

Measurement of CP-violating Wtb couplings using EFT in single top t -channel production with CMS Run 2 data

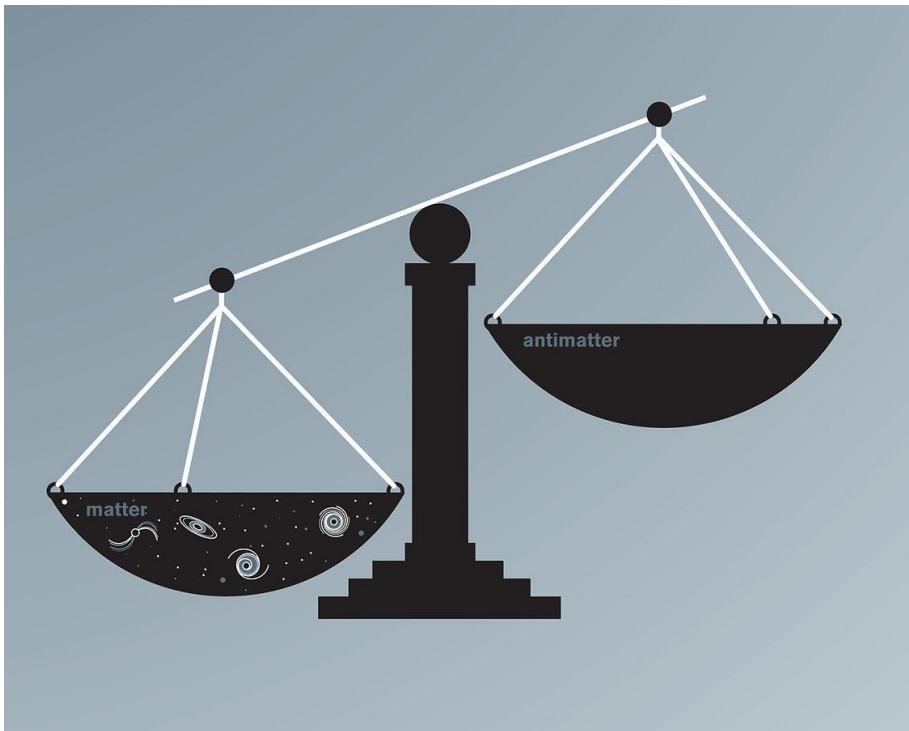
Top LHC France
27/05/2026

Nicolas Chanon, Arnab Purohit, Enzo Fillaudeau (PhD)

CMS group at IP2I Lyon

Motivation : Matter-antimatter asymmetry

Asymmetry of the baryon-number in the Universe

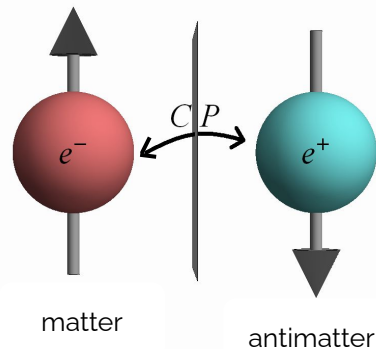


SM predictions : $\eta_{SM} = \frac{n_B - n_{\bar{B}}}{n_\gamma} \propto 10^{-20}$

Observations : $\eta_{obs} = \frac{n_B - n_{\bar{B}}}{n_\gamma} \propto 10^{-10}$

$$\Rightarrow \frac{\eta_{SM}}{\eta_{obs}} \propto 10^{-10}$$

Discrepancy between the SM prediction and observations

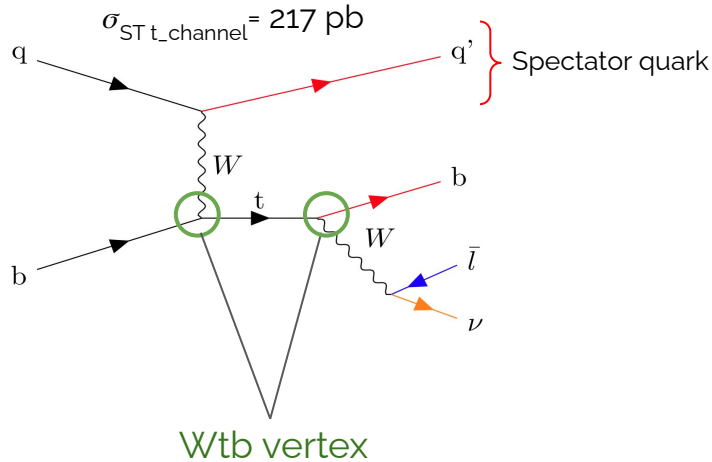


C : Charge conjugation $Q \longrightarrow -Q$

P : Parity transformation $\vec{r} \longrightarrow -\vec{r}$

This asymmetry can be explained by the existence of new sources of CP violation during baryogenesis.

Measure CP violation through the t-channel production of single-top quarks and their subsequent decay to Wb . *This process has the advantage of involving the Wtb vertex in production and decay*



Analysis performed in the context of Effective Field Theory

$$\mathcal{L}_{eff}^{(6)} = \mathcal{L}_{SM} + \sum_i \frac{C_i^{(6)}}{\Lambda_i^2} O_i^{(6)} + h.c.$$

Non-Hermitian Operators
involving two quarks and a boson

$$O_{uW}^{(6)} = (\bar{q}\sigma^{\mu\nu}\tau^I t)\tilde{\varphi}W_{\mu\nu}^I$$

$$O_{dW}^{(6)} = (\bar{q}\sigma^{\mu\nu}\tau^I b)\varphi W_{\mu\nu}^I$$

$$O_{qud}^{(6)} = (\tilde{\varphi}^\dagger iD_\mu\varphi)(\bar{u}\gamma^\mu d)$$

Wilson Coefficients

$$c_{tW} + ic_{tW}^I$$

$$c_{bW} + ic_{bW}^I$$

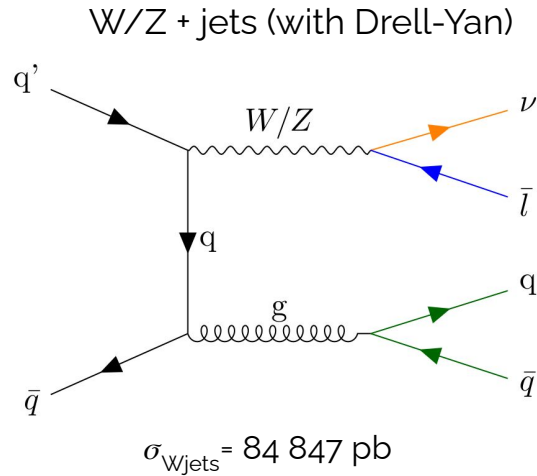
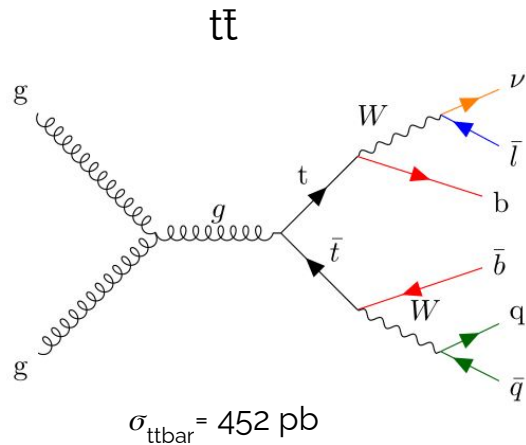
$$c_{\varphi tb} + ic_{\varphi tb}^I$$

CP violation if

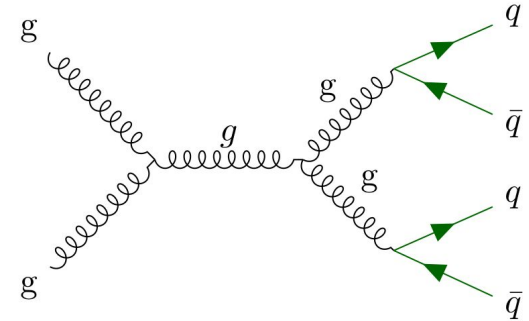
$$c_{tW}^i \neq 0, c_{bW}^i \neq 0 \text{ or } c_{\varphi tb}^i \neq 0$$

- EFT impacts both the production and decay of top quark
- This vertex can be modified by CP violation

Main background processes :



QCD processes



Other backgrounds :

- Diboson
- $t\bar{t}X$
- Single top s-channel and tW process

- **Single Top/AntiTop t-channel (SM)** : POWHEG + MadSpin + Pythia8, scheme 4F, TuneCP5
- **t $\bar{t}$ semileptonic** : POWHEG + Pythia8, TuneCP5
- **tW top and anti-top** : POWHEG + Pythia8, scheme 5F, inclusive decays, TuneCP5
- **W + Jets** :
 - *old sample* : MG5_aMC@NLO FxFx + Pythia8, jet-binned 0,1,2j, TuneCP5
 - *new sample* : POWHEG MiNNLO + Pythia8 + PHOTOS (W⁺/W⁻ split, TuneCP5)
 - The new sample is the one used in the W mass measurement analysis and has a larger statistical sample size.

EFT samples :

- Generated using the dim6Top LO UFO model from Madgraph5
- Using dim6=2 to insert EFT in top production and/or decay (including interferences) while maintaining quadratic behaviour
- A reweighting technique is used to get all the EFT combination

The **single isolated muon** HLT trigger with $p_T > 24$ GeV (2016, 2018) and $p_T > 27$ GeV (2017) are used in this analysis

Jet selection

	Good jets
$ \eta $	< 4.7
p_T 2016 (GeV)	> 40
p_T 2017 and 2018 (GeV)	> 40 ($ \eta < 2.4$) > 60 ($2.4 < \eta < 4.7$)
Overlap	Removed overlap between jets and leptons in a $\Delta R < 0.4$ cone
Jet Id	Tight (discriminate real jets from fake lepton, pile up and detector noise)

b-jets :

- Must be good jets
- $|\eta| < 2.5$ (2017 & 2018)
- $|\eta| < 2.4$ (2016)
- Tight/medium working point of DeepJet tagger used

Muon selection

	Isolated muon	Loose muon
$ \eta $	< 2.4	< 2.4
p_T 2017 (GeV)	> 30	> 10
p_T 2016 and 2018 (GeV)	> 26	> 10
Relative isolation	$< 15\%$	$< 25\%$
Id	tight	loose

A reversed isolated muon is also defined by reverting the isolation : $> 15\%$

→ All recommended CMS corrections are applied (including pileup and b-tag weights, muon Rochester momentum correction, JEC, MET phi modulation)

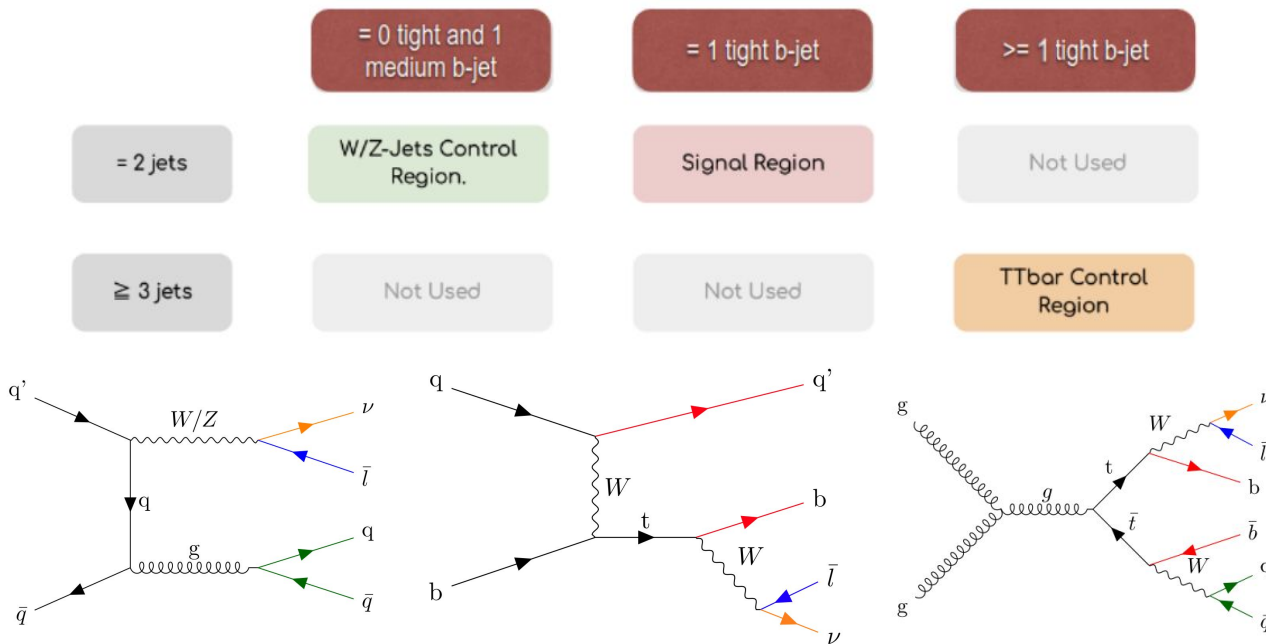
Loose preselection applied to reduce fake lepton contribution :

- **MET** > 20 GeV
- **Transverse mass of W boson with fixed lepton p_T at 45 GeV ($M_T^{W,fix}$)** > 25 GeV

Lepton selection :

- Exactly one isolated tight muon
- Veto events with additional loose muons
- Veto events with veto electrons

