$t\bar{t}X$ (W, Higgs) measurements at CMS

CPAN Days
Santander
October 3rd 2023

Clara Ramón Álvarez*
On behalf of the CMS Collaboration

clara.ramon.alvarez@cern.ch

Gobierno del Principado de Asturias
Consejería de Ciencia, Innovación y Universidad

*Partially funded by *Consejería de Ciencia, Innovación y Universidad (Gobierno del Principado de Asturias)*, through "*Ayudas del Programa Severo Ochoa*".

Grant PID2020-113341RB-100 funded by
Gobierno de España
Ministerio de Ciencia e Innovación
Agencia Estatal de Investigación

PCTI Asturias

European Union
European Regional Development Fund

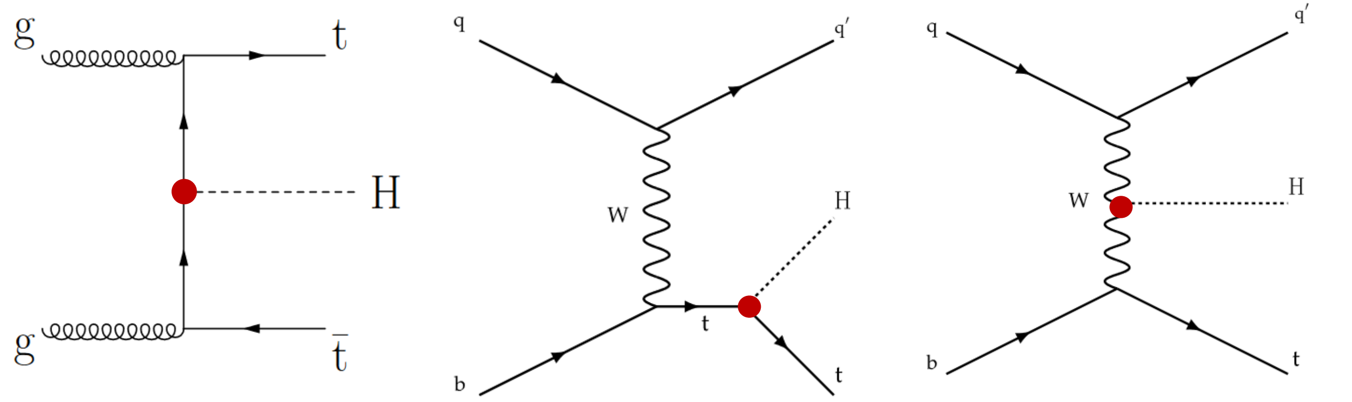
Universidad de Oviedo

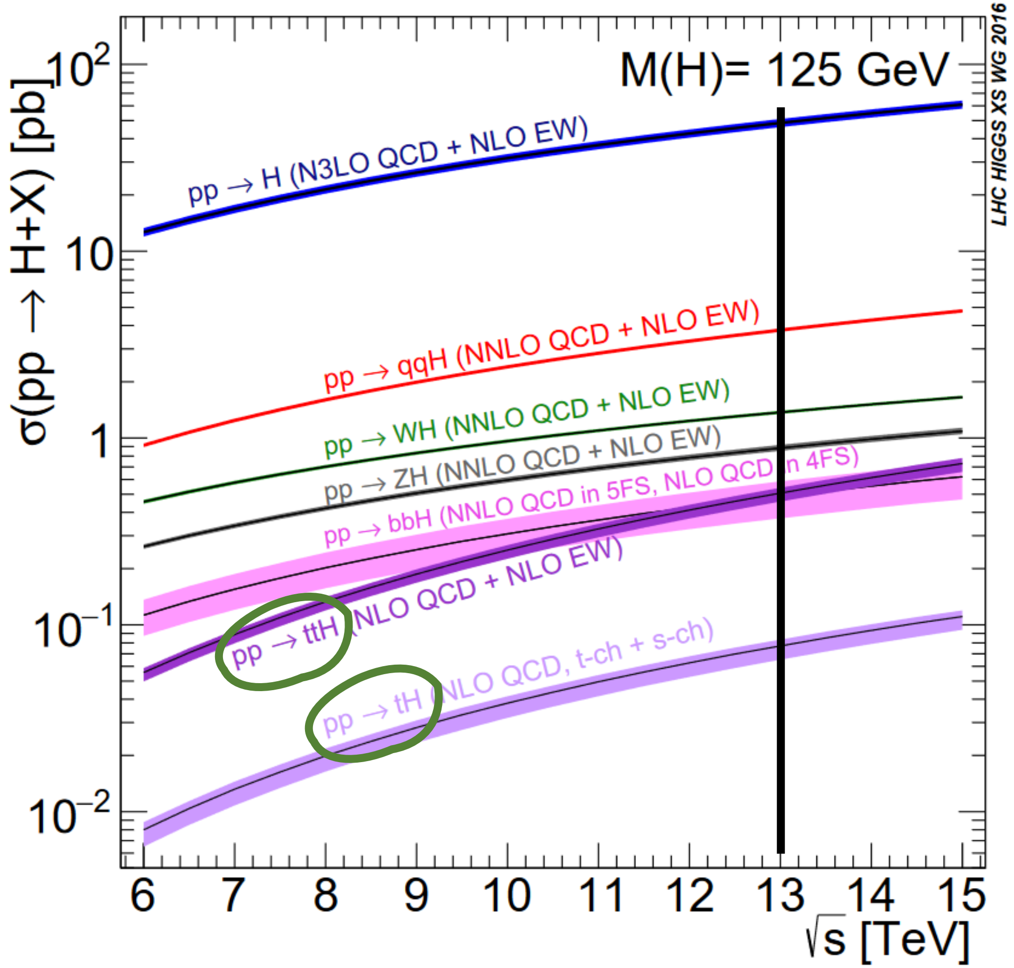
ICTEA
Instituto Universitario de Ciencias y Tecnologías Espaciales de Asturias

CMS
Compact Muon Solenoid

Why to study $t\bar{t}H$…

At LHC **gg fusion is the dominant** H production mode
Top associated production, has much lower rates





Challenging and interesting

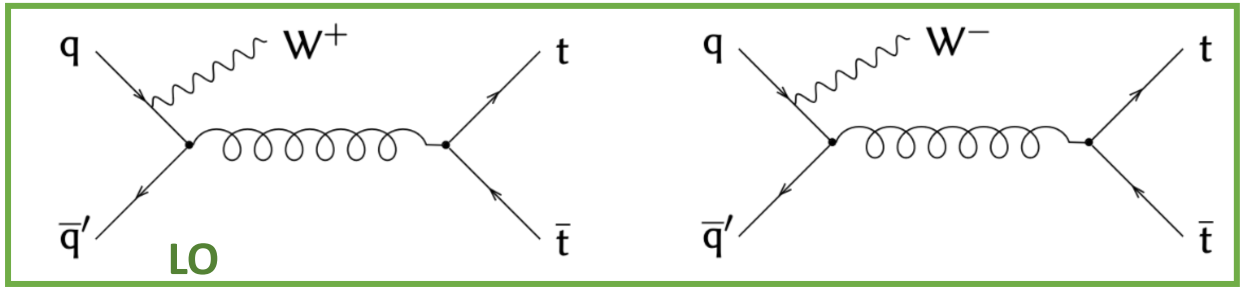
- **Direct probe** of top-Higgs coupling, $y_t \sim 1$ (largest in the SM)
- BSM y_t =-1→ constructive interference → $\sigma_{tH} \sim 0.8$ pb

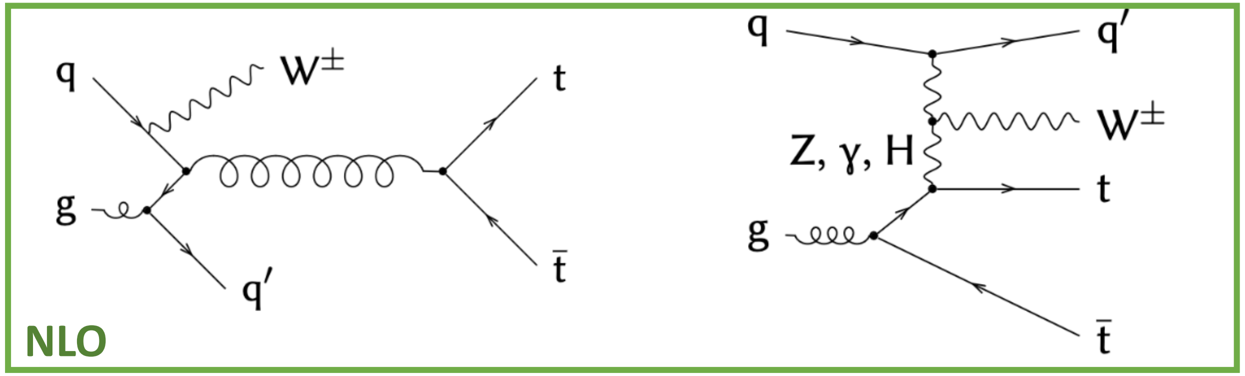
Will summarize the studies from [EPJC 81, 378 (2021)] and [JHEP 07 (2023) 092]

