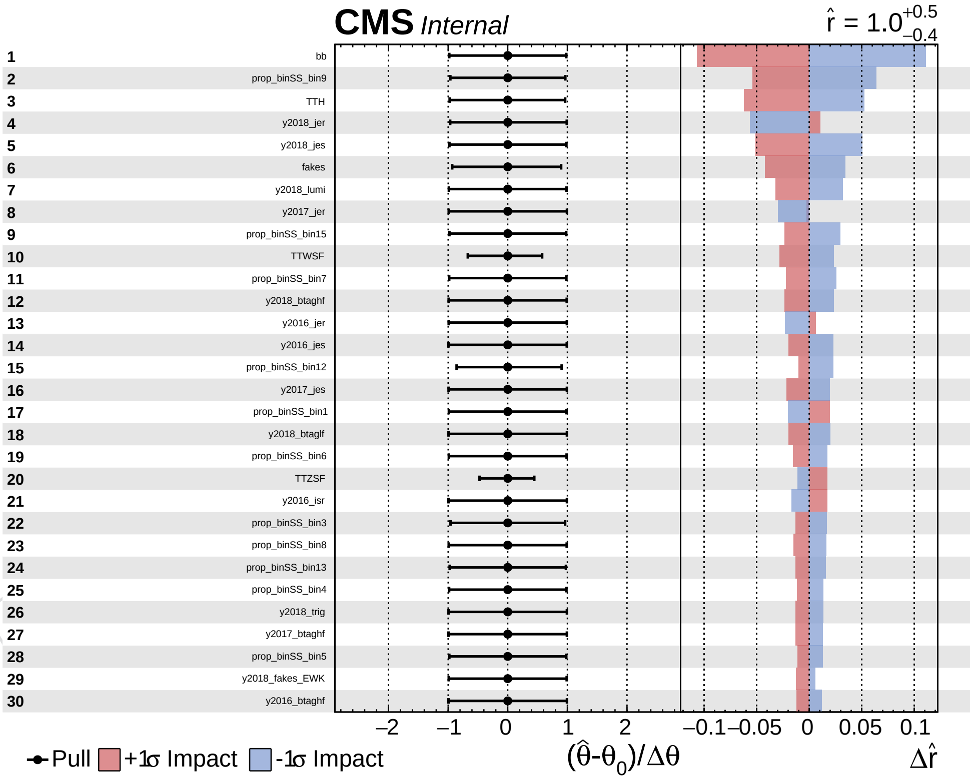


Figure 80: Expected nuisance impacts for the cut-based analysis.

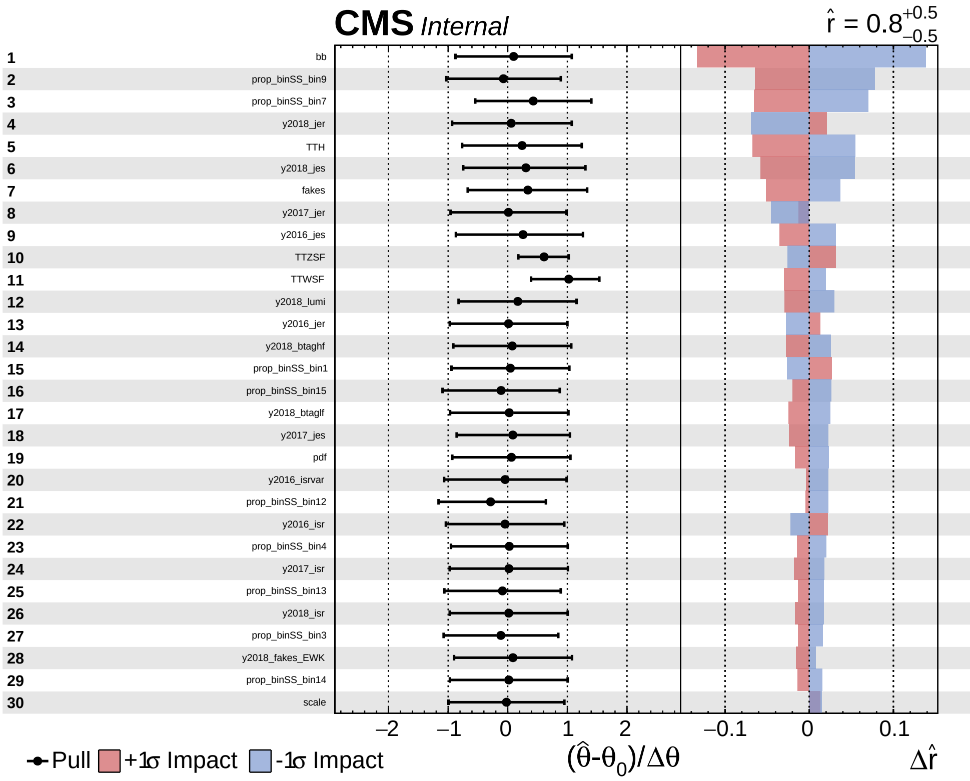


Figure 81: Observed nuisance impacts for the cut-based analysis.

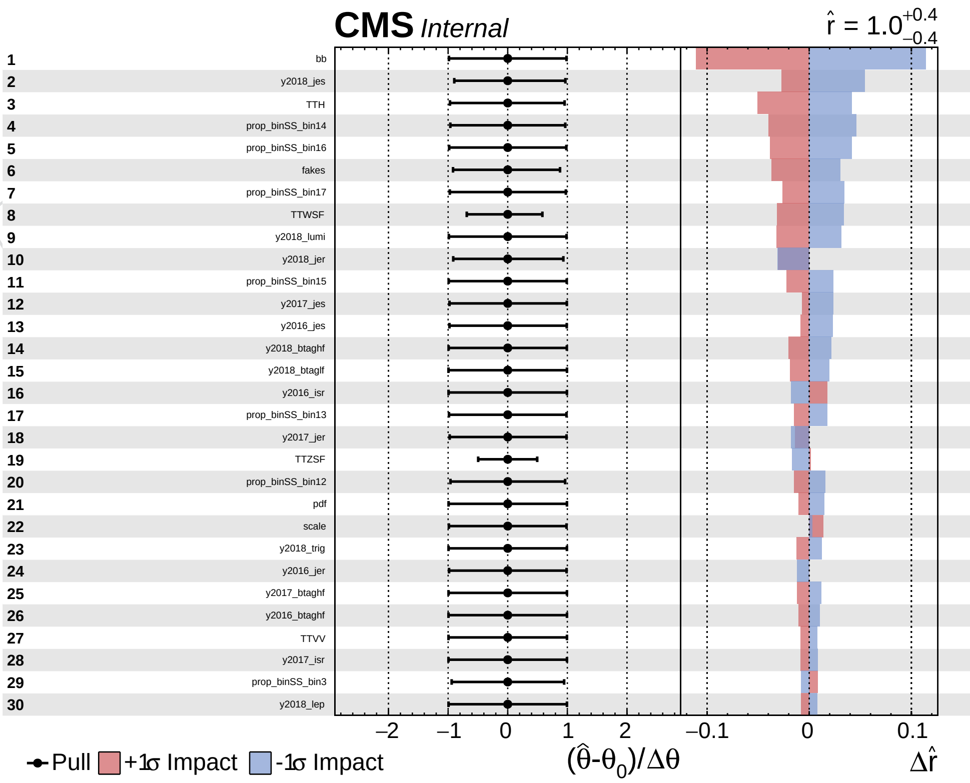


Figure 82: Expected nuisance impacts for the BDT-based analysis.

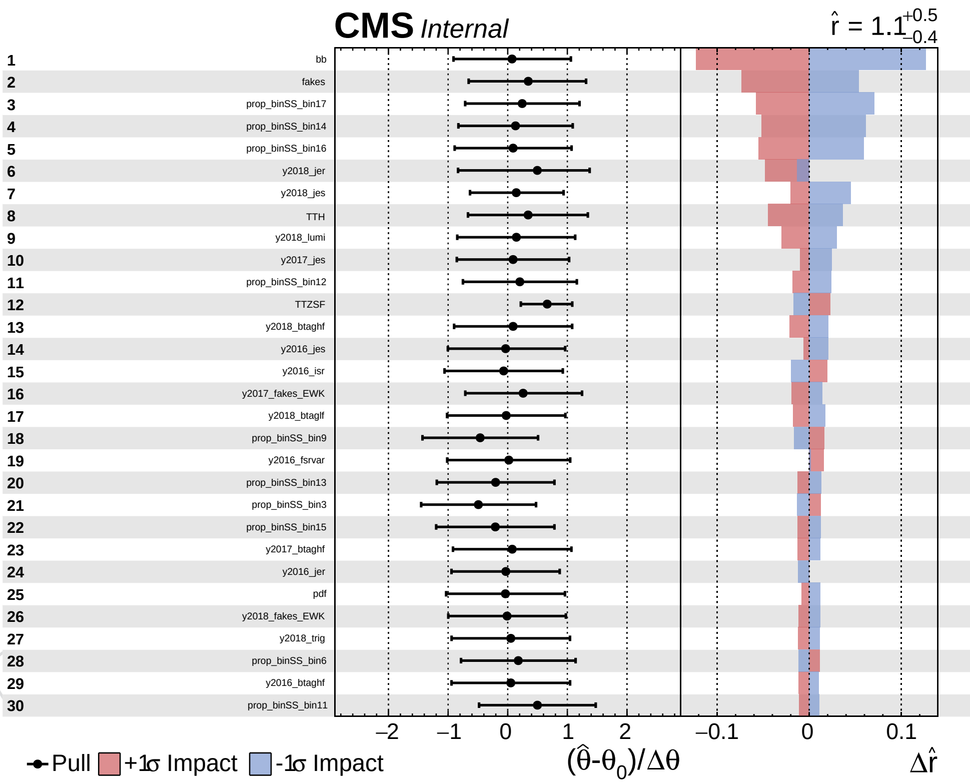


Figure 83: Observed nuisance impacts for the BDT-based analysis.

A.2 Nuisance forms

Nuisance functional forms and widths are tabulated in Table 30.

Table 30: Rows preceeded by (s) apply to signal as well. For "shape" uncertainties, width is not applicable as up and down variation envelopes are taken as the systematic uncertainties.

name	function	width
lumi	lnN	1.021-1.05
(s) jes	shape	-
(s) jer	shape	-
isr	shape	-
bb	shape	-
(s) lep	shape	-
(s) trig	shape	-
(s) btaghf	shape	-
(s) btaglf	shape	-
(s) isrvar	shape	-
(s) fsrvar	shape	-
(s) scale	shape	-
(s) pdf	shape	-
(s) alphas	shape	-
(s) pu	shape	-
TTWSF	lnN	1.4
TTZSF	lnN	1.4
TTH	lnN	1.25
TTVV	lnN	1.11
XG	lnN	1.11
rares	lnN	1.2
fakes	lnN	1.3
fakes_EWK	shape	-
flips	lnN	1.2

A.3 Nuisances

Two sets of nuisance pull values, expected and observed for cut-based and BDT analyses, are tabulated in Table 62 (expected cut-based analysis), Table 63 (observed cut-based analysis), Table 64 (expected BDT analysis), and Table 65 (observed BDT analysis).

The most constrained nuisances correspond to normalization parameters for ttW and ttZ (“TTWSF” and “TTZSF”) due to high statistics in control regions and in the bulk (BDT). Their input normalization uncertainty is 40%.

Table 31: Expected nuisance pulls for the cut-based analysis. The final column indicates the correlation between the nuisance and the signal strength

	b-only fit	$s+b$ fit	
name	$\Delta x/\sigma_{\rm in}, \sigma_{\rm out}/\sigma_{\rm in}$	$\Delta x/\sigma_{\rm in}, \sigma_{\rm out}/\sigma_{\rm in}$	$\rho(\theta,\mu)$
TTH	**+0.38, 0.99**	+0.00, 0.97	-0.12
TTVV	+0.05, 1.00	+0.00, 0.99	-0.02
TTWSF	**+0.09, 0.61**	**+0.00, 0.61**	-0.06
TTZSF	**-0.02, 0.45**	**+0.00, 0.45**	+0.02
XG	-0.01, 0.99	+0.00, 0.99	+0.00
alphas	+0.00, 0.99	-0.00, 0.99	+0.00
bb	**+0.74, 0.94**	+0.00, 0.98	-0.24
fakes	+0.24, 0.92	+0.00, 0.92	-0.08
pdf	+0.07, 0.99	-0.00, 0.99	-0.02
rares	-0.01, 0.99	+0.00, 0.99	+0.00
scale	-0.19, 0.96	-0.00, 0.99	+0.00
y2016_btaghf	+0.04, 0.99	+0.00, 0.99	-0.03
y2016_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2016_fakes_EWK	+0.04, 1.00	+0.00, 0.99	-0.01
y2016_flips	-0.02, 0.99	+0.00, 0.99	+0.00
y2016_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.02
y2016_isr	-0.10, 0.99	-0.00, 0.99	+0.04
y2016_isrvar	+0.00, 0.99	+0.00, 0.99	-0.02
y2016_jer	**-0.13, 1.12**	-0.00, 0.99	+0.03
y2016_jes	+0.07, 0.97	+0.00, 0.99	-0.05
y2016_lep	+0.01, 0.99	+0.00, 0.99	-0.01
y2016_lumi	+0.02, 0.99	+0.00, 0.99	-0.02
y2016_prefire	-0.00, 0.99	-0.00, 0.99	+0.00
y2016_pu	+0.01, 0.99	-0.00, 0.99	+0.00
y2016_trig	+0.01, 0.99	+0.00, 0.99	-0.02
y2017_btaghf	+0.04, 0.99	+0.00, 0.99	-0.03
y2017_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2017_fakes_EWK	+0.00, 0.99	-0.00, 0.98	-0.01
y2017_flips	-0.02, 0.99	+0.00, 0.99	+0.00
y2017_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2017_isr	+0.07, 0.99	+0.00, 0.99	-0.03
y2017_isrvar	+0.00, 0.99	-0.00, 0.99	-0.01
y2017_jer	**-0.12, 1.20**	-0.00, 0.99	+0.03
y2017_jes	+0.07, 0.99	+0.00, 0.99	-0.05
y2017_lep	+0.01, 0.99	+0.00, 0.99	-0.01
y2017_lumi	+0.01, 0.99	+0.00, 0.99	-0.02
y2017_prefire	-0.01, 0.99	-0.00, 0.99	+0.01
y2017_pu	-0.01, 1.00	+0.00, 0.99	+0.00
y2017_trig	+0.01, 0.99	+0.00, 0.99	-0.02
y2018_btaghf	+0.09, 0.99	+0.00, 0.99	-0.05
y2018_btaglf	+0.12, 0.99	-0.00, 0.99	-0.04
y2018_fakes_EWK	+0.08, 1.03	+0.00, 0.99	-0.02
y2018_flips	-0.03, 0.99	+0.00, 0.99	+0.01
y2018_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2018_isr	+0.08, 0.99	+0.00, 0.99	-0.03
y2018_isrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2018_jer	**-0.33, 1.08**	-0.00, 0.99	+0.08
y2018_jes	+0.18, 0.96	-0.00, 0.98	-0.11
y2018_lep	+0.02, 0.99	+0.00, 0.99	-0.02
y2018_lumi	+0.03, 0.99	+0.00, 0.99	-0.04
y2018_pu	+0.00, 1.00	-0.00, 0.99	+0.00
y2018_trig	+0.02, 0.99	+0.00, 0.99	-0.03

Table 32: Observed nuisance pulls for the cut-based analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{in}, \sigma_{out}/\sigma_{in}$	$s+b$ fit $\Delta x/\sigma_{in}, \sigma_{out}/\sigma_{in}$	$\rho(\theta, \mu)$
TTH	**+0.48, 1.02**	+0.24, 1.00	-0.12
TTVV	+0.06, 1.00	+0.03, 1.00	-0.02
TTWSF	**+1.03, 0.57**	**+0.99, 0.56**	-0.06
TTZSF	**+0.57, 0.41**	**+0.60, 0.42**	+0.05
XG	+0.04, 1.00	+0.04, 1.00	+0.00
alphas	+0.00, 0.99	+0.00, 0.99	+0.00
bb	**+0.62, 0.92**	+0.10, 0.97	-0.27
fakes	**+0.54, 1.02**	**+0.33, 1.00**	-0.09
pdf	+0.15, 0.98	+0.06, 1.00	-0.04
rares	+0.06, 1.00	+0.07, 1.00	+0.01
scale	-0.08, 0.91	-0.02, 0.96	-0.00
y2016_btaghf	+0.07, 0.99	+0.04, 0.99	-0.03
y2016_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2016_fakes_EWK	+0.08, 1.00	+0.05, 0.99	-0.01
y2016_flips	+0.00, 0.99	+0.01, 1.00	+0.01
y2016_fsrvar	+0.00, 0.99	-0.01, 1.04	+0.01
y2016_isr	-0.12, 0.99	-0.04, 0.99	+0.04
y2016_isrvar	+0.00, 0.99	-0.04, 1.06	-0.03
y2016_jer	-0.07, 1.04	+0.01, 0.99	+0.04
y2016_jes	**+0.38, 1.08**	**+0.25, 1.12**	-0.08
y2016_lep	+0.04, 0.99	+0.03, 0.99	-0.01
y2016_lumi	+0.05, 1.00	+0.04, 1.00	-0.02
y2016_prefire	-0.01, 0.99	-0.00, 0.99	+0.00
y2016_pu	+0.03, 0.99	+0.01, 0.99	-0.01
y2016_trig	+0.04, 0.99	+0.03, 0.99	-0.01
y2017_btaghf	+0.10, 1.00	+0.06, 0.99	-0.03
y2017_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2017_fakes_EWK	+0.23, 0.98	+0.22, 0.97	-0.01
y2017_flips	+0.00, 0.99	+0.01, 1.00	+0.01
y2017_fsrvar	+0.00, 0.99	-0.01, 0.99	+0.00
y2017_isr	+0.08, 0.99	+0.02, 0.99	-0.04
y2017_isrvar	+0.00, 0.99	-0.01, 0.99	-0.01
y2017_jer	**-0.07, 1.24**	+0.01, 0.96	+0.03
y2017_jes	+0.15, 0.93	+0.09, 0.93	-0.05
y2017_lep	+0.04, 0.99	+0.03, 0.99	-0.01
y2017_lumi	+0.05, 1.00	+0.04, 1.00	-0.01
y2017_prefire	-0.02, 0.99	-0.02, 0.99	+0.01
y2017_pu	+0.03, 1.00	+0.01, 1.00	-0.00
y2017_trig	+0.04, 0.99	+0.04, 0.99	-0.01
y2018_btaghf	+0.15, 0.99	+0.07, 0.99	-0.05
y2018_btaglf	+0.12, 0.99	+0.02, 0.99	-0.05
y2018_fakes_EWK	+0.15, 1.01	+0.09, 0.99	-0.02
y2018_flips	+0.01, 0.99	+0.02, 1.00	+0.01
y2018_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2018_isr	+0.08, 0.99	+0.02, 0.99	-0.04
y2018_isrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2018_jer	**-0.19, 1.17**	+0.06, 1.05	+0.09
y2018_jes	**+0.48, 1.02**	**+0.31, 1.08**	-0.13
y2018_lep	+0.06, 0.99	+0.05, 0.99	-0.02
y2018_lumi	+0.10, 0.99	+0.08, 0.99	-0.03
y2018_pu	+0.00, 1.00	-0.00, 0.99	+0.00
y2018_trig	+0.08, 0.99	+0.07, 0.99	-0.02

Table 33: Expected nuisance pulls for the BDT analysis. The final column indicates the correlation between the nuisance and the signal strength

name	*b*-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta, \mu)$
TTH	**+0.32, 0.98**	+0.00, 0.96	-0.11
TTVV	+0.06, 1.00	+0.00, 0.99	-0.02
TTWSF	**+0.08, 0.63**	**+0.00, 0.62**	-0.08
TTZSF	**-0.04, 0.51**	**+0.00, 0.48**	+0.03
XG	-0.01, 0.99	+0.00, 0.99	+0.00
alphas	-0.00, 0.99	-0.00, 0.99	+0.00
bb	**+0.91, 0.93**	+0.00, 0.99	-0.27
fakes	+0.28, 0.91	**+0.00, 0.90**	-0.08
pdf	+0.12, 0.97	-0.00, 0.99	-0.03
rares	-0.02, 0.99	+0.00, 0.99	+0.00
scale	-0.25, 0.97	+0.00, 0.99	+0.01
y2016_btaghf	+0.04, 0.99	+0.00, 0.99	-0.03
y2016_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2016_fakes_EWK	+0.03, 0.99	+0.00, 0.99	-0.01
y2016_flips	-0.02, 0.99	+0.00, 0.99	+0.00
y2016_fsrvar	-0.00, 0.99	+0.00, 0.99	+0.01
y2016_isr	-0.13, 0.99	+0.00, 0.99	+0.04
y2016_isrvar	-0.00, 0.99	-0.00, 0.99	-0.01
y2016_jer	-0.05, 1.03	+0.00, 0.99	+0.01
y2016_jes	+0.07, 0.90	+0.00, 0.99	-0.04
y2016_lep	+0.01, 0.99	+0.00, 0.99	-0.01
y2016_lumi	+0.02, 0.99	+0.00, 0.99	-0.02
y2016_prefire	-0.00, 0.99	-0.00, 0.99	+0.00
y2016_pu	+0.04, 0.99	-0.00, 0.99	-0.01
y2016_trig	+0.01, 0.99	+0.00, 0.99	-0.02
y2017_btaghf	+0.03, 0.99	+0.00, 0.99	-0.03
y2017_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2017_fakes_EWK	+0.03, 0.98	+0.00, 0.97	-0.02
y2017_flips	-0.02, 0.99	+0.00, 0.99	+0.01
y2017_fsrvar	-0.00, 0.99	-0.00, 0.99	+0.00
y2017_isr	+0.07, 0.99	-0.00, 0.99	-0.02
y2017_isrvar	-0.00, 0.99	+0.00, 0.99	-0.01
y2017_jer	**-0.01, 1.19**	+0.00, 0.99	+0.00
y2017_jes	+0.07, 0.90	+0.00, 0.99	-0.04
y2017_lep	+0.00, 0.99	+0.00, 0.99	-0.02
y2017_lumi	+0.01, 0.99	+0.00, 0.99	-0.02
y2017_prefire	-0.00, 0.99	-0.00, 0.99	+0.01
y2017_pu	+0.03, 1.00	-0.00, 0.99	-0.01
y2017_trig	+0.00, 0.99	+0.00, 0.99	-0.02
y2018_btaghf	+0.08, 0.99	+0.00, 0.99	-0.05
y2018_btaglf	+0.12, 0.99	-0.00, 0.99	-0.05
y2018_fakes_EWK	+0.06, 0.98	+0.00, 0.98	-0.01
y2018_flips	-0.03, 0.99	+0.00, 0.99	+0.01
y2018_fsrvar	-0.00, 0.99	+0.00, 0.99	+0.00
y2018_isr	+0.04, 0.99	-0.00, 0.99	-0.02
y2018_isrvar	-0.00, 0.99	+0.00, 0.99	+0.00
y2018_jer	**-0.01, 1.56**	+0.00, 0.95	+0.00
y2018_jes	**+0.15, 0.80**	-0.00, 0.95	-0.10
y2018_lep	+0.01, 0.99	+0.00, 0.99	-0.02
y2018_lumi	+0.03, 0.99	+0.00, 0.99	-0.04
y2018_pu	+0.02, 1.01	-0.00, 0.99	-0.01
y2018_trig	+0.02, 0.99	+0.00, 0.99	-0.03

Table 34: Observed nuisance pulls for the BDT analysis. The final column indicates the correlation between the nuisance and the signal strength

name	*b*-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	**+0.57, 1.02**	**+0.33, 1.00**	-0.09
TTVV	+0.08, 1.00	+0.04, 1.00	-0.02
TTWSF	**+0.64, 0.66**	**+0.75, 0.60**	-0.01
TTZSF	**+0.61, 0.42**	**+0.65, 0.43**	+0.04
XG	+0.02, 0.99	+0.04, 1.00	+0.00
alphas	+0.00, 0.99	+0.00, 1.00	+0.00
bb	**+0.87, 0.95**	+0.07, 0.98	-0.27
fakes	**+0.93, 0.99**	**+0.32, 0.99**	-0.14
pdf	+0.05, 0.99	-0.03, 1.01	-0.02
rares	+0.03, 0.99	+0.04, 0.99	+0.00
scale	-0.24, 0.94	-0.04, 0.94	+0.00
y2016_btaghf	+0.08, 0.99	+0.05, 0.99	-0.02
y2016_btaglf	+0.04, 1.00	+0.01, 0.99	-0.01
y2016_fakes_EWK	+0.16, 0.99	+0.08, 0.99	-0.02
y2016_flips	-0.01, 0.99	+0.00, 1.00	+0.00
y2016_fsrvar	+0.00, 0.99	+0.02, 1.07	+0.02
y2016_isr	-0.18, 0.99	-0.06, 0.99	+0.04
y2016_isrvar	+0.00, 0.99	+0.01, 1.00	-0.01
y2016_jer	**-0.06, 0.89**	**-0.03, 0.87**	+0.01
y2016_jes	+0.03, 1.00	**-0.04, 1.11**	-0.03
y2016_lep	+0.04, 0.99	+0.03, 0.99	-0.01
y2016_lumi	+0.05, 1.00	+0.04, 1.00	-0.02
y2016_prefire	-0.01, 0.99	-0.00, 0.99	+0.00
y2016_pu	-0.01, 0.99	-0.05, 1.00	-0.01
y2016_trig	+0.05, 0.99	+0.04, 0.99	-0.02
y2017_btaghf	+0.09, 0.99	+0.07, 0.99	-0.03
y2017_btaglf	+0.05, 0.99	+0.02, 0.99	-0.01
y2017_fakes_EWK	**+0.43, 1.04**	+0.25, 0.97	-0.04
y2017_flips	-0.01, 0.99	+0.01, 1.00	+0.00
y2017_fsrvar	+0.00, 0.99	+0.02, 1.00	+0.01
y2017_isr	+0.08, 0.99	+0.02, 0.99	-0.02
y2017_isrvar	+0.00, 0.99	+0.00, 1.00	-0.01
y2017_jer	**+0.09, 0.87**	**+0.07, 0.83**	-0.01
y2017_jes	**+0.14, 0.89**	+0.09, 0.99	-0.04
y2017_lep	+0.05, 0.99	+0.04, 0.99	-0.01
y2017_lumi	+0.05, 1.00	+0.05, 1.00	-0.02
y2017_prefire	-0.01, 0.99	-0.01, 0.99	+0.01
y2017_pu	+0.05, 1.00	+0.03, 0.99	-0.01
y2017_trig	+0.04, 0.99	+0.04, 0.99	-0.01
y2018_btaghf	+0.13, 0.99	+0.09, 0.99	-0.05
y2018_btaglf	+0.05, 0.99	-0.03, 0.99	-0.04
y2018_fakes_EWK	+0.10, 0.98	-0.01, 0.99	-0.03
y2018_flips	-0.02, 0.99	+0.00, 0.99	+0.01
y2018_fsrvar	+0.00, 0.99	+0.00, 0.99	-0.00
y2018_isr	+0.04, 0.99	+0.01, 0.99	-0.02
y2018_isrvar	+0.00, 0.99	+0.00, 0.99	-0.00
y2018_jer	**+0.74, 0.77**	**+0.50, 0.97**	-0.14
y2018_jes	**+0.21, 0.67**	**+0.14, 0.81**	-0.07
y2018_lep	+0.05, 0.99	+0.05, 0.99	-0.02
y2018_lumi	+0.08, 0.99	+0.07, 0.99	-0.03
y2018_pu	+0.10, 1.01	+0.06, 0.99	-0.02
y2018_trig	+0.05, 0.99	+0.05, 0.99	-0.03

1407 ## A.4 Nuisance correlation matrix

1408 Using the combine tool (via `combine -M FitDiagnostics card.txt --robustFit=1`
1409 `--saveShapes --saveWithUncertainties --saveOverallShapes --numToysForShapes`
1410 `200`), we can show the correlations between different nuisance parameters in the fit, in Figs. 84-
1411 85.

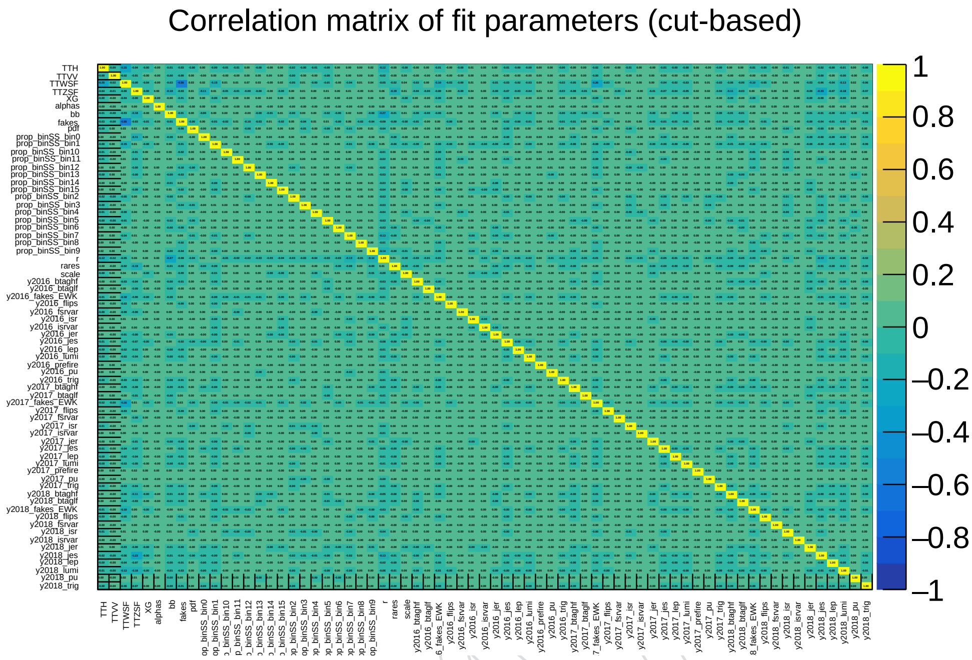


Figure 84: Correlation plot of nuisance-vs-nuisance for the s+b fit for the cut-based regions

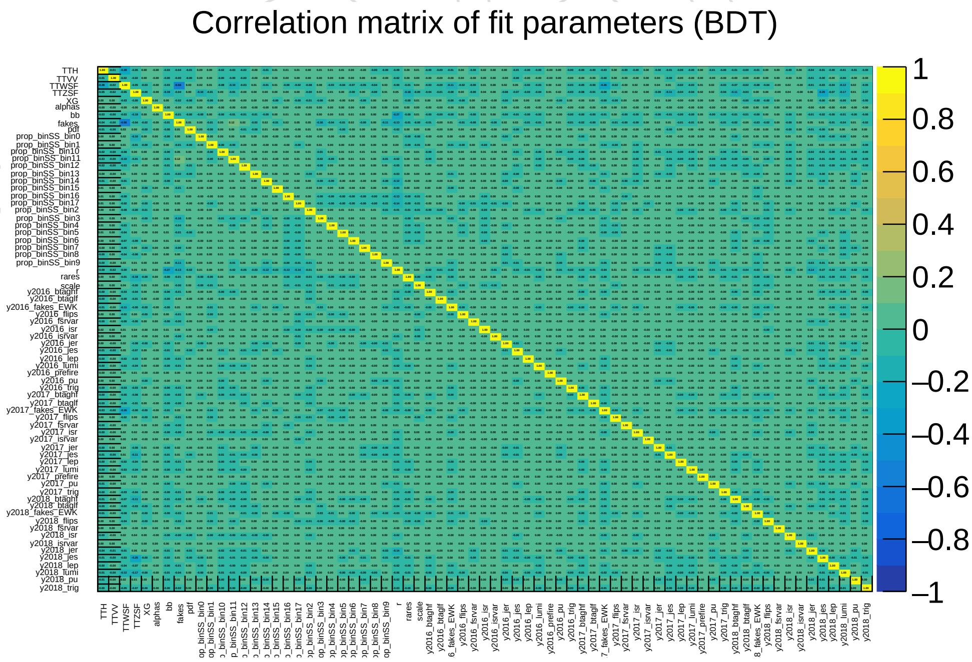


Figure 85: Correlation plot of nuisance-vs-nuisance for the s+b fit for the BDT regions

A.5 Goodness of fits

The goodness of fit distributions (using the saturated, Kolmogorov-Smirnov, and Anderson-Darling test statistics) with the signal+background fit to data for the cut-based and BDT analyses are shown in Figure 121. We note that the observation is generally within the bulk of the expected distributions for both analyses and all three metrics.