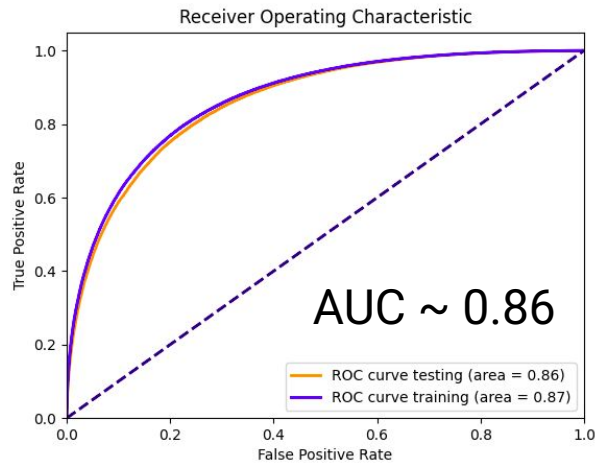
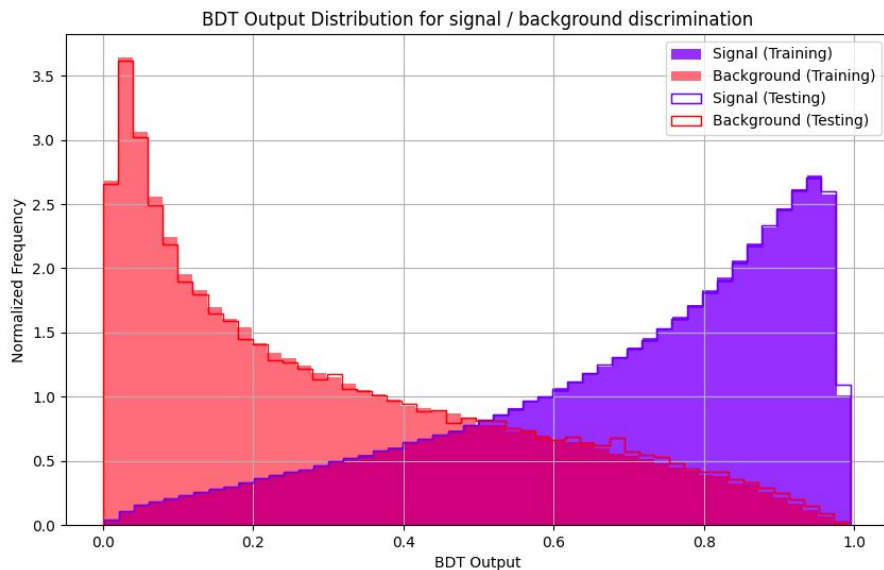
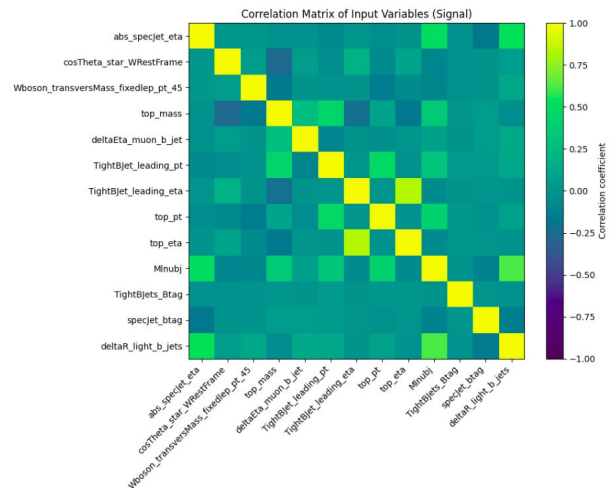
Event categorization based on the number of Jets and b-tagged jets : Signal region (SR) , W/Z-Jets control region (CR) and $t\bar{t}$ CR



Analysis strategy - MVA in the **Signal Region**

Trained a **BDT** in the SR to discriminate **single top (ST) t-channel** from **W+jets and QCD processes**

Input variables for the BDT : $\cos(\theta^*)$, η of the spectator jet, spec jet btagging score, $M_T W^{fix}$, top p_T , top η , M_{top} , $\Delta\eta(\mu, \text{b-jet})$, $\Delta R(\text{spec jet}, \text{b-jet})$, b-jet p_T , b-jet η , b-jet btagging score, lepton p_T , M_{lvbj}



A QCD measurement region (sideband region) is defined for each signal and control region, with the same selection but reverted muon isolation cut.

Observed limitation :

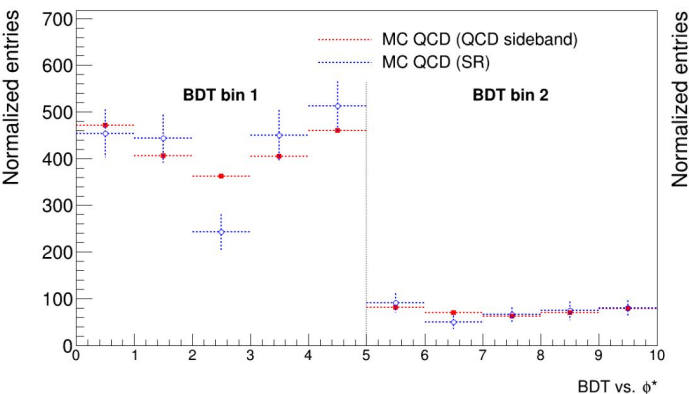
- Large impact of the MC stat uncertainty on the sensitivity of the analysis
- Traced back to the QCD process (see table of the relative contribution to the total MC stat uncertainty from each process in the SR)
- QCD sideband needs to be further populated to reduce uncertainties

Process	σ_{MC}	Fraction (%)	σ_{tot}
qcd ₂₀₁₈	3763.5583	97.82	3805.1868
vjets	522.6826	1.89	
ttbar	145.8065	0.15	
singletop	114.6261	0.09	
sm	80.4689	0.04	
diboson	31.7001	0.01	

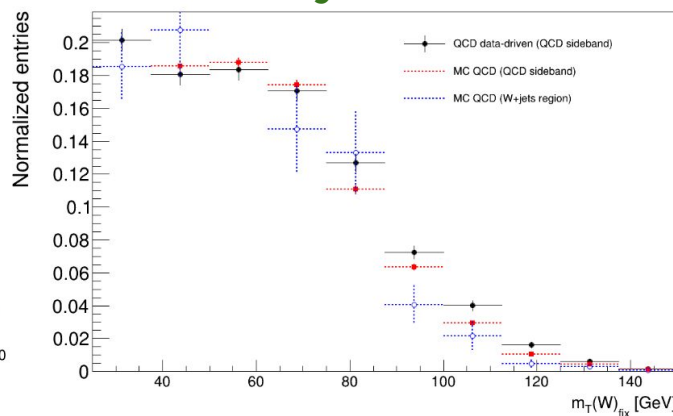
Applied modifications :

- ❖ Any **BDT cut** applied in the SR is **removed in the QCD sideband**
- ❖ The shape of data-driven QCD with and without BDT cut for the ϕ^* distribution agrees within uncertainties

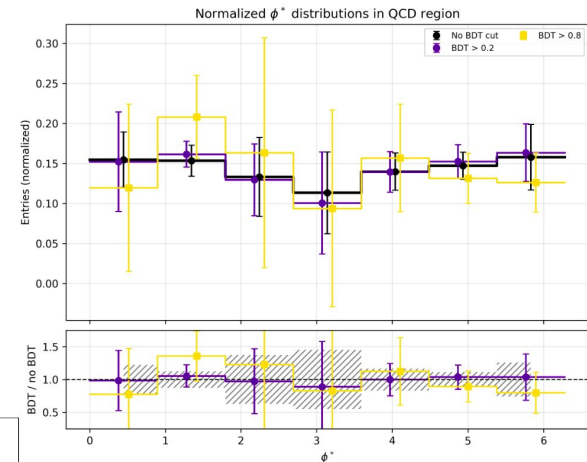
SR



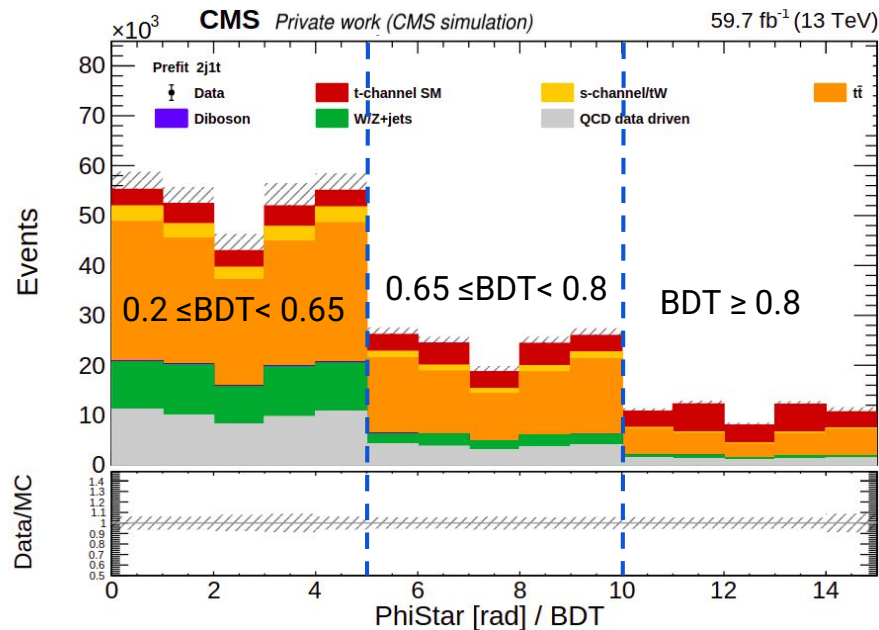
W+jets CR



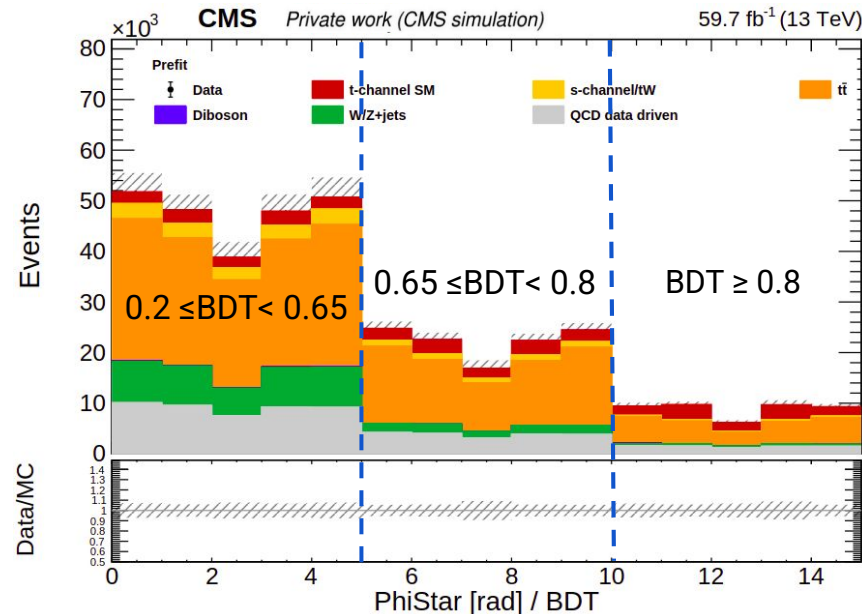
QCD MC in the QCD sideband matches the QCD MC in the SR and CR, validating the method



- The SR is divided into Top and Antitop channels (determined from the lepton charge)
- As the measurement observable we are using the ϕ^* distribution in bins of the BDT