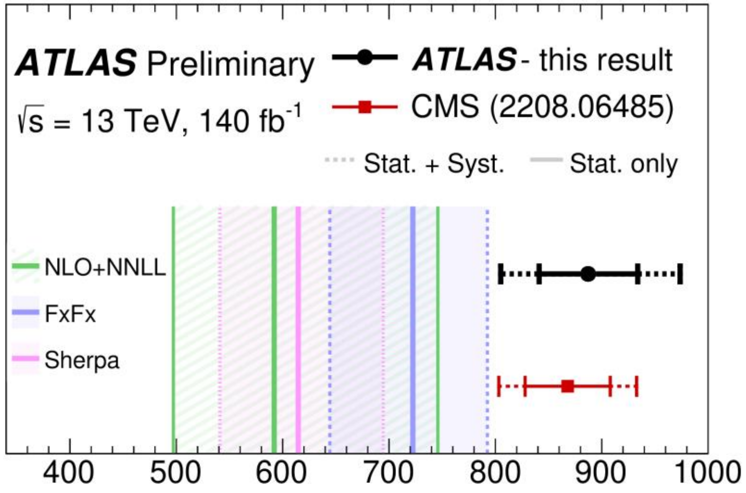
… and $t\bar{t}W$

- **Main background for $t\bar{t}H$, $t\bar{t}t\bar{t}$ …**
- EWK induced process at LO:
sizable difference between $t\bar{t}W^{+}$and $t\bar{t}W^{-}$ production
- Measured $\sigma(t\bar{t}W)$ **consistently above theory** value (both in ATLAS and CMS)
- Active discussion about the modelling of this process
- Sensitive to top-electroweak coupling
- Handle to study the back-forward asymmetry in $t\bar{t}$ production at pp colliders

Will summarized the studies from [JHEP 07 (2023) 219]







ATLAS-CONF-2023-019

Strategy

Using **multilepton final states**, categorize events depending on the lepton multiplicity

- For $t\bar{t}H$ analysis, 12 categories (focusing on the 3 most sensitive)
- For $t\bar{t}W$ analysis, 2 categories (2 same-sign (ss) leptons (ℓ) and 3ℓ)

Dedicated MVA to select isolated leptons from H, W and τ decays is crucial in this analysis → Reduce one of the main backgrounds

Dedicated **selection on each category**. Using Jet and b-tagging multiplicities and vetoing opposite charge leptons within Z peak

Backgrounds:

Reducible backgrounds:

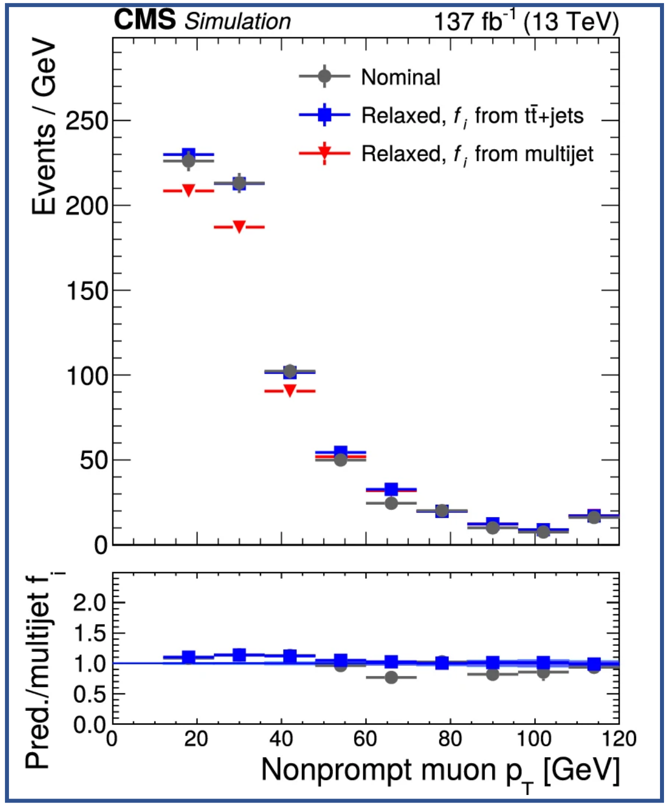
- ➢ Non prompt leptons
- ➢ Electron charge flips
- ➢ Photon Conversions

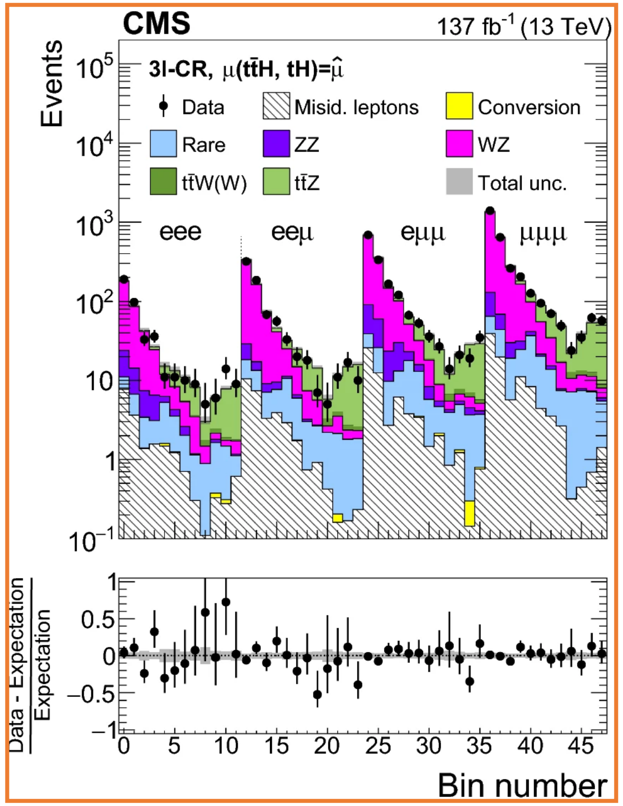
} Estimated with data-driven techniques

Irreducible backgrounds:

- ➢ $t\bar{t}Z$, $t\bar{t}W$ ($t\bar{t}H$)
- ➢ Dibosons

Control regions to constrain these backgrounds





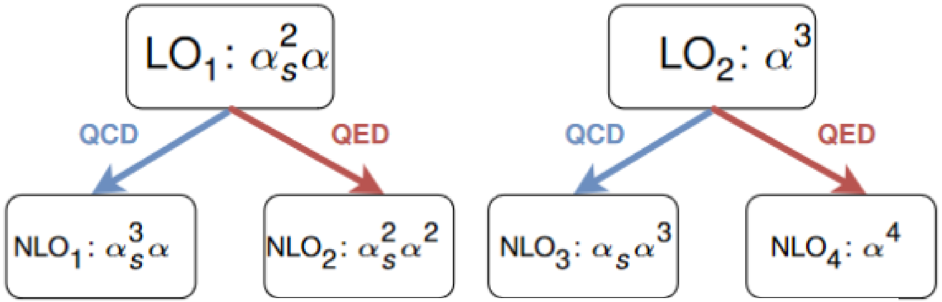
$t\bar{t}W$ MC

[Hand book of LHC cross sections](#) (Yellow Report 4):

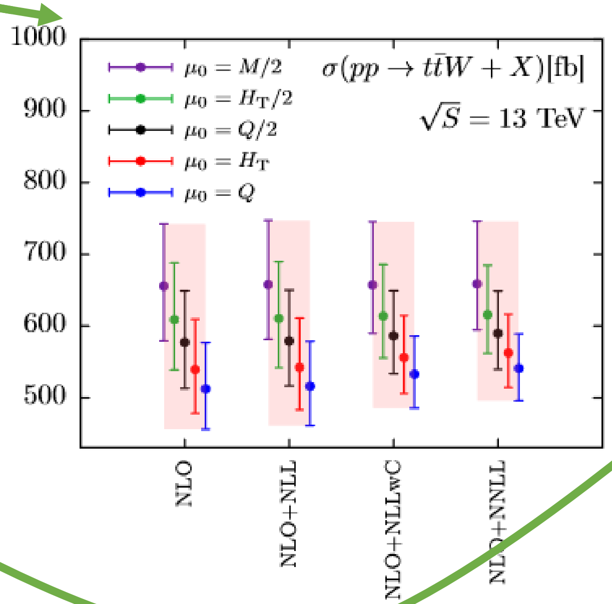
- $t\bar{t}W$ @ NLO QCD
- NLO EWK term $\alpha_s\alpha^3$ considered negligible → in fact found to be ~10%