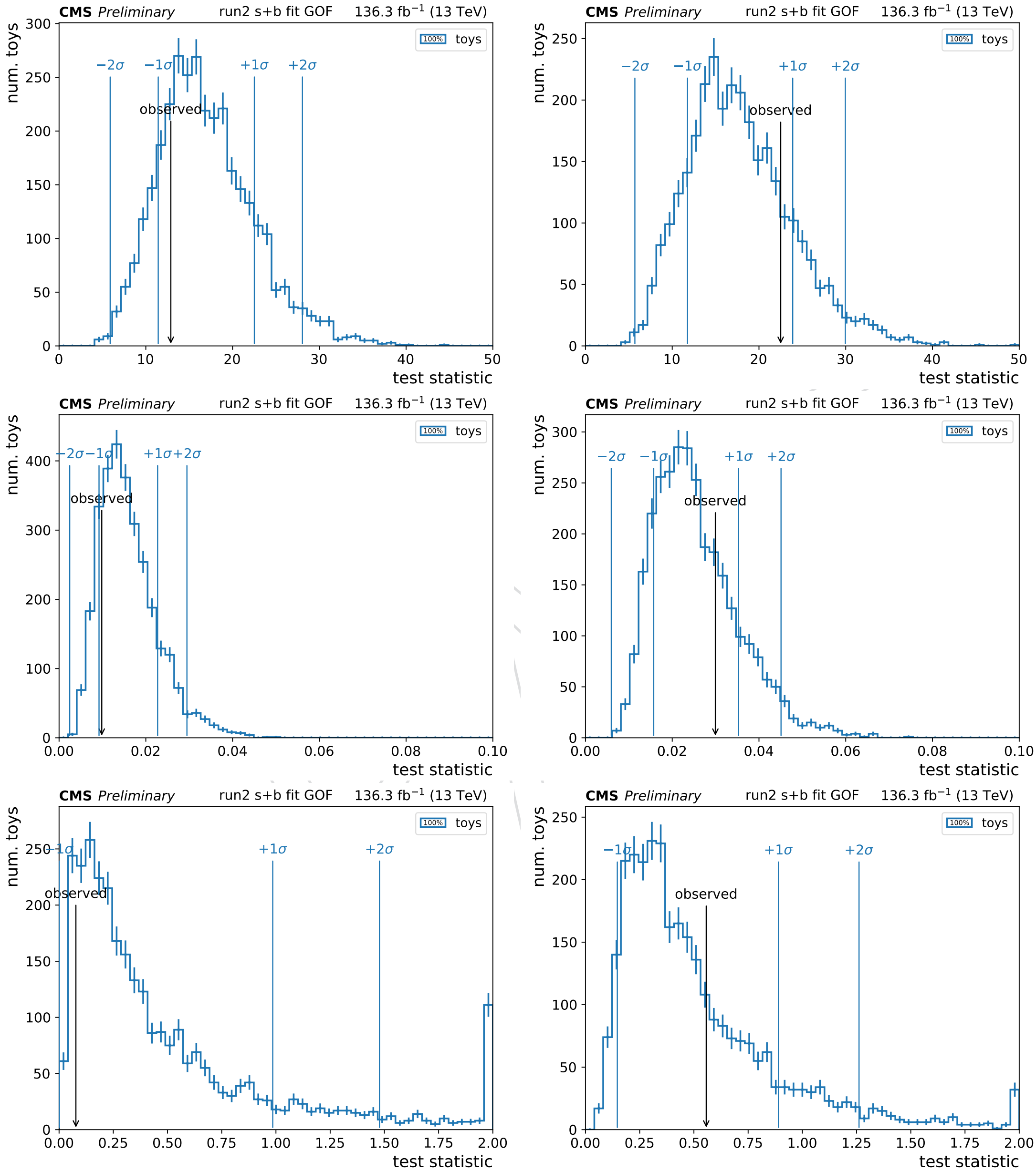


Figure 86: GOF test for the cut-based analysis (left) and BDT-based analysis (right) using the saturated (top), KS (middle), AD (bottom) test statistics.

B Shape variations

B.1 BDT discriminator

Shape variations for the BDT discriminator for the largest backgrounds and signal, for b-tagging, JES, JER, PU, prefiring, ISR/FSR parton shower, and ISR jet reweighting systematics are shown in Figures 87, 88, 89, and 90. Note that complete MC is used with events normalized to 136.3 fb^{-1}.

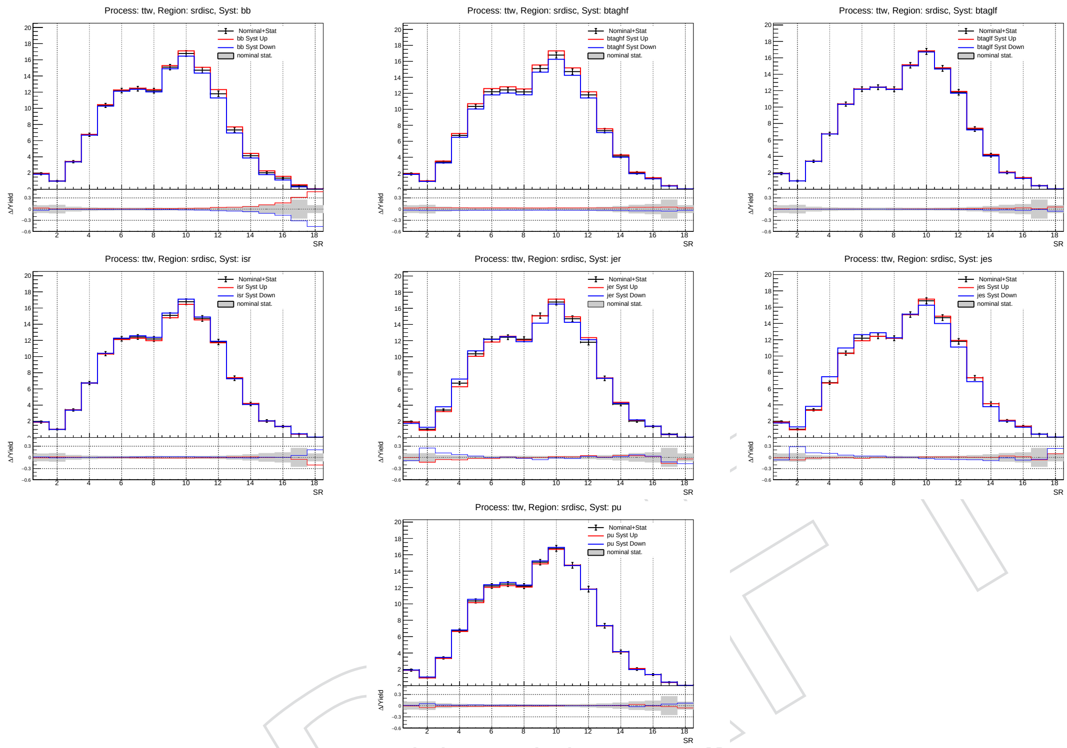


Figure 87: Systematic variations for ttW

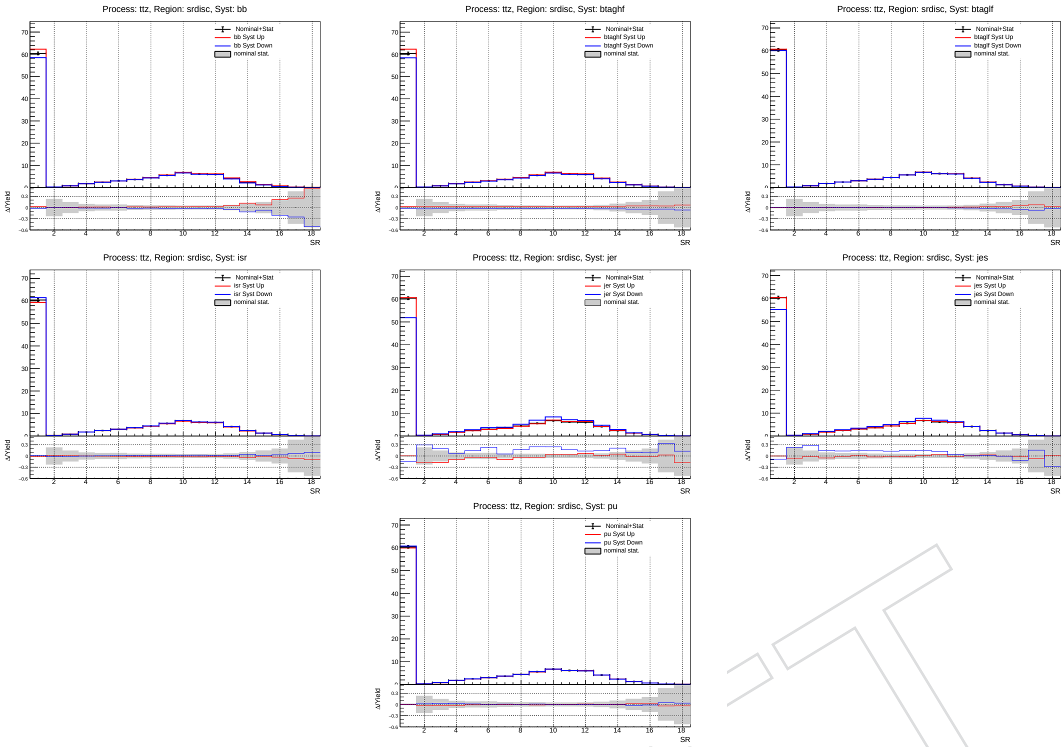


Figure 88: Systematic variations for ttZ

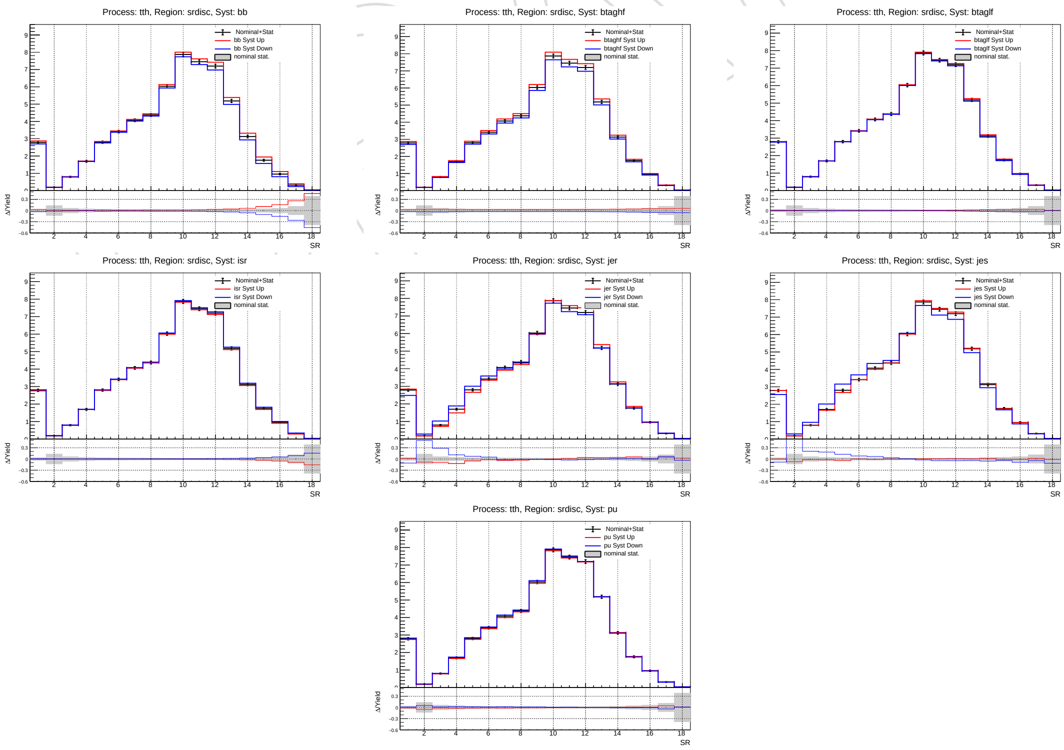


Figure 89: Systematic variations for ttH

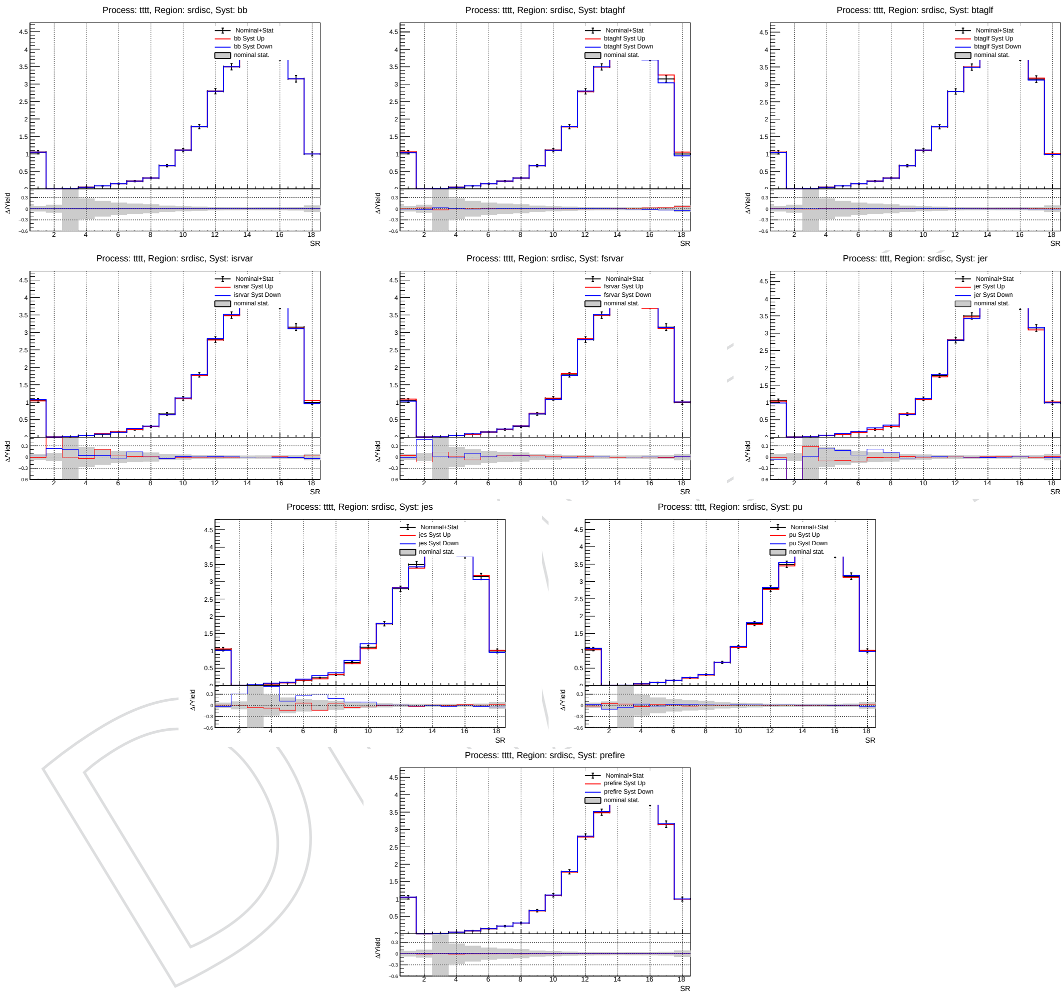


Figure 90: Systematic variations for tttt signal

C Studies which did not result in analysis changes

C.1 Jet and b-jet thresholds

We studied whether the $t\bar{t}t\bar{t}$ signal significance would improve by lowering the N_{jets} and $N_{\text{b jets}}$ counting thresholds from the 2016 ones which were 40 and 25 GeV, respectively. To do this, we used 2016 MC to estimate the expected significance of the analysis with an expected luminosity of 75 fb^{-1}, using the 2016+2017 cut based signal region binning. The result, in Figure 91, shows that the current configuration is close to optimal, with only a 3% increase in significance which could be obtained from reducing the $N_{\text{b jets}}$ threshold from 25 to 20 GeV. Given the minor improvement, and the potential of larger b-tagging scale factors at low p_{T} in the 2017 data, we decided to keep the 2016 thresholds.

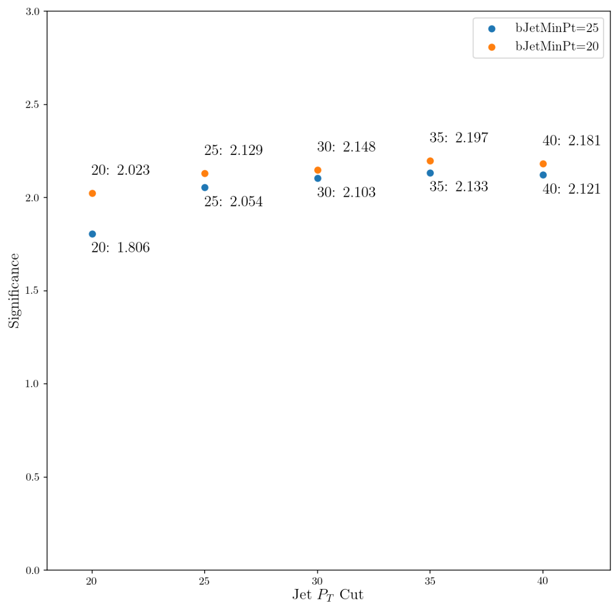


Figure 91: Signal significance for different transverse momentum requirements used in counting N_{jets} and $N_{\text{b jets}}$.

C.2 Taus

We studied whether the $t\bar{t}t\bar{t}$ signal significance would improve by including taus in addition to electrons and muons. For this study, we used an expected luminosity of 75 fb^{-1}, and we plotted the expected yields for events with reconstructed tau leptons, separating the truth-matched and non-truth-matched ones.

We focused on events with only one tau, and the other 1 or 2 leptons being e/μ, to avoid large yields of fake taus. Taus were selected with the following identification requirements: `decayModeFinding`, `againstMuonTight3`, `againstElectronTightMVA6`, `byTightIsolationMVArun2v1DBdR03oldDMwLT`. The last requirement was developed by the ttH analysis to be used in environments with high jet multiplicity [36]. The performance of tau identification is shown in Figure 92 for the $t\bar{t}t\bar{t}$ signal and the ttW/Z/H background samples, where it is clear that there are large off-diagonal terms due to inefficiency and impurity of the tau selection. Figure 93 shows the expected yields for each sample as a function of signal region for 1-tau events, split between $t\bar{t}t\bar{t}$ events where the tau is truth-matched, and the sum of ttW/Z/H events and $t\bar{t}t\bar{t}$ events with fake taus. The plot shows some potential in the 3-lepton regions ($ee\tau/e\mu\tau/\mu\mu\tau$), particularly 14-15-16, which integrate to about 0.5 $t\bar{t}t\bar{t}$ events and 1 background event, but this doesn't account for fake taus from non-ttW/Z/H backgrounds. Given these small yields and large backgrounds, we decided not to include signal regions with taus for the 2016+2017 analysis.

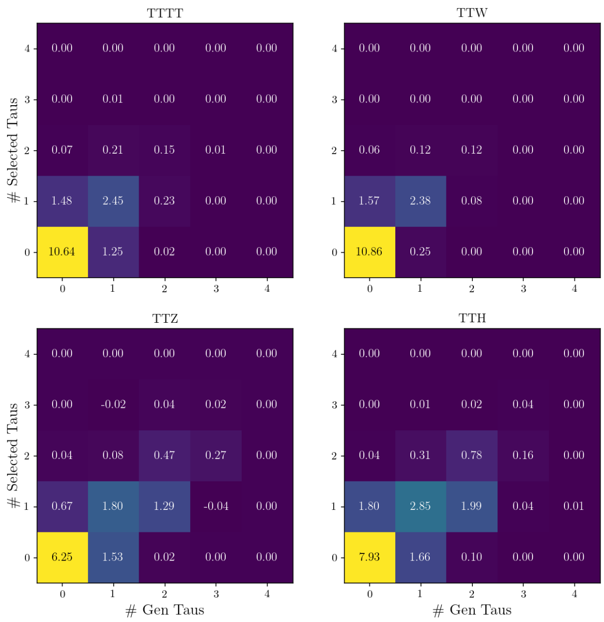


Figure 92: Number of events as a function of the number of true and reconstructed tau leptons, for different samples.

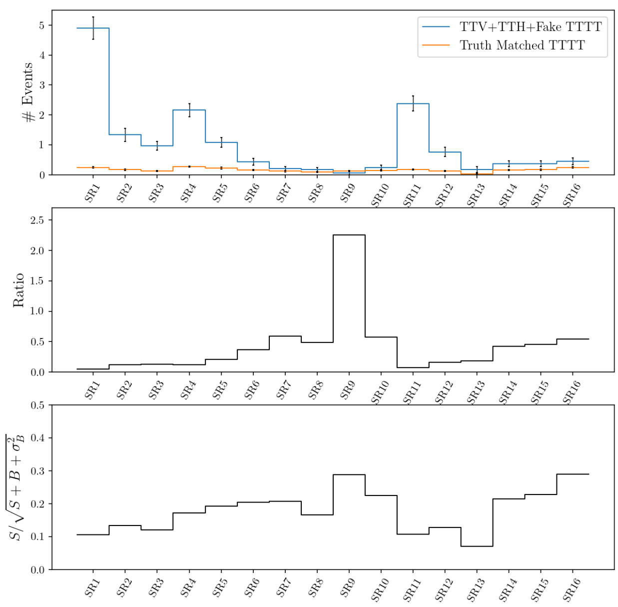


Figure 93: Top: number of events with 1 tau lepton and 1 or 2 e/μ leptons, as a function of the signal regions defined in Table 10, with SR11 and higher including events with 3 leptons (eeτ/e$\mu\tau$/$\mu\mu\tau$), showing truth-matched $t\bar{t}t\bar{t}$ signal (orange) and backgrounds from ttW/Z/H and non-truth-matched $t\bar{t}t\bar{t}$ (blue). Middle and bottom: S/B and and significance per bin, considering only the ttW/Z/H and non-truth-matched $t\bar{t}t\bar{t}$ events as background.

C.3 Top-tagging

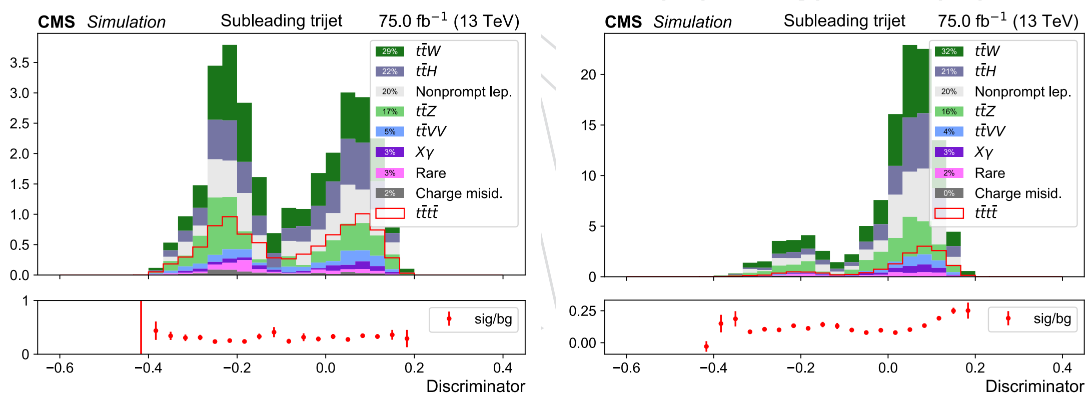
We explored resolved and merged top tagging as additional discrimination handles.

C.3.1 Resolved

Pairs of 3 jets, where one jet is a b-jet candidate, are fed into a pre-trained BDT from the single lepton and opposite-sign analysis, AN2017-146-v17. The BDT uses 6 inputs (b-tag discriminant, top candidate mass, W candidate mass, the ratio of the top p_T to the trijet p_T, and the $\Delta\phi$ between (top,W) and (top,b)). We can find the highest and second highest trijet discriminant in the event. For this analysis, the leading backgrounds can all have a hadronic top, so we focus on the subleading trijet discriminant.

Distributions of this quantity for two jet threshold schemes are shown in Figure 94. The nominal analysis uses jet thresholds of 40 GeV and b-tagged jet thresholds of 25 GeV. To explore the possibility that the high jet threshold is reducing the inputs to the BDT, we compare with looser jet thresholds of 20 GeV, 20 GeV, respectively.

In both cases, the ratio of signal to background shows only a slight trend. To test this more quantitatively, the subleading discriminator for the 20,20 scheme was put into a 19+1 variable BDT and was only ranked 9th. Compared to without the variable, the maximum significance only increased by 1%, so we do not pursue such tagging in the analysis. This is not unexpected since the same-sign and multilepton final state has one or two less hadronic tops to tag.



(a) (b)

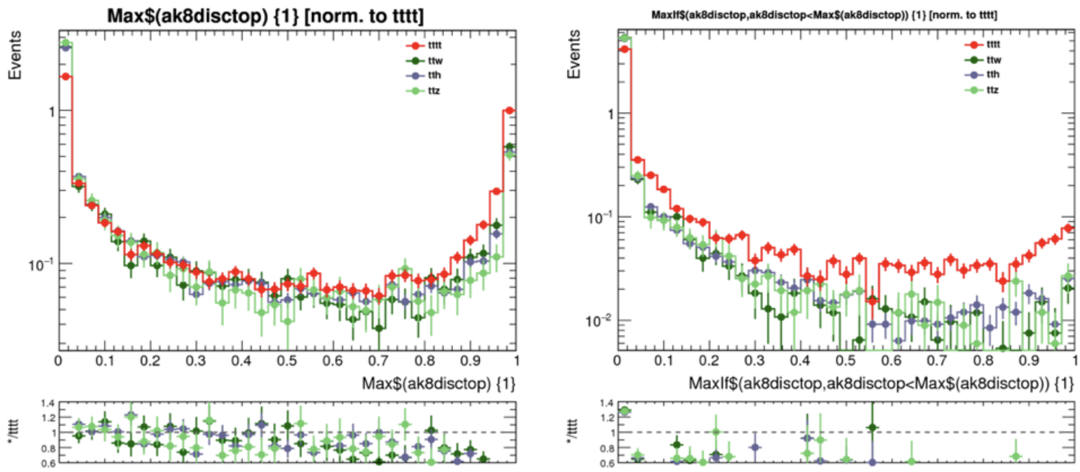
Figure 94: Distributions of subleading trijet discriminator in the signal regions for (jet,b-jet) p_T thresholds of (40,25) [left] and (20,20) [right].

C.3.2 Merged

Top decay products may become merged through accidental overlaps or boosts. The latter is not a large fraction in this final state due to lack of high MET, however, we explored the DeepAK8 tagger developed by the SUSY Heavy Object Tagging group, which provides numerical probabilities for objects such as top, W, Z, H, etc. for AK8 jets through a deep neural network acting on constituent particle flow candidates.

Following the same strategy as for the resolved top tagger, we look at the leading and subleading top discriminants for events passing the baseline selection for a few leading backgrounds in Figure 95.

Again, while there is a small trend toward higher values for the signal sample, the quantity of these events is miniscule (on the order of a few percent), so this tagging method was not pursued further.



(a) (b)

Figure 95: Maximum AK8 top discriminant in the event (left). Subleading AK8 top discriminant in the event (right). Histograms are normalized to the signal (red) cross-section.

D Impact of the L1 ECAL prefiring(2016/2017) and HEM15/16 loss (2018) on the results

D.1 L1 prefiring(2016/2017)

The L1 prefiring issue impacts 2016 and 2017 data collection periods. Due to mistiming and trigger rules, events with high η energy deposits can be preferentially lost. The impact of this is checked on 80X fastsim T1tttt signal samples. Inefficiency maps to be applied to simulation are taken from `https://ncsmith.web.cern.ch/ncsmith/PrefireEfficiencyMaps/Preliminary/Jet_L1FinOReff_bxm1_looseJet_SingleMuon_Run2016B-H.pdf` (Jet map 2016B-H) and `https://lathomas.web.cern.ch/lathomas/TSGStuff/L1Prefiring/PrefiringMaps/` (Jet map 2017B-F) and shown in Figure 96. For each event, consider all jets and electrons that pick up a non-zero scale factor from the chosen inefficiency map, parameterized by jet p_T and η and obtain a multiplicative scale factor (< 1) to apply to MC.

The procedure matches the tool given in `https://twiki.cern.ch/twiki/bin/view/CMS/L1ECALPrefiringWeightRecipe`.

Figures 97 and 98 show the average SFs as a function of (selected) electron and jet multiplicities per event, separately for 2016 and 2017. The average scale factors are about 2.5% (4%) below unity for 2016 (2017). Compared to the central value of the scale factors, the variation for increasing jet or electron multiplicity is small. Note that as a function of raw (not analysis/selected) jets, the trend may be larger, as the analysis jet selection criteria requires $|\eta| < 2.5$.

Scale factors for the prefiring inefficiency will be applied to the 2016 and 2017 MC samples.

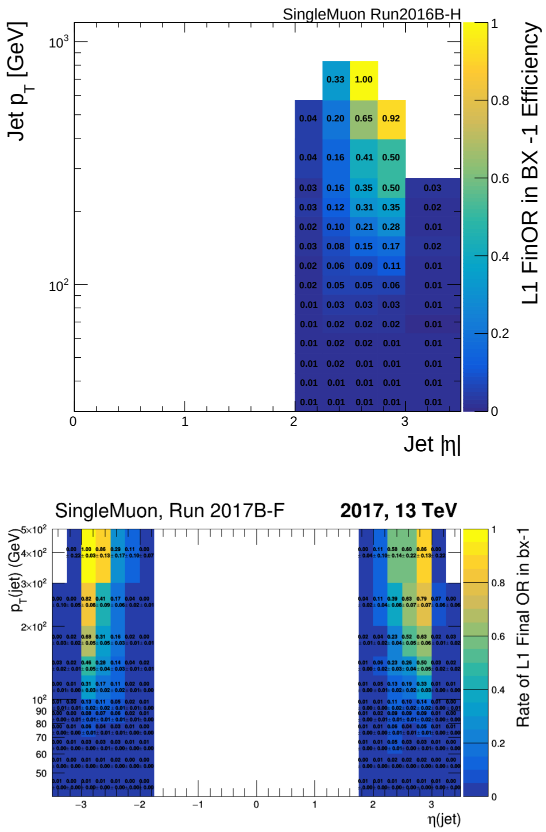


Figure 96: Prefiring inefficiency maps for 2016 (left) and 2017 (right)

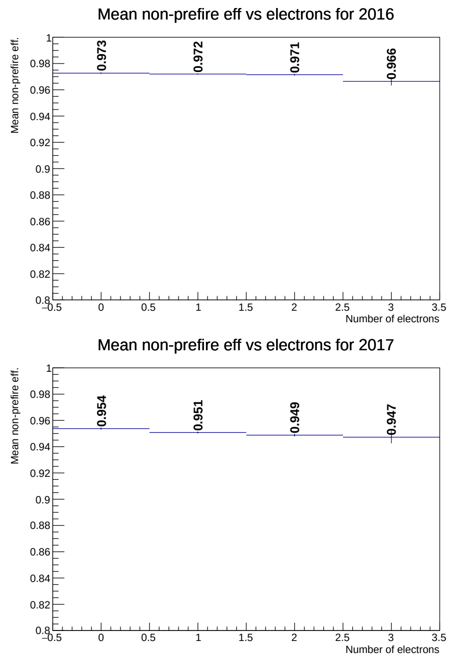


Figure 97: Average prefiring inefficiency as a function of electron multiplicity for 2016 (left) and 2017 (right)

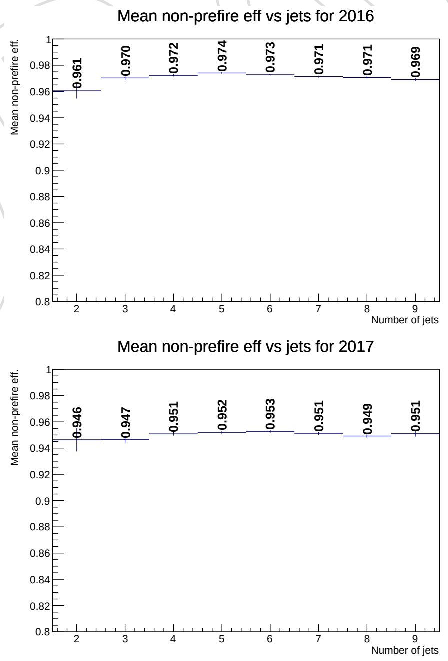


Figure 98: Average prefiring inefficiency as a function of jet multiplicity for 2016 (left) and 2017 (right)

D.2 HEM15/16 loss in 2018

Beginning in Run 319077 for the 2018 data collection period, HEM 15 and 16 sectors are switched off due to a power cut issue. This results in a gap of HCAL info for $\eta \in [-3, 1.3]$ and $\phi \in [-1.57, -0.87]$. This has the potential to increase electron fakes due to lower H/E and isolation values. The impact was checked in MC and in data.

The relative increase of fake electrons passing tight selection requirements was found to be approximately 12% in MC, while muons showed no increase. Since 23% of the Run2 dataset is impacted by the HEM issue, and considering fakes inclusively (e+μ), this increase corresponds to a 1% overall increase in fakes across the full dataset. Data checks showed similar increases to MC.

D.2.1 HEM impact in MC

Dedicated samples were produced to perform an apples-to-apples comparison of the HEM15/16 issue. In particular, the dataset

```
/TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8
/RunIISpring18MiniAOD-HEMPremix_100X_upgrade2018_realistic_v10-v3
/MINIAODSIM
```

was produced with HEM sectors disabled. A sample without "HEM" in the dataset name corresponds to the nominal HEM "in" case. For both samples, truth-matched fake single lepton efficiencies were calculated with nominal analysis tight ID+Iso as the numerator and reco-leptons from miniAOD as the denominator. Leptons were required to have $p_T > 25\,\text{GeV}$. Labeling HEM "in" and "out" samples as "good" and "bad", both good and bad samples were checked inclusively and also restricting to the η-ϕ region impacted by the HEM loss. The efficiencies are listed in Table 35. Muons are not impacted, while electrons show a $(12 \pm 5)\%$ relative increase in fakes inclusively, or a factor of 2 increase in the HEM region specifically.