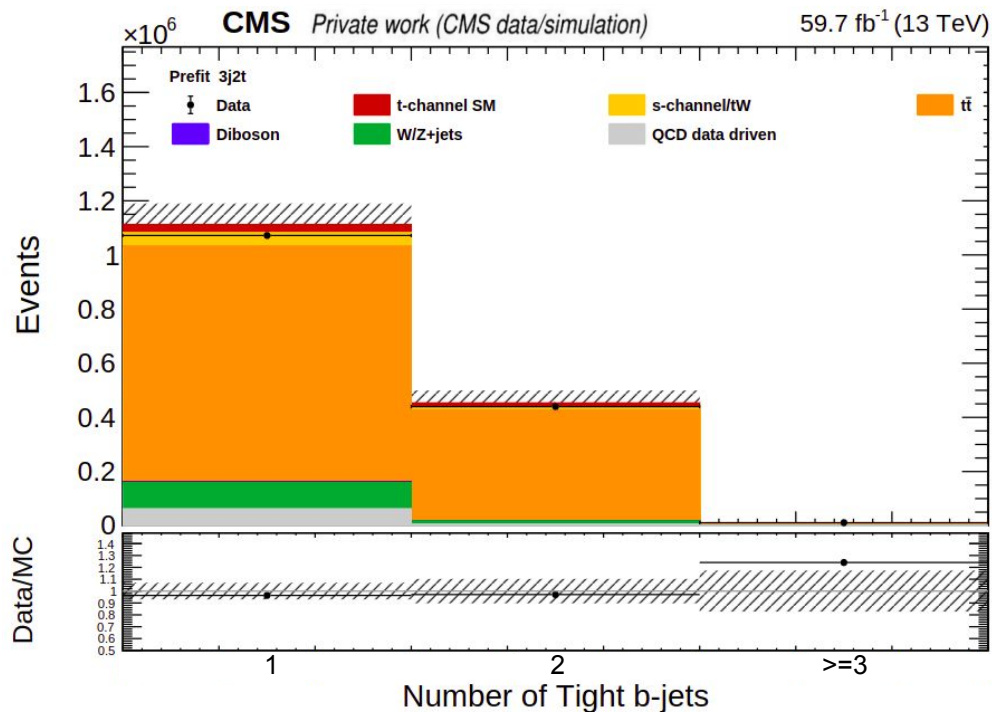


Top channel

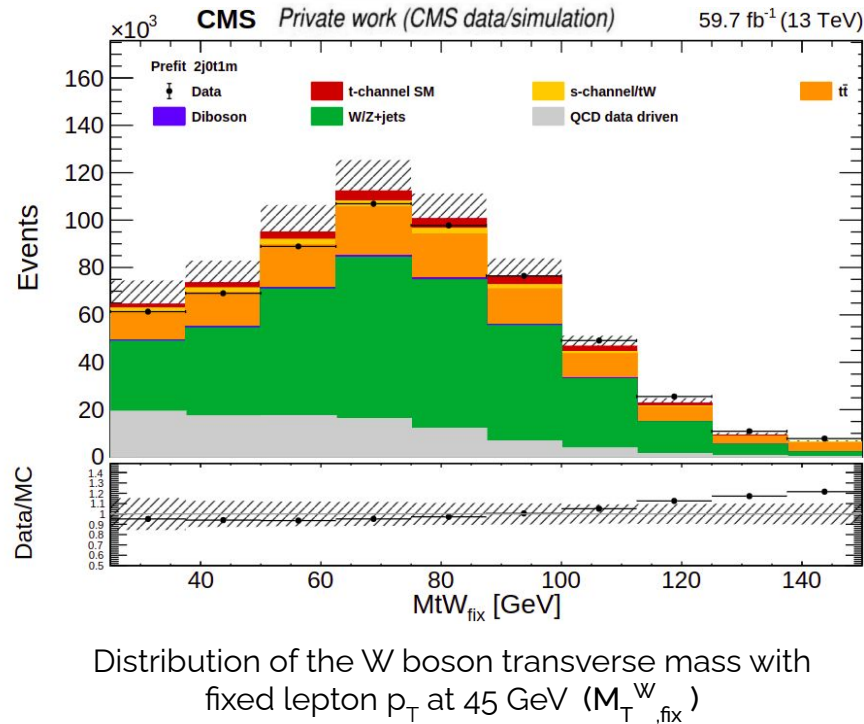


Antitop channel

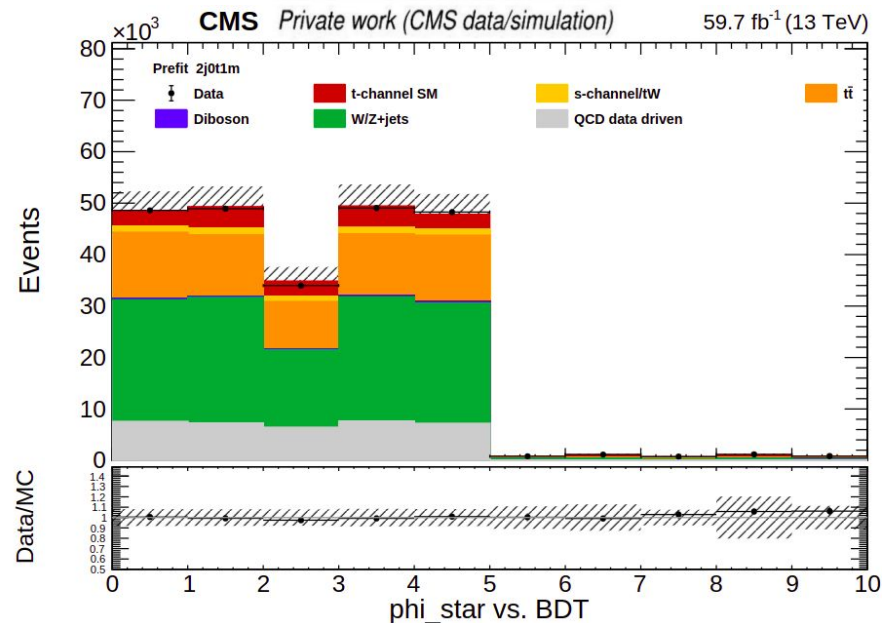
In the $t\bar{t}$ control region, we are using the **number of tight b-jets** (3 bins) as our discriminating distribution.



Current strategy



Modeling of the ϕ^* in bins of BDT observable in the W +jets CR where the data is unblinded

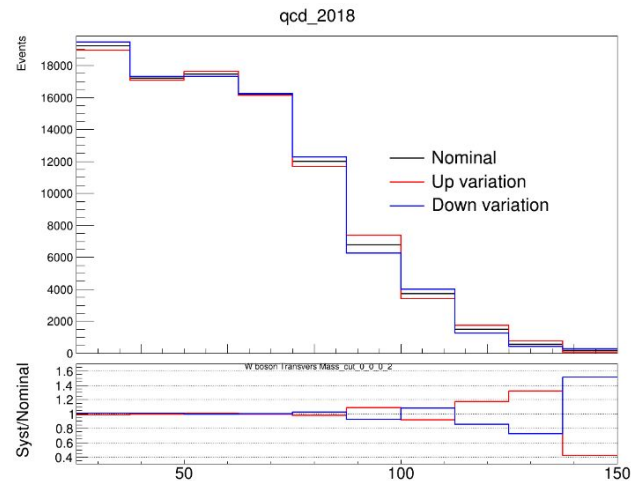


Systematic uncertainties

NAMES	USED	CORRELATION
Muon Reco	Y	All channel
Muon HLT	Y	All channel
Muon ID	Y	All channel
Muon Iso	Y	All channel
B Tagging SF	Y	All channel
Prefiring	Y	All channel
Pile Up	Y	All channel
Top pt	Y	Only TTbar
Lumi	Y	All channel
JES/JER	Y	As per Recom
ISR	Y	Not correlated
FSR	Y	Not correlated
PDF + alpha_s	Y	All channels
QCD Scale	Y	All channels

We introduce 3 uncertainties dedicated to this analysis :

- A shape uncertainty on QCD data-driven distribution, constructed by varying the **run era** (uncorrelated between years)
- A shape **difference between QCD data-driven and MC QCD** (uncorrelated between years and regions)
- Difference between POWHEG ST t-channel in the **4fs and 5fs** (fully correlated)



Down : 2018 A, B, C era
Up : 2018 D era

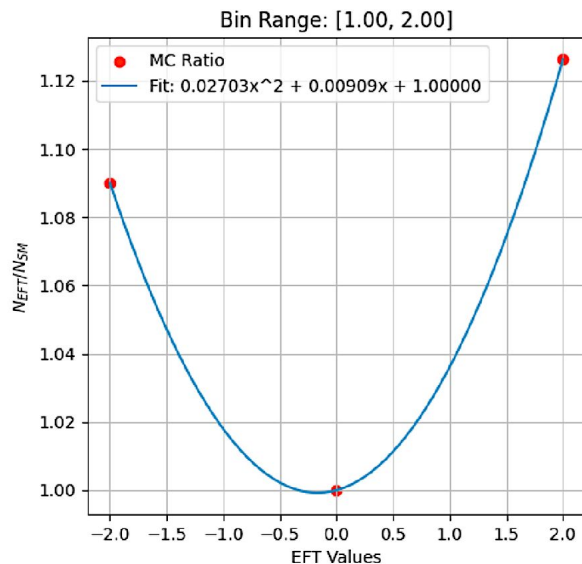
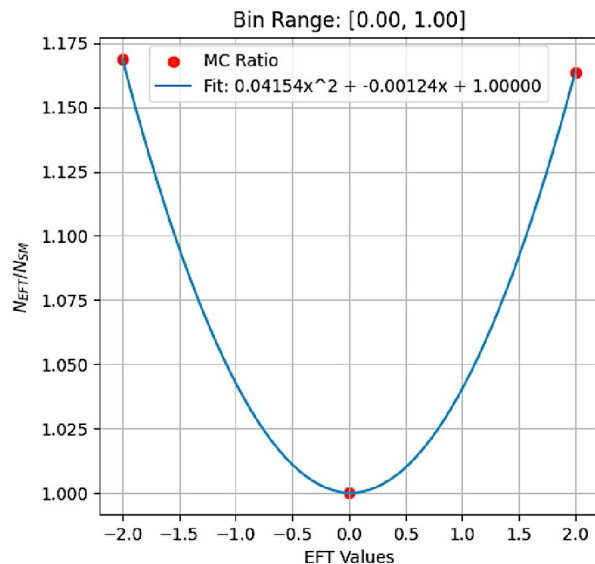
$M_T^{W,fix}$ distribution with Up and Down variations of QCD era uncertainty

Evaluating the expected statistical sensitivity by fitting the SM+EFT model to the Asimov (SM) dataset using a profile likelihood statistical method (CMS Combine)

Using AnalyticAnomalousCoupling plugging to model the EFT template :

<https://github.com/amassiro/AnalyticAnomalousCoupling>

This framework provides a Combine-based model for EFT fits overcoming the issues of negative templates that may arise from interference terms



- Fitting polynomial of order 2 in each bin of our observable
- Using the three parameters of the fit to produce the templates

$$Sm = |\mathcal{A}_{SM}|^2,$$

$$Lin_\alpha = 2\Re(\mathcal{A}_{SM}\mathcal{A}_{Q_\alpha}^\dagger)$$

$$Quad_\alpha = |\mathcal{A}_{Q_\alpha}|^2$$

A profile likelihood scan is performed with Asimov dataset. A best-fit value (at 68% CL) including expected statistical and systematic uncertainties is obtained.

This analysis

Source	C_{tW}^I	C_{bW}^I
Best fit	-0.00	+0.00
Total unc.	+0.09 / - 0.09	+1.08 / - 1.13
Stat. unc.	+0.08 / - 0.08	+0.44 / - 0.44
MC stat. unc.	+0.04 / - 0.04	+0.43 / - 0.48
Other syst. unc.	+0.01 / - 0.02	+0.88 / - 0.92

Expected sensitivity by ATLAS analysis is still better by a factor ~ 1.5 for C_{tW}^I and C_{bW}^I compared to our analysis.

MC stat is also the largest systematic in the ATLAS analysis.

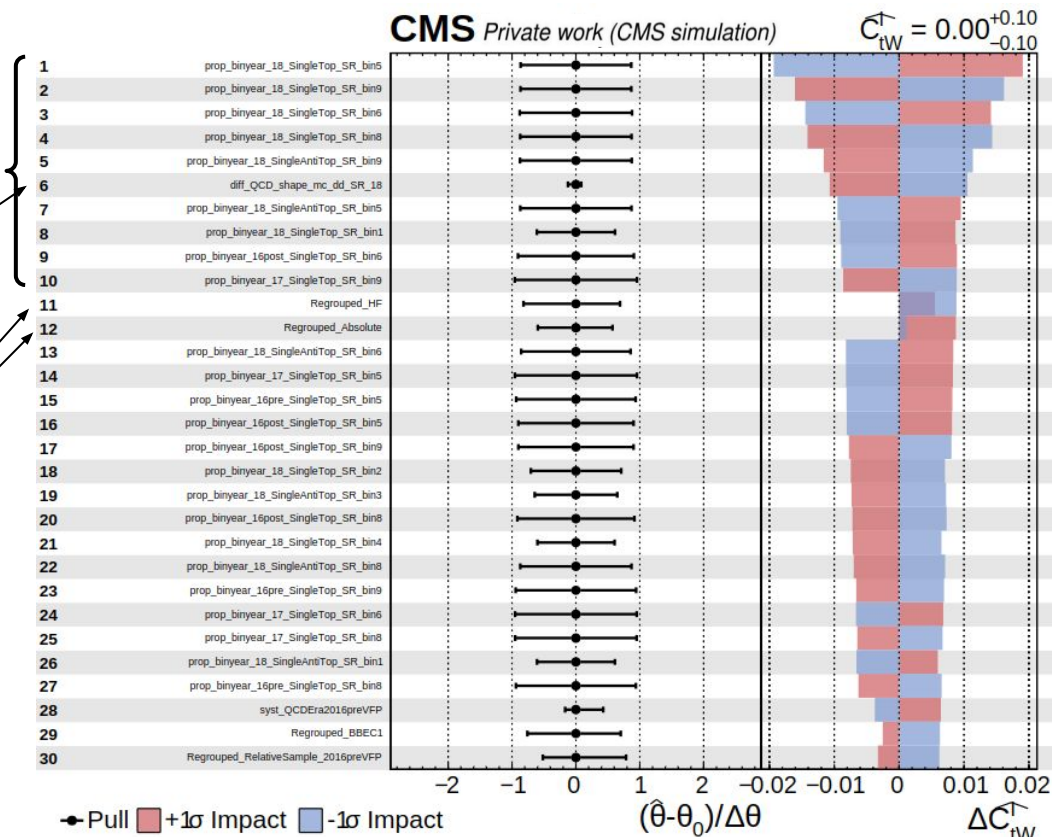
ATLAS analysis [arXiv:2510.23372](https://arxiv.org/abs/2510.23372)

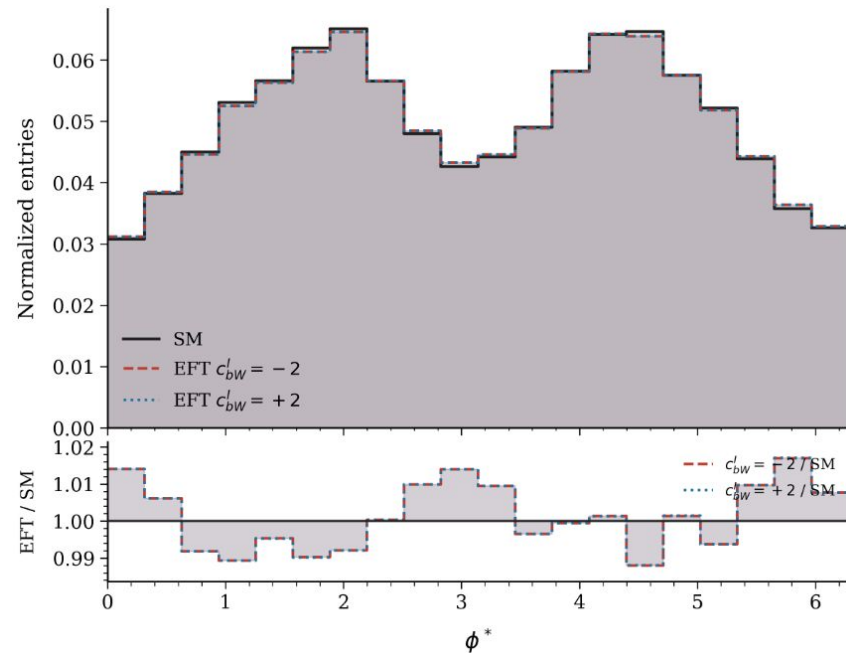
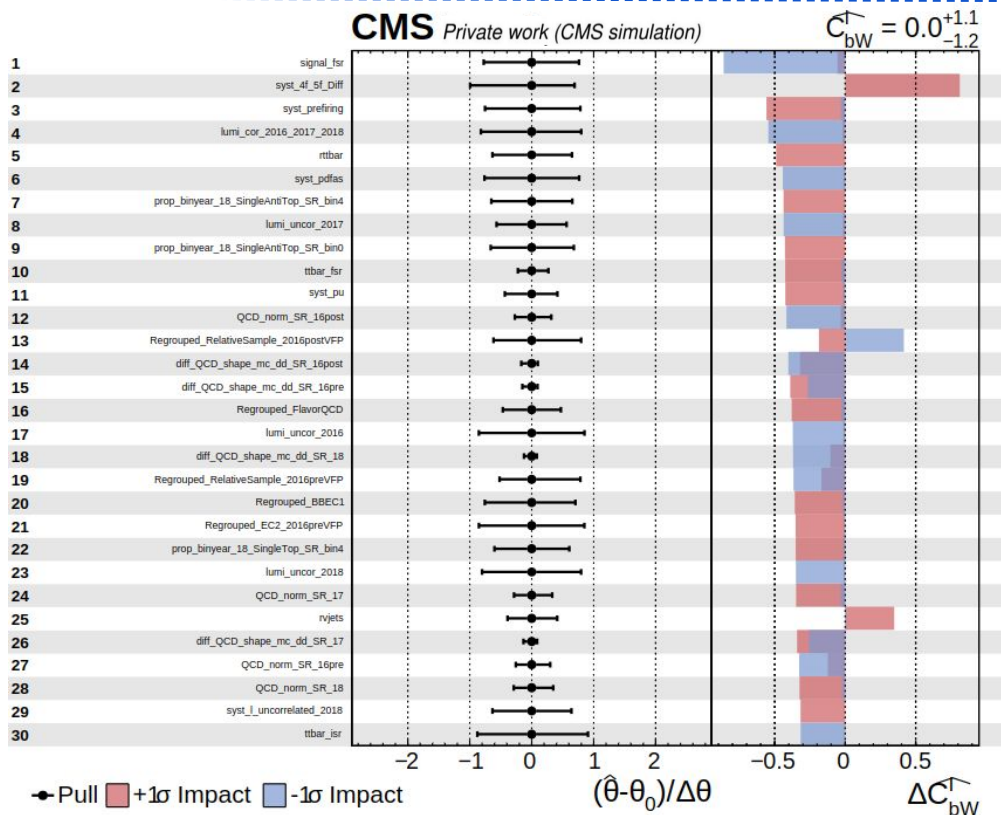
Source of uncertainty	C_{tW}^I	C_{bW}^I
Jet energy scale	-0.01, +0.01	-0.3, +0.3
Pile-up	-	-0.1, -
b -tagging	-0.01, +0.01	-0.1, +0.1
Jet energy resolution	-0.01, +0.01	-0.1, +0.1
PDF	-0.01, +0.01	-0.1, +0.1
Muon	-	-0.1, +0.1
E_T^{miss}	-	-0.2, +0.2
e/γ	-	-0.1, +0.1
Jet vertex tagging	-	-
Lumi. & cross-section	-	-0.1, +0.1
τ_{had} energy scale	-	-
Non-prompt leptons	-	-0.1, +0.1
t -channel modelling	-	-
$t\bar{t}$ modelling	-0.01, -	-0.1, +0.1
W +jets modelling	-0.01, +0.01	-, +0.1
MC sample sizes	-0.03, +0.03	-0.2, +0.2
Data sample size	-0.04, +0.04	-0.5, +0.5
Total	-0.06, +0.06	-0.7, +0.7