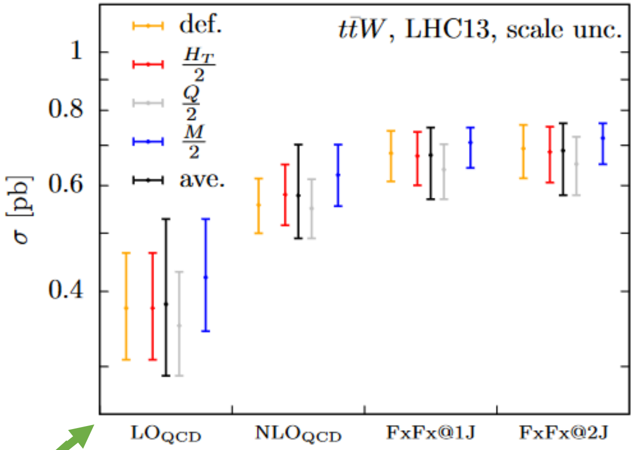
The MC used in both studies presented in this talk include these EWK corrections:

MG5 aMC@NLO 2.6.0+Pythia8 FxFx NLO QCD + $\alpha_s\alpha^3$ EW (+0,1j@NLO)



- $t\bar{t}W$ shows large scale dependence even after including NLO+NNLL (different from $t\bar{t}H$) → new structures at NLO [arXiv:2001.03031]

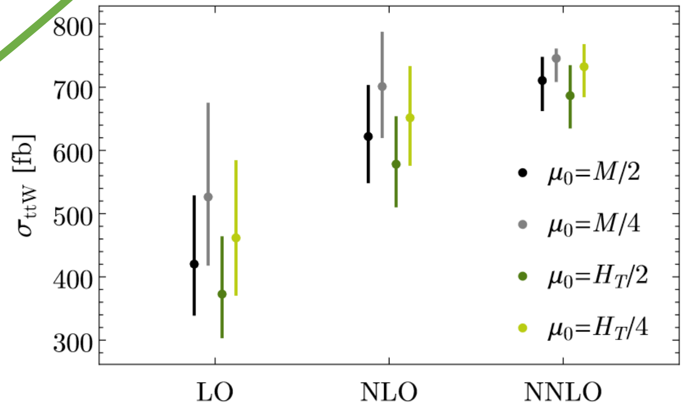




- **NEW FxFx** merging procedure [JHEP 11 (2021) 29]:
  - Treats QCD and EW jets separately
  - Stabilization of scale dependence

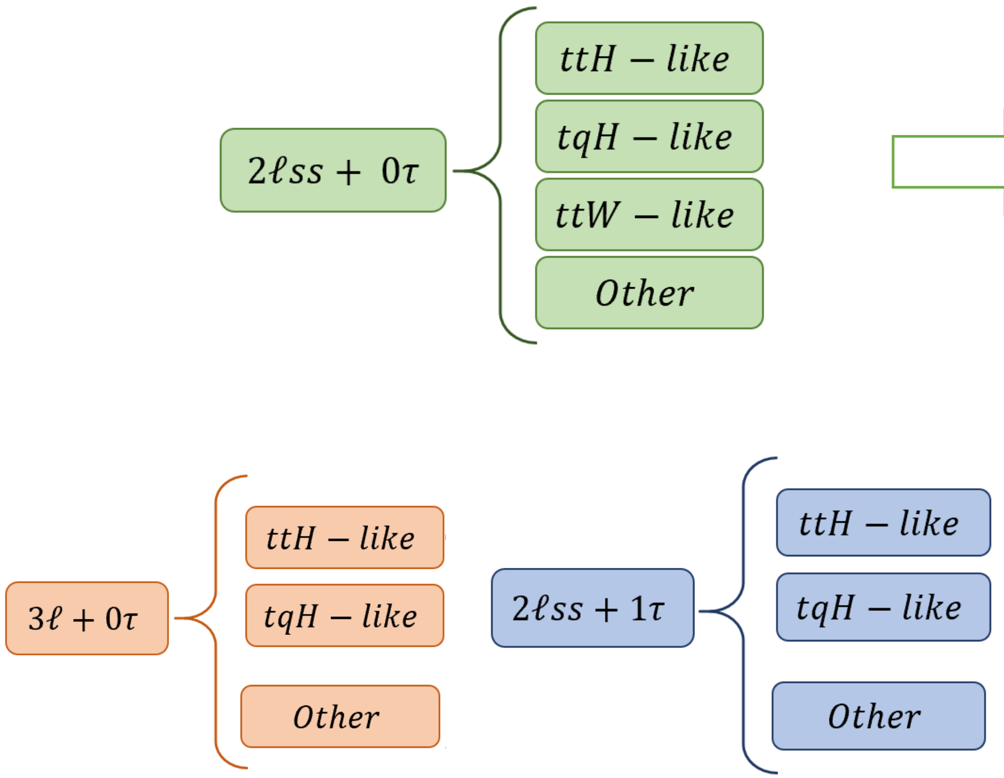
- Latest σ at NNLO (QCD) + NLO (EWK) [2306.16311]:

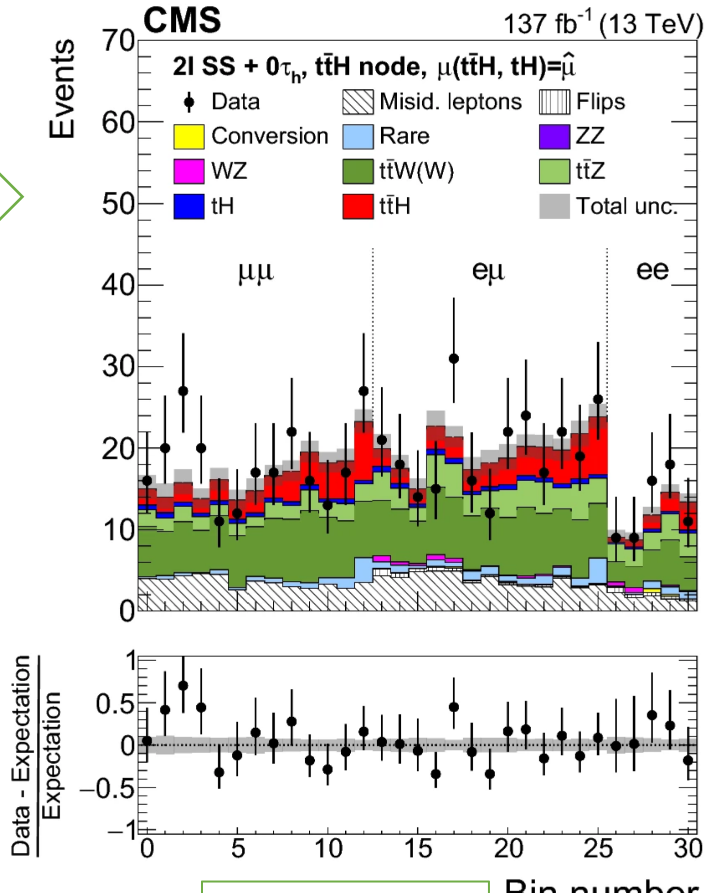
  745.3 ± 6.7% ± 1.8% fb



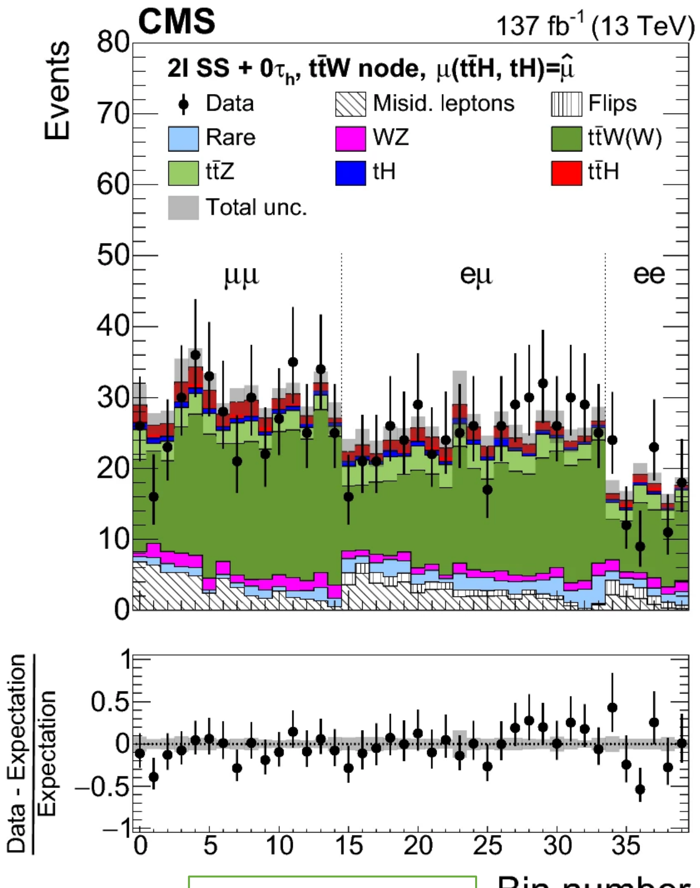
$t\bar{t}$H event classification

- **Dedicated NN** in each of the 12 signal categories in order to discriminate signal from background
- Dedicated **node to target $t\bar{t}W$** in $2\ell ss+0\tau$ category
- Further classification depending on flavor, b-tag multiplicity…