Table 35: Single lepton efficiencies for good and bad HEM samples.

region	scenario	flavor	sig. eff. (%)	bkg. eff. (%)
all	good	e	65.8 ± 0.1	0.15 ± 0.01
		μ	84.6 ± 0.1	0.31 ± 0.01
	bad	e	65.9 ± 0.1	0.17 ± 0.0
		μ	84.7 ± 0.1	0.29 ± 0.01
hemregion	good	e	53.9 ± 0.8	0.21 ± 0.04
		μ	80.7 ± 1.0	0.21 ± 0.06
	bad	e	59.7 ± 0.7	0.47 ± 0.03
		μ	82.3 ± 0.9	0.33 ± 0.07

D.2.2 HEM impact in data

To assess the impact in data, the 2018 dataset was split into two subsections: "before" (runs before 319077, corresponding to 20.25 fb^{-1}) and "after" (runs including and after 319077, corresponding to 30.73 fb^{-1}). After normalizing both sections to the same luminosity, plots of lepton flavor ($\mu\mu$,$e\mu$,ee) for 3 kinematic regions (Z$\rightarrow \ell\ell$-dominated, $t\bar{t} \rightarrow \ell\ell$-dominated, and the (fake-dominated) tight-loose control region) are shown in Figure 99. While the first two prompt-enriched regions show a relative increase afterwards, this could be attributed to a variety of data-taking condition differences between the two periods. The tight-loose region shows similar counts for dimuon events, but after/before ratios of 1.13 ± 0.08 for $e\mu$ and 1.37 ± 0.20 for dielectron events. Such an increase is consistent with the 12% increase found from simulation.

Figure 100 shows the lepton ϕ distribution for Z-dominated and tight-loose regions. Muons show no relative changes. When focusing on $\phi \in [-1.57, -0.87]$, no increase is observed for prompt electrons, while a $(59 \pm 23)\%$ increase is seen for the tight-loose electron counts.

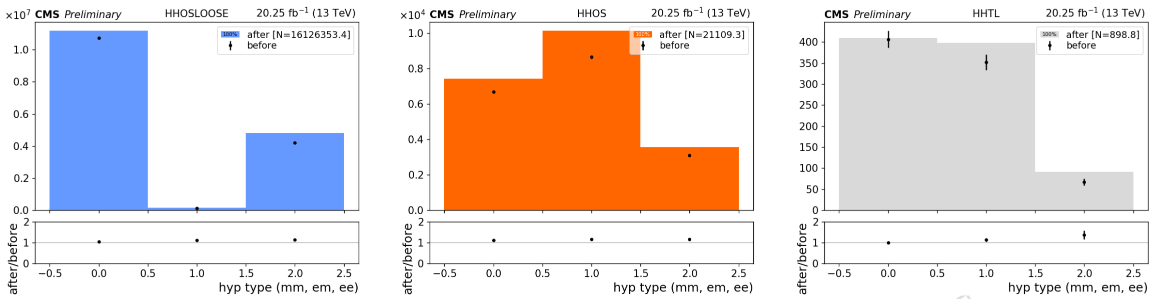


Figure 99: Z-dominated (left), $t\bar{t}$-dominated (center), tight-loose (right).

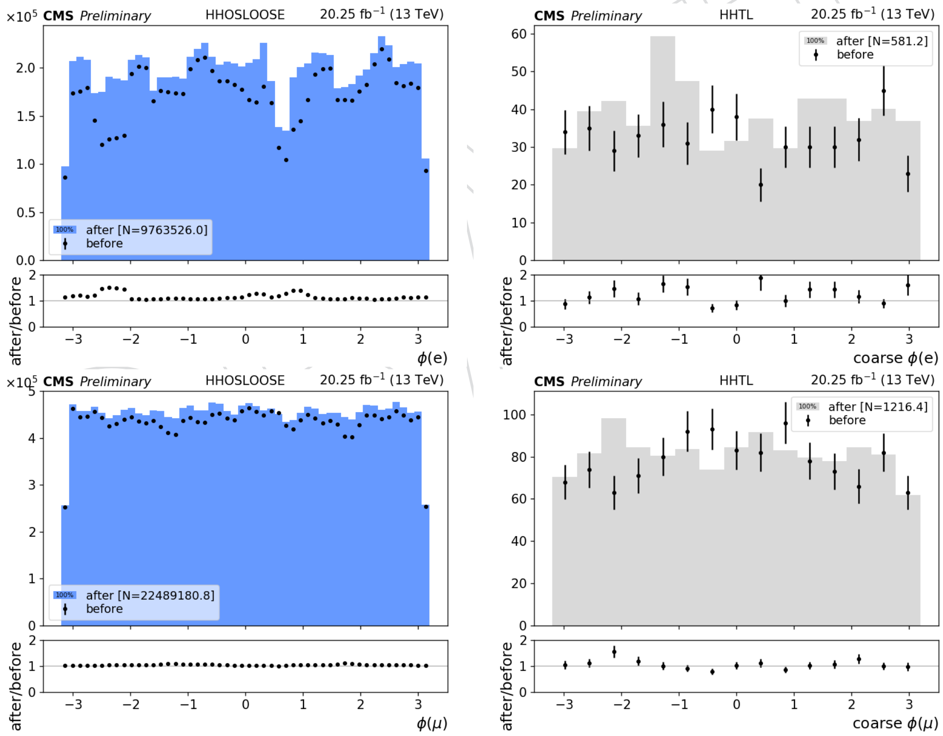


Figure 100: Z-dominated region (left), tight-loose region (right), for electrons (top) and muons (bottom)

E Analysis changes

E.1 Changes from ANv7

This section details some changes and updates with respect to V7 of this note. Different items are accompanied by the relative change in expected significance from the ANV7 number (cut-based gave 3.069σ, and BDT gave 3.328σ).

However, it is worth noting that JECs for 2018 data and MC are not yet available, so 2017 JECs are being used instead. Several of the changes enumerated below and their conclusions are sensitive to JECs.

The compounded effect of these changes brings the cut-based significance to 2.371σ and the BDT-based significance to 2.591σ.

- V32 JEC
    - With respect to the previous version of JECs (V6), V32 reparameterizes the L1 correction for better stability. While the product of L1,L2L3 remained the same from V6 to V32, the relative ratios did not. Our lepton isolation quantities, $p_{\mathrm{T}}^{\mathrm{ratio}}$ and $p_{\mathrm{T}}^{\mathrm{rel}}$ depend on these two corrections separately, not only on their product. Upgrading to V32 caused a per lepton efficiency loss of a few percent, so we re-tuned our isolation working points for 2017 and 2018 to approximately match the signal and background efficiencies from V6.
    - The updated working point values can be found in Table 7.
    - Relevant discussion in `https://hypernews.cern.ch/HyperNews/CMS/get/JetMET/1891.html`
    - σ decreases by 4%.
- Prefiring SF
    - The latest recommended prefiring inefficiency maps (from `https://twiki.cern.ch/twiki/bin/view/CMS/SUSRecommendations18#Prefire_Issue`) were incorporated for 2016 and 2017 MC. Application of the maps scales down four top signal by 2.5% (4.7%) in 2016 (2017) and the $\mathrm{t\bar{t}W}$ background by 1.5% (3.3%) in 2016 (2017).
    - After application of the inefficiency maps, disagreements in high-$|\eta|$ tails of the opposite-sign control region are largely cured.
    - For reference, the affected datasets (2016+2017) are only 58% of the Run2 dataset by integrated luminosity.
    - σ decreases by 1%.
- Merging 4b signal region bins
    - Signal regions 8, 9, and 10 (now just 8), corresponding to at least 4 b-tagged jets and at least 5 jets, were merged (removing the jet multiplicity boundaries) to be more conservative when dealing with the tight-loose control region yields for the fake background prediction
    - The cut-based signal region definition now has 14 signal regions (+2 $\mathrm{t\bar{t}W}/\mathrm{t\bar{t}Z}$ dominated regions, as before).
    - σ decreases by 4%.
- Single card
    - Card inputs to the HiggsCombine tool are no longer treated as 3 sepa-

  rate channels in a simultaneous fit. Instead, yields (including control region yields) are summed into single histograms (2016+2017+2018). Uncorrelated nuisances, for example, prefiring uncertainties affecting only 2016/2017, are considered as shape variations only on the 2016/2017 component of the summed nominal histogram. This has the benefit of not being as susceptible to MC statistics issues for each of the 3 years. It also is more conservative and takes advantage of larger control region statistics for the fake prediction.
  - σ decreases by 5%.
- 2017/2018 MC
  - The Autumn18 MC campaign is nearly complete, so we have switched to taking relevant 2018 predictions from the correct MC rather than using 2017 as a stand-in.
  - Explicitly, we have all the desired samples in the 2018 campaign except the signal sample is still lacking 25% MC statistics. This does not matter as we don't suffer MC statistics issues for $t\bar{t}t\bar{t}$.
  - In 2017, we are still lacking a very small Z+γ background sample, which we take from the 2018 campaign instead.
  - σ decreases by 3%.
- 2017/2018 ISR reweighting
  - For the 2016 dataset, we reweight the $t\bar{t}W$ and $t\bar{t}Z$ background predictions in order to match the data $N_J^{ISR/FSR}$ distribution, as detailed in Section 2. This had the effect of scaling down high jet multiplicity events, decreasing the background prediction. In 2017 and 2018, however, the Pythia tune was changed to CP5, so we requested dedicated samples with configurations matching $t\bar{t}W$ and $t\bar{t}Z$ in order to derive new weights. This results in the opposite trend to 2016, i.e., high jet multiplicity events can get scaled up by almost 30%, increasing the background prediction.
  - In 2017/2018, the samples for $t\bar{t}Z$ and $t\bar{t}W$ with 0 and 1 extra parton, respectively, are
    - /TT_DiLept_TuneCP5_13TeV-amcatnlo-pythia8/
    - /TTPlus1Jet_DiLept_TuneCP5_13TeV-amcatnloFXFX-pythia8/

    which are completed by the appropriate processing string for each campaign. In the case of 2017, full event statistics are split between a nominal sample and an extension ("ext1") sample.
  - σ decreases by 2%.
- 2018 b-tag WPs
  - The 2018 b-tag medium WP for DeepCSV preserves the b tagging and light mis-tagging efficiencies compared to 2017. However, it appears the charm mis-tagging efficiency relatively increased by about 15%. In the 3 and 4 b-tag bins, $t\bar{t}W$ can enter if there is additional heavy flavor contribution (e.g., $t\bar{t}W$ +bb via gluon splitting), or through mistags of W decays containing a charm quark.
  - Relevant discussion in `https://hypernews.cern.ch/HyperNews/CMS/get/btag/1637.html`
  - σ decreases by 3%.

- ttW+bb, ttZ+bb, ttH+bb scaling
  - As a response to a 7 Dec 2018 question from Otto about Figs 12, 13 (`https://twiki.cern.ch/twiki/bin/viewauth/CMS/TOP18003`), we showed that scaling the gluon splitting (events with extra $b\bar{b}$) components up by the systematic uncertainty of 1.35 brings better agreement in the b-tag multiplicity distribution.
  - We now revert to the procedure of the previous iteration of this analysis for 2016 by scaling ($t\bar{t}W/t\bar{t}Z/t\bar{t}H$)+$b\bar{b}$ up by 1.7 ± 0.6 with a systematic uncertainty corresponding to the specified error, from the result of TOP-16-010, which found a Data/MC discrepancy in the ttbb/ttjj ratio.
  - σ decreases by 6%.

F Unblinding of 2016 dataset

Following the pre-approval talk for the analysis on Feb. 15, 2019, the analysis was unblinded using the 2016 dataset. The yields and results are shown in Section F.1 with consistency checks of cut-based and BDT results in Section F.2. We include the results of various statistical checks, including nuisance pulls (Section F.3), and nuisance impacts (Section F.4).

F.1 Yields and results

Plots for the unblinded prefit and postfit event yields for the 2016 data, with a total luminosity of 35.9 fb^{-1}, are shown for both the cut-based and BDT based analysis in Figure 101. Numerical yields are also tabulated in Table 36 (prefit cut-based analysis), Table 37 (postfit cut-based analysis), Table 38 (prefit BDT analysis), Table 39 (postfit BDT analysis).

A binned likelihood fit is performed over signal regions using the Higgs Combine tool [22]; exclusion limits at 95% CL are calculated with the Asymptotic CLs method.

With the cut-based analysis, an observed (expected) upper limit on the production cross section of 33.67 fb ($20.48^{+11.18}_{-6.83}$ fb), assuming the signal process does not exist. The observed (expected) significance is 1.012 (1.374) standard deviations, corresponding to a measured signal strength parameter of $0.812^{+1.054}_{-0.800}$.

With the BDT analysis, an observed (expected) upper limit on the production cross section of 42.59 fb ($19.36^{+10.11}_{-6.46}$ fb), assuming the signal process does not exist. The observed (expected) significance is 1.812 (1.424) standard deviations, corresponding to a measured signal strength parameter of $1.538^{+1.137}_{-0.935}$.

F.2 Cut-based and BDT consistency checks

One noteworthy excess for the cut-based is SR4, which has 8 observed events with about 2.2 ± 0.5 predicted background. This is a 2-lepton, 5 jet, 3 b-tagged jet bin. In the 2016 result, this corresponded to part of oldSR4 (2-lepton, 5 or 6 jets, 3 b-tagged jets), and had 8 observed events. With the new binning which splits oldSR4 into SR4 and SR5, the data events remain in the 5 jet bin, SR4, leaving the slightly higher s/b bin SR5 with 0 data.

It is also worth pointing out the mildness of the disagreement between cut-based and BDT significances (both have differing directions for observed significance with respect to their expected significance). We have checked with toy pseudo-datasets that the probability of this happening is close to 50%.

The asimov-like toys were constructed in a way that would preserve the correlation between the cut-based and BDT significance results. The procedure for creating one toy is to consider all background events for the analysis, storing the cut-based and BDT SR bin number and event weight. For each background/signal process, calculate the total yield y in the cut-based/BDT signal regions and sample a number N from a Poisson distribution with parameter y. Next, draw N random events from the set of events for a particular process with relative probability given by the absolute value of the event weight (so there is a slight bias due to negatively weighted events). These N sampled events then are used to fill a histogram for the cut-based and BDT shapes. In this way, we create $\mu \cdot s + b$ toys with $\mu = 1$, which is then an input to HiggsCombine to produce a pair of significances (cut-based σ and BDT σ).

Several thousand pairs of values from the toy pseudo-datasets are binned into 2D histogram with the 1,2, and 3-σ contours overlaid. The threshold for the contours is calculated with a 2-sample χ^2 metric. Overlaying the observed significances from above, we find it lies within the 0.7-σ contour with a p-value of 0.5, as shown in Figure 108. As described in the caption, the observed point is "corrected" (only within the scope of this toy study) before plotting it.

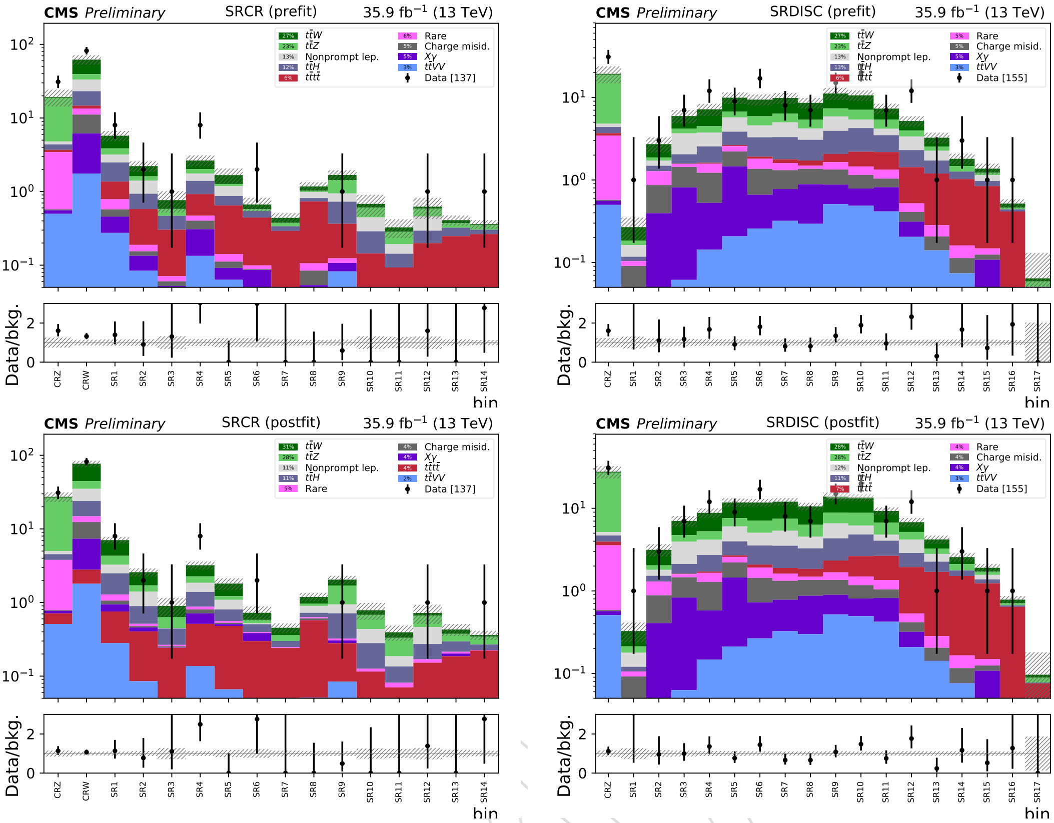


Figure 101: 2016: data yields compared to prefit (top) and postfit (bottom) SM predictions in CRZ, CRW, and the cut-based signal regions (left), and CRZ with the BDT signal regions (right). Note that we plan to retrain the BDT with latest JECs/corrections for 2018 (with identical settings)

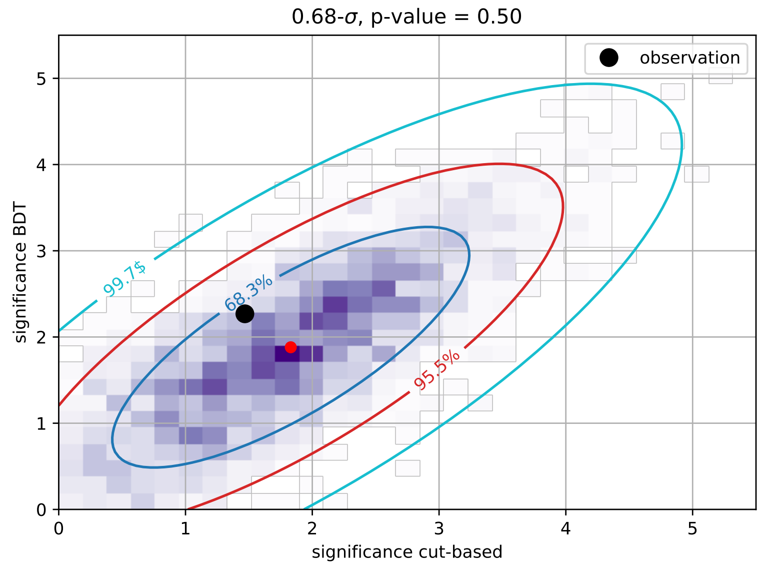


Figure 102: Comparison of correlated expected significances between cut-based and BDT fits with observed point overlaid. Because significances from toy datasets are used, we get a correction factor for expected BDT/cut-based significances with the full fit and the mean of the significances from the toy datasets. This correction factor is then applied to the observed point such that the expected significance values from the full analysis would match the mean values from the toy pseudo-datasets.

Table 36: Prefit event yields in cut-based regions for 2016.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.50 ± 0.18	14.01 ± 4.84	0.69 ± 0.17	0.50 ± 0.06	0.06 ± 0.01	2.87 ± 0.58	0.02 ± 0.00	0.42 ± 0.18	19.06 ± 4.95	31	0.25 ± 0.03
CRW	22.56 ± 7.66	6.31 ± 2.14	8.44 ± 2.07	1.76 ± 0.21	4.41 ± 0.63	2.33 ± 0.48	4.96 ± 0.95	10.05 ± 3.79	60.82 ± 9.06	82	1.23 ± 0.11
SR1	1.87 ± 0.68	0.68 ± 0.24	1.12 ± 0.30	0.27 ± 0.04	0.18 ± 0.08	0.21 ± 0.05	0.12 ± 0.02	0.70 ± 0.32	5.16 ± 0.93	8	0.58 ± 0.06
SR2	0.61 ± 0.24	0.20 ± 0.09	0.35 ± 0.09	0.08 ± 0.02	0.05 ± 0.02	0.03 ± 0.01	0.02 ± 0.00	0.48 ± 0.23	1.83 ± 0.37	2	0.39 ± 0.03
SR3	0.18 ± 0.10	0.11 ± 0.06	0.17 ± 0.06	0.05 ± 0.01	0.00 ± 0.01	0.01 ± 0.00	0.01 ± 0.00	0.00 ± 0.17	0.53 ± 0.25	1	0.23 ± 0.05
SR4	0.63 ± 0.24	0.26 ± 0.11	0.47 ± 0.12	0.13 ± 0.02	0.17 ± 0.07	0.07 ± 0.01	0.09 ± 0.02	0.36 ± 0.21	2.19 ± 0.46	8	0.46 ± 0.06
SR5	0.40 ± 0.18	0.09 ± 0.03	0.22 ± 0.07	0.06 ± 0.02	0.03 ± 0.01	0.03 ± 0.01	0.02 ± 0.00	0.31 ± 0.30	1.17 ± 0.39	0	0.51 ± 0.05
SR6	0.09 ± 0.06	0.03 ± 0.04	0.10 ± 0.04	0.01 ± 0.01	0.07 ± 0.05	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.07	0.32 ± 0.16	2	0.34 ± 0.04
SR7	0.06 ± 0.04	0.04 ± 0.02	0.05 ± 0.02	0.01 ± 0.00	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.01	0.16 ± 0.07	0	0.28 ± 0.08
SR8	0.15 ± 0.08	0.04 ± 0.02	0.08 ± 0.03	0.05 ± 0.01	0.00 ± 0.00	0.02 ± 0.00	0.03 ± 0.01	0.17 ± 0.11	0.55 ± 0.17	0	0.63 ± 0.06
SR9	0.25 ± 0.09	0.49 ± 0.17	0.36 ± 0.09	0.08 ± 0.01	0.02 ± 0.00	0.02 ± 0.01	0.00 ± 0.00	0.22 ± 0.11	1.44 ± 0.27	1	0.24 ± 0.07
SR10	0.08 ± 0.03	0.16 ± 0.07	0.14 ± 0.04	0.02 ± 0.00	0.00 ± 0.00	0.01 ± 0.01	0.00 ± 0.00	0.16 ± 0.19	0.57 ± 0.22	0	0.11 ± 0.02
SR11	0.04 ± 0.02	0.09 ± 0.05	0.05 ± 0.02	0.01 ± 0.00	0.00 ± 0.00	0.01 ± 0.01	0.00 ± 0.00	0.05 ± 0.04	0.26 ± 0.09	0	0.07 ± 0.02
SR12	0.04 ± 0.02	0.13 ± 0.05	0.09 ± 0.03	0.02 ± 0.00	0.00 ± 0.00	0.02 ± 0.00	0.00 ± 0.00	0.17 ± 0.18	0.46 ± 0.20	1	0.17 ± 0.02
SR13	0.04 ± 0.02	0.05 ± 0.03	0.08 ± 0.02	0.01 ± 0.00	0.02 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.02	0.20 ± 0.07	0	0.21 ± 0.02
SR14	0.01 ± 0.02	0.05 ± 0.03	0.04 ± 0.02	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.01	0.11 ± 0.05	1	0.25 ± 0.04

Table 37: Postfit event yields in cut-based regions for 2016.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.72 ± 0.18	21.58 ± 4.13	0.75 ± 0.17	0.51 ± 0.06	0.06 ± 0.01	3.00 ± 0.59	0.02 ± 0.00	0.47 ± 0.21	27.09 ± 4.12	31	0.20 ± 0.22
CRW	32.12 ± 7.40	9.64 ± 1.87	9.08 ± 2.03	1.80 ± 0.22	4.57 ± 0.66	2.45 ± 0.50	5.03 ± 0.84	11.13 ± 4.44	75.82 ± 7.26	82	0.99 ± 1.03
SR1	2.72 ± 0.68	1.04 ± 0.21	1.21 ± 0.27	0.28 ± 0.03	0.19 ± 0.06	0.22 ± 0.05	0.12 ± 0.02	0.77 ± 0.35	6.56 ± 0.71	8	0.47 ± 0.52
SR2	0.87 ± 0.23	0.31 ± 0.07	0.38 ± 0.09	0.09 ± 0.01	0.05 ± 0.01	0.04 ± 0.01	0.02 ± 0.00	0.51 ± 0.23	2.26 ± 0.33	2	0.32 ± 0.30
SR3	0.27 ± 0.10	0.18 ± 0.05	0.18 ± 0.05	0.05 ± 0.01	0.00 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.00 ± 0.18	0.71 ± 0.26	1	0.19 ± 0.21
SR4	0.94 ± 0.25	0.42 ± 0.09	0.52 ± 0.12	0.14 ± 0.02	0.20 ± 0.05	0.07 ± 0.01	0.10 ± 0.02	0.46 ± 0.24	2.83 ± 0.40	8	0.37 ± 0.35
SR5	0.59 ± 0.16	0.13 ± 0.03	0.24 ± 0.07	0.07 ± 0.01	0.03 ± 0.00	0.03 ± 0.01	0.02 ± 0.00	0.28 ± 0.26	1.39 ± 0.33	0	0.41 ± 0.38
SR6	0.14 ± 0.06	0.06 ± 0.03	0.11 ± 0.03	0.01 ± 0.01	0.08 ± 0.05	0.01 ± 0.00	0.00 ± 0.00	0.01 ± 0.08	0.44 ± 0.16	2	0.29 ± 0.26
SR7	0.09 ± 0.05	0.06 ± 0.02	0.05 ± 0.02	0.01 ± 0.00	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.01	0.22 ± 0.07	0	0.23 ± 0.19
SR8	0.23 ± 0.08	0.06 ± 0.02	0.09 ± 0.03	0.05 ± 0.01	0.00 ± 0.00	0.02 ± 0.00	0.03 ± 0.01	0.18 ± 0.11	0.67 ± 0.16	0	0.52 ± 0.50
SR9	0.35 ± 0.09	0.75 ± 0.16	0.39 ± 0.09	0.08 ± 0.01	0.03 ± 0.00	0.02 ± 0.01	0.00 ± 0.00	0.23 ± 0.13	1.85 ± 0.26	1	0.20 ± 0.16
SR10	0.11 ± 0.03	0.24 ± 0.06	0.15 ± 0.04	0.02 ± 0.00	0.00 ± 0.00	0.01 ± 0.01	0.00 ± 0.00	0.15 ± 0.17	0.69 ± 0.20	0	0.09 ± 0.11
SR11	0.06 ± 0.02	0.15 ± 0.04	0.05 ± 0.02	0.01 ± 0.00	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.05 ± 0.05	0.34 ± 0.09	0	0.06 ± 0.07
SR12	0.06 ± 0.02	0.20 ± 0.05	0.10 ± 0.02	0.02 ± 0.00	0.00 ± 0.00	0.02 ± 0.00	0.00 ± 0.00	0.19 ± 0.19	0.58 ± 0.21	1	0.14 ± 0.14
SR13	0.05 ± 0.03	0.08 ± 0.04	0.08 ± 0.02	0.01 ± 0.00	0.02 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.02	0.26 ± 0.08	0	0.17 ± 0.16
SR14	0.02 ± 0.02	0.07 ± 0.03	0.04 ± 0.01	0.01 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.01	0.15 ± 0.05	1	0.21 ± 0.21

Table 38: Prefit event yields in BDT regions for 2016.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.50± 0.17	14.01± 4.41	0.69± 0.17	0.50± 0.06	0.06± 0.01	2.87± 0.56	0.02± 0.00	0.42± 0.17	19.06± 4.49	31	0.25± 0.03
SR1	0.08± 0.04	0.02± 0.01	0.01± 0.01	0.00± 0.00	0.02± 0.01	0.01± 0.01	0.07± 0.01	0.05± 0.05	0.27± 0.08	1	0.00± 0.00
SR2	0.84± 0.30	0.16± 0.06	0.21± 0.05	0.02± 0.00	0.38± 0.19	0.41± 0.18	0.47± 0.10	0.23± 0.15	2.71± 0.58	3	0.00± 0.01
SR3	1.75± 0.65	0.52± 0.19	0.47± 0.15	0.06± 0.01	0.75± 0.21	0.13± 0.05	0.62± 0.13	1.62± 0.69	5.93± 1.09	7	0.01± 0.01
SR4	2.70± 0.89	0.72± 0.32	0.94± 0.24	0.14± 0.02	0.38± 0.19	0.35± 0.10	0.69± 0.14	1.21± 0.57	7.14± 1.20	12	0.04± 0.01
SR5	3.19± 1.07	1.00± 0.45	1.20± 0.31	0.21± 0.03	1.25± 0.29	0.36± 0.20	0.75± 0.15	1.88± 0.73	9.84± 1.60	9	0.08± 0.01
SR6	3.49± 1.15	1.33± 0.45	1.37± 0.35	0.26± 0.03	0.40± 0.21	0.44± 0.12	0.70± 0.14	1.29± 0.48	9.29± 1.45	17	0.11± 0.05
SR7	3.47± 1.14	1.37± 0.48	1.53± 0.37	0.32± 0.04	0.46± 0.15	0.29± 0.06	0.54± 0.11	1.68± 0.70	9.65± 1.50	8	0.18± 0.03
SR8	3.22± 1.06	1.16± 0.43	1.35± 0.34	0.29± 0.06	0.59± 0.15	0.12± 0.10	0.48± 0.10	1.14± 0.48	8.35± 1.26	7	0.23± 0.02
SR9	3.78± 1.40	1.48± 0.51	1.87± 0.48	0.51± 0.06	0.37± 0.22	0.31± 0.07	0.46± 0.09	1.95± 0.92	10.72± 1.91	15	0.42± 0.04
SR10	3.52± 1.21	1.68± 0.56	2.05± 0.52	0.49± 0.06	0.30± 0.07	0.31± 0.06	0.36± 0.07	1.15± 0.52	9.85± 1.56	20	0.74± 0.07
SR11	1.94± 0.70	1.21± 0.41	1.31± 0.36	0.42± 0.06	0.40± 0.09	0.30± 0.08	0.22± 0.04	0.69± 0.39	6.48± 1.07	7	0.84± 0.05
SR12	1.19± 0.46	0.62± 0.22	0.87± 0.23	0.20± 0.03	0.11± 0.02	0.11± 0.03	0.10± 0.02	1.07± 0.45	4.26± 0.75	12	0.91± 0.06
SR13	0.70± 0.28	0.29± 0.11	0.52± 0.18	0.14± 0.03	0.00± 0.03	0.08± 0.02	0.07± 0.01	0.54± 0.32	2.33± 0.48	1	0.91± 0.04
SR14	0.34± 0.16	0.20± 0.08	0.24± 0.08	0.07± 0.01	0.00± 0.03	0.05± 0.01	0.04± 0.01	0.00± 0.11	0.94± 0.28	3	0.86± 0.07
SR15	0.14± 0.08	0.09± 0.04	0.13± 0.05	0.04± 0.01	0.06± 0.06	0.02± 0.01	0.02± 0.00	0.17± 0.12	0.68± 0.21	1	0.69± 0.06
SR16	0.05± 0.03	0.02± 0.02	0.03± 0.01	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.03	0.12± 0.07	1	0.39± 0.05
SR17	0.01± 0.01	0.01± 0.01	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.06	0.02± 0.07	0	0.05± 0.01

Table 39: Postfit event yields in BDT regions for 2016.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.66± 0.19	21.93± 4.77	0.74± 0.17	0.51± 0.06	0.06± 0.01	2.98± 0.55	0.02± 0.00	0.47± 0.18	27.37± 4.71	31	0.38± 0.24
SR1	0.11± 0.04	0.03± 0.01	0.01± 0.01	0.00± 0.00	0.02± 0.01	0.01± 0.01	0.07± 0.01	0.06± 0.06	0.33± 0.09	1	0.00± 0.00
SR2	1.10± 0.33	0.24± 0.08	0.22± 0.05	0.02± 0.00	0.39± 0.18	0.41± 0.18	0.48± 0.09	0.27± 0.17	3.14± 0.54	3	0.00± 0.01
SR3	2.30± 0.71	0.80± 0.21	0.51± 0.14	0.06± 0.01	0.77± 0.17	0.13± 0.05	0.63± 0.11	1.82± 0.80	7.02± 1.04	7	0.02± 0.02
SR4	3.54± 0.97	1.11± 0.40	1.01± 0.22	0.15± 0.02	0.43± 0.18	0.37± 0.09	0.70± 0.13	1.44± 0.71	8.75± 1.15	12	0.06± 0.04
SR5	4.18± 1.18	1.56± 0.57	1.28± 0.28	0.21± 0.03	1.24± 0.26	0.36± 0.16	0.76± 0.14	2.06± 0.77	11.66± 1.41	9	0.12± 0.08
SR6	4.64± 1.26	2.06± 0.48	1.47± 0.32	0.27± 0.04	0.46± 0.19	0.48± 0.11	0.71± 0.13	1.50± 0.62	11.58± 1.24	17	0.17± 0.16
SR7	4.55± 1.22	2.13± 0.52	1.63± 0.35	0.32± 0.04	0.46± 0.15	0.29± 0.06	0.54± 0.10	1.83± 0.76	11.75± 1.23	8	0.27± 0.17
SR8	4.23± 1.12	1.79± 0.49	1.43± 0.31	0.30± 0.05	0.57± 0.13	0.13± 0.08	0.48± 0.09	1.26± 0.57	10.20± 1.11	7	0.36± 0.21
SR9	5.00± 1.48	2.30± 0.56	2.00± 0.45	0.52± 0.06	0.37± 0.17	0.32± 0.07	0.46± 0.08	2.24± 1.07	13.22± 1.68	15	0.66± 0.39
SR10	4.69± 1.32	2.64± 0.62	2.18± 0.49	0.49± 0.06	0.31± 0.06	0.32± 0.06	0.37± 0.07	1.36± 0.64	12.37± 1.37	20	1.14± 0.67
SR11	2.59± 0.74	1.88± 0.46	1.40± 0.34	0.42± 0.06	0.40± 0.09	0.31± 0.07	0.22± 0.04	0.76± 0.43	7.98± 0.96	7	1.30± 0.76
SR12	1.62± 0.52	1.00± 0.27	0.93± 0.22	0.21± 0.03	0.11± 0.02	0.12± 0.03	0.10± 0.02	1.30± 0.65	5.39± 0.84	12	1.41± 0.80
SR13	0.94± 0.29	0.45± 0.13	0.56± 0.17	0.14± 0.02	0.00± 0.02	0.08± 0.01	0.07± 0.01	0.54± 0.33	2.78± 0.48	1	1.41± 0.83
SR14	0.45± 0.18	0.33± 0.10	0.26± 0.08	0.08± 0.01	0.00± 0.02	0.05± 0.01	0.04± 0.01	0.00± 0.11	1.22± 0.30	3	1.35± 0.76
SR15	0.19± 0.09	0.15± 0.06	0.14± 0.05	0.05± 0.01	0.06± 0.05	0.02± 0.00	0.02± 0.00	0.18± 0.12	0.82± 0.22	1	1.09± 0.60
SR16	0.07± 0.04	0.04± 0.03	0.03± 0.01	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.03	0.16± 0.08	1	0.62± 0.36
SR17	0.01± 0.01	0.01± 0.02	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.08	0.02± 0.08	0	0.07± 0.06

F.3 Nuisances

Four sets of nuisance pull values, expected and observed for cut-based and BDT analyses, are tabulated in Table 40 (expected cut-based analysis), Table 41 (observed cut-based analysis), Table 42 (expected BDT analysis), Table 43 (observed BDT analysis).