MC statistical uncertainties remain among the highest-ranked nuisance parameters, but their overall impact is significantly reduced compared to the previous strategy

Uncertainty on the shape **difference between QCD data-driven and MC QCD** in the SR

Need to move to full set of JEC uncertainties (reduced one is constrained)





Many nuisances are one sided because $C_{bW}^I = +2$ and $C_{bW}^I = -2$ effect goes in the same direction.

Overview of the analysis :

- Produced private EFT samples
- Divided the analysis into signal (top and antitop channel) and control regions (W+jets and $t\bar{t}$ CR)
- Trained a BDT to improve the sensitivity of the analysis on EFT measurement
- Developed a QCD estimation strategy
- Performed the fit for 2 Wilson coefficients : C_{tW}^I and C_{bW}^I

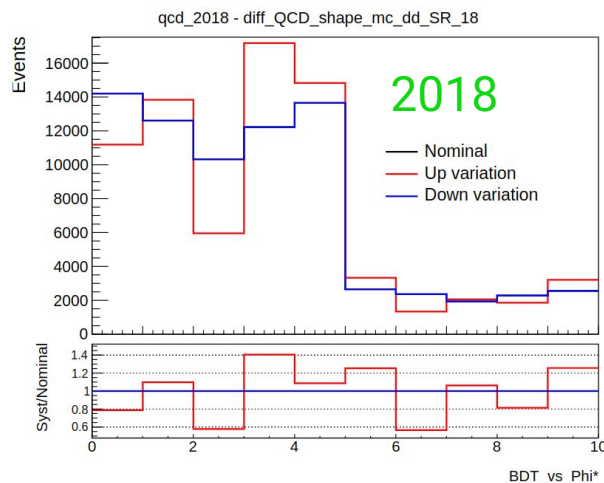
Plans :

- The binning of the BDT in the SR still needs to be optimized
- Consider moving to tight lepton MVA (lepton identification with machine learning technique)
- Add the measurement of the third coefficient : $C_{\phi tb}^I$

Thank you for your attention

Backup

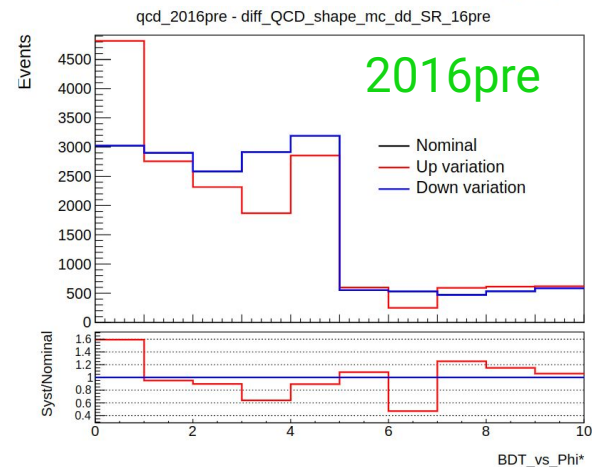
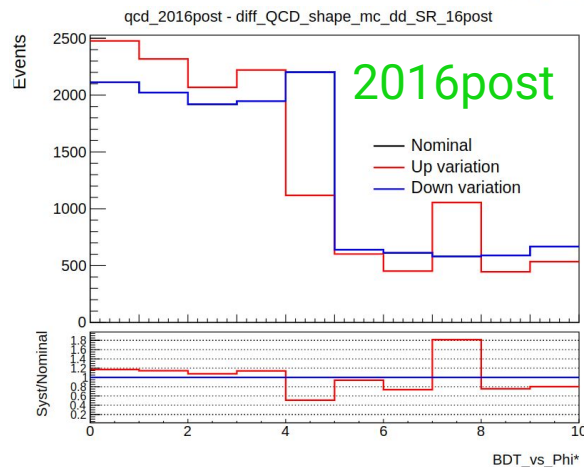
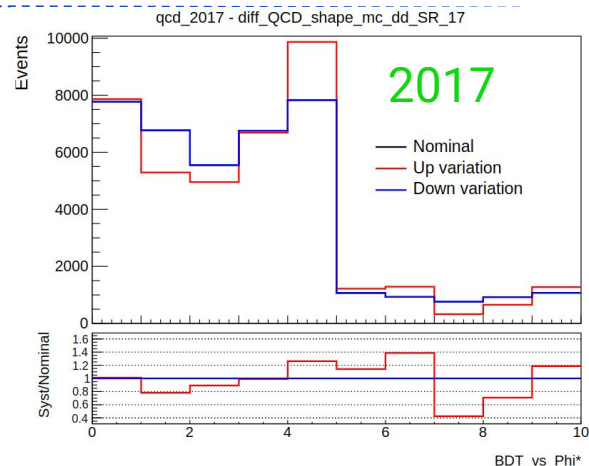
Systematic variation of difference between MC and data-driven QCD



Down : data-driven shape

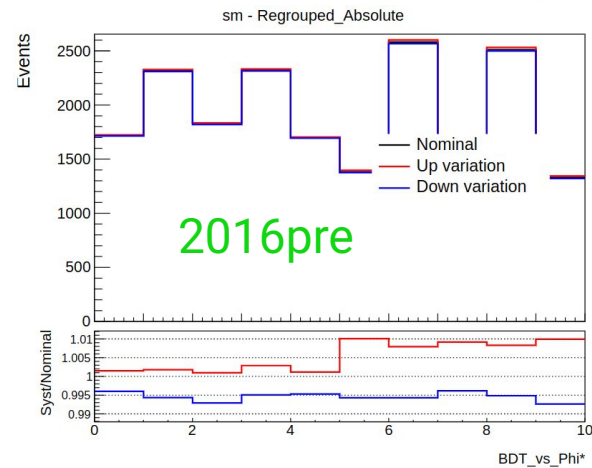
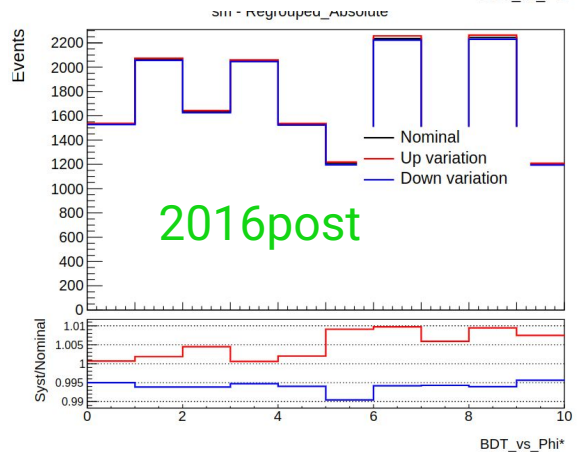
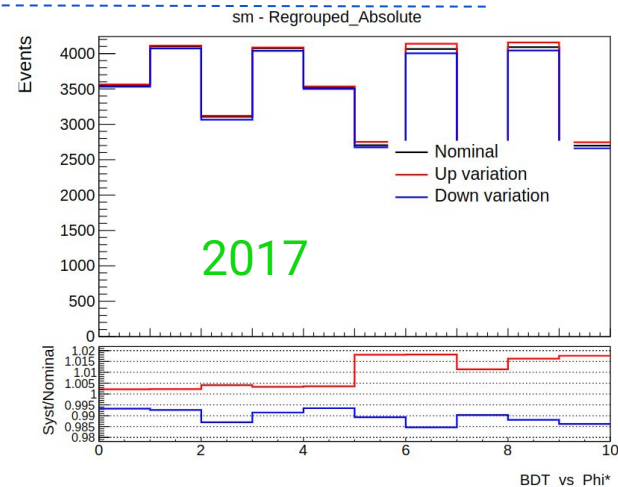
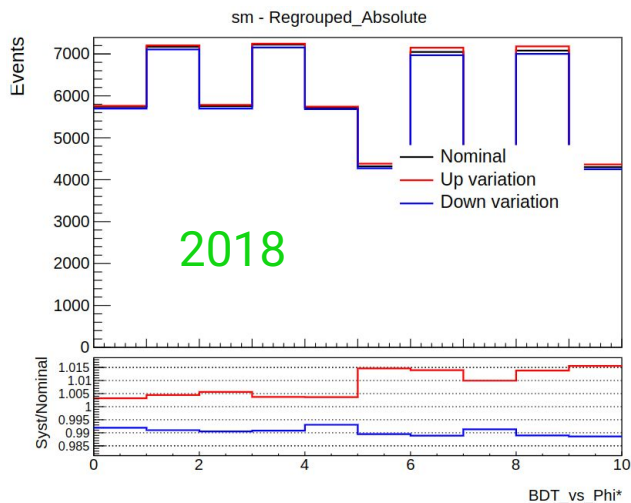
Up : MC QCD shape

SR

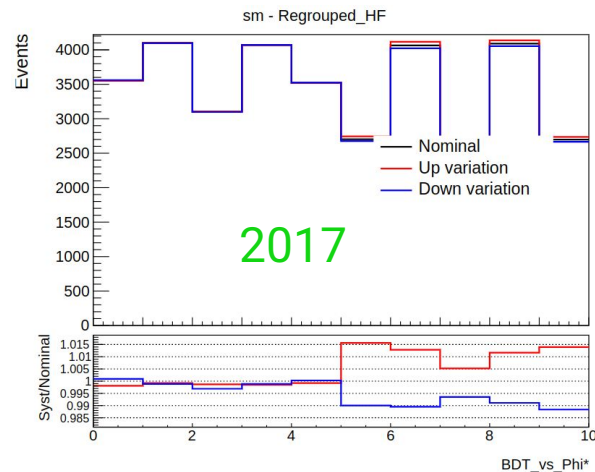
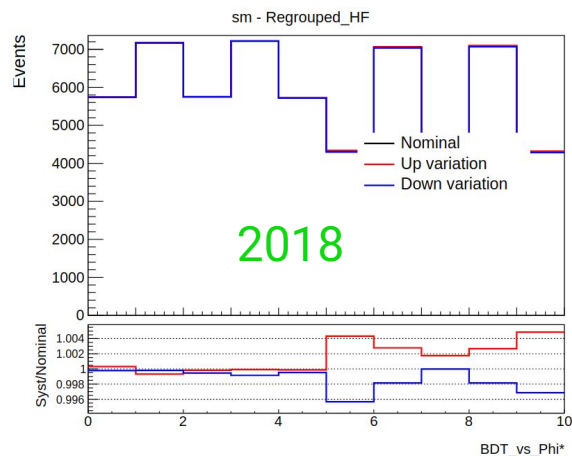


Systematic variation of Regrouped Absolute uncertainty for ST t-channel process

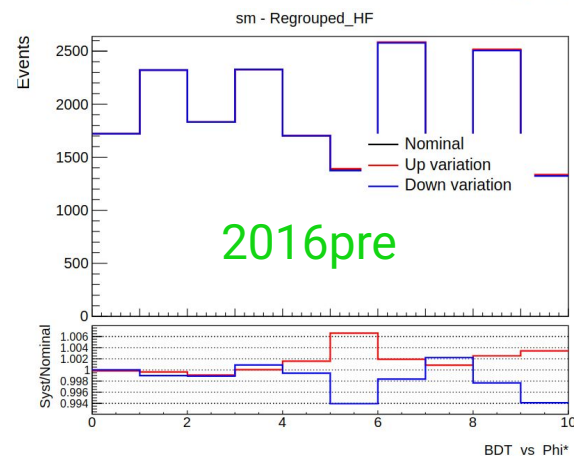
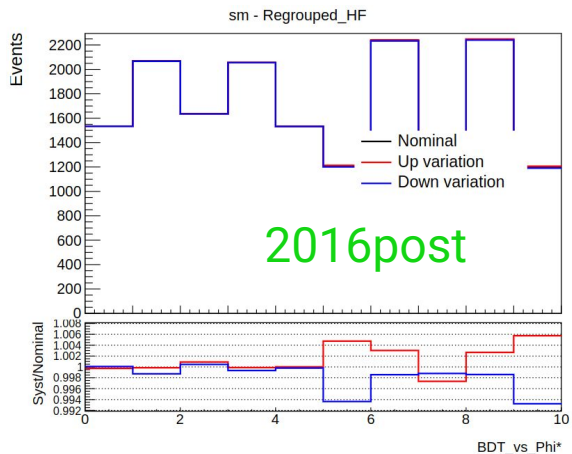
SR