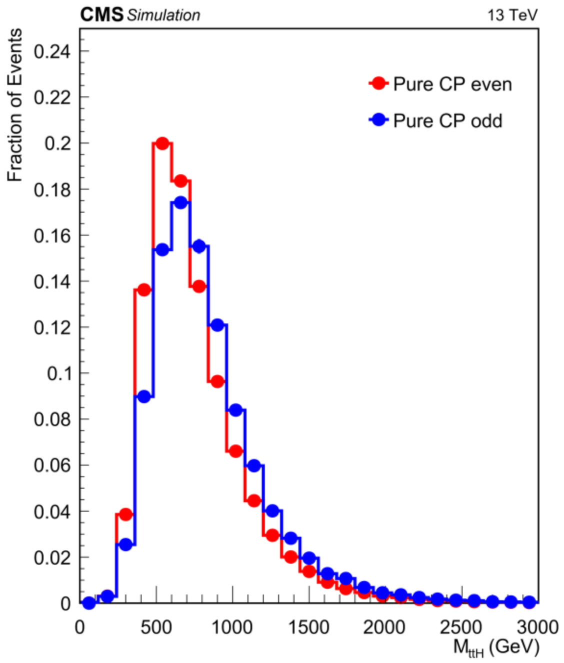
ttH node

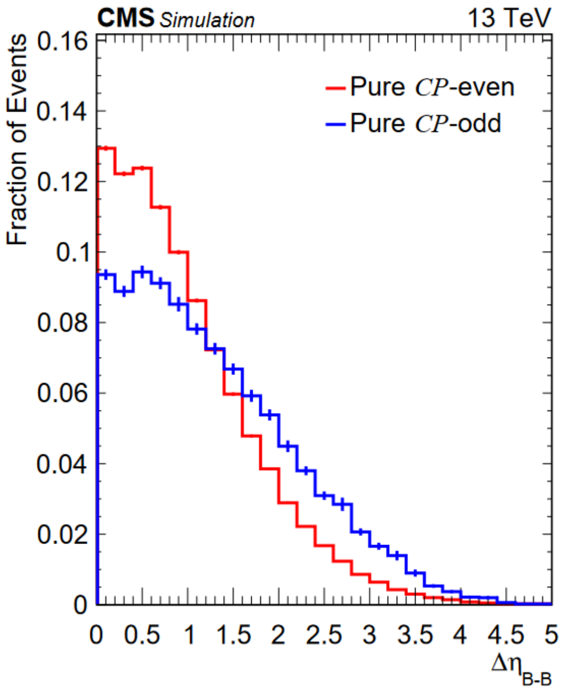


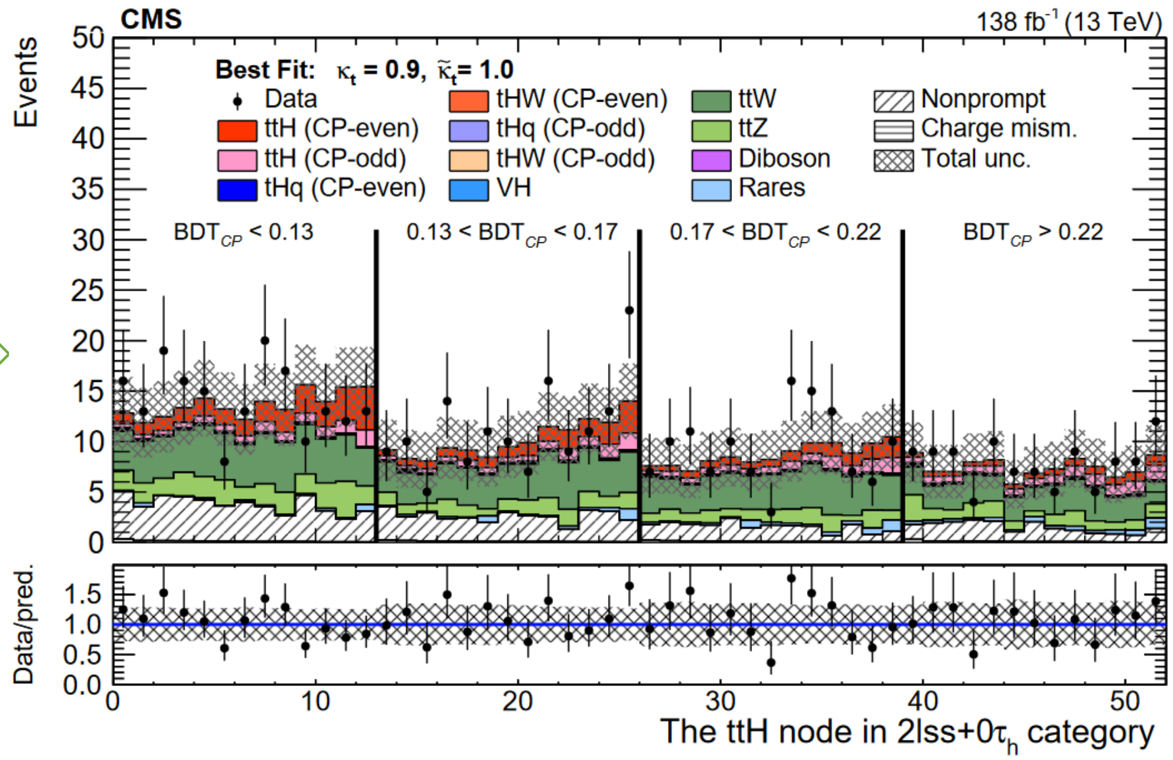
ttW node

$t\bar{t}H$ - CP interpretation

- Lagrangian can be parametrized as: $$\mathcal{L}_{t\bar{t}H} = \frac{-y_t}{2}\bar{\psi}_t(\kappa_t + i\gamma_5\tilde{\kappa}_t)\psi_t H$$
- Modifications on cross section and kinematic variables
- Focuses on three most sensitive categories ($2\ell ss+0\tau$, 3ℓ and $2\ell ss+1\tau$)
- BDT trained in each category to exploit kinematic differences between $t\bar{t}H$ CP-even and CP-odd, outputs used to further classify the events in the $t\bar{t}H$ node