Table 40: Expected nuisance pulls for the cut-based analysis. The final column indicates the correlation between the nuisance and the signal strength

	b-only fit	$s+b$ fit	
name	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	+0.14, 0.99	+0.00, 0.98	-0.08
TTVV	+0.02, 1.00	+0.00, 0.99	-0.01
TTWSF	**+0.16, 0.76**	**+0.00, 0.78**	-0.14
TTZSF	**+0.06, 0.69**	**+0.00, 0.69**	-0.05
XG	+0.01, 0.99	+0.00, 0.99	-0.00
alphas	+0.00, 0.99	-0.00, 0.99	+0.00
bb	+0.25, 0.96	+0.00, 0.99	-0.13
btaghf	+0.07, 0.99	+0.00, 0.99	-0.07
btaglf	+0.05, 0.99	+0.00, 0.99	-0.03
fakes	+0.09, 0.97	+0.00, 0.97	-0.07
fakes_EWK	+0.07, 0.96	+0.00, 0.96	-0.05
flips	-0.01, 0.99	+0.00, 0.99	+0.00
fsrvar	+0.00, 0.99	-0.00, 0.99	+0.04
isr	-0.14, 0.98	-0.00, 0.99	+0.09
isrvar	+0.00, 0.99	+0.00, 0.99	-0.03
jer	-0.09, 0.97	-0.00, 0.99	+0.05
jes	+0.10, 0.91	+0.00, 0.99	-0.10
lep	+0.02, 0.99	+0.00, 0.99	-0.03
lumi	+0.04, 0.99	+0.00, 0.99	-0.02
pdf	+0.09, 0.96	-0.00, 0.99	-0.05
prefire	-0.01, 0.99	-0.00, 0.99	+0.01
pu	+0.00, 0.99	-0.00, 0.99	+0.02
rares	-0.00, 0.99	+0.00, 0.99	-0.00
scale	-0.11, 1.00	+0.00, 0.99	+0.04
trig	+0.03, 0.99	+0.00, 0.99	-0.04

Table 41: Observed nuisance pulls for the cut-based analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	**+0.32, 1.02**	+0.23, 1.01	-0.08
TTVV	+0.04, 1.00	+0.03, 1.00	-0.01
TTWSF	**+1.09, 0.73**	**+0.97, 0.74**	-0.17
TTZSF	**+1.23, 0.61**	**+1.18, 0.62**	-0.06
XG	+0.09, 1.00	+0.07, 1.00	-0.01
alphas	-0.00, 0.99	+0.00, 0.99	+0.00
bb	+0.29, 0.97	+0.17, 0.98	-0.11
btaghf	**+0.33, 0.99**	+0.27, 0.99	-0.07
btaglf	+0.14, 0.99	+0.09, 0.99	-0.04
fakes	+0.25, 1.02	+0.21, 1.01	-0.04
fakes_EWK	+0.17, 0.98	+0.14, 0.95	-0.02
flips	+0.07, 1.00	+0.08, 1.00	+0.01
fsrvar	-0.00, 0.99	**-0.03, 1.15**	+0.00
isr	-0.24, 0.97	-0.15, 0.98	+0.09
isrvar	-0.00, 0.99	+0.06, 1.01	+0.02
jer	**-0.05, 0.52**	**-0.01, 0.50**	+0.04
jes	**+0.13, 0.66**	**+0.07, 0.62**	-0.09
lep	+0.16, 0.99	+0.14, 0.99	-0.03
lumi	+0.25, 0.99	+0.22, 0.99	-0.03
pdf	+0.23, 0.94	+0.10, 0.97	-0.12
prefire	-0.04, 0.99	-0.04, 0.99	+0.01
pu	-0.11, 0.98	-0.10, 0.99	+0.01
rares	+0.11, 1.00	+0.12, 1.00	+0.00
scale	+0.01, 0.96	+0.07, 0.95	+0.04
trig	+0.21, 0.99	+0.19, 0.99	-0.04

Table 42: Expected nuisance pulls for the BDT analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	+0.12, 0.99	-0.00, 0.98	-0.08
TTVV	+0.02, 1.00	+0.00, 0.99	-0.01
TTWSF	**+0.20, 0.76**	**-0.00, 0.78**	-0.16
TTZSF	**+0.04, 0.69**	**+0.00, 0.69**	-0.05
XG	-0.00, 0.99	+0.00, 0.99	+0.00
alphas	+0.00, 0.99	-0.00, 0.99	+0.00
bb	**+0.33, 0.95**	+0.00, 0.99	-0.16
btaghf	+0.07, 0.99	+0.00, 0.99	-0.07
btaglf	+0.05, 0.99	+0.00, 0.99	-0.03
fakes	+0.07, 0.97	-0.00, 0.96	-0.06
fakes_EWK	+0.06, 0.94	+0.00, 0.95	-0.06
flips	-0.00, 0.99	-0.00, 0.99	+0.00
fsrvar	+0.00, 0.99	-0.00, 0.99	+0.02
isr	-0.17, 0.97	-0.00, 0.99	+0.10
isrvar	+0.00, 0.99	-0.00, 0.99	-0.03
jer	-0.07, 1.07	+0.00, 0.99	+0.04
jes	+0.11, 0.98	+0.00, 0.97	-0.09
lep	+0.02, 0.99	+0.00, 0.99	-0.03
lumi	+0.04, 0.99	+0.00, 0.99	-0.02
pdf	+0.09, 0.97	-0.00, 0.99	-0.05
prefire	-0.01, 0.99	-0.00, 0.99	+0.01
pu	+0.04, 0.99	-0.00, 0.99	-0.01
rares	+0.00, 0.99	+0.00, 0.99	-0.00
scale	-0.14, 1.01	+0.00, 0.99	+0.04
trig	+0.03, 0.99	+0.00, 0.99	-0.05

Table 43: Observed nuisance pulls for the BDT analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	**+0.33, 1.02**	+0.21, 1.01	-0.07
TTVV	+0.03, 1.00	+0.02, 1.00	-0.01
TTWSF	**+1.05, 0.74**	**+0.74, 0.78**	-0.20
TTZSF	**+1.28, 0.62**	**+1.24, 0.62**	-0.05
XG	-0.02, 0.99	-0.01, 0.99	+0.00
alphas	+0.00, 0.99	+0.00, 1.00	+0.00
bb	**+0.63, 0.95**	+0.18, 0.98	-0.16
btaghf	+0.25, 0.99	+0.16, 0.99	-0.08
btaglf	+0.12, 0.99	+0.04, 0.99	-0.04
fakes	**+0.30, 1.01**	+0.26, 1.00	-0.05
fakes_EWK	+0.24, 0.99	+0.21, 0.96	-0.04
flips	+0.05, 1.00	+0.06, 1.00	+0.01
fsrvar	+0.00, 0.99	-0.04, 1.01	+0.00
isr	**-0.44, 0.97**	-0.17, 0.98	+0.12
isrvar	+0.00, 0.99	+0.01, 1.01	-0.03
jer	-0.15, 0.93	**-0.03, 0.83**	+0.06
jes	**+0.08, 0.78**	**-0.02, 0.75**	-0.09
lep	+0.15, 0.99	+0.13, 0.99	-0.03
lumi	+0.23, 0.99	+0.18, 0.99	-0.02
pdf	+0.23, 0.95	+0.08, 0.97	-0.08
prefire	-0.05, 0.99	-0.04, 0.99	+0.02
pu	-0.18, 1.04	-0.23, 1.03	-0.00
rares	+0.10, 1.00	+0.11, 1.00	+0.00
scale	+0.00, 0.96	+0.13, 0.95	+0.02
trig	+0.19, 0.99	+0.16, 0.99	-0.05

F.4 Impacts

The leading 30 nuisance impacts for four sets of impacts, expected and observed for cut-based and BDT analyses, are shown in Figure 103 (expected cut-based analysis), Figure 104 (observed cut-based analysis), Figure 105 (expected BDT analysis), Figure 106 (observed BDT analysis). The most constrained nuisances correspond to normalization parameters for ttW and ttZ; as we would expect from the control regions, "TTWSF" and "TTZSF", are moved by approximately 1σ with respect to the input nuisance sizes.

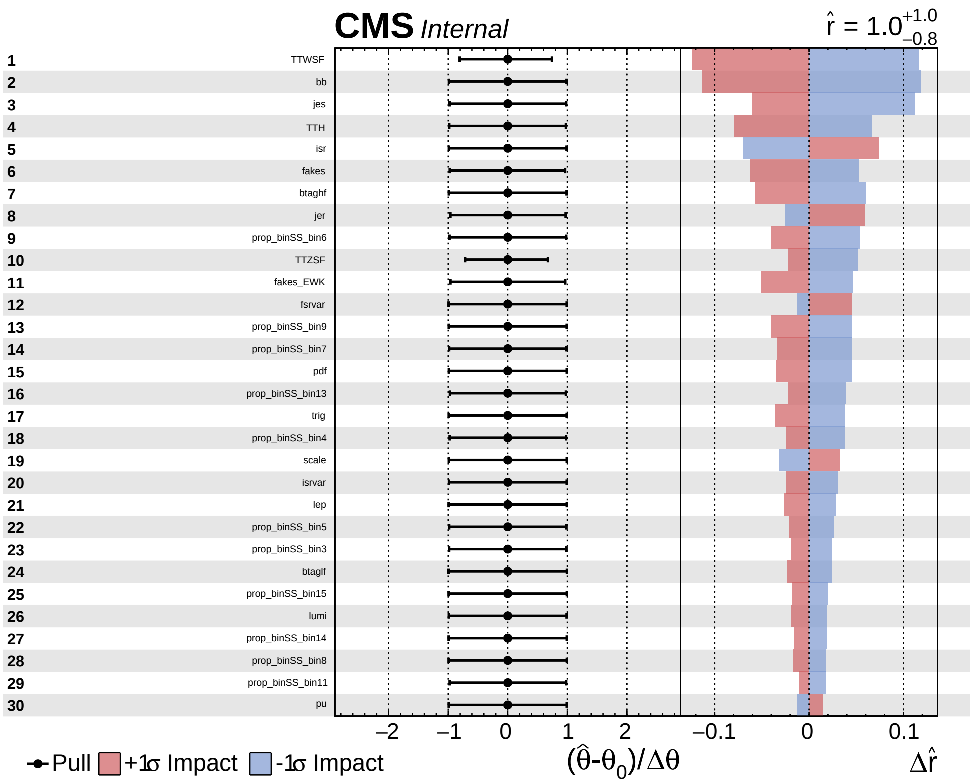


Figure 103: Expected nuisance impacts for the cut-based analysis.

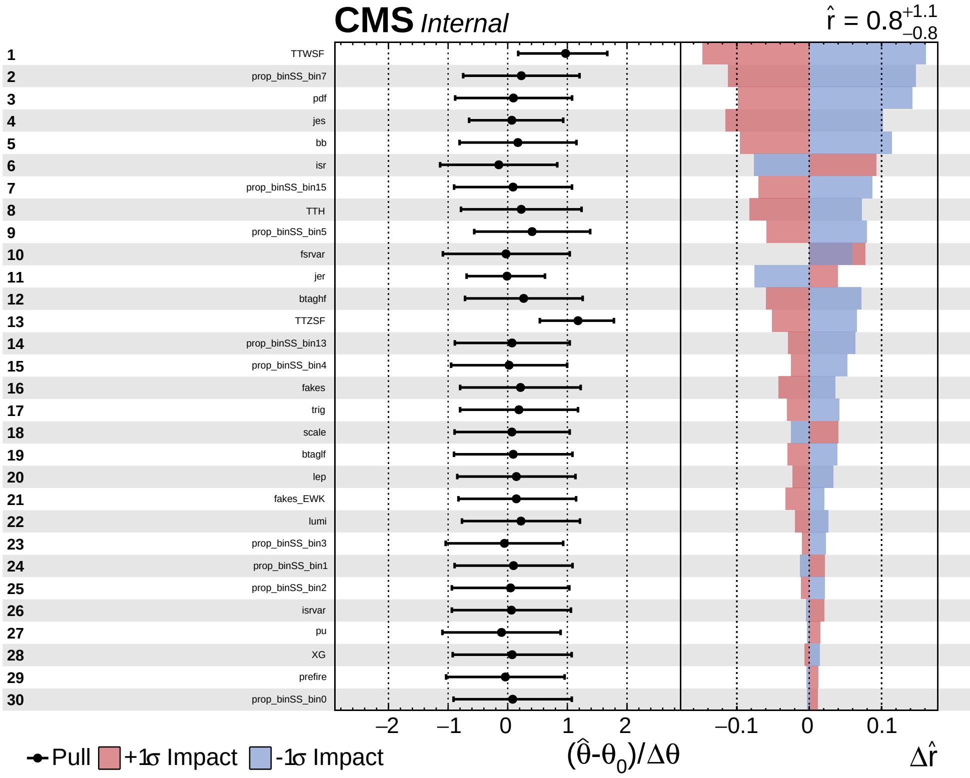


Figure 104: Observed nuisance impacts for the cut-based analysis.

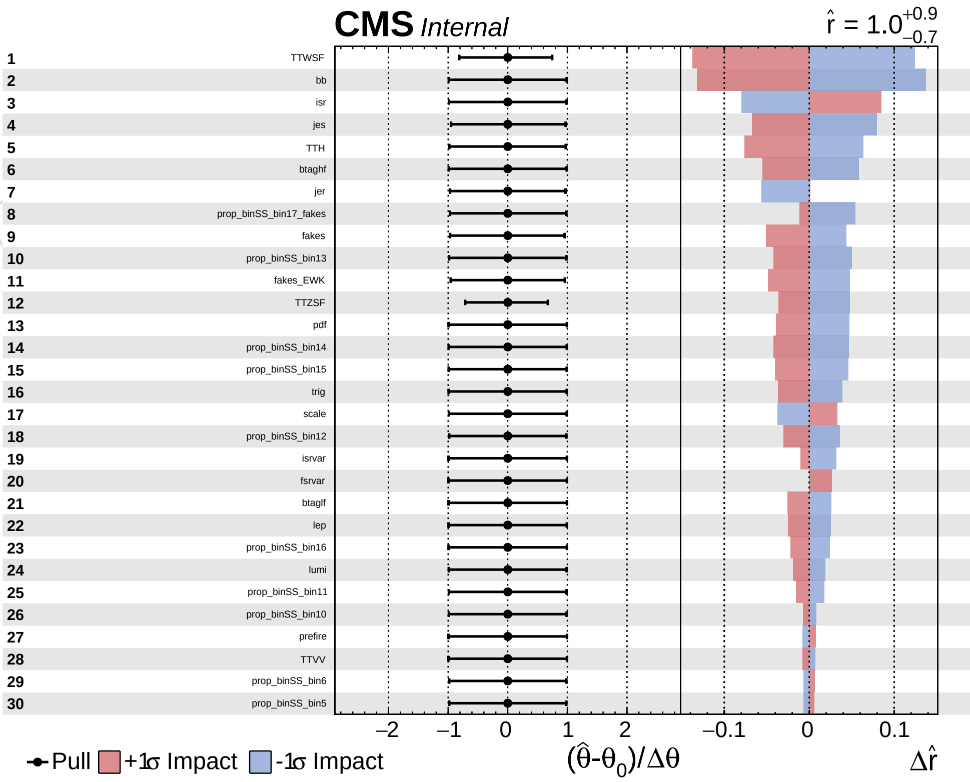


Figure 105: Expected nuisance impacts for the BDT-based analysis.

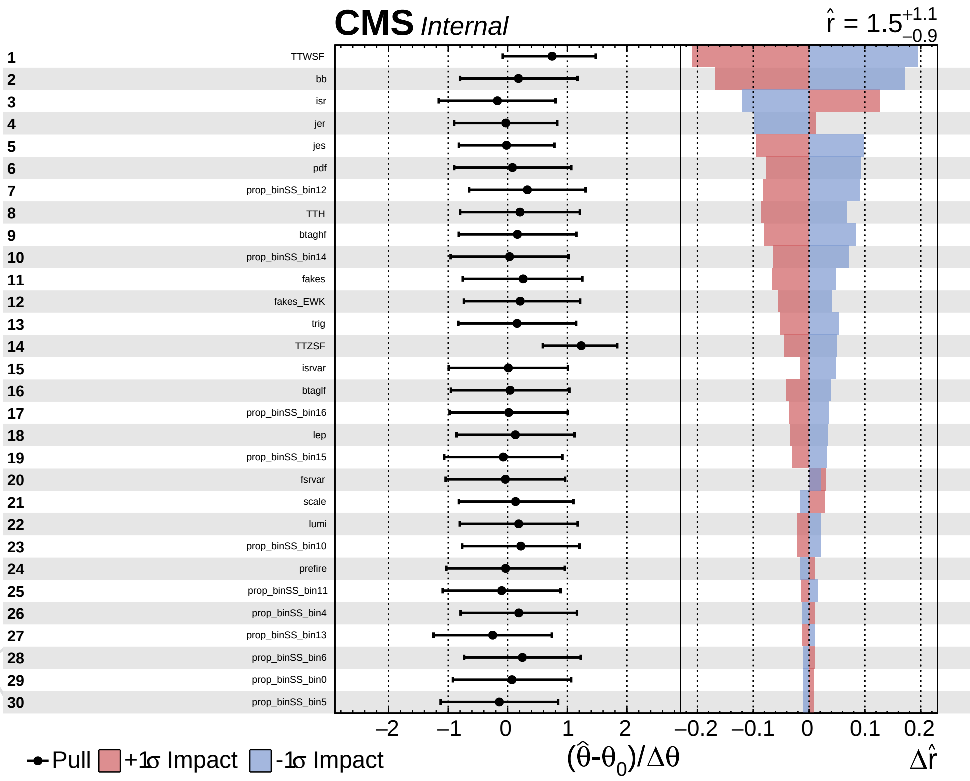


Figure 106: Observed nuisance impacts for the BDT-based analysis.

G Unblinding of 2017 dataset

Following the 2016 unblinding, the analysis was unblinded using the 2017 dataset. The yields and results are shown in Section G.1 with consistency checks of cut-based and BDT results in Section G.2. We include the results of various statistical checks, including nuisance pulls (Section G.3), nuisance impacts (Section G.4), and goodness of fits (Section G.5). Additionally, Section G.6 shows information about the combination of unblinded 2017 with unblinded 2016 data.

G.1 Yields and results

Plots for the unblinded prefit and postfit event yields for the 2017 data, with a total luminosity of 41.5 fb^{-1}, are shown for both the cut-based and BDT based analysis in Figure 107. Numerical yields are also tabulated in Table 44 (prefit cut-based analysis), Table 45 (postfit cut-based analysis), Table 46 (prefit BDT analysis), Table 47 (postfit BDT analysis).

A binned likelihood fit is performed over signal regions using the Higgs Combine tool [22]; exclusion limits at 95% CL are calculated with the Asymptotic CLs method.

With the cut-based analysis, an observed (expected) upper limit on the production cross section of 30.18 fb ($17.77^{+9.63}_{-5.93}$ fb), assuming the signal process does not exist. The observed (expected) significance is 0.853 (1.543) standard deviations, corresponding to a measured signal strength parameter of $0.703^{+0.985}_{-0.703}$.

With the BDT analysis, an observed (expected) upper limit on the production cross section of 29.85 fb ($15.94^{+8.71}_{-5.36}$ fb), assuming the signal process does not exist. The observed (expected) significance is 1.105 (1.756) standard deviations, corresponding to a measured signal strength parameter of $0.806^{+0.924}_{-0.736}$.

G.2 Cut-based and BDT consistency checks

In 2016, cut-based SR4 had a high data yield with respect to the SM prediction. In 2017, this region is in agreement. The situation is reversed for cut-based SR3, for example.

For the 2016 unblinding, we quantified the mildness of the disagreement between cut-based and BDT significances (both had differing directions for observed significance with respect to their expected significance), using fits of toy pseudodatasets correlated between cut-based and BDT regions.

Although 2017 shows much better agreement between the two fits, we repeat the same procedure for 2017. The observed values lie within a 0.3-σ contour, with a p-value of nearly 0.8, as shown in Figure 108.

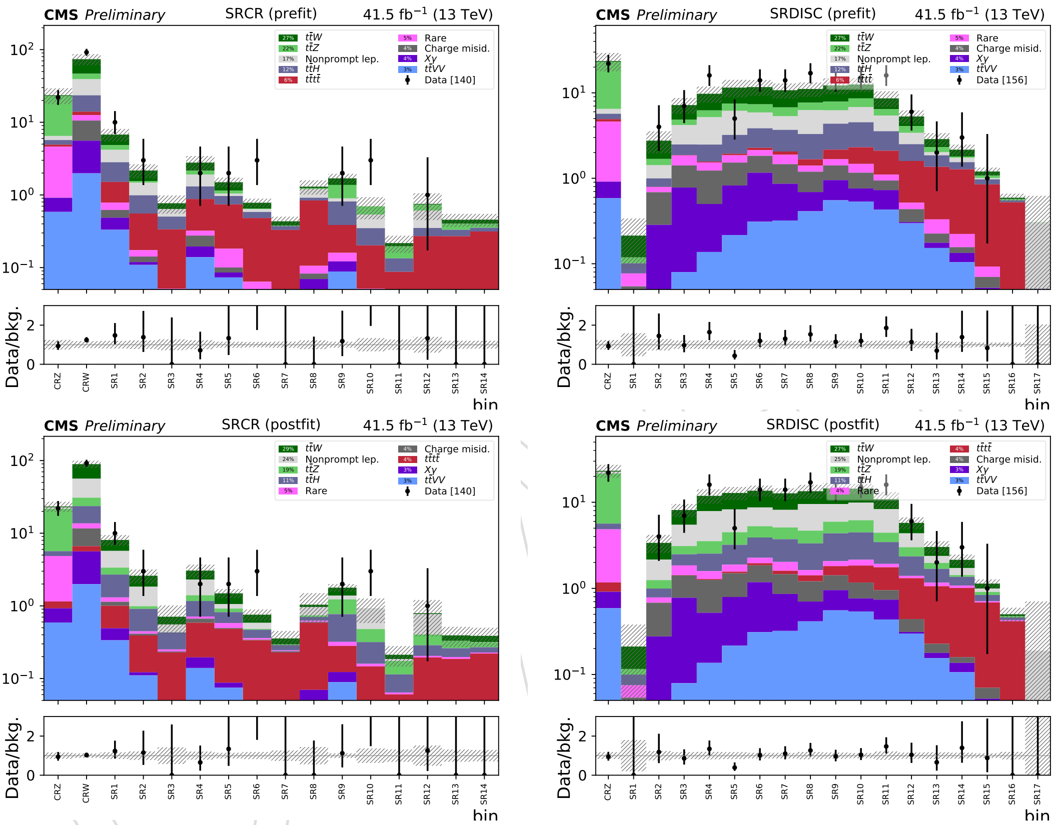


Figure 107: 2017: data yields compared to prefit (top) and postfit (bottom) SM predictions in CRZ, CRW, and the cut-based signal regions (left), and CRZ with the BDT signal regions (right). Note that we plan to retrain the BDT with latest JECs/corrections for 2018 (with identical settings)

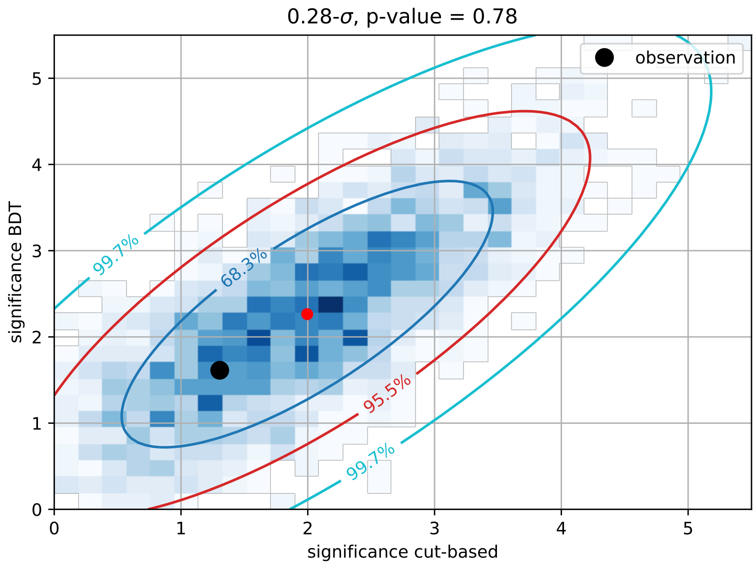


Figure 108: Comparison of correlated expected significances between cut-based and BDT fits with observed point overlaid. Because significances from toy datasets are used, we get a correction factor for expected BDT/cut-based significances with the full fit and the mean of the significances from the toy datasets. This correction factor is then applied to the observed point such that the expected significance values from the full analysis would match the mean values from the toy pseudo-datasets.

Table 44: Prefit event yields in cut-based regions for 2017.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.53± 0.19	16.56± 5.59	0.77± 0.21	0.58± 0.07	0.33± 0.13	3.69± 0.78	0.00± 0.00	0.79± 0.42	23.26± 5.81	22	0.32± 0.02
CRW	27.27± 9.20	7.44± 2.50	9.52± 2.48	1.98± 0.25	3.60± 0.50	2.01± 0.41	4.98± 0.95	15.64± 8.38	72.45±12.99	92	1.43± 0.10
SR1	1.92± 0.66	0.68± 0.24	1.31± 0.35	0.33± 0.04	0.15± 0.04	0.16± 0.05	0.14± 0.03	1.36± 1.04	6.05± 1.33	10	0.73± 0.06
SR2	0.62± 0.25	0.08± 0.07	0.43± 0.12	0.11± 0.01	0.01± 0.01	0.03± 0.01	0.02± 0.00	0.48± 0.37	1.79± 0.47	3	0.38± 0.03
SR3	0.13± 0.07	0.00± 0.02	0.17± 0.06	0.03± 0.01	0.00± 0.00	0.02± 0.00	0.01± 0.00	0.14± 0.16	0.48± 0.21	0	0.28± 0.05
SR4	0.62± 0.24	0.25± 0.12	0.43± 0.13	0.14± 0.02	0.06± 0.01	0.05± 0.01	0.08± 0.02	0.60± 0.50	2.23± 0.65	2	0.56± 0.07
SR5	0.35± 0.18	0.10± 0.04	0.23± 0.07	0.07± 0.01	0.01± 0.01	0.08± 0.02	0.01± 0.00	0.08± 0.09	0.94± 0.23	2	0.56± 0.04
SR6	0.13± 0.06	0.01± 0.03	0.10± 0.03	0.03± 0.00	0.01± 0.00	0.02± 0.00	0.01± 0.00	0.05± 0.06	0.36± 0.11	3	0.42± 0.04
SR7	0.05± 0.03	0.01± 0.01	0.04± 0.02	0.01± 0.00	0.01± 0.00	0.01± 0.00	0.01± 0.00	0.00± 0.04	0.14± 0.07	0	0.30± 0.03
SR8	0.06± 0.04	0.04± 0.02	0.08± 0.03	0.05± 0.01	0.02± 0.01	0.02± 0.00	0.01± 0.00	0.28± 0.27	0.57± 0.29	0	0.73± 0.08
SR9	0.32± 0.13	0.48± 0.16	0.43± 0.11	0.09± 0.01	0.03± 0.00	0.04± 0.01	0.00± 0.00	0.08± 0.07	1.47± 0.26	2	0.23± 0.02
SR10	0.00± 0.03	0.16± 0.07	0.15± 0.05	0.03± 0.00	0.01± 0.01	0.01± 0.00	0.00± 0.00	0.19± 0.21	0.55± 0.24	3	0.15± 0.01
SR11	0.02± 0.01	0.06± 0.03	0.05± 0.02	0.01± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.03	0.14± 0.06	0	0.08± 0.01
SR12	0.02± 0.01	0.12± 0.05	0.08± 0.03	0.03± 0.00	0.01± 0.00	0.01± 0.00	0.00± 0.00	0.26± 0.30	0.53± 0.30	1	0.23± 0.03
SR13	0.05± 0.03	0.07± 0.03	0.06± 0.02	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.07	0.21± 0.09	0	0.25± 0.02
SR14	0.06± 0.03	0.05± 0.03	0.04± 0.02	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.07	0.16± 0.09	0	0.29± 0.03

Table 45: Postfit event yields in cut-based regions for 2017.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.65± 0.20	16.04± 3.53	0.81± 0.20	0.59± 0.08	0.34± 0.13	3.70± 0.77	0.00± 0.00	0.98± 0.72	23.11± 3.77	22	0.22± 0.23
CRW	32.98± 9.73	7.23± 1.61	9.98± 2.40	2.00± 0.26	3.63± 0.51	2.01± 0.41	5.00± 0.84	25.44±12.71	88.27±10.26	92	1.01± 1.04
SR1	2.36± 0.74	0.67± 0.15	1.39± 0.33	0.34± 0.04	0.16± 0.04	0.17± 0.04	0.14± 0.02	2.35± 1.56	7.56± 1.37	10	0.51± 0.53
SR2	0.76± 0.29	0.08± 0.06	0.46± 0.11	0.11± 0.01	0.01± 0.01	0.03± 0.01	0.02± 0.00	0.85± 0.59	2.33± 0.57	3	0.27± 0.29
SR3	0.15± 0.07	0.00± 0.01	0.18± 0.05	0.03± 0.01	0.00± 0.00	0.02± 0.00	0.01± 0.00	0.13± 0.27	0.51± 0.30	0	0.20± 0.19
SR4	0.75± 0.25	0.24± 0.07	0.45± 0.12	0.14± 0.02	0.06± 0.01	0.05± 0.01	0.08± 0.01	0.91± 0.63	2.67± 0.64	2	0.39± 0.39
SR5	0.43± 0.18	0.10± 0.03	0.24± 0.07	0.07± 0.01	0.01± 0.00	0.08± 0.02	0.01± 0.00	0.14± 0.12	1.09± 0.22	2	0.39± 0.41
SR6	0.17± 0.07	0.01± 0.03	0.11± 0.03	0.03± 0.00	0.01± 0.00	0.02± 0.00	0.01± 0.00	0.10± 0.11	0.46± 0.13	3	0.29± 0.28
SR7	0.07± 0.03	0.01± 0.01	0.04± 0.02	0.01± 0.00	0.01± 0.00	0.01± 0.00	0.01± 0.00	0.00± 0.08	0.15± 0.10	0	0.21± 0.22
SR8	0.07± 0.05	0.04± 0.01	0.08± 0.03	0.05± 0.01	0.02± 0.01	0.02± 0.00	0.01± 0.00	0.23± 0.45	0.52± 0.46	0	0.52± 0.52
SR9	0.40± 0.15	0.46± 0.11	0.45± 0.11	0.09± 0.01	0.03± 0.01	0.04± 0.01	0.00± 0.00	0.15± 0.12	1.62± 0.22	2	0.16± 0.16
SR10	0.00± 0.03	0.16± 0.05	0.16± 0.04	0.03± 0.00	0.01± 0.00	0.01± 0.00	0.00± 0.00	0.44± 0.35	0.81± 0.34	3	0.11± 0.11
SR11	0.03± 0.01	0.06± 0.02	0.05± 0.01	0.01± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.01± 0.06	0.16± 0.07	0	0.05± 0.06
SR12	0.02± 0.01	0.12± 0.03	0.08± 0.03	0.03± 0.00	0.01± 0.00	0.01± 0.00	0.00± 0.00	0.37± 0.41	0.63± 0.41	1	0.16± 0.16
SR13	0.07± 0.03	0.07± 0.02	0.06± 0.02	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.11	0.22± 0.12	0	0.17± 0.17
SR14	0.07± 0.03	0.05± 0.02	0.04± 0.02	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.10	0.18± 0.11	0	0.21± 0.21