Systematic variation of Regrouped HF uncertainty for ST t-channel process

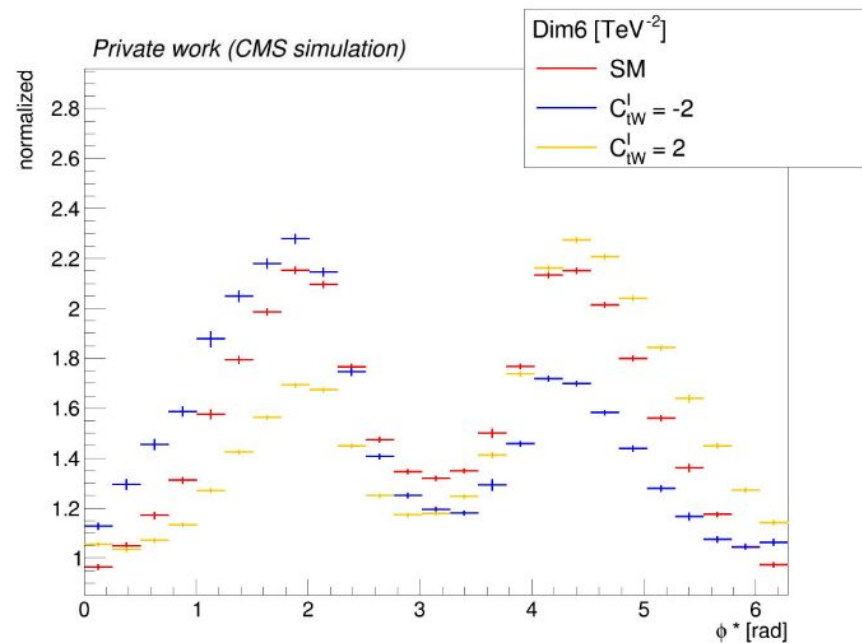
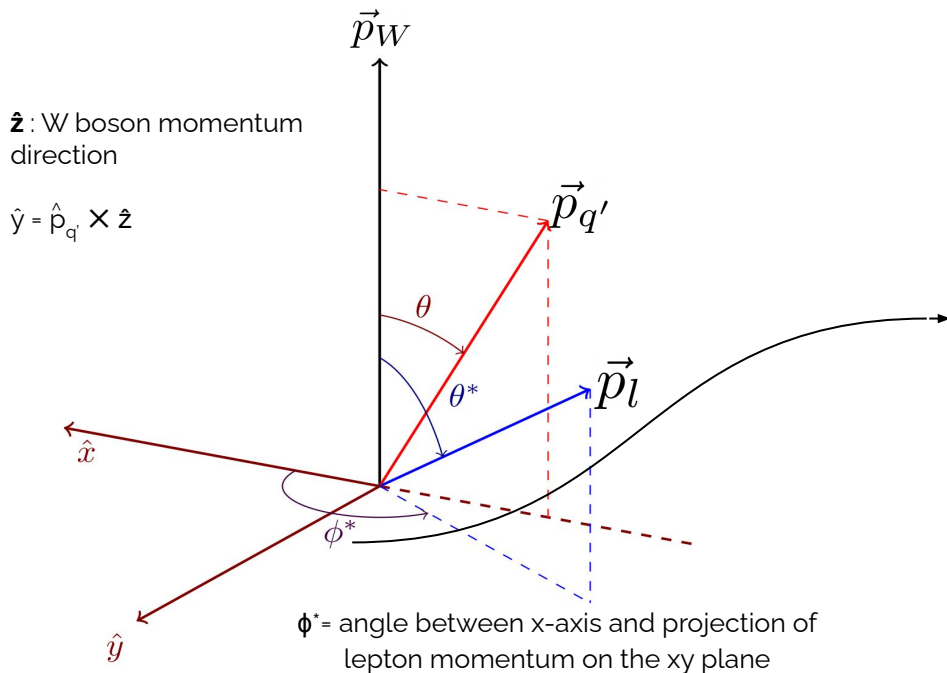


SR



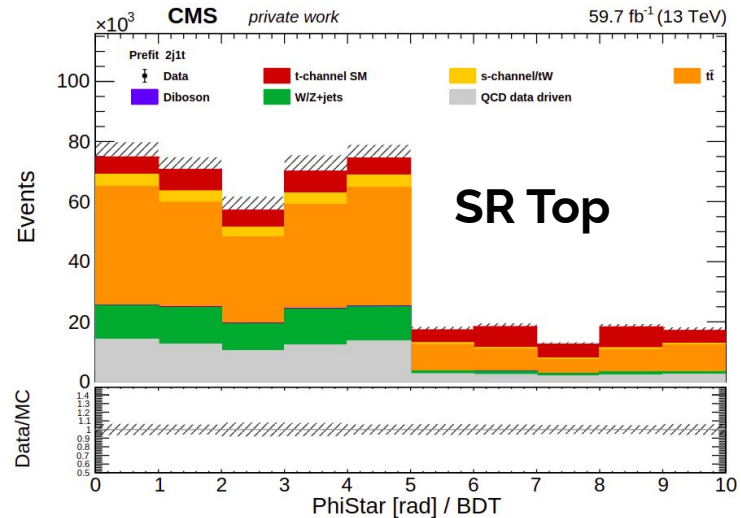
Top quark rest frame

Reference frame used in ATLAS 8 TeV [[arXiv:1707.05393](https://arxiv.org/abs/1707.05393)]

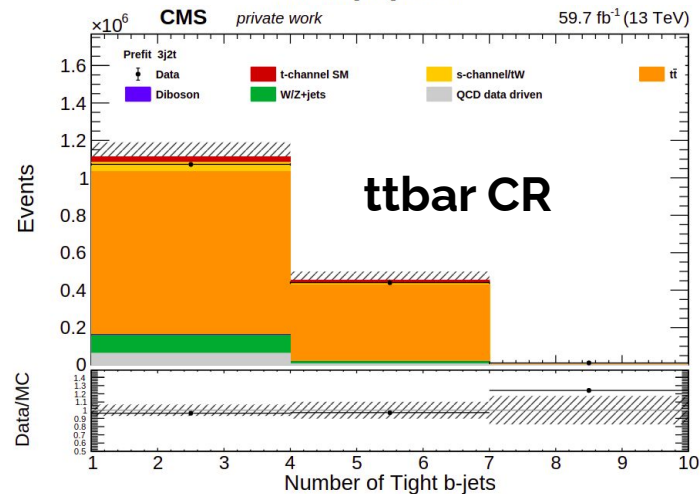
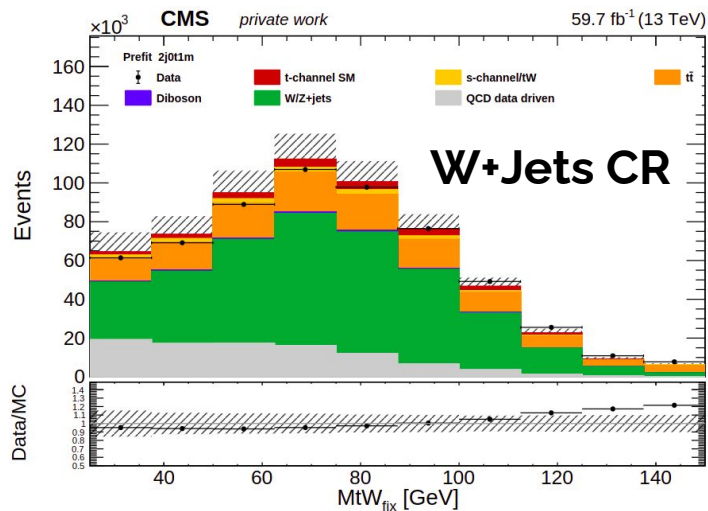
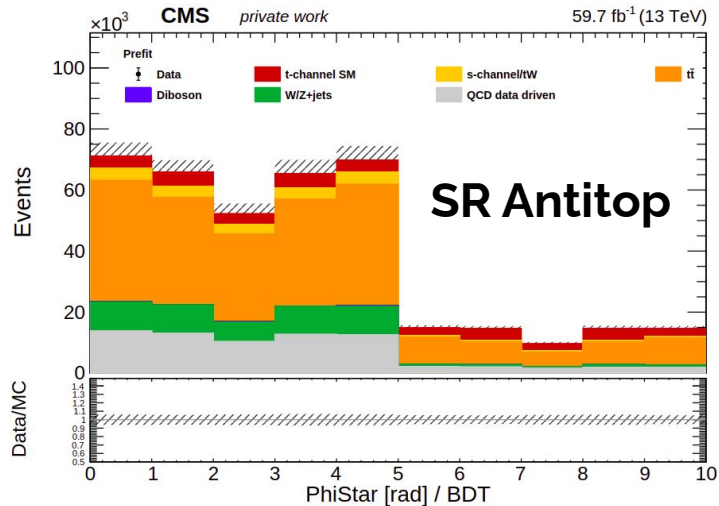


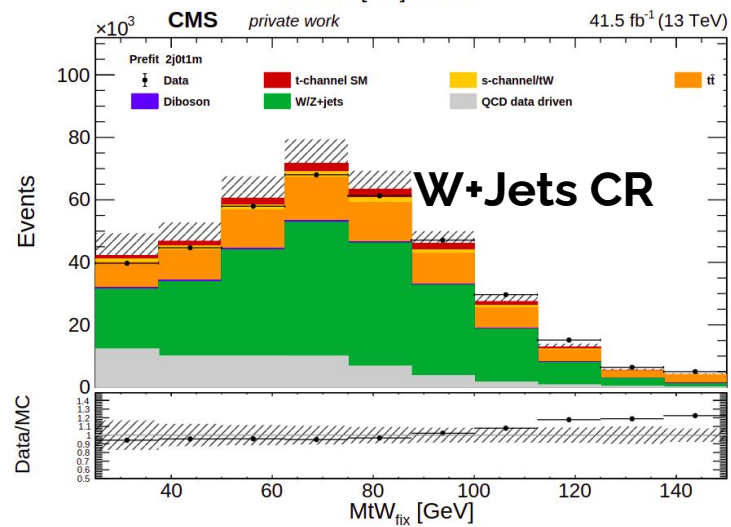
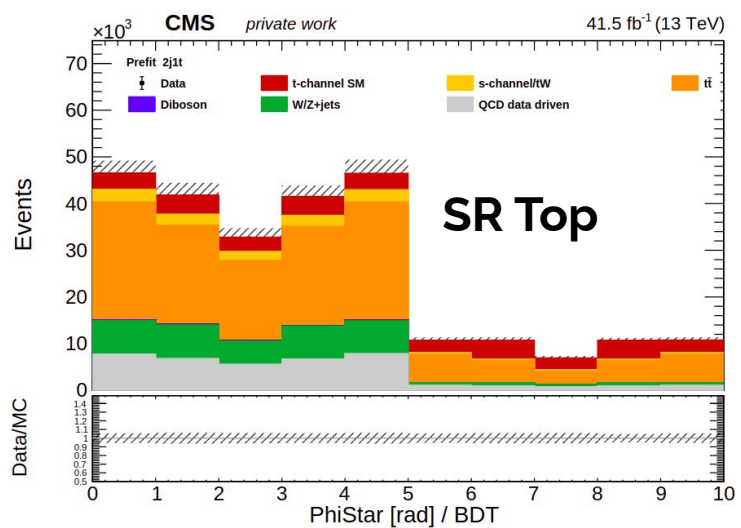
The shape of the distribution varies depending on the value of the EFT coefficient

The value of EFT coefficients can be extracted using such angular distributions

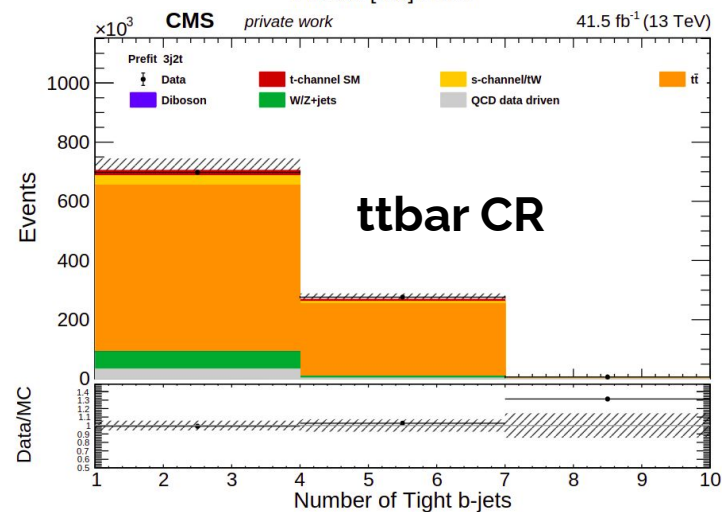
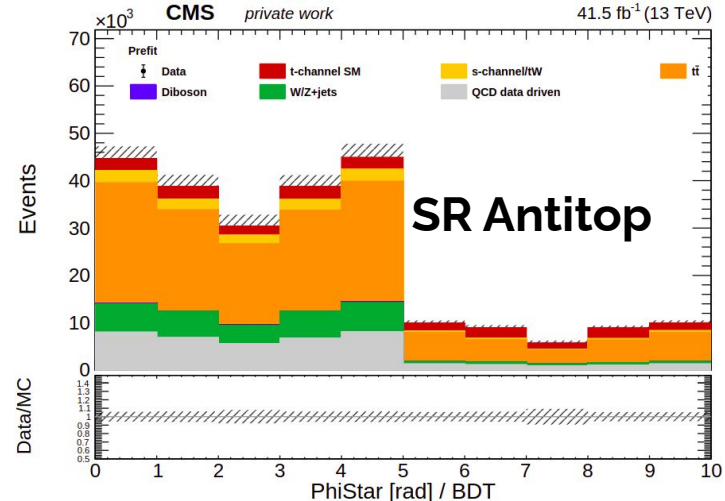