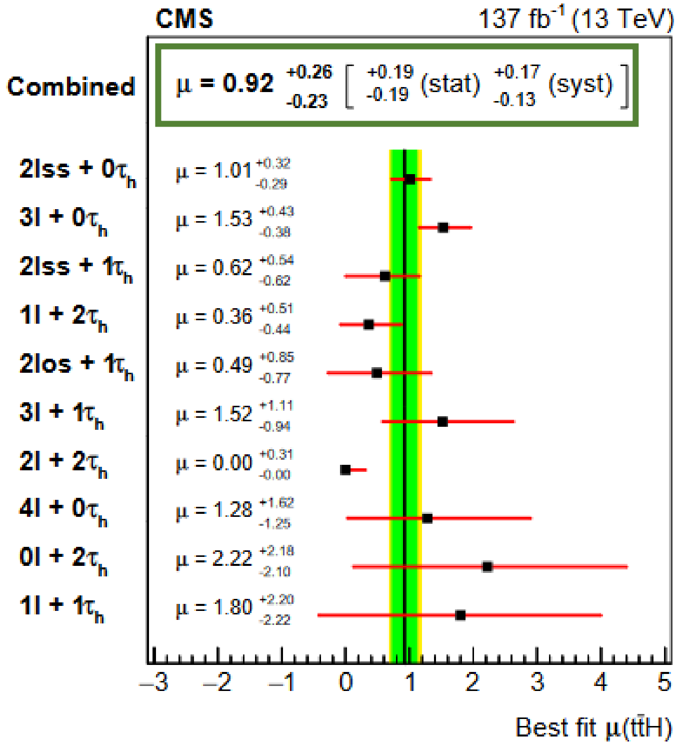


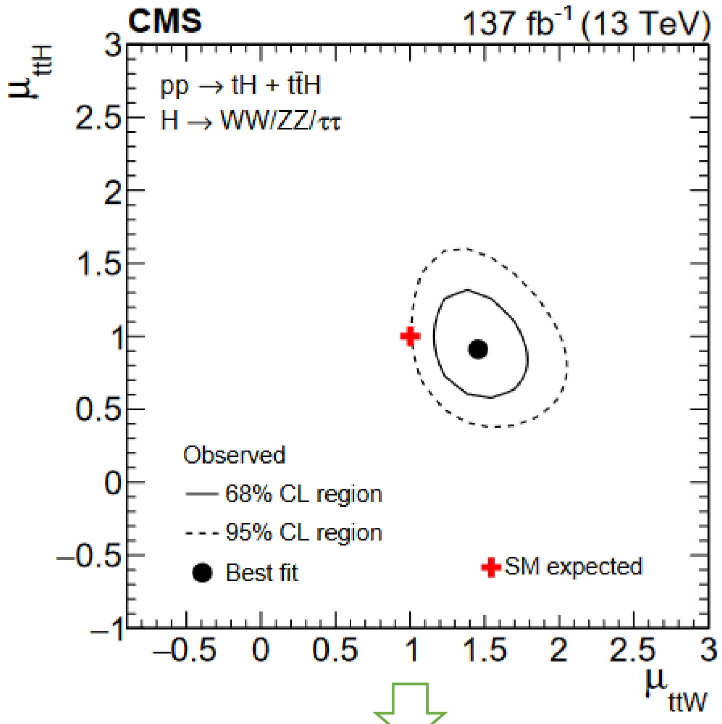




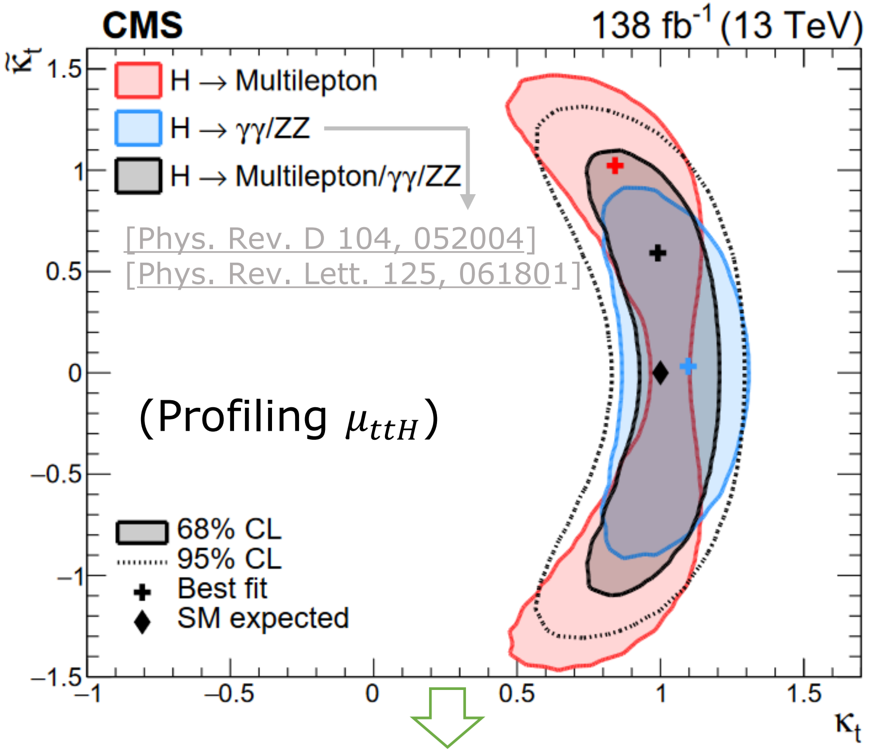
$t\bar{t}H$ result

- Simultaneous **maximum likelihood fit** in the signal region categories as well as the control regions
  - $t\bar{t}W$ and $t\bar{t}Z$ signal strengths (μ) freely floated in the fit
- **CP interpretation:** using kinematic differences between $t\bar{t}H$ CP-even and CP-odd components
  - Yields parametrized using: κ_t and $\tilde{\kappa}_t$ (ratio of the CP-even and CP-odd terms to SM expectation, respectively)