Table 46: Prefit event yields in BDT regions for 2017.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.53± 0.21	16.56± 5.54	0.77± 0.16	0.58± 0.08	0.33± 0.15	3.69± 0.76	0.00± 0.00	0.79± 0.44	23.26± 5.72	22	0.32± 0.02
SR1	0.09± 0.05	0.02± 0.01	0.02± 0.01	0.00± 0.00	0.01± 0.01	0.02± 0.03	0.05± 0.01	0.00± 0.09	0.21± 0.13	0	0.00± 0.00
SR2	1.07± 0.46	0.26± 0.15	0.20± 0.06	0.03± 0.00	0.26± 0.12	0.10± 0.06	0.40± 0.08	0.42± 0.53	2.75± 0.80	4	0.00± 0.00
SR3	2.35± 1.01	0.65± 0.31	0.62± 0.14	0.08± 0.02	0.70± 0.24	0.44± 0.17	0.63± 0.12	1.71± 1.01	7.18± 1.52	7	0.01± 0.01
SR4	3.52± 1.38	1.05± 0.53	0.91± 0.21	0.14± 0.03	0.36± 0.20	0.30± 0.14	0.73± 0.14	2.70± 1.31	9.71± 2.10	16	0.04± 0.01
SR5	4.02± 1.69	1.34± 0.60	1.28± 0.29	0.22± 0.03	0.61± 0.19	0.33± 0.17	0.70± 0.13	2.93± 1.88	11.44± 2.67	5	0.08± 0.01
SR6	4.33± 1.73	1.45± 0.64	1.58± 0.34	0.31± 0.04	0.85± 0.21	0.30± 0.11	0.67± 0.13	2.05± 1.52	11.54± 2.46	14	0.14± 0.02
SR7	3.97± 1.63	1.44± 0.65	1.63± 0.35	0.32± 0.05	0.54± 0.08	0.42± 0.15	0.60± 0.11	1.63± 1.41	10.56± 2.17	14	0.17± 0.03
SR8	3.42± 1.37	1.41± 0.57	1.63± 0.36	0.41± 0.05	0.28± 0.09	0.21± 0.05	0.50± 0.10	2.97± 1.60	10.85± 2.13	17	0.26± 0.02
SR9	4.33± 1.74	1.65± 0.73	2.19± 0.46	0.55± 0.06	0.40± 0.09	0.26± 0.06	0.46± 0.09	1.88± 1.83	11.73± 2.57	14	0.49± 0.06
SR10	3.92± 1.59	1.78± 0.67	2.22± 0.49	0.53± 0.07	0.23± 0.10	0.31± 0.07	0.40± 0.08	2.30± 1.48	11.70± 2.39	15	0.83± 0.05
SR11	2.18± 1.00	1.01± 0.45	1.45± 0.33	0.43± 0.05	0.30± 0.08	0.16± 0.04	0.21± 0.04	1.87± 1.45	7.62± 1.82	16	1.00± 0.04
SR12	1.21± 0.55	0.68± 0.25	0.91± 0.23	0.30± 0.04	0.01± 0.04	0.08± 0.02	0.13± 0.02	0.89± 0.60	4.21± 0.95	6	1.08± 0.05
SR13	0.56± 0.30	0.30± 0.14	0.49± 0.13	0.15± 0.02	0.02± 0.01	0.10± 0.02	0.05± 0.01	0.16± 0.28	1.84± 0.51	2	1.03± 0.05
SR14	0.36± 0.21	0.06± 0.04	0.27± 0.08	0.10± 0.02	0.03± 0.01	0.07± 0.01	0.02± 0.00	0.19± 0.17	1.11± 0.32	3	1.05± 0.05
SR15	0.12± 0.07	0.07± 0.04	0.12± 0.04	0.05± 0.01	0.00± 0.00	0.02± 0.01	0.02± 0.00	0.04± 0.06	0.45± 0.13	1	0.76± 0.06
SR16	0.02± 0.02	0.02± 0.01	0.03± 0.01	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.05	0.10± 0.07	0	0.49± 0.03
SR17	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.25± 0.32	0.26± 0.32	0	0.05± 0.01

Table 47: Postfit event yields in BDT regions for 2017.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.61± 0.20	15.95± 3.76	0.80± 0.18	0.59± 0.07	0.32± 0.12	3.68± 0.77	0.00± 0.00	1.00± 0.74	22.96± 4.00	22	0.26± 0.22
SR1	0.11± 0.04	0.02± 0.01	0.02± 0.01	0.00± 0.00	0.01± 0.01	0.02± 0.03	0.05± 0.01	0.00± 0.15	0.21± 0.17	0	0.00± 0.00
SR2	1.22± 0.43	0.25± 0.10	0.21± 0.07	0.03± 0.00	0.25± 0.11	0.10± 0.07	0.40± 0.07	0.91± 0.78	3.37± 0.81	4	0.00± 0.00
SR3	2.66± 0.95	0.61± 0.21	0.64± 0.17	0.08± 0.02	0.70± 0.24	0.42± 0.15	0.63± 0.11	2.38± 1.37	8.13± 1.43	7	0.01± 0.02
SR4	4.00± 1.31	1.00± 0.36	0.95± 0.26	0.14± 0.03	0.38± 0.18	0.30± 0.16	0.73± 0.13	4.35± 2.37	11.85± 2.25	16	0.03± 0.03
SR5	4.56± 1.57	1.28± 0.38	1.32± 0.35	0.22± 0.03	0.58± 0.15	0.32± 0.20	0.70± 0.12	3.78± 1.98	12.76± 2.24	5	0.07± 0.07
SR6	4.92± 1.61	1.38± 0.39	1.63± 0.40	0.31± 0.04	0.87± 0.18	0.30± 0.13	0.67± 0.11	3.44± 2.08	13.52± 2.00	14	0.11± 0.11
SR7	4.54± 1.52	1.37± 0.39	1.70± 0.41	0.32± 0.05	0.54± 0.09	0.43± 0.17	0.60± 0.10	3.07± 2.10	12.56± 1.85	14	0.14± 0.15
SR8	3.90± 1.28	1.36± 0.34	1.70± 0.43	0.41± 0.05	0.29± 0.10	0.21± 0.05	0.51± 0.09	4.86± 2.79	13.24± 2.49	17	0.21± 0.19
SR9	4.95± 1.59	1.59± 0.45	2.29± 0.53	0.56± 0.06	0.40± 0.10	0.26± 0.08	0.46± 0.08	3.58± 2.60	14.09± 2.31	14	0.40± 0.32
SR10	4.51± 1.47	1.72± 0.42	2.32± 0.57	0.54± 0.07	0.23± 0.10	0.31± 0.08	0.40± 0.07	3.71± 2.04	13.74± 1.97	15	0.67± 0.63
SR11	2.54± 0.93	0.98± 0.29	1.52± 0.38	0.43± 0.05	0.31± 0.07	0.17± 0.04	0.21± 0.04	3.89± 2.56	10.06± 2.31	16	0.81± 0.72
SR12	1.41± 0.48	0.65± 0.16	0.96± 0.25	0.30± 0.03	0.01± 0.03	0.08± 0.02	0.13± 0.02	1.35± 0.84	4.89± 0.92	6	0.87± 0.76
SR13	0.64± 0.26	0.29± 0.09	0.52± 0.14	0.15± 0.02	0.02± 0.01	0.10± 0.02	0.05± 0.01	0.41± 0.45	2.19± 0.51	2	0.83± 0.73
SR14	0.42± 0.16	0.07± 0.04	0.28± 0.09	0.11± 0.02	0.03± 0.01	0.07± 0.01	0.02± 0.00	0.30± 0.30	1.29± 0.32	3	0.85± 0.74
SR15	0.14± 0.06	0.07± 0.03	0.13± 0.05	0.05± 0.01	0.00± 0.00	0.02± 0.01	0.02± 0.00	0.09± 0.11	0.52± 0.14	1	0.61± 0.53
SR16	0.03± 0.02	0.02± 0.01	0.03± 0.01	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.08	0.10± 0.10	0	0.40± 0.35
SR17	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.14± 0.52	0.15± 0.52	0	0.04± 0.03

G.3 Nuisances

Four sets of nuisance pull values, expected and observed for cut-based and BDT analyses, are tabulated in Table 48 (expected cut-based analysis), Table 49 (observed cut-based analysis), Table 50 (expected BDT analysis), Table 51 (observed BDT analysis).

Table 48: Expected nuisance pulls for the cut-based analysis. The final column indicates the correlation between the nuisance and the signal strength

	b-only fit	$s+b$ fit	
name	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta, \mu)$
TTH	+0.15, 0.99	-0.00, 0.98	-0.08
TTVV	+0.02, 1.00	-0.00, 0.99	-0.01
TTWSF	**+0.13, 0.80**	**-0.00, 0.81**	-0.08
TTZSF	**+0.05, 0.67**	**-0.00, 0.67**	-0.03
XG	-0.00, 0.99	-0.00, 0.99	+0.00
alphas	+0.00, 0.99	-0.00, 0.99	+0.00
bb	+0.25, 0.96	-0.00, 0.99	-0.12
btaghf	+0.10, 0.98	-0.00, 0.98	-0.08
btaglf	+0.05, 0.99	-0.00, 0.99	-0.03
fakes	+0.14, 0.95	-0.00, 0.96	-0.07
fakes_EWK	**+0.10, 0.85**	**-0.00, 0.85**	-0.08
flips	-0.01, 0.99	-0.00, 0.99	+0.00
fsrvar	+0.00, 0.99	-0.00, 0.99	+0.01
isr	+0.10, 0.98	+0.00, 0.99	-0.05
isrvar	+0.00, 0.99	+0.00, 0.99	-0.02
jer	-0.08, 0.95	+0.00, 0.99	+0.05
jes	+0.10, 0.92	+0.00, 0.99	-0.10
lep	+0.03, 0.99	-0.00, 0.99	-0.04
lumi	+0.03, 0.99	-0.00, 0.99	-0.02
pdf	+0.01, 1.01	+0.00, 0.99	+0.01
prefire	-0.02, 0.99	+0.00, 0.99	+0.02
pu	-0.02, 0.99	+0.00, 0.99	+0.01
rares	+0.01, 0.99	-0.00, 0.99	-0.00
scale	-0.05, 0.99	+0.00, 0.99	-0.00
trig	+0.03, 0.99	-0.00, 0.99	-0.04

Table 49: Observed nuisance pulls for the cut-based analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta, \mu)$
TTH	+0.26, 1.02	+0.20, 1.01	-0.07
TTVV	+0.03, 1.00	+0.02, 1.00	-0.01
TTWSF	**+0.59, 0.96**	**+0.56, 0.93**	-0.03
TTZSF	**-0.08, 0.71**	**-0.10, 0.69**	-0.03
XG	+0.01, 0.99	+0.01, 1.00	+0.00
alphas	-0.00, 0.99	+0.00, 0.99	+0.00
bb	+0.05, 0.98	-0.04, 0.99	-0.09
btaghf	+0.13, 0.98	+0.03, 0.98	-0.11
btaglf	+0.00, 0.99	-0.02, 0.99	-0.03
fakes	**+0.70, 0.97**	**+0.58, 0.98**	-0.13
fakes_EWK	**+1.00, 0.95**	**+0.84, 1.01**	-0.17
flips	+0.02, 0.99	+0.03, 1.00	+0.01
fsrvar	-0.00, 0.99	+0.02, 1.01	+0.03
isr	+0.21, 0.99	+0.13, 0.99	-0.09
isrvar	-0.00, 0.99	-0.01, 0.99	-0.02
jer	-0.15, 1.06	**-0.05, 0.82**	+0.08
jes	**+0.17, 0.80**	**+0.13, 0.81**	-0.07
lep	+0.05, 0.99	+0.04, 0.99	-0.02
lumi	+0.07, 0.99	+0.06, 0.99	-0.01
pdf	-0.02, 1.05	-0.02, 1.05	-0.00
prefire	-0.03, 0.99	-0.02, 0.99	+0.02
pu	+0.02, 1.01	-0.01, 1.01	-0.03
rares	-0.01, 0.99	-0.01, 0.99	-0.01
scale	-0.11, 1.05	-0.10, 1.07	-0.01
trig	+0.07, 0.99	+0.06, 0.99	-0.03

Table 50: Expected nuisance pulls for the BDT analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	+0.17, 0.99	+0.00, 0.98	-0.08
TTVV	+0.03, 1.00	+0.00, 0.99	-0.01
TTWSF	**+0.16, 0.80**	**+0.00, 0.82**	-0.09
TTZSF	**+0.04, 0.67**	**+0.00, 0.67**	-0.03
XG	-0.01, 0.99	+0.00, 0.99	+0.00
alphas	+0.00, 0.99	+0.00, 0.99	+0.00
bb	**+0.36, 0.95**	+0.00, 0.99	-0.14
btaghf	+0.11, 0.98	+0.00, 0.98	-0.09
btaglf	+0.05, 0.99	+0.00, 0.99	-0.03
fakes	+0.05, 0.94	+0.00, 0.94	-0.04
fakes_EWK	**+0.12, 0.86**	**+0.00, 0.82**	-0.10
flips	-0.01, 0.99	+0.00, 0.99	+0.00
fsrvar	+0.00, 0.99	+0.00, 0.99	+0.01
isr	+0.08, 0.99	+0.00, 0.99	-0.04
isrvar	+0.00, 0.99	+0.00, 0.99	-0.01
jer	**-0.10, 1.32**	+0.00, 0.99	+0.04
jes	+0.12, 0.97	+0.00, 0.97	-0.08
lep	+0.03, 0.99	+0.00, 0.99	-0.04
lumi	+0.04, 0.99	+0.00, 0.99	-0.02
pdf	+0.03, 1.01	+0.00, 0.99	-0.01
prefire	-0.02, 0.99	+0.00, 0.99	+0.02
pu	+0.05, 0.99	+0.00, 0.99	-0.02
rares	+0.02, 0.99	+0.00, 0.99	-0.00
scale	-0.02, 0.97	+0.00, 0.99	-0.01
trig	+0.04, 0.99	+0.00, 0.99	-0.05

Table 51: Observed nuisance pulls for the BDT analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta, \mu)$
TTH	+0.25, 1.02	+0.17, 1.01	-0.07
TTVV	+0.04, 1.00	+0.02, 1.00	-0.01
TTWSF	**+0.42, 0.98**	**+0.38, 0.95**	-0.03
TTZSF	**-0.11, 0.69**	**-0.13, 0.69**	-0.02
XG	-0.01, 0.99	+0.00, 0.99	+0.00
alphas	+0.00, 0.99	+0.00, 0.99	+0.00
bb	+0.23, 0.97	+0.04, 0.99	-0.14
btaghf	+0.10, 0.98	+0.03, 0.98	-0.09
btaglf	+0.06, 0.99	+0.02, 0.99	-0.04
fakes	**+0.75, 0.97**	**+0.64, 0.97**	-0.09
fakes_EWK	**+1.01, 0.88**	**+0.83, 0.91**	-0.17
flips	+0.00, 0.99	+0.01, 0.99	+0.01
fsrvar	+0.00, 0.99	-0.01, 0.99	-0.00
isr	+0.14, 0.99	+0.07, 0.99	-0.05
isrvar	+0.00, 0.99	-0.01, 0.99	-0.02
jer	**-0.09, 1.15**	-0.05, 1.06	+0.03
jes	+0.24, 1.01	**+0.17, 1.10**	-0.09
lep	+0.03, 0.99	+0.02, 0.99	-0.02
lumi	+0.05, 0.99	+0.04, 0.99	-0.01
pdf	+0.01, 1.03	-0.03, 1.03	-0.01
prefire	-0.01, 0.99	-0.01, 0.99	+0.02
pu	+0.08, 1.00	+0.04, 1.00	-0.03
rares	-0.02, 0.99	-0.03, 0.99	-0.01
scale	+0.03, 0.93	+0.04, 0.94	-0.01
trig	+0.05, 0.99	+0.04, 0.99	-0.03

G.4 Impacts

The leading 30 nuisance impacts for four sets of impacts, expected and observed for cut-based and BDT analyses, are shown in Figure 109 (expected cut-based analysis), Figure 110 (observed cut-based analysis), Figure 111 (expected BDT analysis), Figure 112 (observed BDT analysis). Note that statistical nuisances tend to show up in the impact plots because the binning for the cut-based and BDT analyses were optimized for the full luminosity (3.3 times what is shown here for 2017). As in 2016, the most constrained nuisances correspond to normalization parameters for ttW and ttZ, as we would expect from the control regions. "TTWSF" is moved by approximately 1σ with respect to the input nuisance sizes. Because the ttZ control region has better data, simulation agreement in 2017 compared to 2016, the "TTZSF" nuisance is not pulled up here.

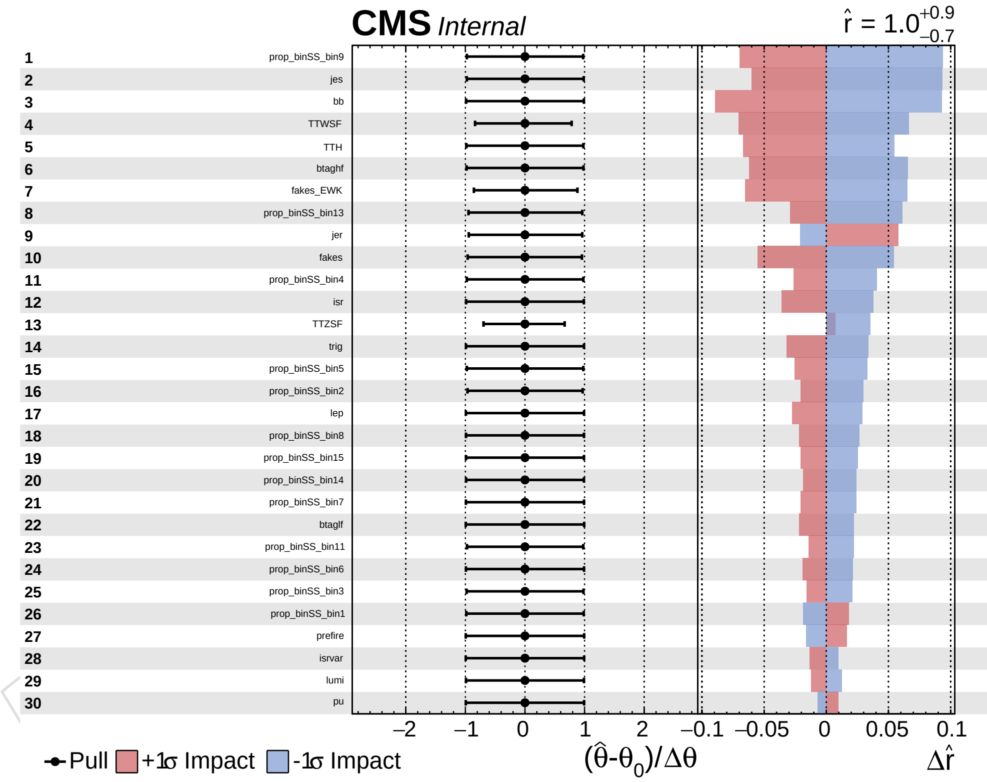


Figure 109: Expected nuisance impacts for the cut-based analysis.

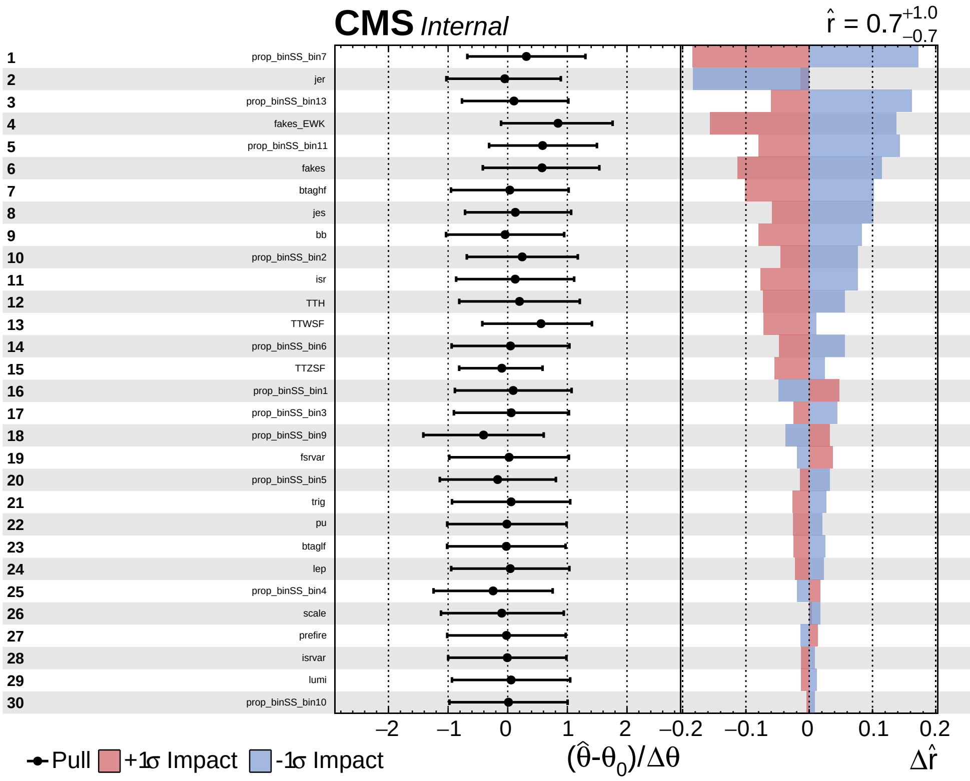


Figure 110: Observed nuisance impacts for the cut-based analysis.

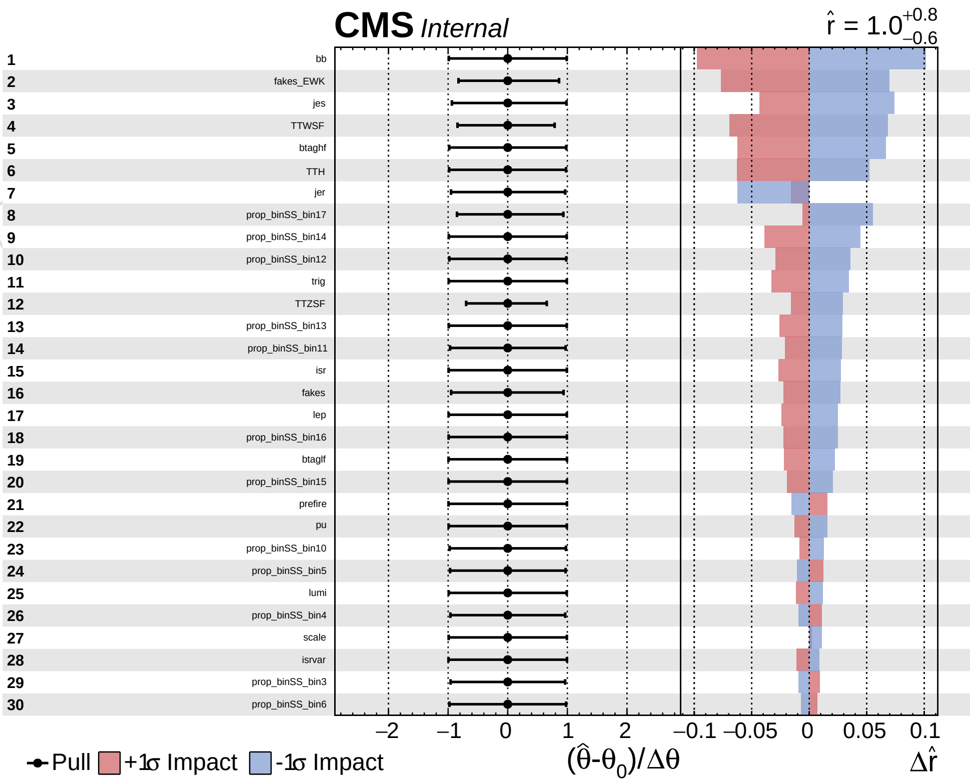


Figure 111: Expected nuisance impacts for the BDT-based analysis.

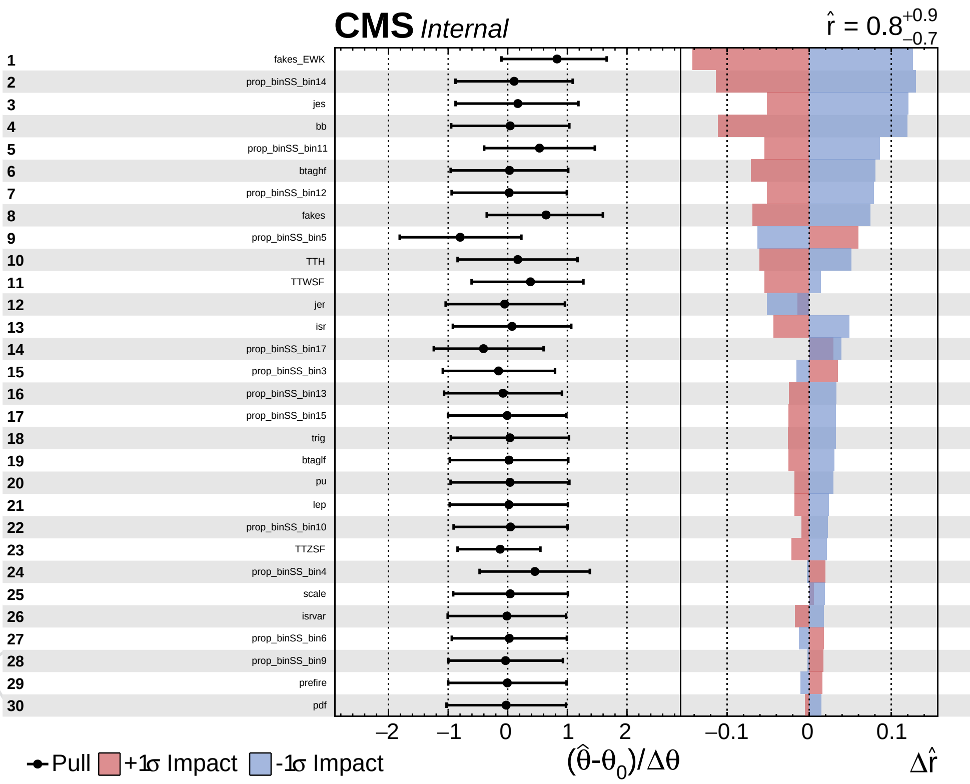


Figure 112: Observed nuisance impacts for the BDT-based analysis.

1743 G.5 Goodness of fits

1744 The goodness of fit distributions (using the saturated test statistic with the signal+background fit to data and asimov toys) for the cut-based and BDT analyses are shown in Figure 113.

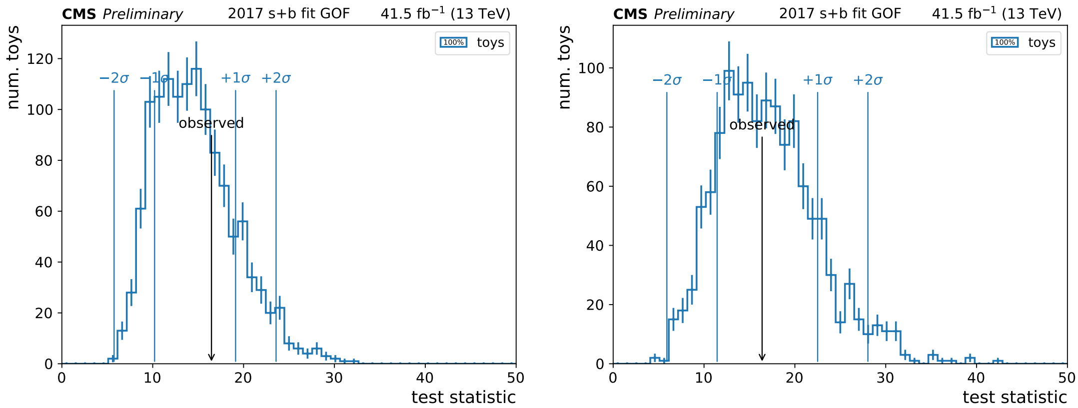


Figure 113: GOF test for the cut-based analysis (left) and BDT-based analysis (right)

1745

G.6 Combination with 2016

This subsection presents the results of the combination of the unblinded 2017 dataset from above with the 2016 dataset from Appendix F following the correlation model from Section 9.

G.6.1 Yields and results

Plots for the unblinded prefit and postfit event yields for the 2016+2017 data, with a total luminosity of 35.9 fb^{-1} +41.5 fb^{-1} =77.4 fb^{-1}, are shown for both the cut-based and BDT based analysis in Figure 114. Numerical yields are also tabulated in Table 52 (postfit cut-based analysis), Table 53 (postfit BDT analysis).

With the cut-based analysis, an observed (expected) upper limit on the production cross section of 23.15 fb ($12.34^{+6.30}_{-4.00}$ fb), assuming the signal process does not exist. The observed (expected) significance is 1.208 (2.074) standard deviations, corresponding to a measured signal strength parameter of $0.696^{+0.690}_{-0.584}$.

With the BDT analysis, an observed (expected) upper limit on the production cross section of 28.47 fb ($11.36^{+5.84}_{-3.71}$ fb), assuming the signal process does not exist. The observed (expected) significance is 2.221 (2.272) standard deviations, corresponding to a measured signal strength parameter of $1.167^{+0.697}_{-0.593}$.