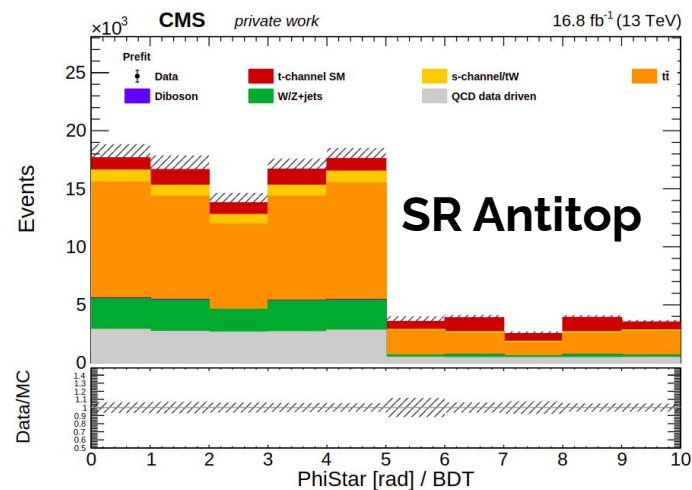
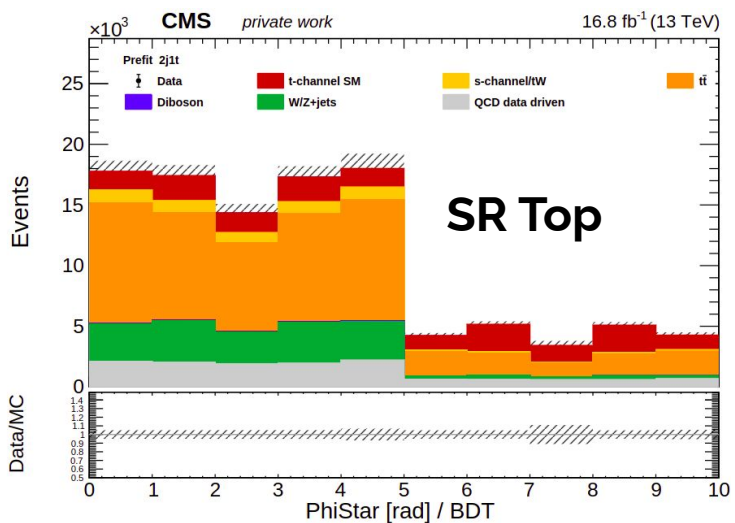
2018



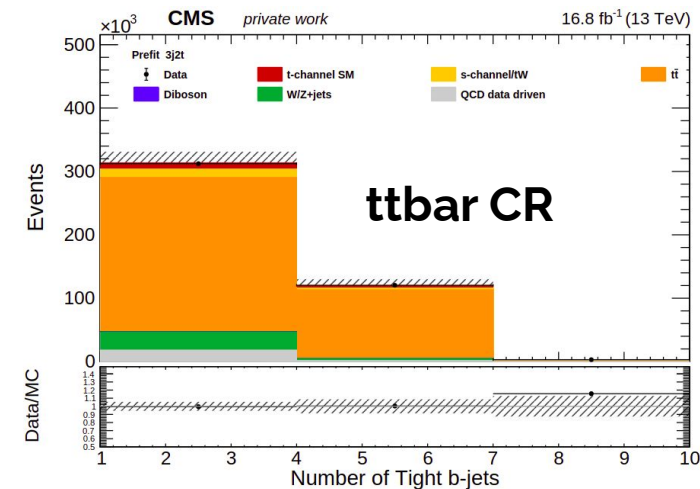
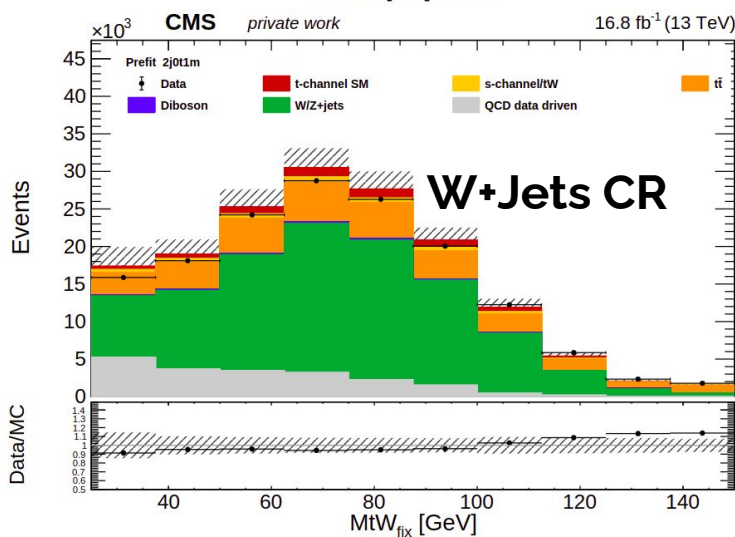


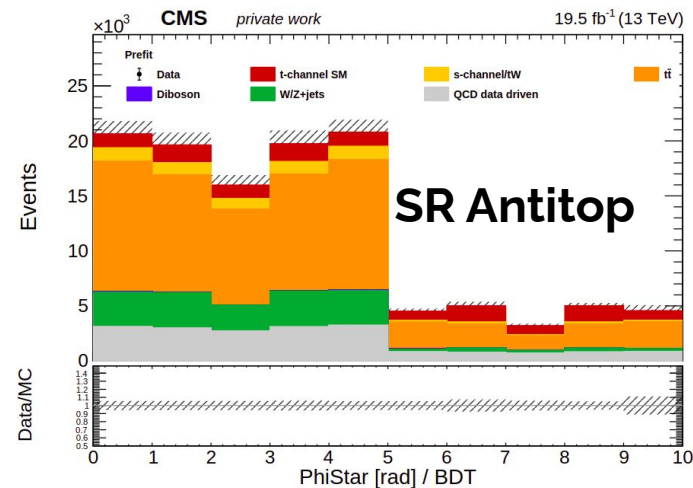
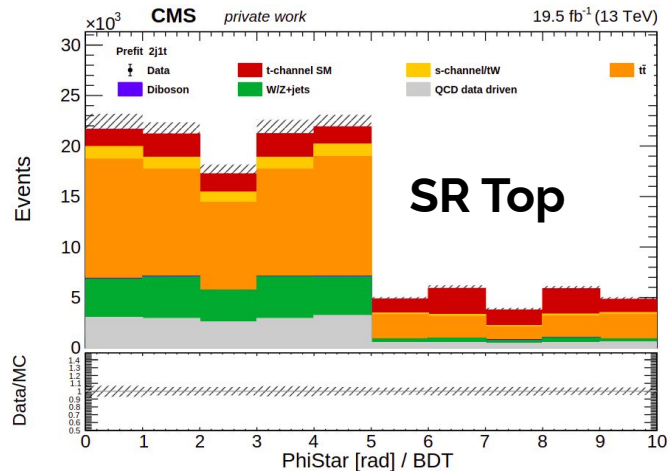
2017



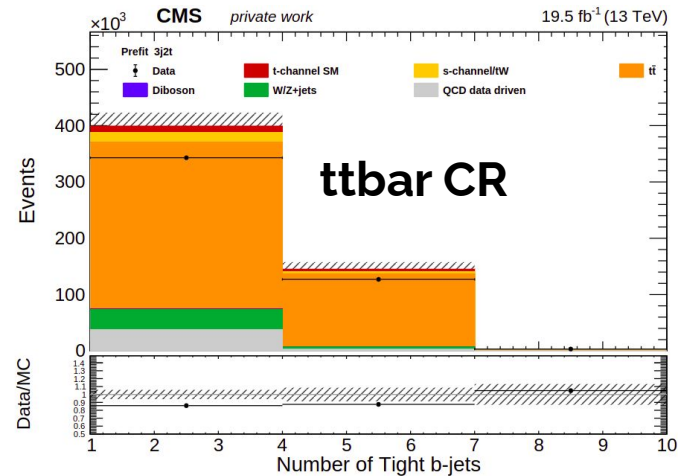
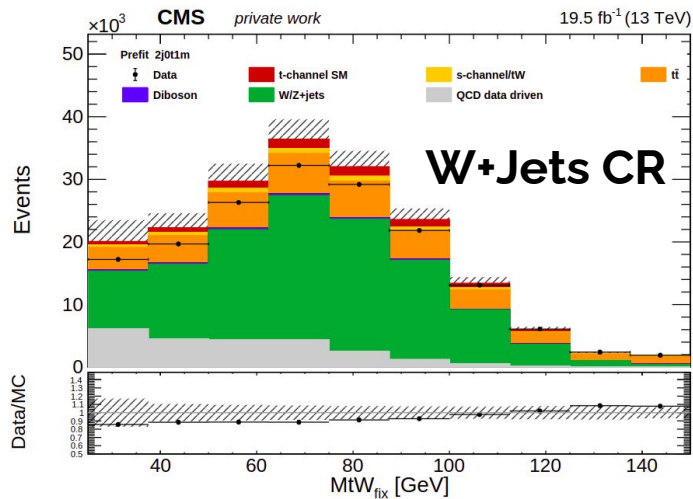