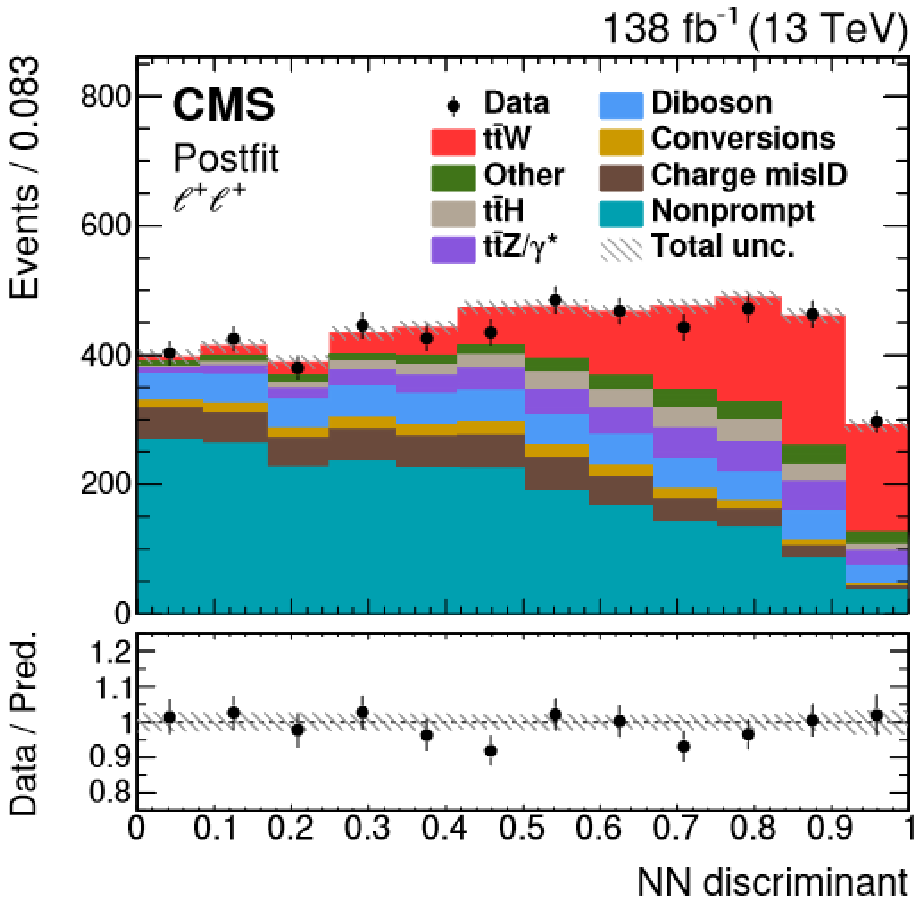
Small tension on $\mu(t\bar{t}W)$

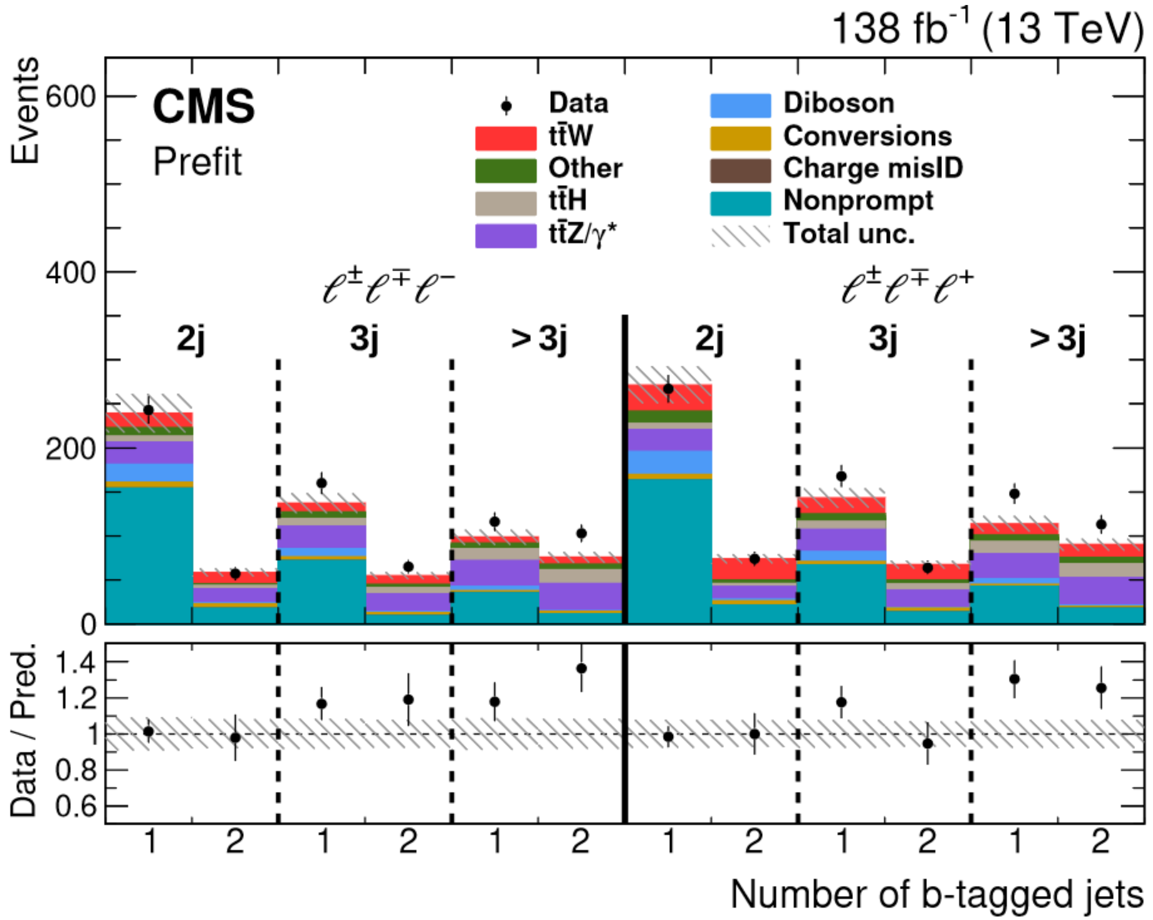


Exclude pure CP-odd with 3.7σ

$t\bar{t}W$ event classification

- **2ℓss**: a NN is used to distinguish signal from background
  - Further classification depending on flavor and **charge** of the leptons
- **3ℓ** : Events categorized using charge of the leptons, jet and b-tag multiplicities.
  - Invariant mass of the 3ℓ is used as discriminating variable



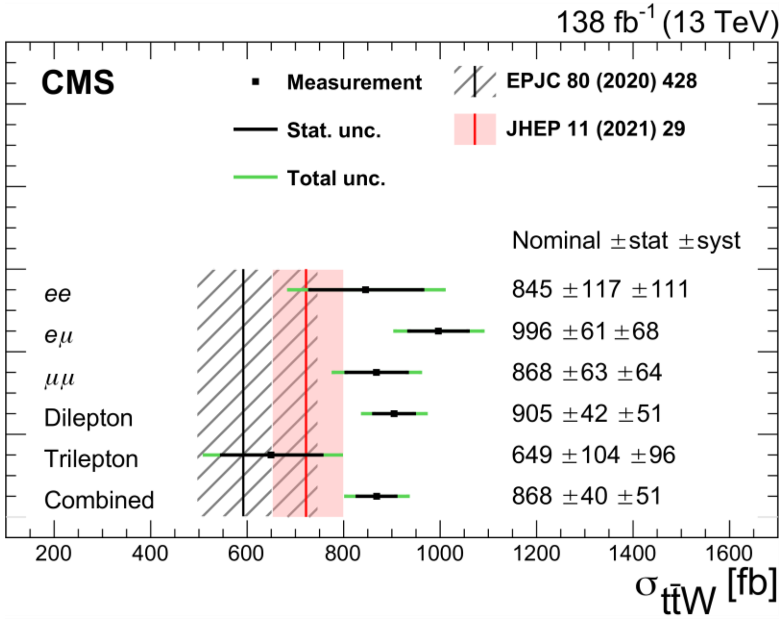


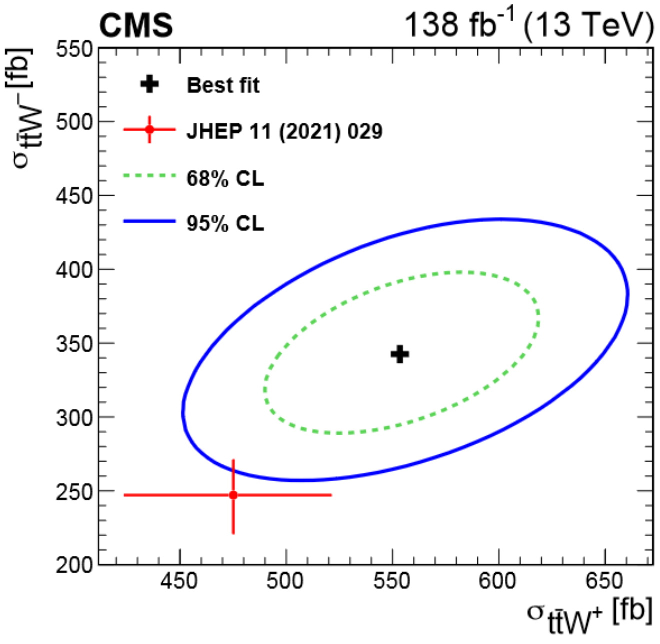
$t\bar{t}W$ Results

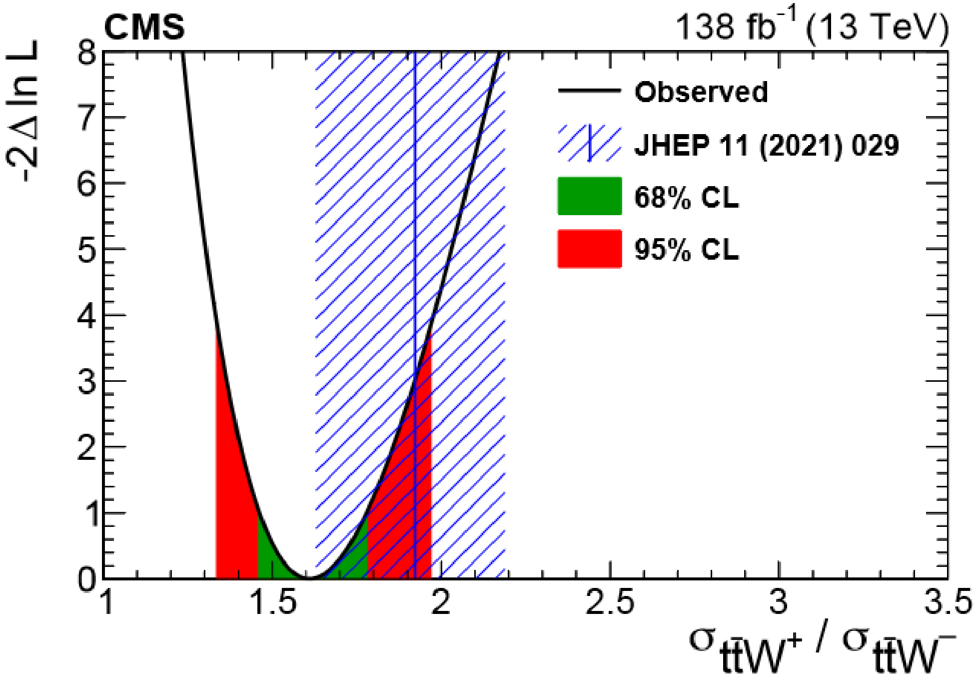
Simultaneous **maximum likelihood fit** performed using signal regions as well as control regions

- $t\bar{t}Z$ normalization is freely floated in the fit
- $\sigma_{t\bar{t}W}$ = 868 ± 40 ± 51 fb (reducing syst. unc. by a factor >2 wrt. previous CMS measurement)
- $\boldsymbol{\mu}(\mathbf{t\bar{t}W}) = \mathbf{1.47} \pm \mathbf{0.11}$ → compatible with SM within 2 s.d.
- Compatible with EPJC 81, 378 (2021) and ATLAS-CONF-2023-019

- Also perform a simultaneous extraction of $t\bar{t}W^+$ and $t\bar{t}W^-$
- The ratio is found to be $1.61\ \pm\ 0.15\ (stat)^{+0.07}_{-0.05}\ (syst)$ → good agreement with SM