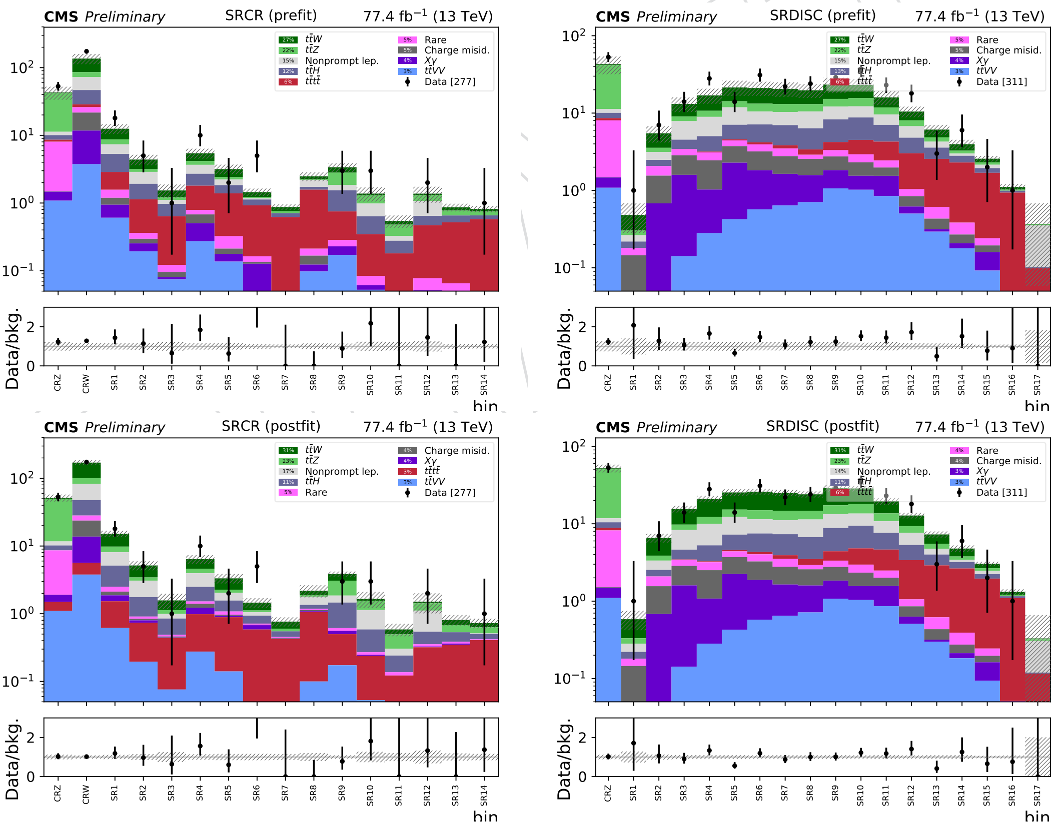


Figure 114: 2016+2017: data yields compared to prefit (top) and postfit (bottom) SM predictions in CRZ, CRW, and the cut-based signal regions (left), and CRZ with the BDT signal regions (right). Note that we plan to retrain the BDT with latest JECs/corrections for 2018 (with identical settings)

Table 52: Postfit event yields in cut-based regions for 2016+2017.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	$X+\gamma$	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	1.45± 0.36	38.30± 6.11	1.57± 0.38	1.10± 0.13	0.40± 0.09	6.71± 1.25	0.02± 0.00	1.50± 0.74	51.05± 6.13	53	0.39± 0.35
CRW	70.12±16.58	17.21± 2.75	19.31± 4.64	3.79± 0.45	8.14± 0.98	4.44± 0.83	9.98± 1.16	35.72±18.98	168.72±14.28	174	1.85± 1.56
SR1	5.38± 1.34	1.71± 0.28	2.63± 0.66	0.62± 0.07	0.34± 0.07	0.39± 0.08	0.25± 0.03	2.96± 1.89	14.28± 1.65	18	0.91± 0.79
SR2	1.72± 0.45	0.35± 0.08	0.85± 0.21	0.20± 0.02	0.06± 0.02	0.07± 0.01	0.04± 0.01	1.32± 0.86	4.62± 0.79	5	0.54± 0.45
SR3	0.44± 0.15	0.14± 0.04	0.36± 0.10	0.08± 0.01	0.01± 0.00	0.03± 0.01	0.02± 0.00	0.14± 0.34	1.21± 0.40	1	0.36± 0.31
SR4	1.79± 0.44	0.64± 0.11	0.98± 0.23	0.28± 0.04	0.24± 0.05	0.12± 0.02	0.18± 0.02	1.47± 1.07	5.69± 0.92	10	0.71± 0.57
SR5	1.07± 0.31	0.23± 0.05	0.49± 0.13	0.14± 0.02	0.04± 0.01	0.11± 0.02	0.04± 0.00	0.46± 0.44	2.58± 0.53	2	0.75± 0.61
SR6	0.34± 0.11	0.06± 0.04	0.22± 0.06	0.05± 0.01	0.09± 0.03	0.03± 0.01	0.01± 0.00	0.12± 0.17	0.92± 0.24	5	0.54± 0.44
SR7	0.16± 0.06	0.06± 0.02	0.10± 0.03	0.02± 0.00	0.01± 0.00	0.01± 0.00	0.01± 0.00	0.00± 0.07	0.36± 0.12	0	0.40± 0.32
SR8	0.30± 0.09	0.10± 0.02	0.17± 0.04	0.10± 0.01	0.03± 0.01	0.05± 0.01	0.04± 0.01	0.44± 0.42	1.22± 0.45	0	0.96± 0.80
SR9	0.79± 0.20	1.21± 0.19	0.85± 0.20	0.17± 0.02	0.06± 0.01	0.06± 0.01	0.00± 0.00	0.39± 0.22	3.53± 0.34	3	0.33± 0.24
SR10	0.11± 0.04	0.41± 0.08	0.32± 0.08	0.05± 0.01	0.01± 0.00	0.02± 0.01	0.00± 0.00	0.55± 0.43	1.47± 0.44	3	0.18± 0.17
SR11	0.09± 0.03	0.19± 0.05	0.10± 0.03	0.02± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.06± 0.09	0.48± 0.12	0	0.10± 0.09
SR12	0.08± 0.03	0.31± 0.05	0.19± 0.04	0.04± 0.01	0.01± 0.00	0.03± 0.01	0.00± 0.00	0.57± 0.50	1.23± 0.49	2	0.28± 0.23
SR13	0.13± 0.05	0.15± 0.04	0.15± 0.04	0.03± 0.00	0.02± 0.00	0.02± 0.00	0.00± 0.00	0.00± 0.10	0.49± 0.15	0	0.32± 0.27
SR14	0.10± 0.03	0.12± 0.03	0.09± 0.02	0.02± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.01± 0.09	0.34± 0.12	1	0.38± 0.33

Table 53: Postfit event yields in BDT regions for 2016+2017.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	$X+\gamma$	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	1.46± 0.36	38.83± 6.01	1.58± 0.36	1.10± 0.12	0.40± 0.10	6.70± 1.18	0.02± 0.00	1.34± 0.52	51.42± 6.02	53	0.66± 0.35
SR1	0.25± 0.07	0.05± 0.02	0.04± 0.02	0.00± 0.00	0.03± 0.01	0.04± 0.03	0.12± 0.02	0.06± 0.13	0.58± 0.16	1	0.00± 0.00
SR2	2.69± 0.68	0.53± 0.14	0.44± 0.12	0.05± 0.01	0.64± 0.22	0.52± 0.18	0.88± 0.12	0.81± 0.57	6.56± 0.86	7	0.00± 0.01
SR3	5.76± 1.52	1.46± 0.31	1.17± 0.31	0.14± 0.02	1.46± 0.28	0.57± 0.15	1.26± 0.17	3.67± 1.37	15.50± 1.55	14	0.03± 0.02
SR4	8.76± 2.13	2.23± 0.54	1.99± 0.49	0.28± 0.03	0.80± 0.21	0.67± 0.18	1.43± 0.20	4.72± 1.82	20.89± 2.05	28	0.09± 0.05
SR5	10.13± 2.50	2.96± 0.68	2.66± 0.65	0.43± 0.04	1.82± 0.28	0.69± 0.30	1.46± 0.20	5.06± 2.04	25.21± 2.50	14	0.18± 0.11
SR6	11.04± 2.67	3.50± 0.60	3.17± 0.77	0.58± 0.06	1.31± 0.26	0.77± 0.18	1.37± 0.19	3.96± 1.86	25.70± 2.27	31	0.29± 0.19
SR7	10.50± 2.58	3.55± 0.66	3.40± 0.81	0.65± 0.07	1.00± 0.18	0.73± 0.18	1.14± 0.15	3.80± 1.91	24.77± 2.33	22	0.41± 0.23
SR8	9.40± 2.32	3.26± 0.60	3.21± 0.80	0.72± 0.09	0.88± 0.13	0.34± 0.08	0.99± 0.13	4.66± 1.92	23.45± 2.21	24	0.58± 0.30
SR9	11.50± 2.89	3.98± 0.71	4.40± 1.08	1.08± 0.11	0.77± 0.20	0.57± 0.12	0.92± 0.13	4.52± 2.34	27.73± 2.73	29	1.07± 0.54
SR10	10.62± 2.62	4.40± 0.70	4.63± 1.15	1.03± 0.12	0.53± 0.11	0.63± 0.12	0.77± 0.10	4.13± 1.95	26.75± 2.45	35	1.84± 1.00
SR11	5.91± 1.57	2.83± 0.51	3.00± 0.76	0.86± 0.10	0.71± 0.10	0.48± 0.10	0.44± 0.06	3.15± 1.88	17.37± 2.03	23	2.16± 1.14
SR12	3.46± 0.91	1.66± 0.29	1.94± 0.48	0.51± 0.06	0.12± 0.04	0.20± 0.05	0.23± 0.03	2.33± 0.96	10.45± 1.08	18	2.33± 1.21
SR13	1.79± 0.49	0.74± 0.13	1.11± 0.29	0.30± 0.04	0.02± 0.02	0.18± 0.03	0.12± 0.02	0.75± 0.50	4.99± 0.67	3	2.28± 1.22
SR14	1.00± 0.33	0.35± 0.08	0.55± 0.14	0.18± 0.02	0.03± 0.03	0.12± 0.02	0.06± 0.01	0.23± 0.23	2.53± 0.46	6	2.25± 1.18
SR15	0.38± 0.11	0.20± 0.04	0.27± 0.07	0.09± 0.01	0.07± 0.04	0.05± 0.01	0.03± 0.00	0.23± 0.13	1.33± 0.19	2	1.71± 0.88
SR16	0.11± 0.04	0.05± 0.03	0.06± 0.02	0.02± 0.00	0.00± 0.00	0.02± 0.00	0.01± 0.00	0.00± 0.06	0.27± 0.10	1	1.05± 0.56
SR17	0.01± 0.00	0.02± 0.01	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.00± 0.00	0.19± 0.33	0.22± 0.33	0	0.11± 0.05

H Unblinding of 2018 dataset, BDT retraining

For reference, a table with "historical" numbers can be found here: `https://docs.google.com/spreadsheets/d/140yqrUwEsJtOJ8OmDdOun-J8iUrV49aAdUkvfNjpn8k/edit#gid=0` The last row of the table corresponds to the results presented below.

H.1 BDT retraining

With the availability of latest recipes for 2018, the BDT was retrained, keeping identical settings to what was used previously, in Section 6.3. In particular, the finalized retraining included two new key components: Autumn18V8 MC JECs and 2018 b-tag scale factors.

H.2 Yields and results

Plots for the prefit and postfit event yields for the 2018 data, with a total luminosity of 59.6 fb^{-1}, are shown for both the cut-based and BDT based analysis in Figure 115. Numerical yields are also tabulated in Table 54 (prefit cut-based analysis), Table 55 (postfit cut-based analysis), Table 56 (prefit BDT analysis), and Table 57 (postfit BDT analysis).

Similarly, plots of yields for combination of the full Run2 dataset, corresponding to 137.2 fb^{-1}, are shown for both the cut-based and BDT-based analyses in Figure 116. Full Run2 dataset numerical yields are also tabulated in Table 58 (prefit cut-based analysis), Table 59 (postfit cut-based analysis), Table 60 (prefit BDT analysis), and Table 61 (postfit BDT analysis).

A binned likelihood fit is performed over signal regions using the Higgs Combine tool [22] to extract signal strength and significance; exclusion limits at 95% CL are calculated with the Asymptotic CLs method.

Using the full run2 data, the *cut-based* analysis sets an observed (expected) upper limit on the production cross section of 20.31 fb ($9.40^{+4.38}_{-2.87}$ fb), assuming the signal process does not exist. The observed (expected) significance is 1.718 (2.461) standard deviations, corresponding to a measured observed (expected) signal strength parameter of $0.795^{+0.520}_{-0.475}$ ($1.000^{+0.485}_{-0.436}$)

The *BDT* analysis sets an observed (expected) upper limit on the production cross section of 22.86 fb ($8.56^{+3.96}_{-2.63}$ fb), assuming the signal process does not exist. The observed (expected) significance is 2.568 (2.680) standard deviations, corresponding to a measured observed (expected) signal strength parameter of $1.069^{+0.490}_{-0.443}$ ($1.000^{+0.447}_{-0.403}$)

To isolate the effect of the updates to 2018 data and MC, we recalculated the expected limit and significance with the latest corrections used above, but with the old BDT function, which gave an expected upper limit on the production cross section of $8.95^{+4.25}_{-2.80}$ (central value 4.6% worse than new training, relative), an expected significance of 2.636 (1.6% worse than new training, relative), and an expected signal strength of $1.000^{+0.470}_{-0.422}$ (measurement precision 5% worse than new training, relative).

Due to the retraining and updated SFs, updated numbers for the BDT analysis observed (expected) significances are as follows: 1.342 (1.496) for 2016, 1.182 (1.740) for 2017, and 2.950 (1.924) for 2018.

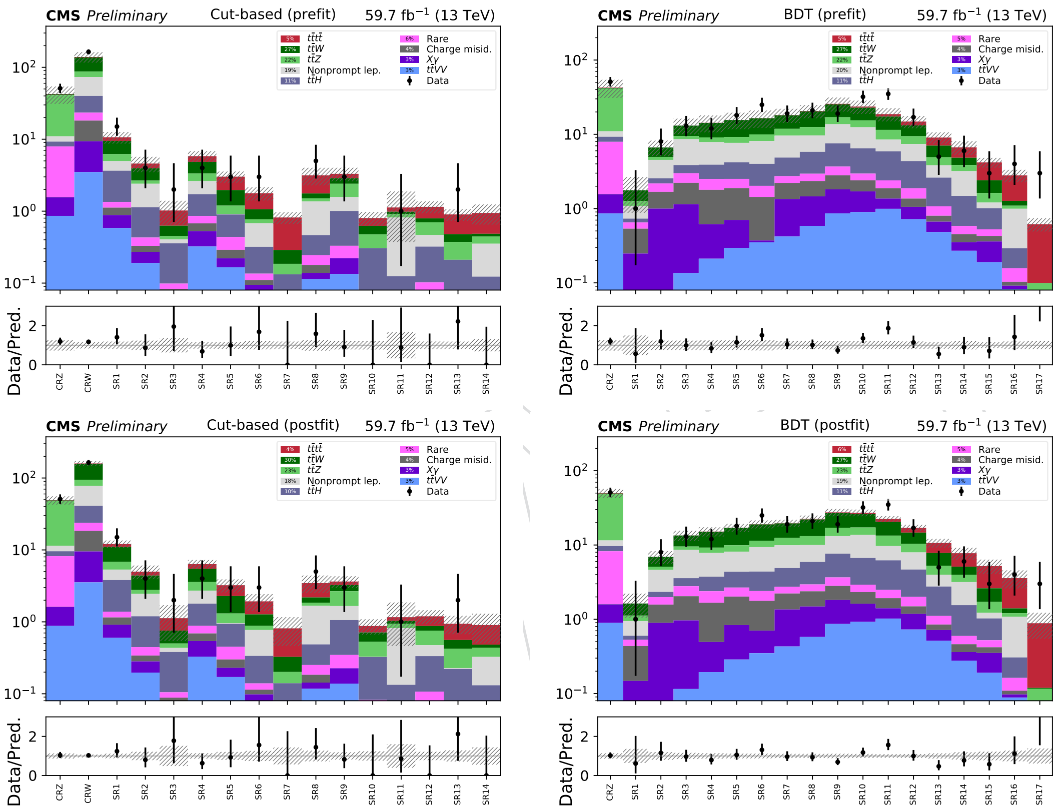


Figure 115: 2018: data yields compared to prefit (top) and postfit (bottom) SM predictions in CRZ, CRW, and the cut-based signal regions (left), and CRZ with the BDT signal regions (right).

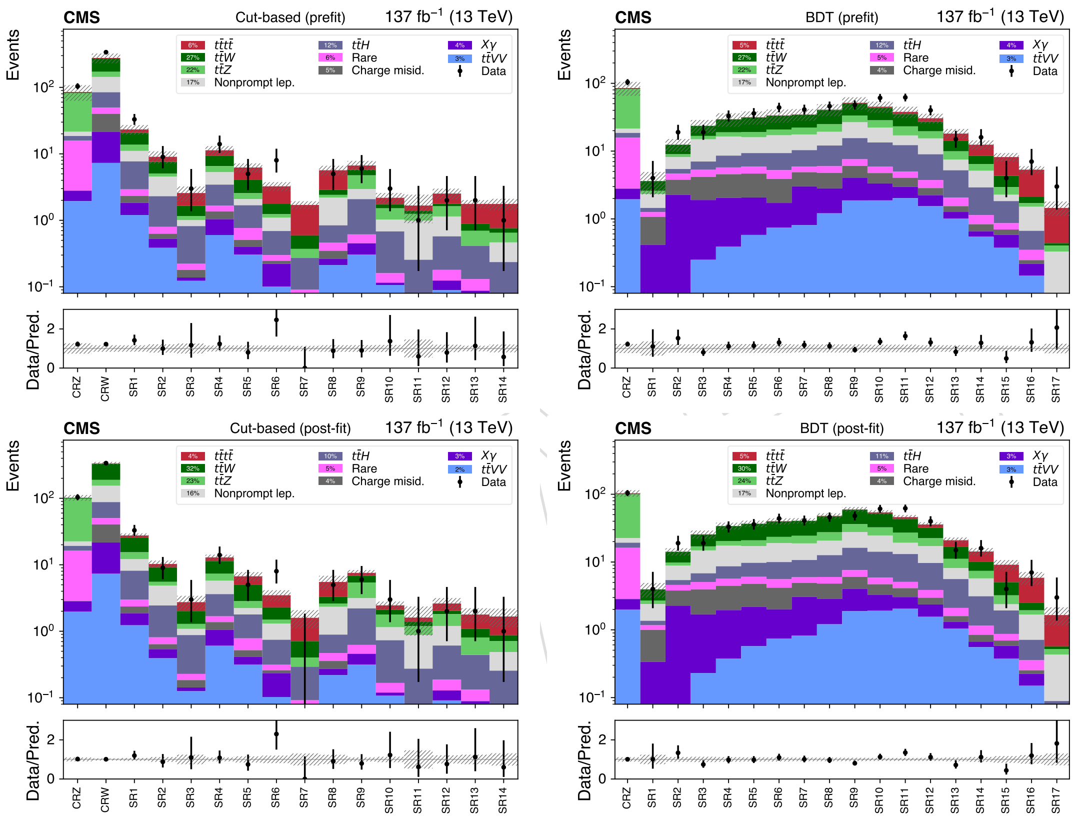


Figure 116: Run2: data yields compared to prefit (top) and postfit (bottom) SM predictions in CRZ, CRW, and the cut-based signal regions (left), and CRZ with the BDT signal regions (right).

Table 54: Prefit event yields in cut-based regions for 2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.87± 0.33	29.83±11.02	1.33± 0.36	0.85± 0.12	0.71± 0.41	6.28± 1.30	0.01± 0.00	1.72± 0.60	41.61±11.49	51	0.48± 0.03
CRW	48.99±18.52	13.77± 4.92	16.49± 4.22	3.45± 0.44	5.75± 1.14	5.18± 1.11	8.79± 1.75	33.41± 9.77	135.83±23.00	163	2.21± 0.08
SR1	3.21± 1.37	1.23± 0.47	2.29± 0.64	0.57± 0.09	0.30± 0.09	0.22± 0.05	0.23± 0.05	1.32± 0.62	9.37± 1.96	15	1.16± 0.07
SR2	1.27± 0.53	0.25± 0.16	0.70± 0.24	0.19± 0.04	0.08± 0.04	0.10± 0.03	0.06± 0.01	1.27± 0.79	3.92± 1.09	4	0.64± 0.04
SR3	0.18± 0.17	0.04± 0.08	0.26± 0.10	0.05± 0.01	0.01± 0.02	0.02± 0.00	0.03± 0.01	0.05± 0.04	0.62± 0.31	2	0.39± 0.07
SR4	1.48± 0.62	0.71± 0.31	0.86± 0.27	0.32± 0.05	0.20± 0.06	0.18± 0.05	0.15± 0.03	0.89± 0.42	4.79± 1.01	4	0.96± 0.05
SR5	0.76± 0.32	0.28± 0.18	0.46± 0.15	0.16± 0.03	0.05± 0.03	0.14± 0.05	0.07± 0.01	0.01± 0.05	1.94± 0.54	3	1.04± 0.05
SR6	0.29± 0.14	0.09± 0.04	0.18± 0.07	0.05± 0.01	0.04± 0.01	0.02± 0.00	0.02± 0.00	0.36± 0.24	1.06± 0.33	3	0.70± 0.08
SR7	0.10± 0.05	0.05± 0.03	0.09± 0.04	0.02± 0.00	0.00± 0.00	0.02± 0.01	0.00± 0.00	0.00± 0.04	0.28± 0.11	0	0.52± 0.07
SR8	0.25± 0.16	0.14± 0.09	0.22± 0.09	0.11± 0.02	0.02± 0.01	0.06± 0.01	0.04± 0.01	0.90± 0.75	1.75± 0.83	5	1.37± 0.13
SR9	0.46± 0.23	0.83± 0.36	0.67± 0.19	0.13± 0.02	0.09± 0.02	0.11± 0.09	0.00± 0.00	0.56± 0.34	2.85± 0.71	3	0.43± 0.04
SR10	0.15± 0.09	0.16± 0.08	0.23± 0.07	0.05± 0.01	0.00± 0.01	0.02± 0.01	0.00± 0.00	0.00± 0.04	0.62± 0.17	0	0.18± 0.04
SR11	0.07± 0.05	0.10± 0.06	0.09± 0.03	0.02± 0.00	0.01± 0.01	0.01± 0.01	0.00± 0.00	0.69± 0.75	0.98± 0.75	1	0.14± 0.02
SR12	0.09± 0.05	0.23± 0.09	0.22± 0.06	0.05± 0.01	0.03± 0.01	0.03± 0.01	0.00± 0.00	0.14± 0.11	0.79± 0.20	0	0.35± 0.03
SR13	0.10± 0.06	0.16± 0.08	0.14± 0.05	0.04± 0.01	0.00± 0.00	0.02± 0.01	0.00± 0.00	0.00± 0.05	0.47± 0.16	2	0.42± 0.04
SR14	0.04± 0.02	0.09± 0.06	0.09± 0.04	0.02± 0.01	0.00± 0.00	0.02± 0.00	0.00± 0.00	0.23± 0.25	0.48± 0.27	0	0.45± 0.06

Table 55: Postfit event yields in cut-based regions for 2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	1.11± 0.28	36.09± 6.39	1.40± 0.36	0.87± 0.11	0.75± 0.54	6.53± 1.39	0.01± 0.00	1.89± 0.67	48.65± 6.64	51	0.45± 0.31
CRW	62.04±14.89	16.59± 2.94	17.26± 4.06	3.52± 0.40	5.90± 1.46	5.36± 1.19	8.87± 1.59	36.98±12.30	156.51±12.23	163	2.02± 1.45
SR1	4.15± 1.27	1.49± 0.29	2.42± 0.62	0.60± 0.09	0.31± 0.10	0.23± 0.05	0.24± 0.04	1.52± 0.84	10.96± 1.46	15	1.06± 0.83
SR2	1.62± 0.57	0.33± 0.16	0.75± 0.24	0.20± 0.04	0.09± 0.07	0.10± 0.04	0.06± 0.01	1.25± 0.88	4.39± 1.10	4	0.59± 0.40
SR3	0.25± 0.28	0.06± 0.11	0.28± 0.11	0.05± 0.02	0.01± 0.02	0.02± 0.00	0.03± 0.00	0.05± 0.04	0.75± 0.48	2	0.37± 0.30
SR4	1.88± 0.60	0.85± 0.23	0.91± 0.26	0.33± 0.05	0.21± 0.10	0.19± 0.05	0.15± 0.03	0.93± 0.38	5.44± 0.87	4	0.88± 0.61
SR5	0.97± 0.30	0.35± 0.16	0.49± 0.15	0.17± 0.03	0.06± 0.06	0.15± 0.05	0.07± 0.01	0.02± 0.05	2.28± 0.48	3	0.95± 0.67
SR6	0.39± 0.16	0.11± 0.04	0.20± 0.08	0.05± 0.01	0.04± 0.02	0.03± 0.01	0.02± 0.00	0.43± 0.30	1.27± 0.36	3	0.66± 0.47
SR7	0.13± 0.06	0.06± 0.03	0.10± 0.04	0.02± 0.00	0.00± 0.00	0.02± 0.01	0.00± 0.00	0.00± 0.05	0.32± 0.12	0	0.49± 0.36
SR8	0.34± 0.19	0.17± 0.10	0.24± 0.10	0.12± 0.02	0.03± 0.02	0.07± 0.02	0.04± 0.01	1.18± 0.68	2.18± 0.75	5	1.28± 0.92
SR9	0.58± 0.29	1.00± 0.24	0.71± 0.18	0.14± 0.02	0.09± 0.02	0.12± 0.15	0.00± 0.00	0.60± 0.37	3.24± 0.64	3	0.39± 0.26
SR10	0.18± 0.14	0.20± 0.11	0.24± 0.07	0.06± 0.01	0.00± 0.02	0.02± 0.01	0.00± 0.00	0.00± 0.04	0.71± 0.22	0	0.17± 0.15
SR11	0.09± 0.08	0.13± 0.07	0.10± 0.03	0.02± 0.01	0.01± 0.01	0.01± 0.01	0.00± 0.00	0.68± 0.69	1.03± 0.70	1	0.13± 0.08
SR12	0.12± 0.07	0.28± 0.10	0.23± 0.07	0.05± 0.01	0.03± 0.01	0.03± 0.01	0.00± 0.00	0.14± 0.11	0.88± 0.22	0	0.32± 0.21
SR13	0.14± 0.08	0.20± 0.10	0.15± 0.05	0.05± 0.01	0.00± 0.00	0.02± 0.01	0.00± 0.00	0.01± 0.05	0.56± 0.16	2	0.39± 0.27
SR14	0.05± 0.03	0.11± 0.06	0.10± 0.04	0.02± 0.01	0.00± 0.00	0.02± 0.00	0.00± 0.00	0.19± 0.27	0.48± 0.29	0	0.42± 0.30

Table 56: Prefit event yields in BDT regions for 2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	0.87± 0.31	29.83±11.11	1.33± 0.35	0.85± 0.13	0.71± 0.43	6.28± 1.56	0.01± 0.00	1.72± 0.66	41.61±11.56	51	0.48± 0.03
SR1	0.49± 0.38	0.18± 0.10	0.08± 0.06	0.01± 0.01	0.24± 0.20	0.11± 0.17	0.29± 0.05	0.36± 0.36	1.75± 0.87	1	0.00± 0.00
SR2	1.66± 0.86	0.43± 0.28	0.38± 0.20	0.03± 0.02	0.95± 0.49	0.49± 0.19	0.68± 0.13	1.97± 1.04	6.60± 1.58	8	0.00± 0.01
SR3	3.54± 1.50	0.89± 0.48	0.86± 0.31	0.13± 0.04	0.97± 0.29	0.76± 0.24	1.08± 0.20	4.74± 2.28	12.98± 3.15	13	0.02± 0.01
SR4	5.09± 2.29	1.23± 0.70	1.38± 0.42	0.21± 0.04	0.40± 0.47	0.71± 0.30	1.17± 0.22	3.95± 1.41	14.14± 3.61	12	0.04± 0.01
SR5	6.26± 2.38	1.53± 1.03	1.68± 0.50	0.29± 0.06	0.38± 0.32	0.61± 0.35	1.19± 0.22	3.47± 1.57	15.40± 3.92	17	0.06± 0.01
SR6	6.38± 2.61	1.83± 0.89	1.96± 0.54	0.35± 0.05	0.03± 0.62	0.59± 0.21	1.06± 0.20	4.10± 1.97	16.30± 4.32	25	0.10± 0.02
SR7	6.05± 2.30	2.23± 1.25	2.11± 0.58	0.42± 0.08	0.91± 0.20	0.51± 0.13	0.87± 0.16	4.74± 2.16	17.83± 3.90	19	0.15± 0.03
SR8	7.65± 2.83	2.85± 1.70	2.84± 0.77	0.58± 0.08	0.86± 0.24	0.59± 0.26	0.89± 0.17	3.86± 1.57	20.12± 4.17	21	0.29± 0.04
SR9	8.18± 3.15	3.26± 2.41	3.83± 0.98	0.85± 0.12	0.93± 0.24	0.86± 0.27	0.99± 0.19	6.14± 2.59	25.05± 5.46	19	0.45± 0.06
SR10	7.28± 2.75	2.93± 1.82	3.49± 0.95	0.89± 0.11	0.79± 0.33	0.55± 0.28	0.69± 0.13	5.99± 2.66	22.60± 4.84	32	0.82± 0.04
SR11	5.77± 2.33	2.94± 1.59	3.47± 0.93	0.98± 0.12	0.37± 0.07	0.51± 0.15	0.52± 0.10	2.60± 1.35	17.14± 3.76	35	1.36± 0.06
SR12	3.71± 1.35	2.00± 1.04	2.47± 0.68	0.71± 0.09	0.33± 0.21	0.51± 0.15	0.31± 0.06	3.04± 1.69	13.08± 2.83	17	1.63± 0.07
SR13	2.10± 0.87	1.07± 0.68	1.53± 0.42	0.48± 0.07	0.18± 0.04	0.27± 0.09	0.13± 0.03	1.22± 0.51	6.98± 1.54	5	1.97± 0.09
SR14	0.97± 0.44	0.66± 0.33	0.89± 0.26	0.27± 0.04	0.08± 0.02	0.13± 0.03	0.10± 0.02	1.71± 1.10	4.81± 1.39	6	1.81± 0.10
SR15	0.78± 0.37	0.41± 0.20	0.48± 0.14	0.19± 0.02	0.16± 0.11	0.10± 0.02	0.07± 0.01	0.19± 0.10	2.37± 0.59	3	1.75± 0.14
SR16	0.17± 0.10	0.10± 0.09	0.13± 0.05	0.08± 0.01	0.01± 0.00	0.05± 0.01	0.01± 0.00	0.70± 0.47	1.26± 0.51	4	1.51± 0.11
SR17	0.00± 0.02	0.05± 0.05	0.02± 0.01	0.01± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.00± 0.03	0.10± 0.08	3	0.51± 0.07

Table 57: Postfit event yields in BDT regions for 2018.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	1.01± 0.25	36.74± 6.82	1.43± 0.31	0.89± 0.12	0.71± 0.41	6.60± 1.34	0.01± 0.00	1.86± 0.67	49.25± 6.88	51	0.70± 0.29
SR1	0.50± 0.23	0.18± 0.06	0.07± 0.04	0.01± 0.00	0.14± 0.15	0.09± 0.09	0.29± 0.05	0.34± 0.33	1.62± 0.59	1	0.00± 0.00
SR2	1.83± 0.58	0.45± 0.17	0.35± 0.14	0.03± 0.01	0.85± 0.32	0.40± 0.15	0.68± 0.12	2.30± 1.17	6.89± 1.34	8	0.00± 0.01
SR3	3.82± 1.08	0.95± 0.29	0.79± 0.26	0.11± 0.03	0.82± 0.28	0.73± 0.22	1.08± 0.19	5.04± 1.99	13.34± 2.15	13	0.02± 0.02
SR4	5.79± 1.59	1.42± 0.37	1.36± 0.35	0.19± 0.04	0.30± 0.37	0.71± 0.23	1.17± 0.21	4.08± 1.53	15.02± 2.15	12	0.05± 0.03
SR5	7.02± 1.78	1.76± 0.52	1.69± 0.42	0.29± 0.05	0.53± 0.24	0.65± 0.24	1.19± 0.21	3.74± 1.42	16.86± 2.16	17	0.08± 0.04
SR6	7.41± 1.92	2.04± 0.52	1.99± 0.47	0.35± 0.05	0.37± 0.46	0.63± 0.17	1.06± 0.19	4.97± 2.09	18.83± 2.37	25	0.14± 0.08
SR7	6.97± 1.75	2.56± 0.66	2.15± 0.51	0.43± 0.07	0.91± 0.14	0.51± 0.11	0.87± 0.15	5.03± 1.92	19.43± 2.11	19	0.19± 0.10
SR8	8.77± 2.15	3.18± 0.92	2.96± 0.69	0.57± 0.08	0.90± 0.21	0.56± 0.18	0.89± 0.16	4.09± 1.71	21.94± 2.44	21	0.42± 0.19
SR9	9.62± 2.42	3.95± 1.19	3.99± 0.89	0.86± 0.12	0.94± 0.18	0.84± 0.23	0.99± 0.17	5.54± 2.32	26.73± 3.16	19	0.65± 0.30
SR10	8.65± 2.14	3.56± 0.91	3.76± 0.86	0.92± 0.12	0.69± 0.21	0.60± 0.20	0.69± 0.12	7.02± 2.90	25.88± 2.88	32	1.19± 0.52
SR11	7.01± 1.78	3.70± 0.83	3.73± 0.85	1.01± 0.13	0.38± 0.07	0.56± 0.14	0.52± 0.09	3.48± 1.58	20.39± 2.05	35	1.94± 0.85
SR12	4.24± 1.06	2.45± 0.53	2.67± 0.62	0.72± 0.10	0.46± 0.16	0.50± 0.13	0.31± 0.05	3.28± 1.49	14.63± 1.65	17	2.36± 0.99
SR13	2.52± 0.68	1.27± 0.33	1.65± 0.39	0.51± 0.07	0.20± 0.03	0.26± 0.08	0.13± 0.02	1.23± 0.45	7.79± 0.92	5	2.86± 1.24
SR14	1.15± 0.35	0.79± 0.17	0.95± 0.25	0.28± 0.04	0.08± 0.02	0.13± 0.03	0.10± 0.02	1.62± 1.02	5.10± 1.12	6	2.66± 1.13
SR15	0.89± 0.30	0.48± 0.13	0.51± 0.14	0.19± 0.03	0.16± 0.09	0.10± 0.02	0.07± 0.01	0.20± 0.12	2.59± 0.51	3	2.61± 1.13
SR16	0.19± 0.09	0.13± 0.05	0.14± 0.06	0.09± 0.01	0.01± 0.00	0.05± 0.01	0.01± 0.00	0.78± 0.40	1.40± 0.42	4	2.20± 0.93
SR17	0.00± 0.01	0.06± 0.03	0.03± 0.02	0.02± 0.00	0.00± 0.00	0.01± 0.00	0.00± 0.00	0.01± 0.03	0.12± 0.07	3	0.77± 0.34

Table 58: Prefit event yields in cut-based regions for Run2.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	1.91± 0.70	60.40±22.00	2.79± 0.73	1.93± 0.22	0.85± 0.41	12.86± 2.77	0.03± 0.00	2.94± 1.10	83.71±22.75	104	1.05± 0.06
CRW	98.82±35.54	27.53±10.10	34.45± 8.64	7.20± 0.79	13.76± 1.80	9.33± 2.08	18.72± 2.14	59.11±21.59	268.90±45.98	337	4.87± 0.20
SR1	7.00± 2.66	2.59± 0.96	4.72± 1.30	1.18± 0.15	0.63± 0.14	0.60± 0.14	0.48± 0.06	3.37± 1.41	20.57± 3.76	33	2.47± 0.14
SR2	2.49± 0.96	0.53± 0.28	1.49± 0.47	0.38± 0.06	0.14± 0.05	0.16± 0.04	0.10± 0.01	2.23± 1.24	7.53± 1.87	9	1.41± 0.07
SR3	0.49± 0.33	0.15± 0.11	0.59± 0.22	0.12± 0.02	0.01± 0.02	0.04± 0.01	0.04± 0.01	0.19± 0.18	1.64± 0.61	3	0.91± 0.12
SR4	2.74± 1.00	1.22± 0.53	1.76± 0.49	0.59± 0.07	0.43± 0.14	0.29± 0.07	0.32± 0.04	1.86± 0.87	9.21± 1.74	14	1.97± 0.10
SR5	1.51± 0.63	0.47± 0.21	0.91± 0.29	0.30± 0.04	0.10± 0.04	0.26± 0.07	0.10± 0.01	0.40± 0.28	4.05± 0.97	5	2.10± 0.10
SR6	0.52± 0.20	0.13± 0.10	0.39± 0.14	0.10± 0.02	0.12± 0.05	0.05± 0.01	0.02± 0.00	0.41± 0.27	1.75± 0.48	8	1.46± 0.09
SR7	0.22± 0.09	0.10± 0.04	0.18± 0.07	0.04± 0.01	0.01± 0.01	0.03± 0.01	0.01± 0.00	0.00± 0.08	0.59± 0.19	0	1.09± 0.10
SR8	0.47± 0.24	0.21± 0.11	0.38± 0.14	0.21± 0.03	0.05± 0.02	0.11± 0.02	0.08± 0.01	1.36± 0.81	2.86± 0.93	5	2.73± 0.17
SR9	1.03± 0.44	1.79± 0.66	1.46± 0.39	0.30± 0.04	0.15± 0.02	0.16± 0.09	0.00± 0.00	0.86± 0.45	5.75± 1.18	6	0.90± 0.08
SR10	0.23± 0.10	0.48± 0.17	0.52± 0.16	0.11± 0.02	0.01± 0.01	0.04± 0.01	0.00± 0.00	0.35± 0.24	1.73± 0.39	3	0.44± 0.06
SR11	0.13± 0.06	0.25± 0.11	0.19± 0.06	0.04± 0.01	0.01± 0.01	0.02± 0.01	0.00± 0.00	0.74± 0.74	1.38± 0.76	1	0.29± 0.03
SR12	0.15± 0.08	0.48± 0.17	0.39± 0.10	0.09± 0.01	0.04± 0.01	0.06± 0.02	0.00± 0.00	0.57± 0.37	1.77± 0.46	2	0.75± 0.04
SR13	0.20± 0.10	0.28± 0.11	0.28± 0.09	0.07± 0.01	0.02± 0.01	0.04± 0.01	0.00± 0.00	0.00± 0.09	0.89± 0.24	2	0.87± 0.05
SR14	0.11± 0.05	0.18± 0.09	0.17± 0.07	0.04± 0.01	0.00± 0.00	0.03± 0.01	0.00± 0.00	0.23± 0.26	0.76± 0.32	1	1.00± 0.09