2016post



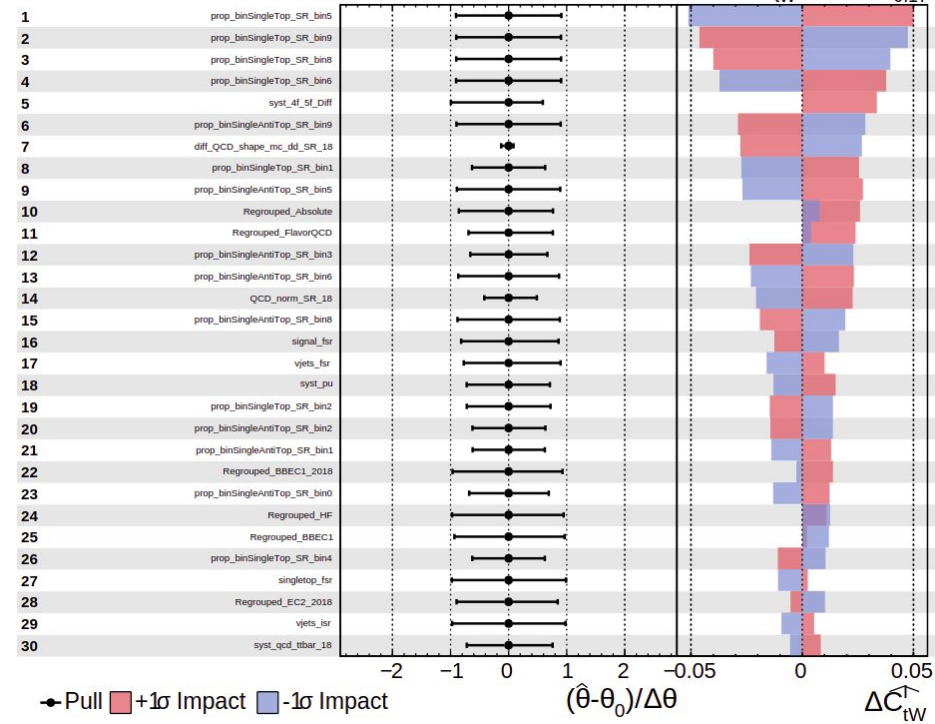


2016pre

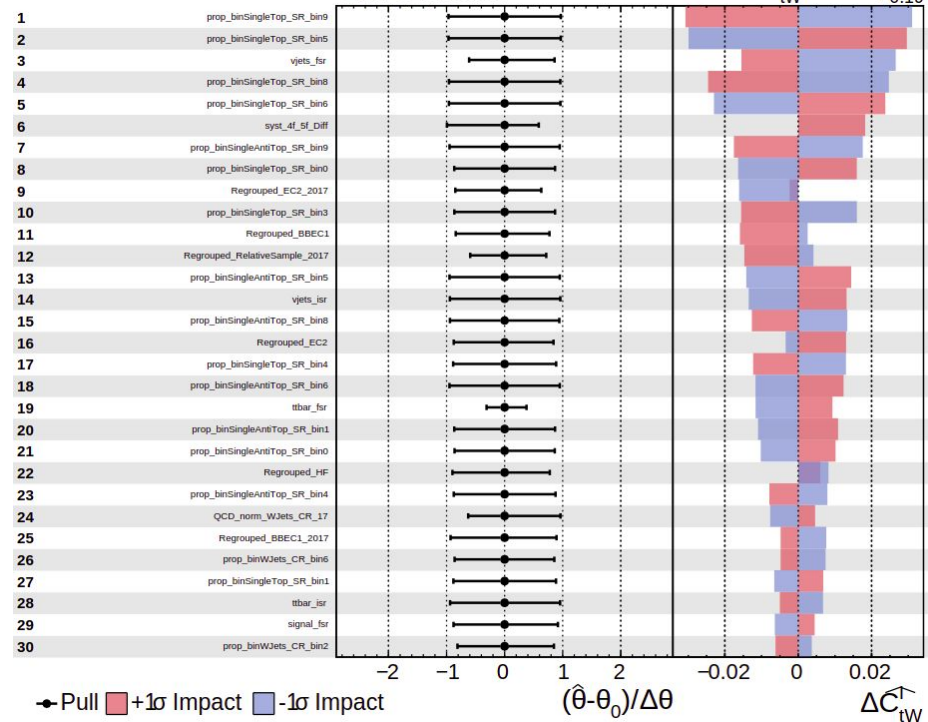


C_t^l w **2018****CMS Internal**

$$\hat{C}_{tw}^\uparrow = -0.00^{+0.17}_{-0.17}$$

**2017****CMS Internal**

$$\hat{C}_{tw}^\uparrow = 0.00^{+0.19}_{-0.19}$$

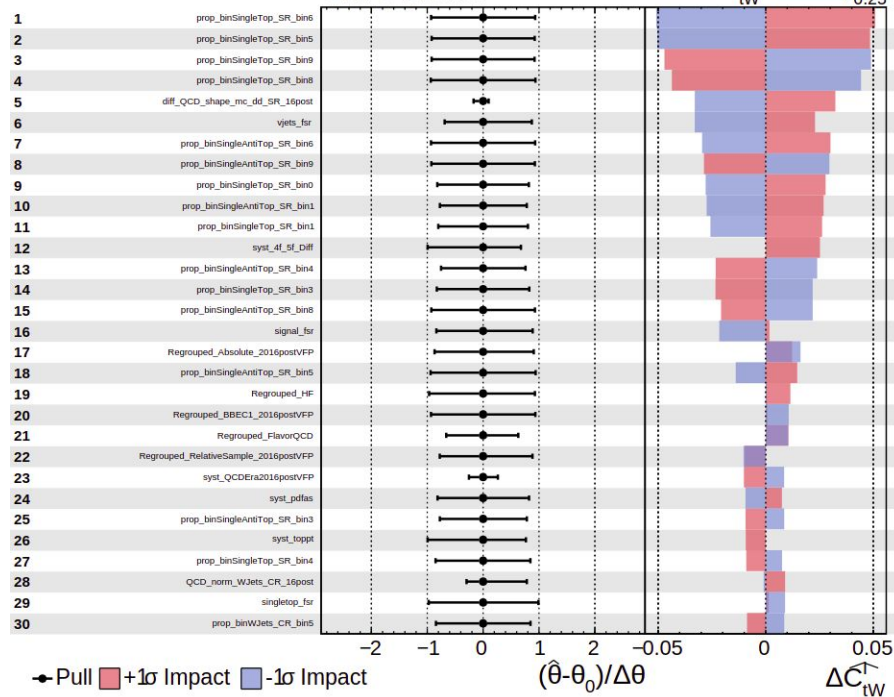


C_t^l
 W

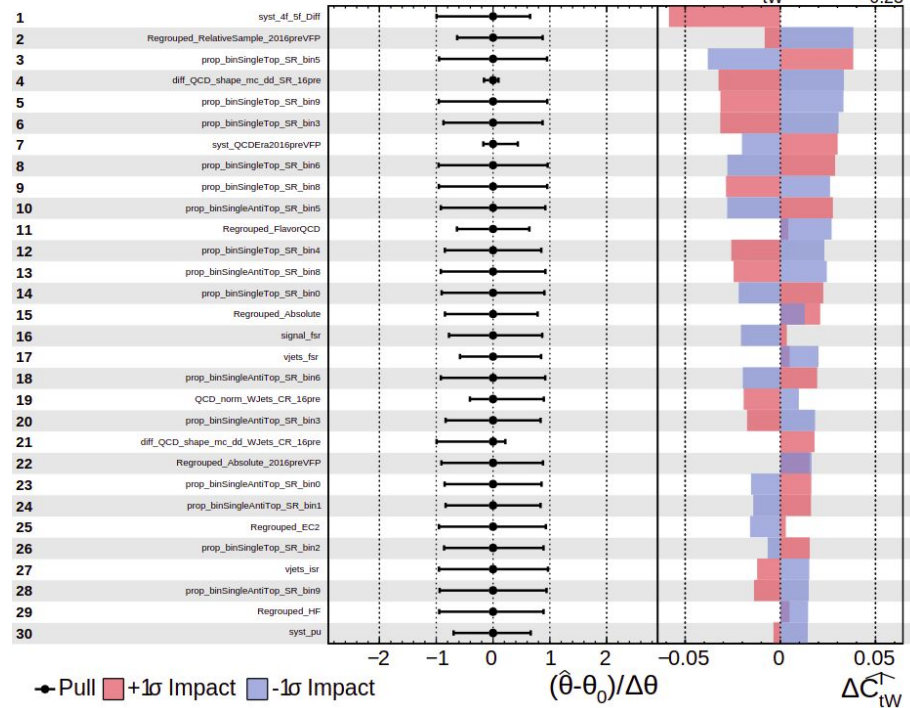
2016post

2016pre

CMS Internal

 $\bar{C}_{tw}^\uparrow = -0.00^{+0.25}_{-0.25}$


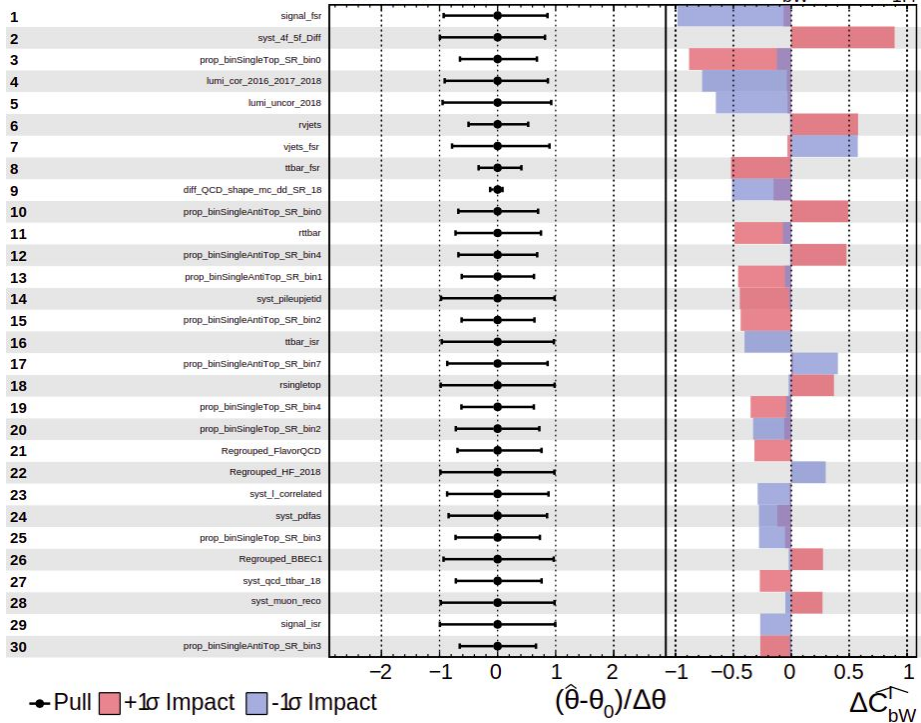
CMS Internal

 $\bar{C}_{tw}^\uparrow = 0.00^{+0.22}_{-0.23}$


2018

CMS Internal

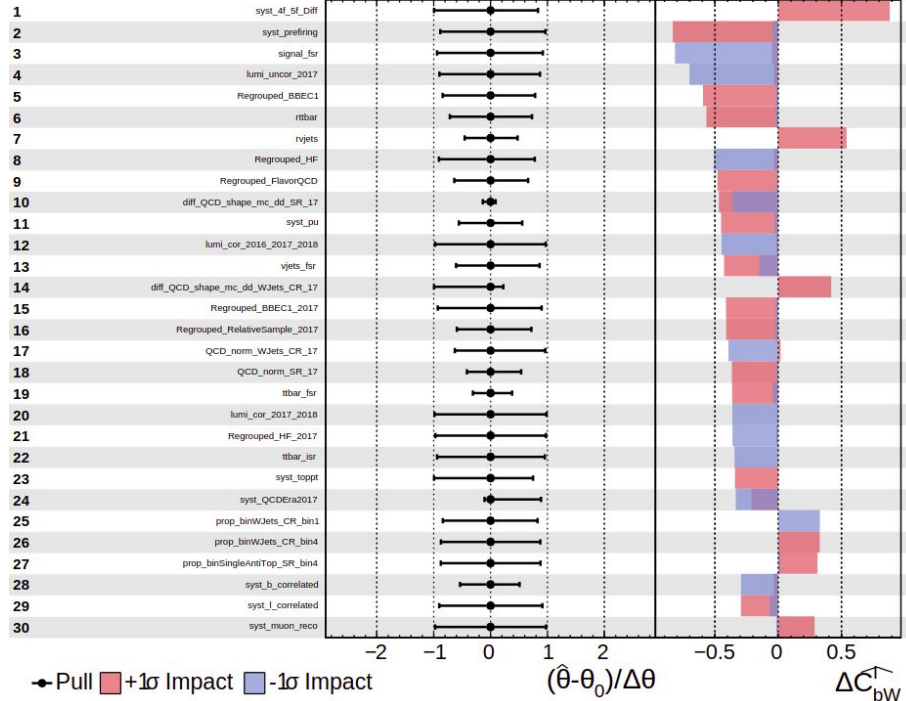
$$C_{bw}^{\uparrow} = -0.0^{+1.4}_{-1.4}$$



2017

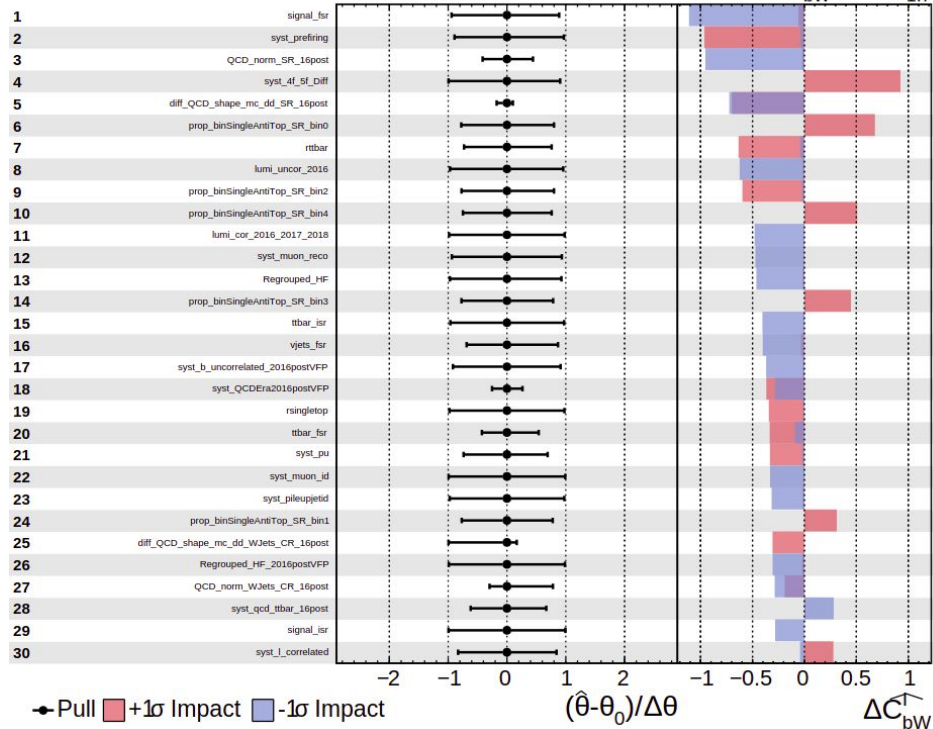
CMS Internal

$$C_{bw}^{\uparrow} = 0.0^{+1.4}_{-1.4}$$



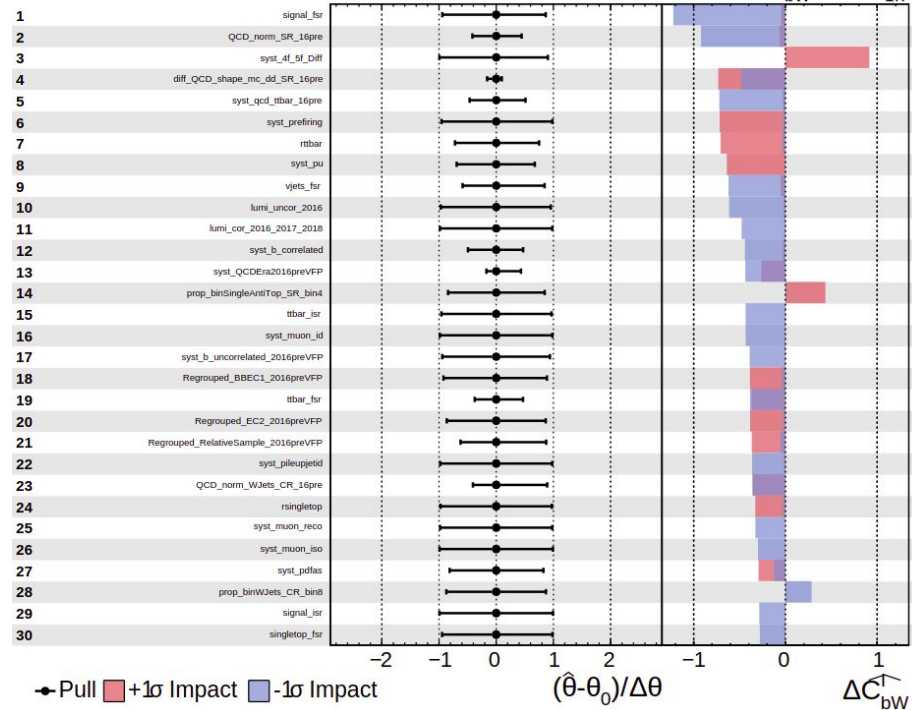
2016post

CMS Internal $\hat{C}_{bW}^I = 0.0^{+1.6}_{-1.7}$



2016pre

CMS Internal $\hat{C}_{bW}^I = 0.0^{+1.6}_{-1.7}$



BDT > 0.2

Uncertainty Breakdown for k_cuWPh

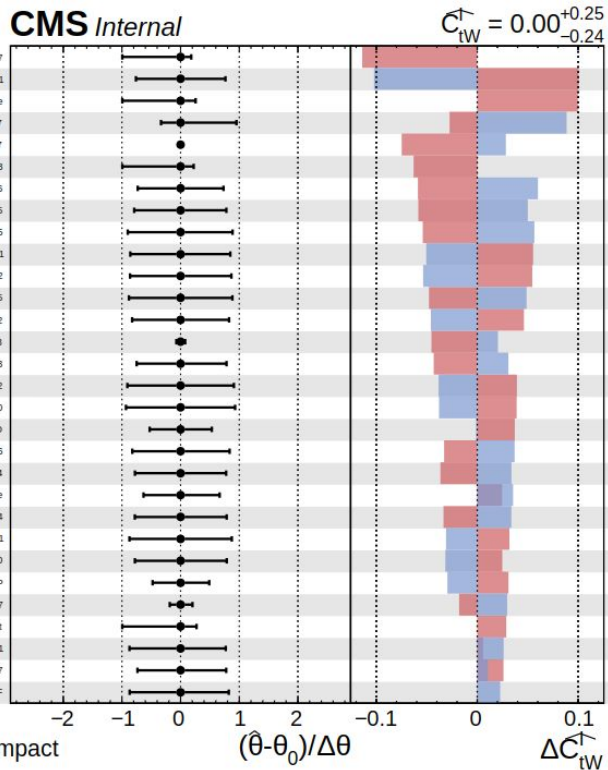
Category	Value
Best fit k_cuWPh	-0.000
Total uncertainty	+0.237 -0.234
Stat. uncertainty	+0.118 -0.118
MC stat. uncertainty	+0.177 -0.175
Other syst. uncert.	+0.105 -0.102