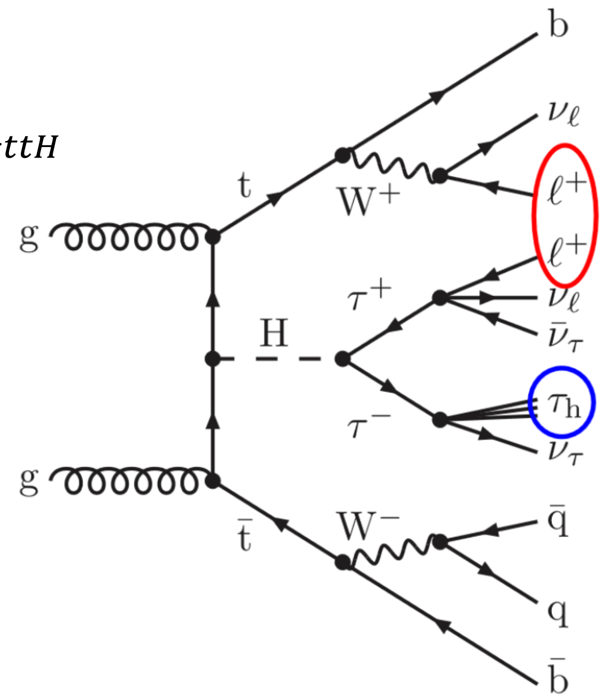
Prospects for $t\bar{t}H$

Work in progress

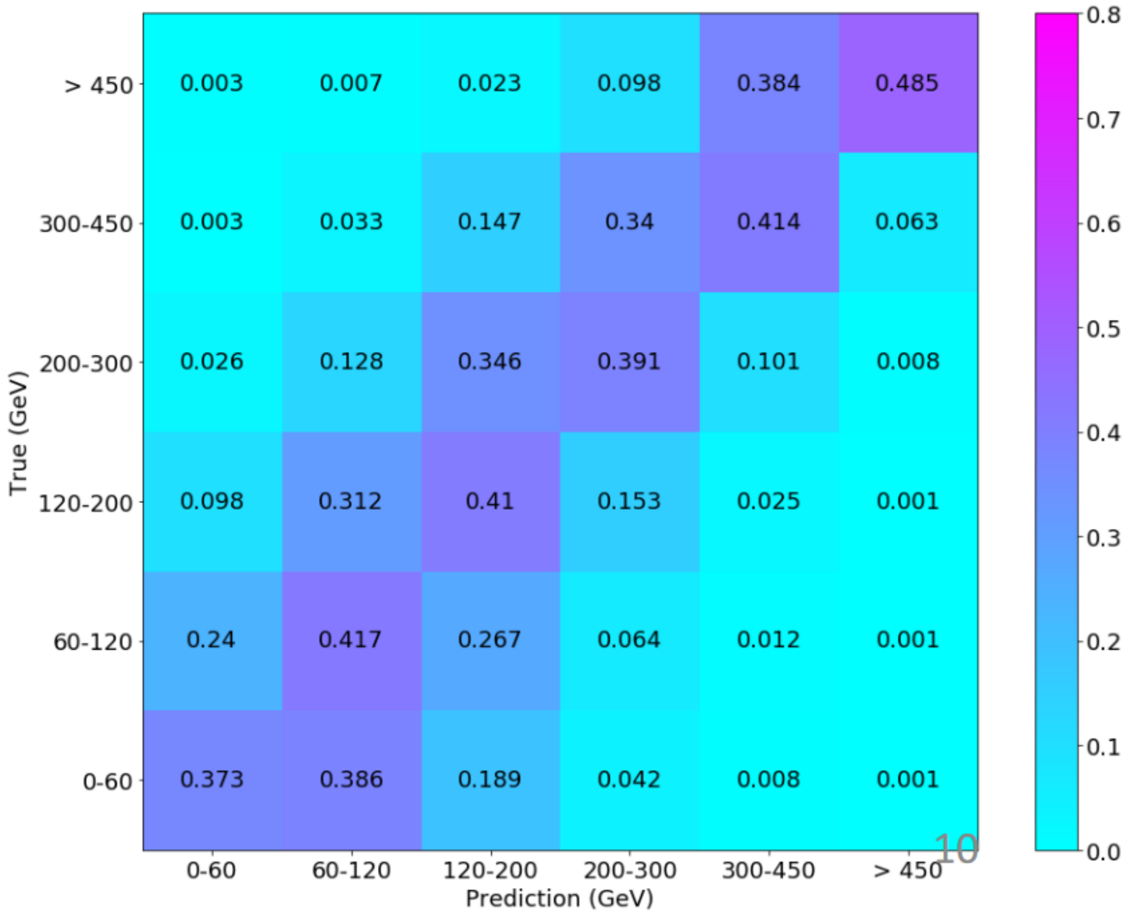
Aim to perform differential measurements using full Run 2 dataset

Study differentially: p_T^H and m_{ttH}

- Use maximum likelihood fit unfolding
- DNN used to regress the p_T (several strategies tried)
  - Kinematic information from objects, top tagger, missing transvers energy, jet multiplicity…
- Proxy variable for m_{ttH}



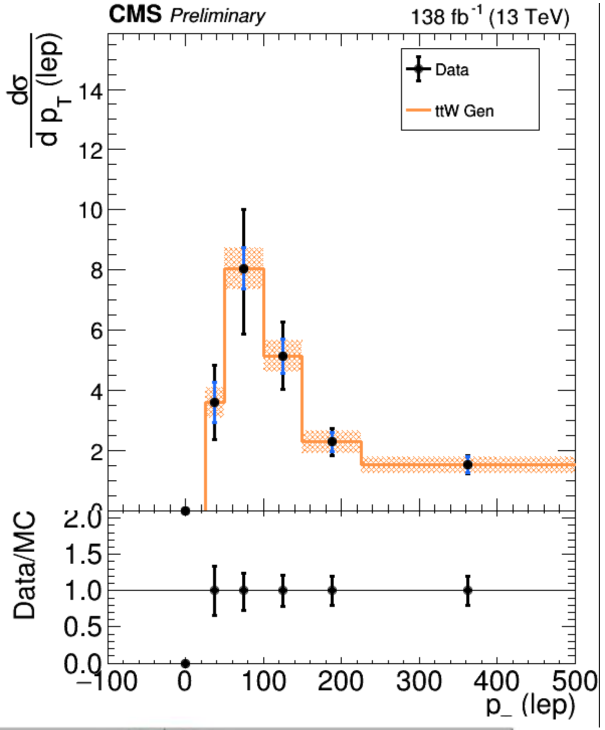
DNN p_T^H regressor in the $2\ell ss+1\tau$



Prospects for $t\bar{t}W$

Work in progress

Aim to perform differential measurements using full Run 2 dataset

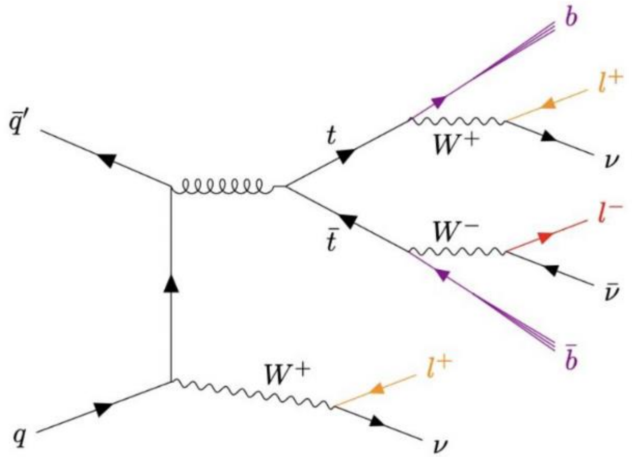


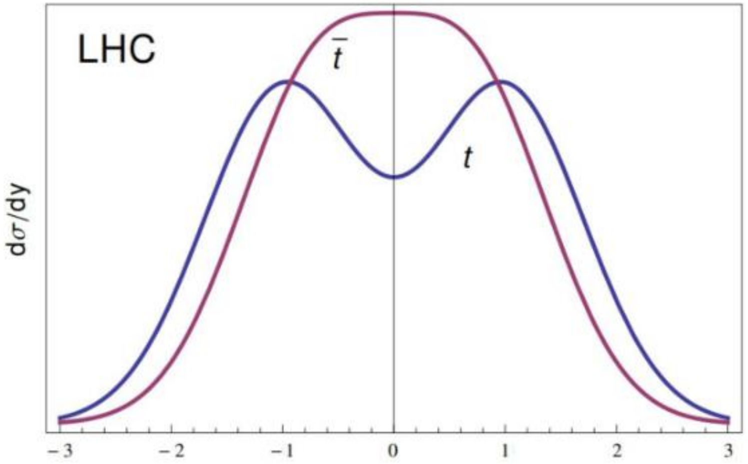
- Study several observables, useful to provide feedback to theorist
  - Use maximum likelihood fit unfolding
  - Will use updated state-of-the-art MC including **new FxFx** merging [JHEP 11 (2021) 29]

$t\bar{t}W$ can be used to measure back-forward $t\bar{t}$ asymmetry in case of $q\bar{q}$ initial state.