Table 59: postfit event yields in cut-based regions for Run2.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	2.74± 0.57	75.94±10.52	3.00± 0.68	1.97± 0.24	0.86± 0.38	13.33± 2.24	0.03± 0.00	3.25± 1.03	101.14±10.08	104	0.84± 0.50
CRW	141.84±27.96	34.54± 4.77	37.01± 8.03	7.31± 0.83	14.04± 1.78	9.71± 1.73	18.78± 2.53	66.93±19.96	330.16±18.88	337	3.88± 2.28
SR1	10.20± 2.18	3.26± 0.47	5.12± 1.17	1.21± 0.15	0.64± 0.12	0.62± 0.12	0.49± 0.07	4.01± 1.61	25.55± 2.11	33	1.98± 1.18
SR2	3.60± 0.86	0.70± 0.22	1.62± 0.41	0.39± 0.06	0.15± 0.05	0.17± 0.03	0.10± 0.01	2.41± 1.05	9.13± 1.27	9	1.13± 0.65
SR3	0.72± 0.34	0.20± 0.10	0.64± 0.20	0.12± 0.02	0.02± 0.02	0.04± 0.01	0.04± 0.01	0.21± 0.17	2.00± 0.59	3	0.73± 0.42
SR4	4.01± 0.93	1.56± 0.34	1.93± 0.46	0.60± 0.07	0.43± 0.13	0.30± 0.06	0.32± 0.04	2.14± 0.85	11.30± 1.26	14	1.58± 0.90
SR5	2.20± 0.58	0.60± 0.13	1.00± 0.27	0.31± 0.05	0.10± 0.04	0.26± 0.05	0.10± 0.02	0.44± 0.27	5.01± 0.77	5	1.69± 0.95
SR6	0.80± 0.21	0.18± 0.08	0.43± 0.13	0.10± 0.02	0.13± 0.05	0.05± 0.01	0.02± 0.00	0.56± 0.29	2.29± 0.40	8	1.20± 0.67
SR7	0.31± 0.11	0.12± 0.04	0.20± 0.06	0.04± 0.01	0.01± 0.01	0.03± 0.01	0.01± 0.00	0.00± 0.08	0.71± 0.20	0	0.89± 0.48
SR8	0.70± 0.28	0.28± 0.12	0.42± 0.15	0.22± 0.03	0.05± 0.02	0.11± 0.02	0.08± 0.01	1.45± 0.82	3.30± 0.95	5	2.21± 1.27
SR9	1.47± 0.42	2.24± 0.34	1.58± 0.36	0.31± 0.05	0.14± 0.02	0.16± 0.09	0.00± 0.00	0.94± 0.46	6.85± 0.80	6	0.72± 0.39
SR10	0.33± 0.11	0.62± 0.14	0.56± 0.14	0.11± 0.02	0.01± 0.01	0.05± 0.01	0.00± 0.00	0.42± 0.26	2.10± 0.31	3	0.36± 0.22
SR11	0.19± 0.07	0.32± 0.07	0.20± 0.06	0.04± 0.01	0.01± 0.01	0.02± 0.01	0.00± 0.00	0.60± 0.72	1.38± 0.75	1	0.23± 0.14
SR12	0.22± 0.08	0.61± 0.12	0.42± 0.10	0.09± 0.01	0.04± 0.01	0.06± 0.01	0.00± 0.00	0.59± 0.40	2.04± 0.48	2	0.59± 0.34
SR13	0.29± 0.11	0.36± 0.12	0.31± 0.09	0.07± 0.01	0.02± 0.01	0.04± 0.01	0.00± 0.00	0.00± 0.11	1.10± 0.28	2	0.70± 0.40
SR14	0.16± 0.05	0.23± 0.07	0.19± 0.06	0.04± 0.01	0.00± 0.00	0.03± 0.01	0.00± 0.00	0.23± 0.27	0.87± 0.30	1	0.81± 0.45

Table 60: Prefit event yields in BDT regions for Run2.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	1.91± 0.76	60.40±19.18	2.79± 0.60	1.93± 0.25	0.85± 0.36	12.86± 2.43	0.03± 0.00	2.94± 0.97	83.71±19.68	104	1.05± 0.06
SR1	1.01± 0.65	0.27± 0.15	0.19± 0.13	0.02± 0.01	0.39± 0.27	0.18± 0.24	0.66± 0.07	0.86± 0.55	3.58± 1.29	4	0.00± 0.00
SR2	3.40± 1.73	0.81± 0.44	0.79± 0.34	0.07± 0.03	2.14± 0.61	0.84± 0.27	1.52± 0.17	2.72± 1.28	12.31± 2.61	19	0.01± 0.01
SR3	6.73± 3.26	1.79± 0.78	1.70± 0.57	0.25± 0.05	1.61± 0.43	1.12± 0.38	2.26± 0.26	7.91± 2.79	23.35± 5.16	19	0.04± 0.02
SR4	10.36± 4.68	2.44± 1.11	2.80± 0.80	0.38± 0.08	1.64± 0.63	1.17± 0.45	2.51± 0.28	7.87± 2.55	29.17± 6.61	33	0.08± 0.03
SR5	12.19± 5.18	3.01± 1.54	3.41± 0.95	0.57± 0.10	1.44± 0.61	1.21± 0.53	2.59± 0.29	6.65± 2.23	31.08± 7.00	35	0.15± 0.03
SR6	12.42± 5.43	3.66± 1.52	4.07± 1.02	0.74± 0.09	0.94± 0.85	1.34± 0.37	2.17± 0.25	7.74± 2.76	33.08± 7.38	44	0.22± 0.06
SR7	12.18± 5.10	4.43± 2.07	4.38± 1.10	0.80± 0.14	2.17± 0.35	1.01± 0.26	1.86± 0.21	7.27± 2.42	34.09± 6.84	41	0.31± 0.07
SR8	15.09± 6.14	5.59± 2.83	6.03± 1.44	1.19± 0.15	1.60± 0.37	1.19± 0.42	1.95± 0.22	7.17± 2.61	39.80± 8.25	46	0.67± 0.05
SR9	16.78± 7.08	6.74± 3.65	7.87± 1.83	1.85± 0.25	2.11± 0.34	1.54± 0.42	2.02± 0.23	11.11± 4.16	50.01±10.62	48	1.11± 0.08
SR10	14.71± 6.01	6.14± 3.04	7.45± 1.70	1.85± 0.24	1.43± 0.44	1.12± 0.36	1.40± 0.16	8.86± 3.31	42.96± 8.47	61	1.78± 0.08
SR11	11.80± 5.11	6.03± 2.64	7.20± 1.66	2.00± 0.26	0.95± 0.19	0.91± 0.22	1.07± 0.12	4.80± 2.15	34.75± 7.12	62	2.80± 0.15
SR12	7.32± 3.04	4.14± 1.72	5.18± 1.21	1.51± 0.18	0.69± 0.22	0.88± 0.26	0.61± 0.07	6.41± 2.60	26.76± 5.08	40	3.50± 0.14
SR13	4.14± 1.85	2.38± 1.16	3.13± 0.77	1.00± 0.13	0.25± 0.06	0.51± 0.13	0.28± 0.03	2.31± 0.88	14.00± 2.99	15	3.97± 0.16
SR14	2.04± 1.00	1.30± 0.50	1.76± 0.47	0.54± 0.07	0.11± 0.08	0.30± 0.05	0.18± 0.02	2.21± 1.13	8.43± 1.93	16	3.88± 0.17
SR15	1.37± 0.73	0.67± 0.29	0.95± 0.27	0.37± 0.05	0.20± 0.10	0.17± 0.03	0.12± 0.01	0.44± 0.24	4.27± 1.12	4	3.79± 0.22
SR16	0.42± 0.24	0.20± 0.12	0.31± 0.13	0.14± 0.02	0.07± 0.04	0.10± 0.02	0.03± 0.00	0.84± 0.48	2.12± 0.66	7	3.15± 0.19
SR17	0.03± 0.03	0.07± 0.05	0.03± 0.02	0.02± 0.00	0.00± 0.00	0.02± 0.00	0.00± 0.00	0.25± 0.30	0.44± 0.33	3	1.00± 0.10

Table 61: postfit event yields in BDT regions for Run2.

	$t\bar{t}W$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}VV$	X+γ	Rares	Charge misid.	Nonprompt lep.	SM expected	Data	$t\bar{t}t\bar{t}$
CRZ	2.51± 0.54	77.27±11.74	3.09± 0.65	1.98± 0.21	0.87± 0.36	13.29± 2.58	0.03± 0.00	3.25± 1.19	102.30±11.59	104	1.12± 0.43
SR1	1.24± 0.49	0.31± 0.11	0.19± 0.10	0.01± 0.01	0.32± 0.25	0.17± 0.19	0.66± 0.07	1.05± 0.65	3.95± 0.95	4	0.00± 0.00
SR2	4.33± 1.28	0.96± 0.32	0.82± 0.29	0.08± 0.02	2.16± 0.53	0.78± 0.23	1.53± 0.17	3.49± 1.46	14.15± 1.77	19	0.01± 0.01
SR3	8.48± 2.41	2.14± 0.50	1.74± 0.53	0.23± 0.05	1.43± 0.47	1.10± 0.38	2.26± 0.26	8.08± 2.79	25.45± 3.53	19	0.04± 0.03
SR4	13.42± 3.36	3.02± 0.68	2.98± 0.77	0.37± 0.08	1.56± 0.49	1.18± 0.40	2.51± 0.29	8.81± 3.10	33.86± 4.00	33	0.08± 0.05
SR5	15.71± 3.69	3.71± 0.93	3.66± 0.93	0.57± 0.08	1.57± 0.51	1.25± 0.45	2.59± 0.29	7.42± 2.61	36.48± 4.00	35	0.15± 0.07
SR6	16.26± 3.87	4.47± 0.99	4.38± 1.03	0.74± 0.08	1.22± 0.67	1.39± 0.33	2.17± 0.25	9.05± 3.47	39.68± 4.14	44	0.23± 0.12
SR7	15.88± 3.61	5.48± 1.28	4.72± 1.10	0.81± 0.12	2.21± 0.31	1.01± 0.27	1.86± 0.21	8.18± 3.08	40.17± 3.72	41	0.32± 0.16
SR8	19.71± 4.39	6.83± 1.75	6.58± 1.47	1.19± 0.14	1.64± 0.32	1.19± 0.34	1.95± 0.22	8.03± 2.83	47.11± 4.34	46	0.72± 0.28
SR9	22.13± 4.94	8.52± 2.23	8.59± 1.90	1.87± 0.21	2.13± 0.30	1.52± 0.43	2.02± 0.24	11.59± 3.95	58.35± 5.27	48	1.18± 0.46
SR10	19.58± 4.27	7.76± 1.85	8.27± 1.79	1.89± 0.21	1.36± 0.36	1.17± 0.29	1.40± 0.16	10.58± 3.98	52.01± 4.30	61	1.92± 0.74
SR11	15.88± 3.65	7.77± 1.53	7.99± 1.74	2.04± 0.23	0.98± 0.17	0.96± 0.22	1.07± 0.12	6.16± 2.65	42.86± 3.51	62	3.00± 1.19
SR12	9.61± 2.22	5.28± 1.01	5.78± 1.26	1.54± 0.18	0.81± 0.21	0.88± 0.24	0.61± 0.07	7.52± 3.06	32.03± 3.07	40	3.76± 1.41
SR13	5.55± 1.35	3.00± 0.66	3.49± 0.77	1.03± 0.12	0.27± 0.05	0.50± 0.12	0.28± 0.03	2.54± 0.94	16.67± 1.62	15	4.28± 1.63
SR14	2.73± 0.74	1.65± 0.28	1.96± 0.48	0.55± 0.07	0.11± 0.07	0.31± 0.06	0.18± 0.02	2.62± 1.24	10.12± 1.25	16	4.19± 1.59
SR15	1.81± 0.53	0.84± 0.24	1.06± 0.26	0.37± 0.04	0.19± 0.10	0.17± 0.03	0.12± 0.01	0.46± 0.22	5.01± 0.77	4	4.11± 1.55
SR16	0.56± 0.21	0.26± 0.10	0.35± 0.12	0.15± 0.02	0.07± 0.04	0.10± 0.02	0.03± 0.00	0.97± 0.54	2.50± 0.61	7	3.39± 1.26
SR17	0.05± 0.03	0.09± 0.05	0.04± 0.02	0.02± 0.00	0.00± 0.00	0.02± 0.00	0.00± 0.00	0.34± 0.35	0.57± 0.36	3	1.08± 0.42

H.3 Nuisances

Two sets of nuisance pull values, expected and observed for cut-based and BDT analyses, are tabulated in Table 62 (expected cut-based analysis), Table 63 (observed cut-based analysis), Table 64 (expected BDT analysis), and Table 65 (observed BDT analysis).

The most constrained nuisances correspond to normalization parameters for ttW and ttZ (“TTWSF” and “TTZSF”) due to high statistics in control regions and in the bulk (BDT). Their input normalization uncertainty is 40%.

Table 62: Expected nuisance pulls for the cut-based analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	**+0.37, 0.99**	+0.00, 0.97	-0.12
TTVV	+0.05, 1.00	+0.00, 0.99	-0.02
TTWSF	**+0.09, 0.62**	**+0.00, 0.62**	-0.06
TTZSF	**-0.02, 0.45**	**+0.00, 0.45**	+0.02
XG	-0.01, 0.99	+0.00, 0.99	+0.00
alphas	+0.00, 0.99	+0.00, 0.99	+0.00
bb	**+0.72, 0.94**	+0.00, 0.98	-0.23
fakes	+0.25, 0.91	+0.00, 0.91	-0.09
pdf	+0.07, 0.99	-0.00, 0.99	-0.02
rares	-0.01, 0.99	+0.00, 0.99	+0.00
scale	-0.20, 0.97	-0.00, 0.99	+0.00
y2016_btaghf	+0.04, 0.99	+0.00, 0.99	-0.03
y2016_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2016_fakes_EWK	+0.04, 1.00	+0.00, 0.99	-0.01
y2016_flips	-0.02, 0.99	+0.00, 0.99	+0.00
y2016_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.02
y2016_isr	-0.10, 0.99	-0.00, 0.99	+0.04
y2016_isrvar	+0.00, 0.99	-0.00, 0.99	-0.02
y2016_jer	**-0.13, 1.12**	+0.00, 0.99	+0.03
y2016_jes	+0.07, 0.97	+0.00, 0.99	-0.05
y2016_lep	+0.01, 0.99	+0.00, 0.99	-0.01
y2016_lumi	+0.02, 0.99	+0.00, 0.99	-0.02
y2016_prefire	-0.00, 0.99	-0.00, 0.99	+0.00
y2016_pu	+0.01, 0.99	-0.00, 0.99	+0.00
y2016_trig	+0.01, 0.99	+0.00, 0.99	-0.02
y2017_btaghf	+0.04, 0.99	+0.00, 0.99	-0.03
y2017_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2017_fakes_EWK	-0.00, 0.99	+0.00, 0.98	-0.00
y2017_flips	-0.02, 0.99	+0.00, 0.99	+0.01
y2017_fsrvar	+0.00, 0.99	-0.00, 0.99	+0.00
y2017_isr	+0.07, 0.99	-0.00, 0.99	-0.03
y2017_isrvar	+0.00, 0.99	-0.00, 0.99	-0.01
y2017_jer	**-0.12, 1.20**	+0.00, 0.99	+0.03
y2017_jes	+0.07, 1.00	+0.00, 0.99	-0.05
y2017_lep	+0.01, 0.99	+0.00, 0.99	-0.01
y2017_lumi	+0.01, 0.99	+0.00, 0.99	-0.02
y2017_prefire	-0.01, 0.99	-0.00, 0.99	+0.01
y2017_pu	-0.01, 0.99	-0.00, 0.99	+0.00
y2017_trig	+0.01, 0.99	+0.00, 0.99	-0.02
y2018_btaghf	+0.09, 0.99	+0.00, 0.99	-0.05
y2018_btaglf	+0.12, 0.99	+0.00, 0.99	-0.04
y2018_fakes_EWK	+0.08, 1.03	+0.00, 0.99	-0.02
y2018_flips	-0.03, 0.99	+0.00, 0.99	+0.01
y2018_fsrvar	+0.00, 0.99	+0.00, 0.99	-0.00
y2018_isr	+0.08, 0.99	-0.00, 0.99	-0.03
y2018_isrvar	+0.00, 0.99	+0.00, 0.99	-0.00
y2018_jer	**-0.32, 1.08**	+0.00, 0.99	+0.08
y2018_jes	+0.15, 0.98	+0.00, 0.99	-0.09
y2018_lep	+0.02, 0.99	+0.00, 0.99	-0.02
y2018_lumi	+0.06, 0.99	+0.00, 0.99	-0.07
y2018_pu	+0.00, 1.00	-0.00, 0.99	+0.00
y2018_trig	+0.02, 0.99	+0.00, 0.99	-0.03

Table 63: Observed nuisance pulls for the cut-based analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	**+0.49, 1.02**	+0.24, 1.00	-0.12
TTVV	+0.06, 1.00	+0.03, 1.00	-0.02
TTWSF	**+1.06, 0.58**	**+1.02, 0.56**	-0.05
TTZSF	**+0.57, 0.42**	**+0.61, 0.42**	+0.06
XG	+0.04, 1.00	+0.04, 1.00	+0.00
alphas	+0.00, 0.99	+0.00, 0.99	+0.00
bb	**+0.62, 0.92**	+0.10, 0.97	-0.27
fakes	**+0.55, 1.02**	**+0.34, 1.00**	-0.09
pdf	+0.15, 0.98	+0.06, 1.00	-0.04
rares	+0.06, 1.00	+0.07, 1.00	+0.01
scale	-0.09, 0.91	-0.02, 0.96	-0.00
y2016_btaghf	+0.07, 0.99	+0.04, 0.99	-0.03
y2016_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2016_fakes_EWK	+0.08, 1.00	+0.06, 0.99	-0.01
y2016_flips	+0.00, 0.99	+0.01, 1.00	+0.01
y2016_fsrvar	+0.00, 0.99	-0.01, 1.04	+0.01
y2016_isr	-0.13, 0.99	-0.04, 0.99	+0.05
y2016_isrvar	+0.00, 0.99	-0.04, 1.06	-0.03
y2016_jer	-0.07, 1.04	+0.01, 0.98	+0.04
y2016_jes	**+0.39, 1.07**	**+0.26, 1.12**	-0.08
y2016_lep	+0.04, 0.99	+0.03, 0.99	-0.01
y2016_lumi	+0.05, 1.00	+0.04, 1.00	-0.02
y2016_prefire	-0.01, 0.99	-0.00, 0.99	+0.00
y2016_pu	+0.03, 0.99	+0.01, 0.99	-0.01
y2016_trig	+0.04, 0.99	+0.04, 0.99	-0.01
y2017_btaghf	+0.10, 1.00	+0.06, 0.99	-0.03
y2017_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2017_fakes_EWK	+0.24, 0.98	+0.22, 0.97	-0.01
y2017_flips	+0.00, 0.99	+0.01, 1.00	+0.01
y2017_fsrvar	+0.00, 0.99	-0.01, 0.99	+0.00
y2017_isr	+0.08, 0.99	+0.02, 0.99	-0.04
y2017_isrvar	+0.00, 0.99	-0.01, 0.99	-0.01
y2017_jer	**-0.07, 1.24**	+0.01, 0.96	+0.03
y2017_jes	+0.14, 0.93	+0.09, 0.93	-0.05
y2017_lep	+0.03, 0.99	+0.03, 0.99	-0.01
y2017_lumi	+0.05, 1.00	+0.04, 1.00	-0.01
y2017_prefire	-0.02, 0.99	-0.02, 0.99	+0.01
y2017_pu	+0.02, 1.00	+0.01, 1.00	-0.00
y2017_trig	+0.05, 0.99	+0.04, 0.99	-0.01
y2018_btaghf	+0.15, 0.99	+0.08, 0.99	-0.05
y2018_btaglf	+0.12, 0.99	+0.03, 0.99	-0.05
y2018_fakes_EWK	+0.15, 1.01	+0.09, 0.99	-0.03
y2018_flips	+0.01, 0.99	+0.02, 1.00	+0.01
y2018_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2018_isr	+0.08, 0.99	+0.02, 0.99	-0.04
y2018_isrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2018_jer	**-0.19, 1.16**	+0.06, 1.05	+0.09
y2018_jes	**+0.48, 1.03**	**+0.30, 1.08**	-0.13
y2018_lep	+0.06, 0.99	+0.05, 0.99	-0.02
y2018_lumi	+0.20, 0.99	+0.17, 0.99	-0.06
y2018_pu	+0.00, 1.00	-0.00, 0.99	+0.00
y2018_trig	+0.09, 0.99	+0.07, 0.99	-0.02

Table 64: Expected nuisance pulls for the BDT analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta,\mu)$
TTH	**+0.31, 0.98**	-0.00, 0.96	-0.11
TTVV	+0.06, 1.00	-0.00, 0.99	-0.02
TTWSF	**+0.07, 0.64**	**-0.00, 0.63**	-0.06
TTZSF	**-0.03, 0.50**	**+0.00, 0.48**	+0.03
XG	-0.01, 0.99	+0.00, 0.99	+0.00
alphas	+0.00, 0.99	+0.00, 0.99	+0.00
bb	**+0.89, 0.94**	-0.00, 0.99	-0.26
fakes	**+0.31, 0.90**	**+0.00, 0.90**	-0.09
pdf	+0.12, 0.97	-0.00, 0.99	-0.03
rares	-0.02, 0.99	+0.00, 0.99	+0.00
scale	-0.25, 0.98	+0.00, 0.99	+0.01
y2016_btaghf	+0.04, 0.99	+0.00, 0.99	-0.02
y2016_btaglf	+0.03, 0.99	-0.00, 0.99	-0.01
y2016_fakes_EWK	+0.03, 0.99	-0.00, 0.99	-0.01
y2016_flips	-0.02, 0.99	+0.00, 0.99	+0.01
y2016_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.01
y2016_isr	-0.13, 0.99	+0.00, 0.99	+0.04
y2016_isrvar	+0.00, 0.99	-0.00, 0.99	-0.01
y2016_jer	-0.04, 1.03	+0.00, 0.99	+0.01
y2016_jes	+0.07, 0.90	+0.00, 0.99	-0.04
y2016_lep	+0.01, 0.99	+0.00, 0.99	-0.01
y2016_lumi	+0.01, 0.99	+0.00, 0.99	-0.02
y2016_prefire	-0.00, 0.99	-0.00, 0.99	+0.00
y2016_pu	+0.04, 0.99	-0.00, 0.99	-0.01
y2016_trig	+0.01, 0.99	+0.00, 0.99	-0.02
y2017_btaghf	+0.03, 0.99	+0.00, 0.99	-0.03
y2017_btaglf	+0.03, 0.99	+0.00, 0.99	-0.01
y2017_fakes_EWK	+0.03, 0.98	+0.00, 0.96	-0.01
y2017_flips	-0.02, 0.99	+0.00, 0.99	+0.00
y2017_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2017_isr	+0.07, 0.99	-0.00, 0.99	-0.02
y2017_isrvar	+0.00, 0.99	-0.00, 0.99	-0.01
y2017_jer	**-0.00, 1.18**	+0.00, 0.99	+0.00
y2017_jes	+0.07, 0.90	+0.00, 0.99	-0.04
y2017_lep	+0.00, 0.99	+0.00, 0.99	-0.01
y2017_lumi	+0.00, 0.99	+0.00, 0.99	-0.02
y2017_prefire	-0.00, 0.99	-0.00, 0.99	+0.01
y2017_pu	+0.03, 1.00	-0.00, 0.99	-0.01
y2017_trig	+0.00, 0.99	+0.00, 0.99	-0.02
y2018_btaghf	+0.08, 0.99	+0.00, 0.99	-0.05
y2018_btaglf	+0.11, 0.99	-0.00, 0.99	-0.04
y2018_fakes_EWK	+0.06, 0.98	+0.00, 0.98	-0.01
y2018_flips	-0.03, 0.99	+0.00, 0.99	+0.01
y2018_fsrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2018_isr	+0.04, 0.99	-0.00, 0.99	-0.02
y2018_isrvar	+0.00, 0.99	+0.00, 0.99	+0.00
y2018_jer	**+0.01, 1.49**	+0.00, 0.95	-0.00
y2018_jes	**+0.12, 0.85**	+0.00, 0.97	-0.08
y2018_lep	+0.01, 0.99	+0.00, 0.99	-0.02
y2018_lumi	+0.05, 0.99	+0.00, 0.99	-0.07
y2018_pu	+0.02, 1.01	-0.00, 0.99	-0.01
y2018_trig	+0.02, 0.99	+0.00, 0.99	-0.03

Table 65: Observed nuisance pulls for the BDT analysis. The final column indicates the correlation between the nuisance and the signal strength

name	b-only fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$s+b$ fit $\Delta x/\sigma_{\text{in}}, \sigma_{\text{out}}/\sigma_{\text{in}}$	$\rho(\theta, \mu)$
TTH	**+0.59, 1.02**	**+0.34, 1.01**	-0.09
TTVV	+0.08, 1.00	+0.04, 1.00	-0.02
TTWSF	**+0.65, 0.67**	**+0.77, 0.60**	-0.00
TTZSF	**+0.62, 0.43**	**+0.66, 0.43**	+0.05
XG	+0.02, 0.99	+0.04, 1.00	+0.01
alphas	+0.00, 0.99	+0.00, 1.00	+0.00
bb	**+0.88, 0.95**	+0.07, 0.98	-0.27
fakes	**+0.95, 0.99**	**+0.34, 0.99**	-0.13
pdf	+0.04, 1.00	-0.04, 1.01	-0.02
rares	+0.03, 0.99	+0.04, 0.99	+0.00
scale	-0.25, 0.94	-0.05, 0.94	+0.00
y2016_btaghf	+0.08, 0.99	+0.05, 0.99	-0.02
y2016_btaglf	+0.04, 1.00	+0.01, 0.99	-0.01
y2016_fakes_EWK	+0.17, 0.99	+0.08, 0.99	-0.02
y2016_flips	-0.01, 0.99	+0.00, 1.00	+0.00
y2016_fsrvar	+0.00, 0.99	+0.02, 1.07	+0.02
y2016_isr	-0.19, 0.99	-0.07, 0.99	+0.04
y2016_isrvar	+0.00, 0.99	+0.01, 1.00	-0.01
y2016_jer	**-0.06, 0.89**	**-0.03, 0.87**	+0.01
y2016_jes	+0.04, 0.99	**-0.04, 1.11**	-0.03
y2016_lep	+0.04, 0.99	+0.03, 0.99	-0.01
y2016_lumi	+0.06, 1.00	+0.05, 1.00	-0.02
y2016_prefire	-0.01, 0.99	-0.00, 0.99	+0.00
y2016_pu	-0.01, 1.00	-0.05, 1.00	-0.01
y2016_trig	+0.05, 0.99	+0.04, 0.99	-0.02
y2017_btaghf	+0.09, 0.99	+0.08, 0.99	-0.03
y2017_btaglf	+0.05, 0.99	+0.02, 0.99	-0.01
y2017_fakes_EWK	**+0.44, 1.03**	+0.26, 0.98	-0.04
y2017_flips	-0.01, 0.99	+0.01, 1.00	+0.01
y2017_fsrvar	+0.00, 0.99	+0.02, 1.00	+0.01
y2017_isr	+0.08, 0.99	+0.02, 0.99	-0.02
y2017_isrvar	+0.00, 0.99	+0.00, 1.00	-0.01
y2017_jer	**+0.09, 0.88**	**+0.07, 0.84**	-0.01
y2017_jes	**+0.15, 0.90**	+0.09, 1.00	-0.04
y2017_lep	+0.05, 0.99	+0.04, 0.99	-0.01
y2017_lumi	+0.05, 1.00	+0.05, 1.00	-0.02
y2017_prefire	-0.01, 0.99	-0.01, 0.99	+0.01
y2017_pu	+0.05, 1.00	+0.03, 1.00	-0.01
y2017_trig	+0.04, 0.99	+0.05, 0.99	-0.01
y2018_btaghf	+0.14, 0.99	+0.09, 0.99	-0.05
y2018_btaglf	+0.06, 0.99	-0.03, 0.99	-0.04
y2018_fakes_EWK	+0.10, 0.98	-0.01, 0.99	-0.03
y2018_flips	-0.02, 0.99	+0.00, 0.99	+0.01
y2018_fsrvar	+0.00, 0.99	+0.00, 0.99	-0.00
y2018_isr	+0.04, 0.99	+0.01, 0.99	-0.02
y2018_isrvar	+0.00, 0.99	+0.00, 0.99	-0.00
y2018_jer	**+0.74, 0.77**	**+0.50, 0.97**	-0.14
y2018_jes	**+0.21, 0.67**	**+0.14, 0.81**	-0.07
y2018_lep	+0.06, 0.99	+0.05, 0.99	-0.02
y2018_lumi	+0.16, 0.99	+0.15, 0.99	-0.06
y2018_pu	+0.09, 1.01	+0.06, 0.99	-0.02
y2018_trig	+0.06, 0.99	+0.05, 0.99	-0.03

H.4 Impacts

The leading 30 nuisance impacts for two sets of impacts, expected and observed for cut-based and BDT analyses, are shown in Figure 117 (expected cut-based analysis), Figure 118 (observed cut-based analysis), Figure 119 (expected BDT analysis), and Figure 120 (observed BDT analysis). The leading expected nuisance in both cases corresponds to the $\sigma(\mathrm{ttbb})/\sigma(\mathrm{ttjj})$ scaling. Note that the "prop binSS" nuisances for MC statistics include (and are dominated by) tight-loose sideband statistics.

The obseved pulls show the most constrained/pulled nuisances correspond to normalization parameters for ttW and ttZ, as we would expect from the control regions. "TTWSF" is moved by approximately 1σ (0.8σ) with respect to the input nuisance sizes for the cut-based (BDT) analysis. "TTZSF" is moved up by approximately 0.6σ (0.7σ) for the cut-based (BDT) analysis.

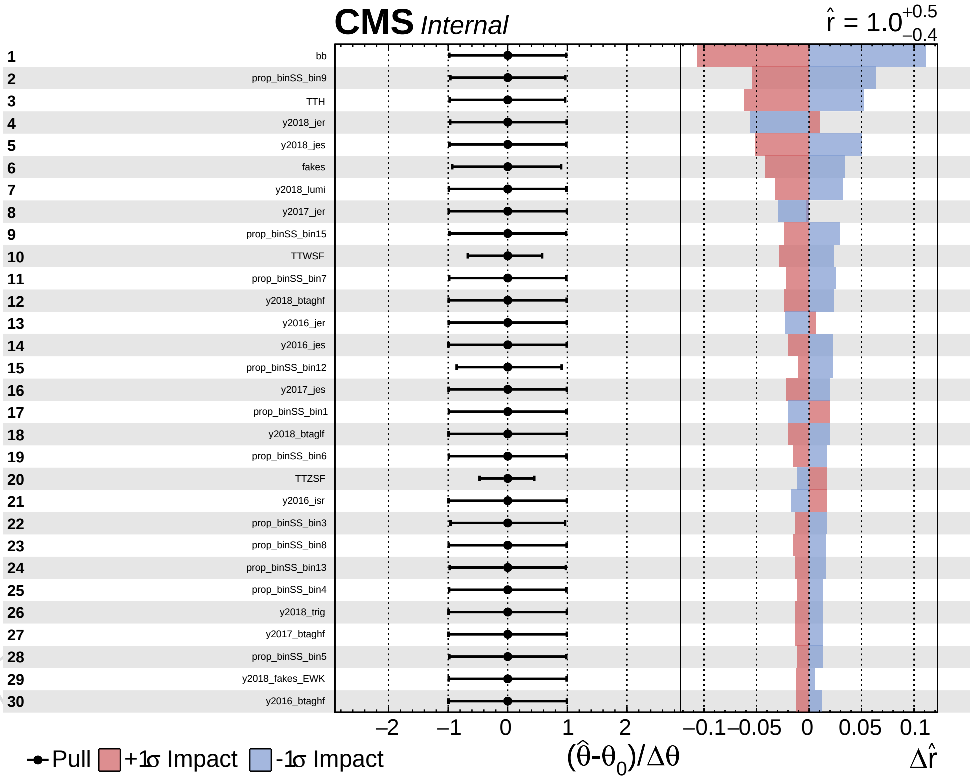


Figure 117: Expected nuisance impacts for the cut-based analysis.