BDT > 0.8

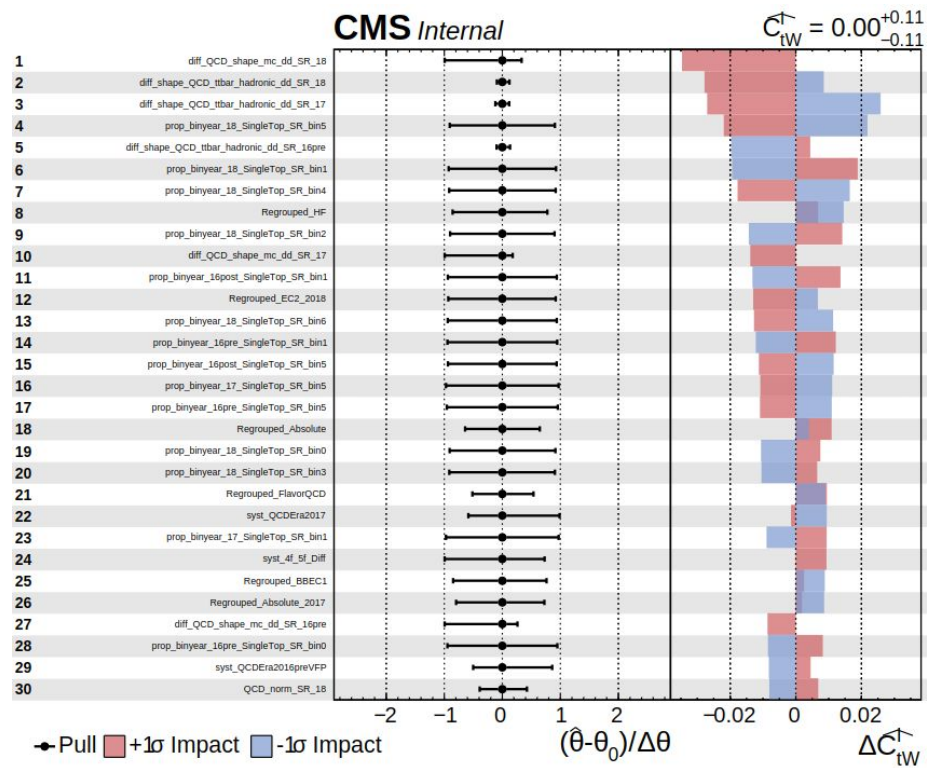
Uncertainty Breakdown for k_cuWPh

Category	Value
Best fit k_cuWPh	0.000
Total uncertainty	+0.113 -0.113
Stat. uncertainty	+0.088 -0.088
MC stat. uncertainty	+0.052 -0.052
Other syst. uncert.	+0.047 -0.049

BDT > 0.2

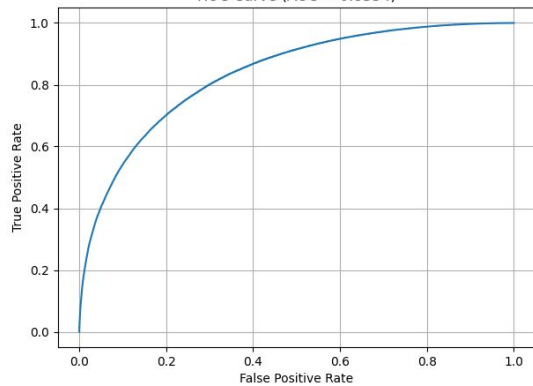


BDT > 0.8



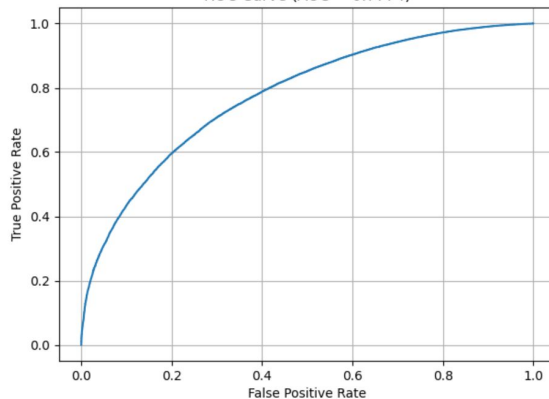
W+jets

ROC Curve (AUC = 0.8354)



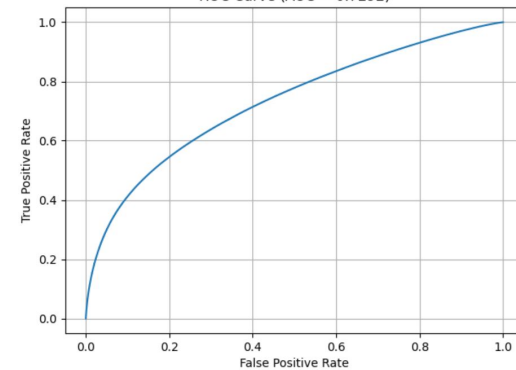
QCD

ROC Curve (AUC = 0.7774)

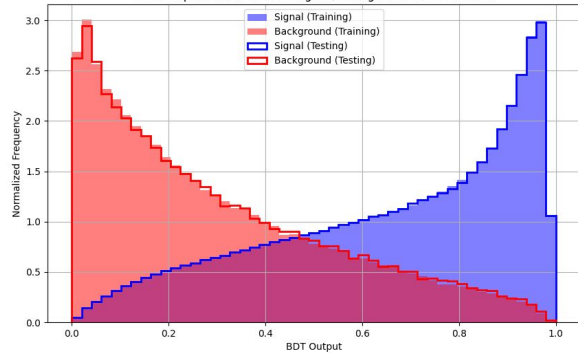


t $\bar{t}$

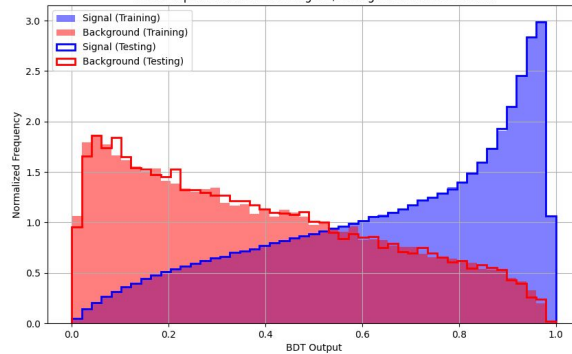
ROC Curve (AUC = 0.7292)



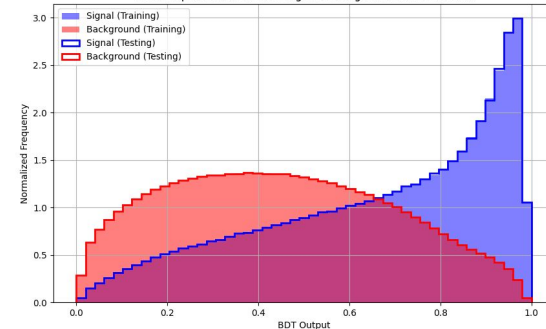
BDT Output Distribution for signal / background discrimination



BDT Output Distribution for signal / background discrimination



BDT Output Distribution for signal / background discrimination



EFT impact on top width and cross section

BSM Matrix Element

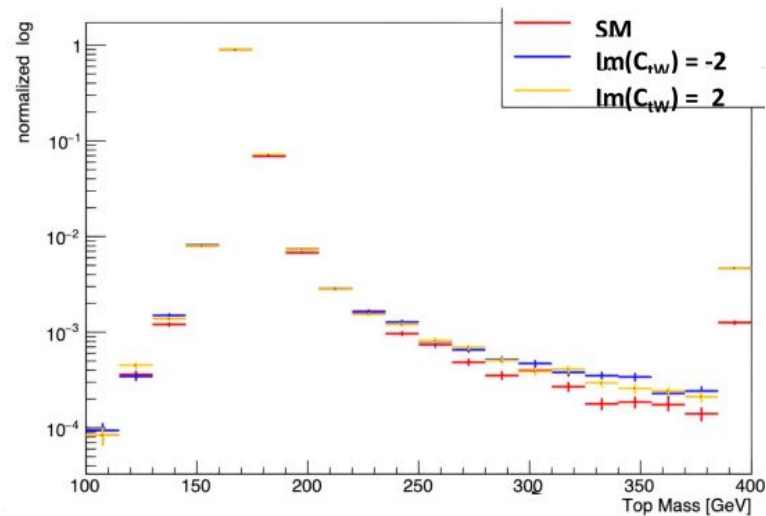
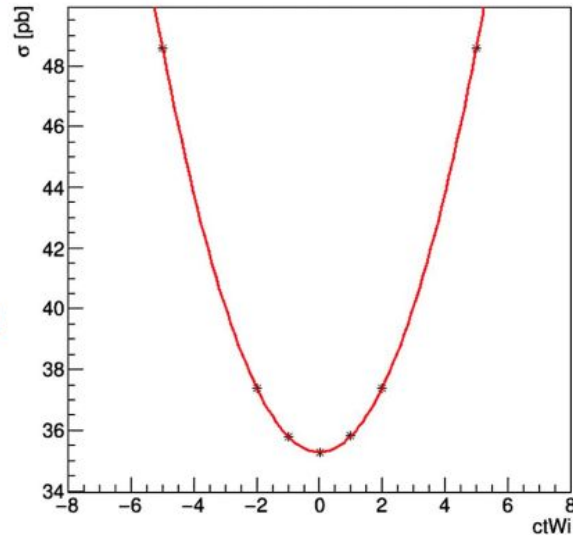
$$\mathcal{M} = \mathcal{M}_{SM} + \sum_i \frac{c_i}{\Lambda^2} \overbrace{\mathcal{M}_i}^{\text{BSM Matrix Element}}$$

$$\sigma \propto |\mathcal{M}|^2$$

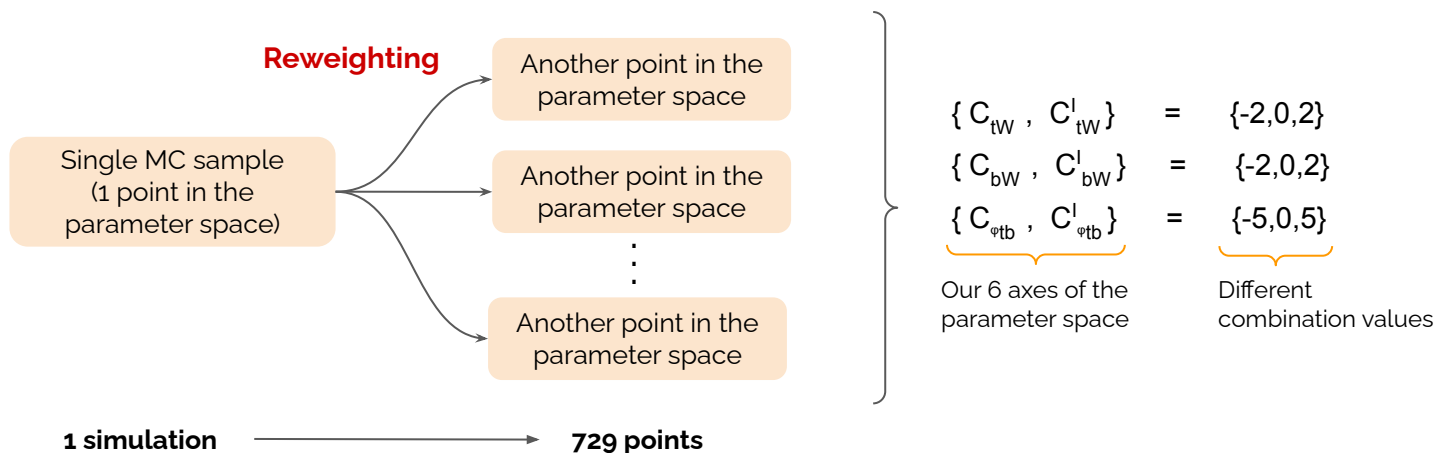
Quadratic behavior on the cross section as expected, [EPJC 78 (2018) 999].

Same quadratic behavior with the top width.

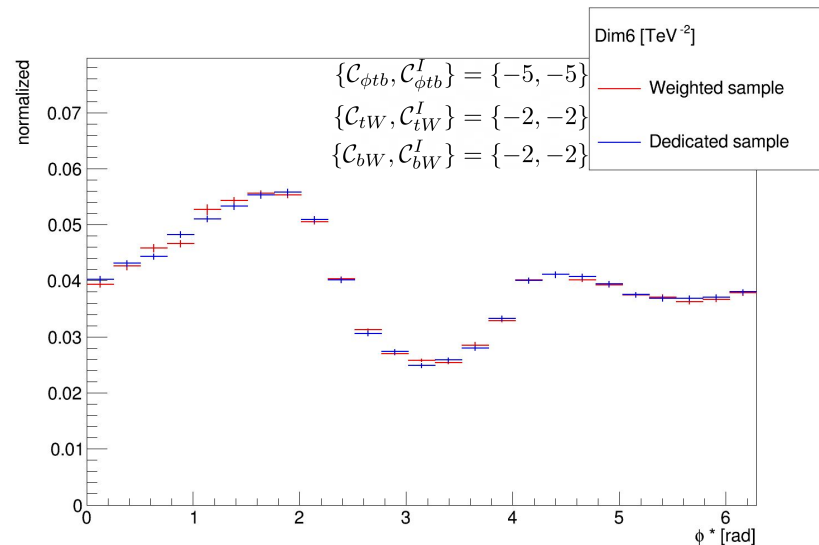
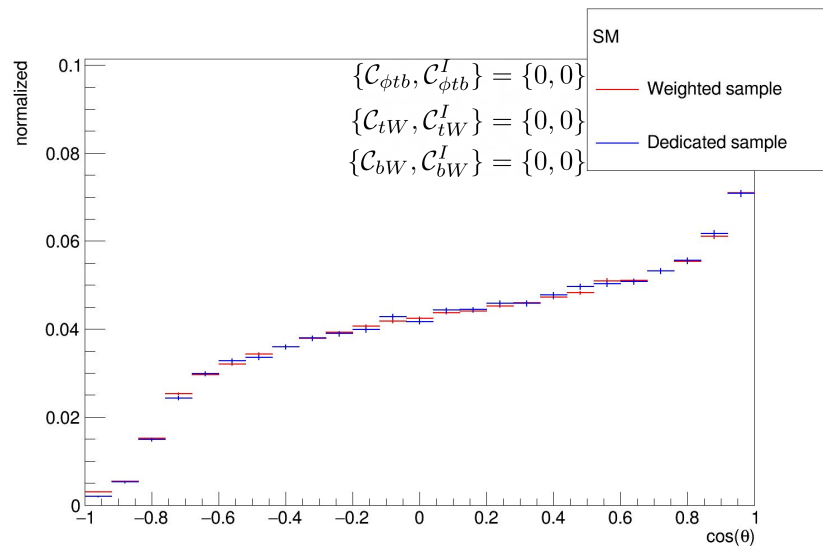
N.B. First sample in CMS that has top width varying with EFT value.



- We produce a simulation sample for single top production including EFT coefficients at top production and decay
- Reweighting method: different regions of the parameter space to be probed with a single Monte Carlo (MC) sample



The reweighting method allows to produce a single sample instead of 729



Comparing reweighted distributions of $\cos(\theta)$ and ϕ^* to dedicated (non-reweighted) samples at two different distant points of the parameter space

→ **Reweightings is validated**