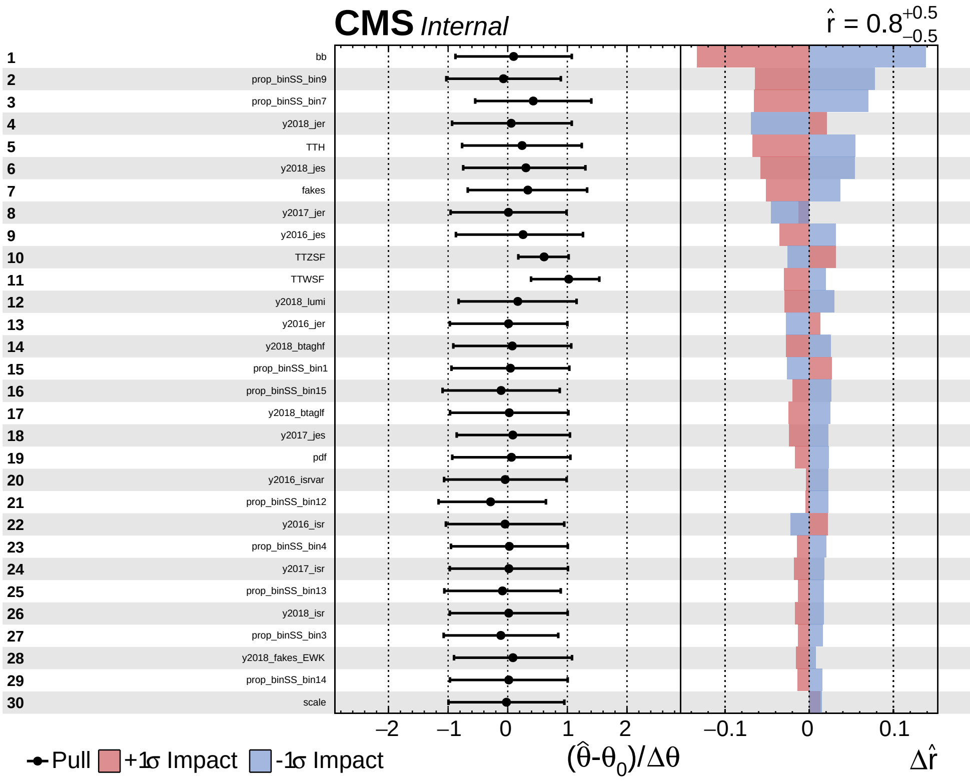


Figure 118: Observed nuisance impacts for the cut-based analysis.

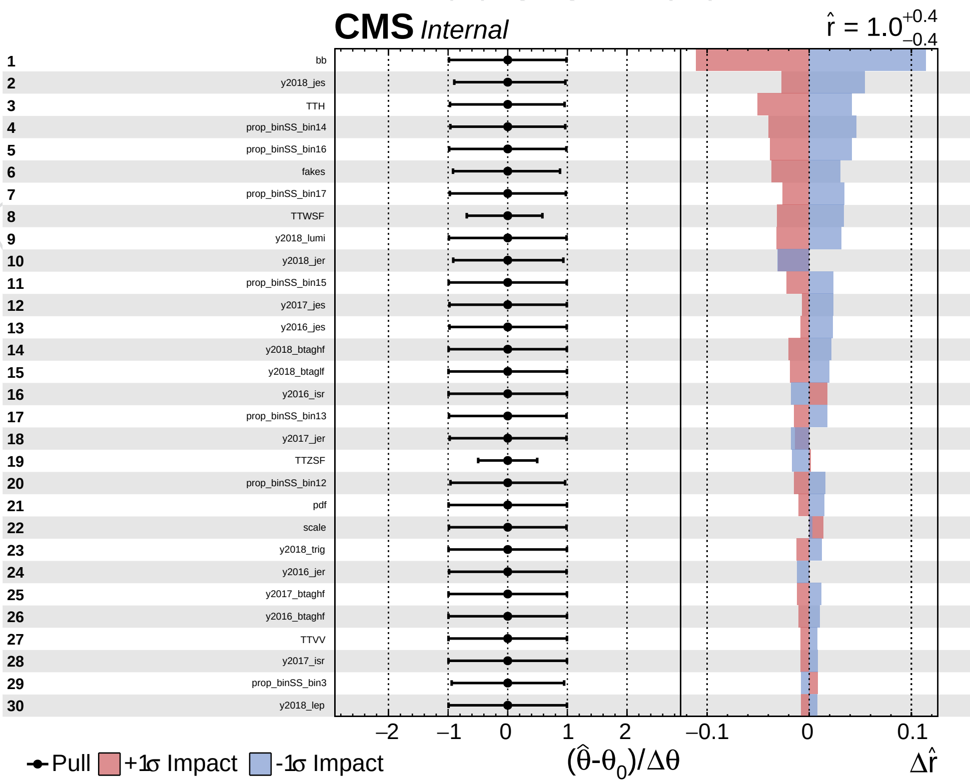


Figure 119: Expected nuisance impacts for the BDT-based analysis.

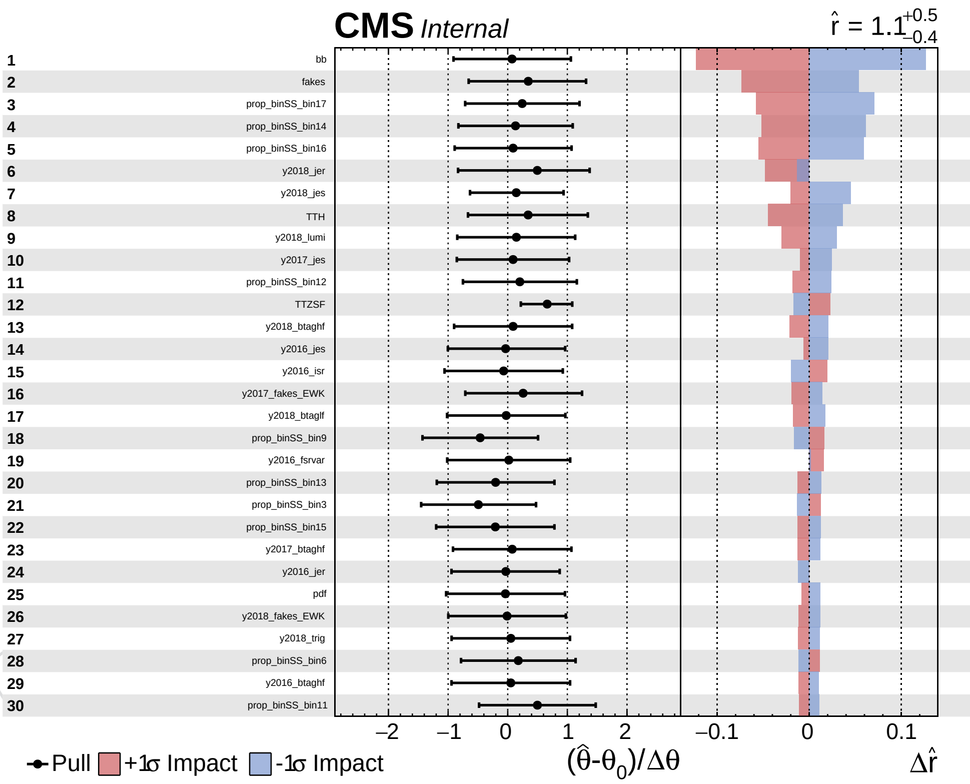


Figure 120: Observed nuisance impacts for the BDT-based analysis.

H.5 Goodness of fits

The goodness of fit distributions (using the saturated, Kolmogorov-Smirnov, and Anderson-Darling test statistics) with the signal+background fit to data for the cut-based and BDT analyses are shown in Figure 121. We note that the observation is generally within the bulk of the expected distributions for both analyses and all three fit types.

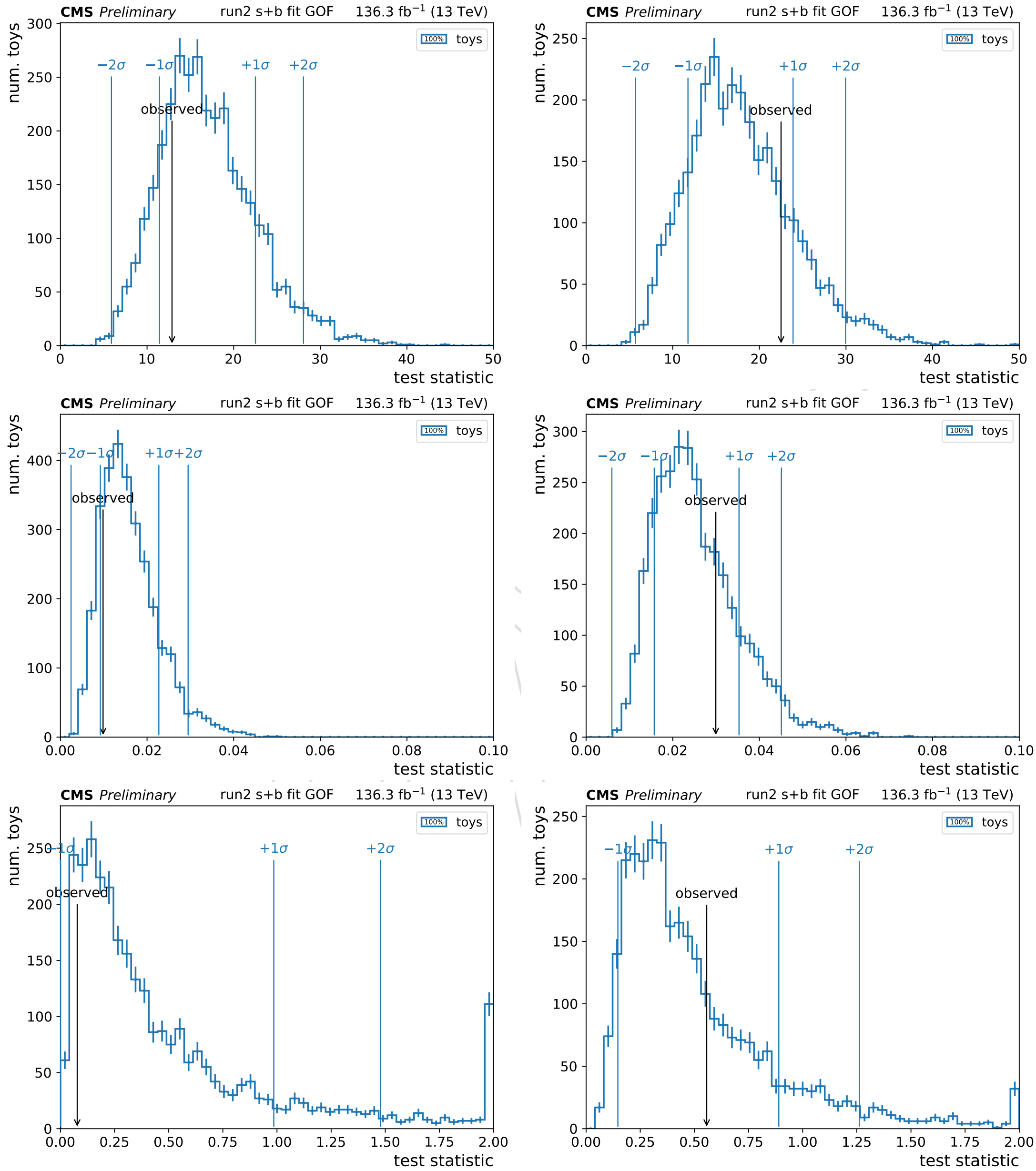


Figure 121: GOF test for the cut-based analysis (left) and BDT-based analysis (right) using the saturated (top), KS (middle), AD (bottom) test statistics.

Table 66: Event kinematics for tail BDT events

run:lumi:event	SRnum	year	nleps	id1	id2	pT1	pT2	eta1	eta2	phi1	phi2	Nj	Nb	MET	HT
276776:1569:2615397660	17	2016	2	11	13	131.313	71.6004	1.54234	0.574022	-0.434035	-1.93809	7	3	256.921	1458.27
283050:58:92826095	17	2016	2	-13	-11	48.3626	39.4692	-0.217067	-2.34033	2.24725	-0.568915	7	3	168.629	844.845
305377:944:1751515966	17	2017	2	13	13	146.593	117.307	-1.86365	0.71062	2.92705	0.0105647	7	3	212.99	836.855
315644:87:107703005	17	2018	2	-13	-13	102.218	24.2424	-1.61099	-0.637295	1.11559	-1.64171	7	5	82.6389	754.944
316240:915:1292896376	17	2018	2	11	13	46.3985	30.2089	0.580867	-1.1041	0.0199213	3.09587	7	4	56.2346	632.974
317435:1373:1969850764	17	2018	3	-11	-11	62.5848	38.0043	0.175213	0.108995	2.11431	-0.932109	9	2	95.8658	1622.7
321051:870:1284078109	18	2018	2	13	13	55.4322	50.6932	1.42346	0.887908	2.50299	-1.42865	8	4	78.4541	558.675
321774:50:85657761	18	2018	2	-11	-11	112.112	107.589	1.17477	0.895216	1.18255	-2.27924	9	4	245.434	1466.6
322356:79:153159025	18	2018	2	-11	-13	54.3049	31.4315	-0.017497	-0.924084	1.47584	1.63024	8	4	174.04	650.043
323727:421:771405322	17	2018	2	11	11	55.2933	42.2259	-0.0188495	-1.35181	2.1227	-1.93872	7	3	126.791	592.848

H.6 Tail BDT events

We included event information about the 7+3 events in the two tail BDT bins in Table 66. As expected, the high number of jets and b-tag multiplicity contribute to a high BDT score. We don't see any localization in the detector (for example, due to HEM).

I Checks from ARC review

I.1 Additional uncertainty from tt+bb

Figure 126 shows the fraction of events affected by the correction to account for the measured ratio of $t\bar{t}b\bar{b}/t\bar{t}jj$ cross sections (1.7 ± 0.6), as discussed in Section 2. $t\bar{t}$ and $t\bar{t}Z$ show agreement in this fraction (approximately 1%), while $t\bar{t}W$ is low by 30%. To study possible effects due to differences in kinematics with respect to $t\bar{t}$, we tested a configuration where an additional 30% uncertainty was included in quadrature, s o that the ratio and corresponding systematic uncertainty become 1.7 ± 0.78. The effect on the analysis was to reduce the observed signal strength of the cut-based and BDT analyses by 0.5%, and the expected (observed) significance of both analyses by 1% (1.5%). This additional uncertainty was eventually not adopted in the analysis.

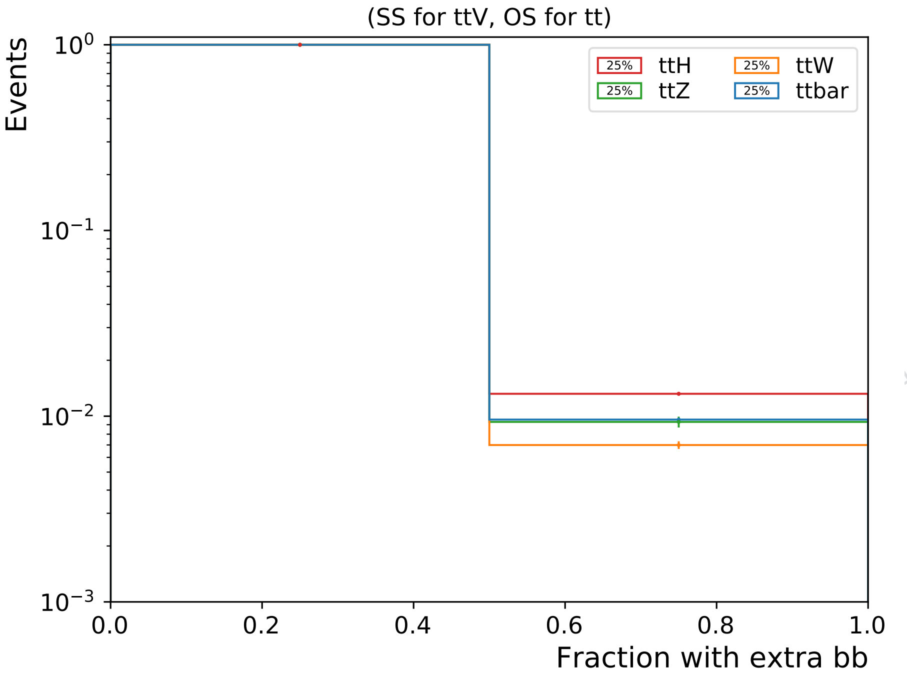


Figure 122: Fraction of events rewighted by measured ratio of $t\bar{t}b\bar{b}/t\bar{t}jj$ for $t\bar{t}$, $t\bar{t}W$, and $t\bar{t}Z$, inclusively. Note that the same-sign requirement has been lifted from $t\bar{t}$, which requires opposite-sign dileptons.

I.2 Distributions of BDT inputs for signal region events

Figure 123 shows distributions of the 19 BDT input variables for the same events that enter the BDT signal regions, except for CRZ. They are shown with "postfit yields", meaning prefit shapes are used, but normalization of processes is adjusted to match postfit/prefit scale factors. Good agreement is observed.

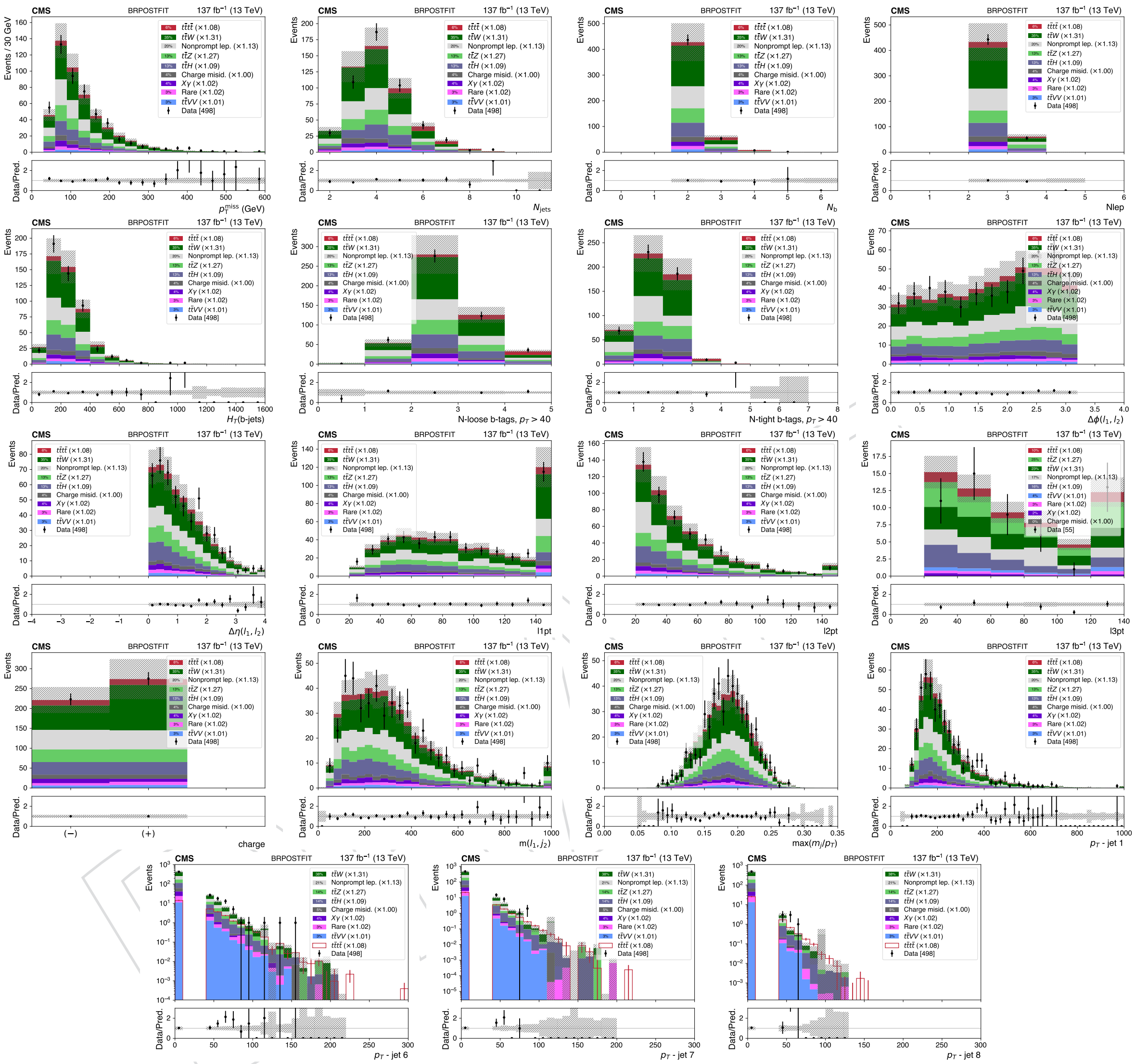


Figure 123: Run2 Data to prediction comparisons for the BDT input variables. From left top to right bottom, $\not{E}_{\mathrm{T}}$, N_{jets}, $N_{\mathrm{b\ jets}}$ N_{leps}, H_T^b, Nlooseb, Ntightb, $\Delta\phi(\ell_1, \ell_2)$, $\Delta\eta(\ell_1, \ell_2)$, $p_T(\ell_1)$, $p_T(\ell_2)$, $p_T(\ell_3)$, q_1, $m(\ell_1, j_1)$, $\max(m(j)/p_T(j))$ and the p_{T} for jets 1,6,7, and 8, are shown for the BDT signal region events excluding CRZ, with scaled $\mathrm{t\bar{t}t\bar{t}}$ signal stacked when in solid red.

I.3 More jets at higher η

The analysis has a jet requirement of $|\eta| < 2.4$ to keep (b)jets almost completely within tracker acceptance. Additionally, b-tag scale factor uncertainties are provided up to $|\eta|$ of 2.4 (2.5) for 2016 (2017, 2018) according to `https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation`, and reflected in the .csv files.

While optimizing the Run 2 analysis, we considered relaxing both the p_T and $|\eta|$ requirements for jets and b-jets. First, we studied the additional acceptance for signal jets when relaxing these cuts. The additional jet acceptance from relaxing the $|\eta|$ requirement is small. Figure 124 shows the $|\eta|$ distribution of $p_T > 30$GeV generator-level jets which are matched to top decay products (light-flavor quarks and b quarks) in $t\bar{t}t\bar{t}$ events with 2 leptonically-decaying W bosons. The left plot shows an additional 3.8% of jets would be gained by completely relaxing the $|\eta|$ requirement up to approximately 3.6. The right plot shows up to 0.8% of jets would be gained by relaxing $|\eta|$ to 2.5, in order to stay within an appropriate range of the b-tag scale factors and allow for well-behaved JEC factors. Figure 125 shows the significantly more impressive gain from relaxing p_T thresholds instead of η.

As the gain from relaxing p_T thresholds was much larger, we put aside the η cut and focused on the p_T cut. This resulted in the studies shown in Appendix C, where the whole analysis was re-run with varying p_T cuts, signal and background counts were re-evaluated, but the significance was not found to improve. For these reasons, we decided not to change our thresholds with respect to the 2016 analysis.

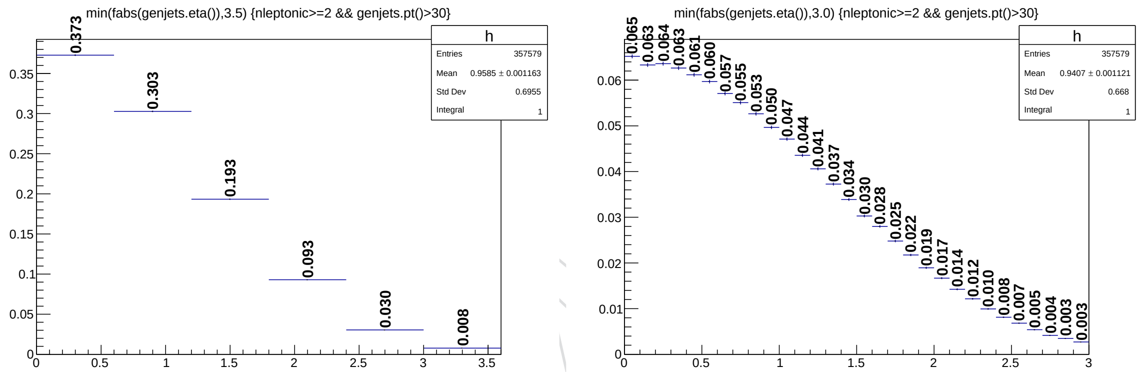


Figure 124: $|\eta|$ distribution of generator-level jets which are matched to top decay products for 2 lepton $t\bar{t}t\bar{t}$ events with coarse-binning (left) and fine-binning (right). The plots are normalized to unit area and bin values are superimposed.

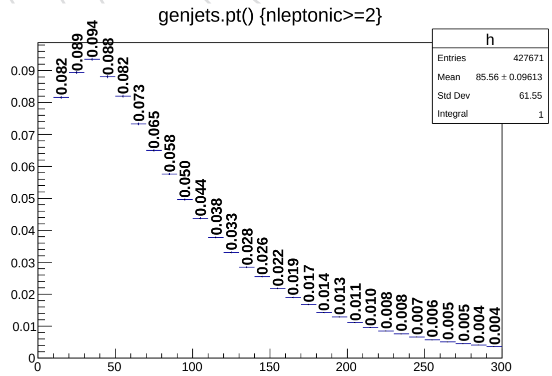


Figure 125: p_T distribution of generator-level jets which are matched to top decay products for 2 lepton $t\bar{t}t\bar{t}$ events.

I.4 Effect of ttH and ttbb/ttjj measurement updates

In this section, we quantify the effects of potential updates to the $t\bar{t}H$ and ttbb measurements on this analysis.

I.4.1 ttH

As we do not scale the $\mathrm{t\bar{t}H}$ background, an updated measurement of $\mathrm{t\bar{t}H}$ would only change the 25% relative uncertainty we place on this process. The effect of modifying this uncertainty can be seen in the impact plots from Appendix A, which shows the $\mathrm{t\bar{t}H}$ normalization uncertainty nuisance is up to a 6% effect on the measured signal strength, approximately. Changes with respect to the 25% normalization uncertainty can then be extrapolated from this 6%.

I.4.2 $\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$

Conversely, an updated measurement of $\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$ can change the central value of background predictions. Currently, the $\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$ ratio is 1.7 ± 0.6. Based on TOP-18-002 (`http://cms.cern.ch/iCMS/analysisadmin/cadilines?line=TOP-18-002`), the $\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$ ratio in the visible phase space is 1.17 ± 0.16 for the lepton+jets channels, and 1.38 ± 0.24 for the dilepton channels. While a numerical combination is not provided, a crude estimate using the average would be around 1.30 ± 0.20. Thus, we use 1.3 ± 0.2 and re-evaluate the results of this analysis. Numbers for both values/uncertainties of $\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$ are tabulated in Table 67. The smaller scale factor from $\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$ results in 10% higher expected and observed significance values on average.

BDT			
$\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$	exp. σ	obs. σ	obs. μ
1.7 ± 0.6	2.699	2.561	$1.052^{+0.483}_{-0.438}$
1.3 ± 0.2	2.904	2.841	$1.108^{+0.470}_{-0.420}$

Cut-based			
$\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$	exp. σ	obs. σ	obs. μ
1.7 ± 0.6	2.478	1.713	$0.784^{+0.514}_{-0.469}$
1.3 ± 0.2	2.629	1.968	$0.852^{+0.498}_{-0.450}$

Table 67

I.5 tt+bb correction applied to ttV samples

Figure 126 shows the fraction of events affected by the correction to account for the measured ratio of $\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}$ cross sections (1.7 ± 0.6), as discussed in Section 2. $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}Z}$ show agreement in this fraction (approximately 1%), while $\mathrm{t\bar{t}W}$ is low by 30%. To study possible effects due to differences in kinematics with respect to $\mathrm{t\bar{t}}$, we tested a configuration where an additional 30% uncertainty was included in quadrature, s o that the ratio and corresponding systematic uncertainty become 1.7 ± 0.78. The effect on the analysis was to reduce the observed signal strength of the cut-based and BDT analyses by 0.5%, and the expected (observed) significance of both analyses by 1% (1.5%). This additional uncertainty was eventually not adopted in the analysis.

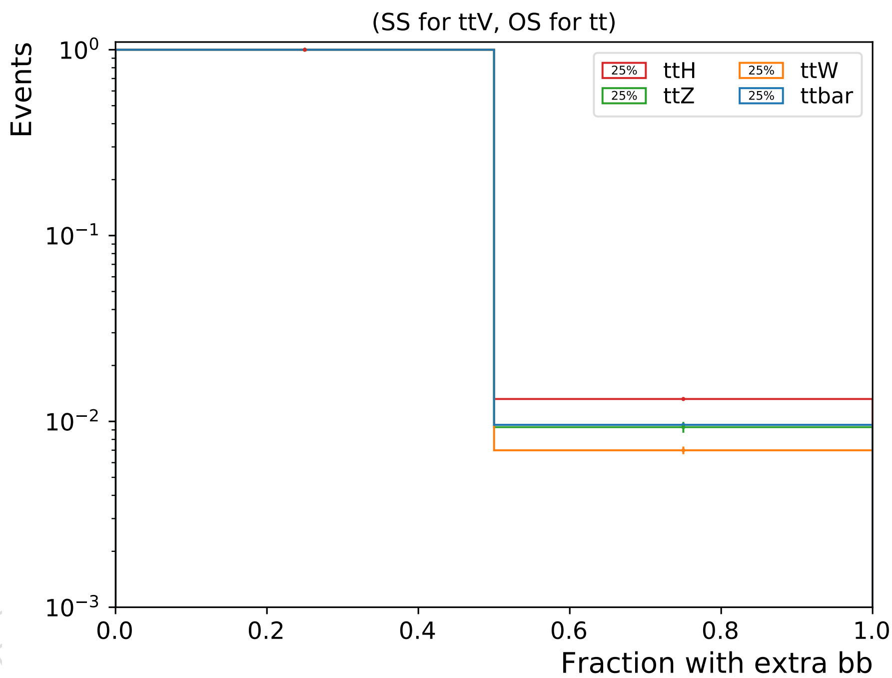


Figure 126: Fraction of events rewighted by measured ratio of $t\bar{t}b\bar{b}/t\bar{t}jj$ for $t\bar{t}$, $t\bar{t}W$, and $t\bar{t}Z$, inclusively. Note that the same-sign requirement has been lifted from $t\bar{t}$, which requires opposite-sign dileptons.

I.6 $N_J^{\mathrm{ISR/FSR}}$ systematic

As introduced in Section 2, $\mathrm{t\bar{t}W}$ and $\mathrm{t\bar{t}Z}$ MC events are reweighted based on $N_J^{\mathrm{ISR/FSR}}$, and a systematic corresponding to half of the reweighting factor deviation from unity is used (Section 9).

The size of the systematic comes from deviations observed in an orthogonal single-lepton channel.

The tunes for 2016 (TuneCUETP8M1) and 2017/2018 (TuneCP5) are different, so both need to be checked separately. For the case of 2016, the study documented in Ref. [8] assesses a systematic of half of the reweighting factor deviation from unity. Below, we check the 2018 dataset explicitly to verify this is also the case for the new tune, TuneCP5.

Following Ref. [8], we make a tight single-lepton selection with analysis jet, b-tagging, and lepton selections. Selected events have lepton $p_{\mathrm{T}} > 25$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV, $H_{\mathrm{T}} > 300$ GeV, and 2 b tagged jets. MC integral is normalized to data to account for trigger and lepton scale factors.

Figure 127 shows the jet multiplicity distribution for 2018 data/MC before (left) and after (right) applying the 2018-derived reweighting factors. The systematic uncertainty of half of the deviation from unity covers variations present in the reweighted single lepton jet multiplicity distributions for the updated TuneCP5.

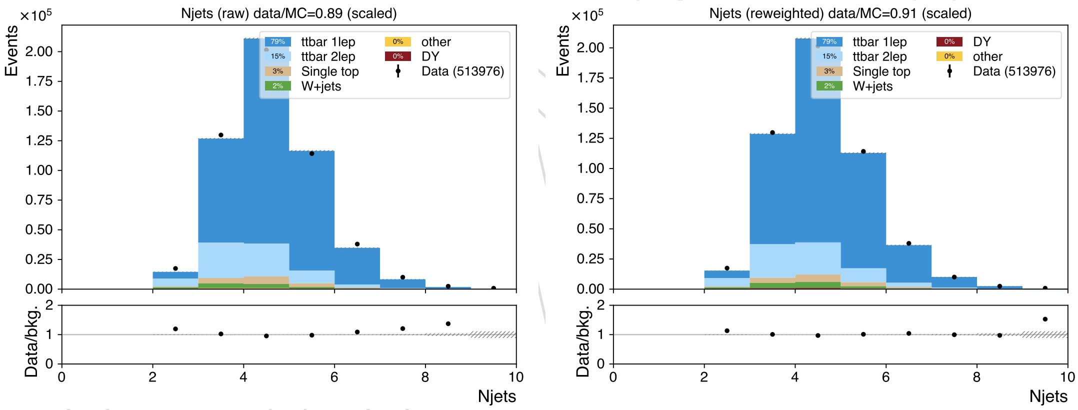


Figure 127: Jet multiplicity distribution for 2018 data/MC before (left) and after (right) applying the 2018-derived reweighting factors. The “other” category includes VV, $\mathrm{t\bar{t}W}$, and $\mathrm{t\bar{t}Z}$ contributions.