- Need to tag the lepton coming form the top and antitop

$$A_C = A_C^{\ell} = \frac{N(\Delta|\eta^{\ell}| > 0) - N(\Delta|\eta^{\ell}| < 0)}{N(\Delta|\eta^{\ell}| > 0) + N(\Delta|\eta^{\ell}| < 0)}$$





Summary

- Run 2 allowed to measure low cross section processes with high precision:
  - Unprecedent amount of data
  - Improvement on nonprompt background rejection
  - Better control of systematic uncertainties

- $\mathbf{t\bar{t}H}$ process allowed to study the top-Higgs interaction, results are in good agreement with the SM
  - CP violation in the Higgs sector also studied

- $\mathbf{t\bar{t}W}$ cross section observed to be above predictions → need to improve modelling of this process

- Amount of data available also allows to perform differential measurements of both $\mathbf{t\bar{t}H}$ and $\mathbf{t\bar{t}W}$ processes → **coming soon!**

Back-up

Strategy

- Data taken by the CMS experiment at 13 TeV during Run 2 (138 fb^{-1}).
- Using multilepton final states, categorize events depending on the lepton multiplicity
  - For ttH analysis 12 categories (focusing on the 3 most sensitive)
- Select isolated muons from H, W and tau decays is crucial in this analysis
  - Use a dedicated MVA to distinguish those leptons form nonpromt leptons -> reduce one of the main backgrounds
- Dedicated selection on each category:

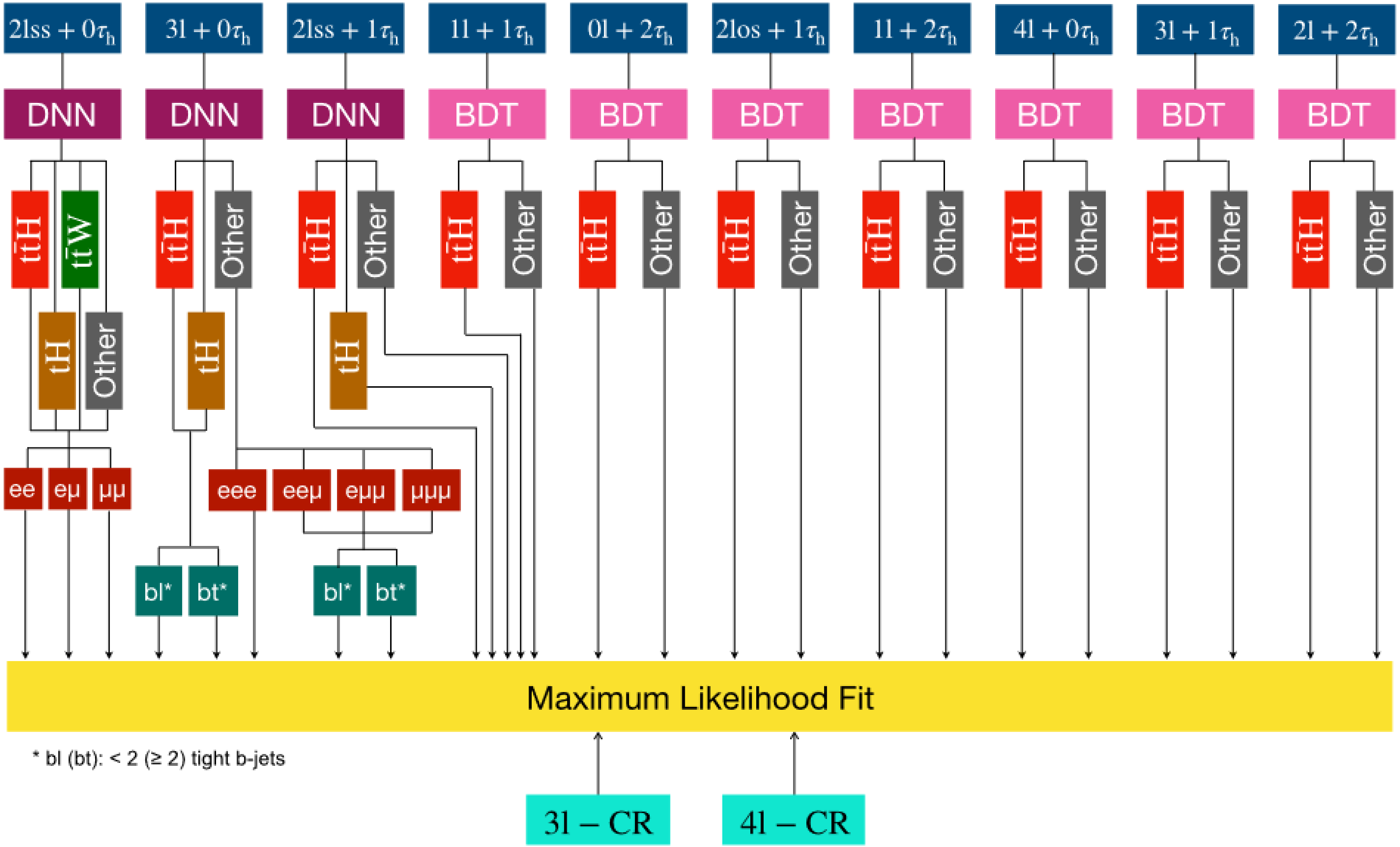
	ttH			**ttW**	
	2lss+0tau	2lss +1 tau	3l	2lss+0tau	3l
nJets	≥ 2		≥ 3	≥ 2	
nbtag	$\geq$ 1 medium b-tagged Jet or $\geq$ 2 loose b-tagged Jet				
MET	> 30 GeV		>30 * > 45 GeV **	>30 GeV	-
$\sum_i q_l^i$			± 1	± 1	± 1
	Veto os leptons within Z peak (10 GeV mass window)				

* If NJet $\leq$ 3 and 1 SFOS

* NJet $\leq$ 3 and 0 SFOS

After selection still dominated by background → Use NN to classify events

ttH categories



* bl (bt): < 2 (≥ 2) tight b-jets

ttH syst

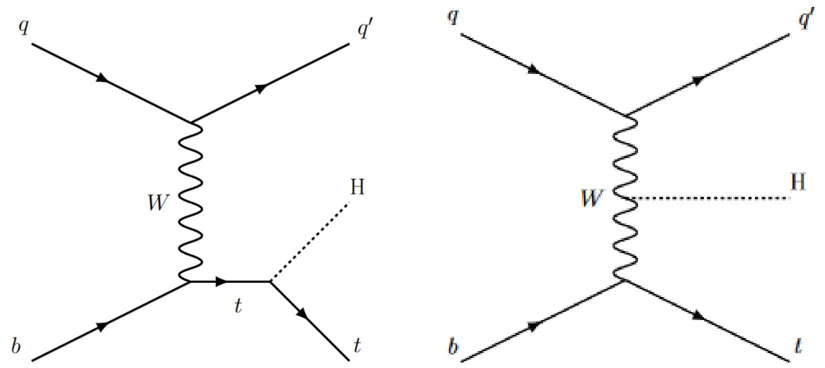
Source	$\Delta\mu_{t\bar{t}H}/\mu_{t\bar{t}H}$ [%]	$\Delta\mu_{tH}/\mu_{tH}$ [%]	$\Delta\mu_{t\bar{t}W}/\mu_{t\bar{t}W}$ [%]	$\Delta\mu_{t\bar{t}Z}/\mu_{t\bar{t}Z}$ [%]
Trigger efficiency	2.3	8.1	1.2	1.9
e, μ reconstruction and identification efficiency	2.9	7.1	1.7	3.2
τ_h identification efficiency	4.6	9.1	1.7	1.3
b tagging efficiency and mistag rate	3.6	13.6	1.3	2.9
Misidentified leptons and flips	6.0	36.8	2.6	1.4
Jet energy scale and resolution	3.4	8.3	1.1	1.2
MC sample and sideband statistical uncertainty	7.1	27.2	2.4	2.3
Theory-related sources affecting acceptance and shape of distributions	4.6	18.2	2.0	4.2
Normalization of MC-estimated processes	13.3	12.3	13.9	11.3
Integrated luminosity	2.2	4.6	1.8	3.1
Statistical uncertainty	20.9	48.0	5.9	5.8

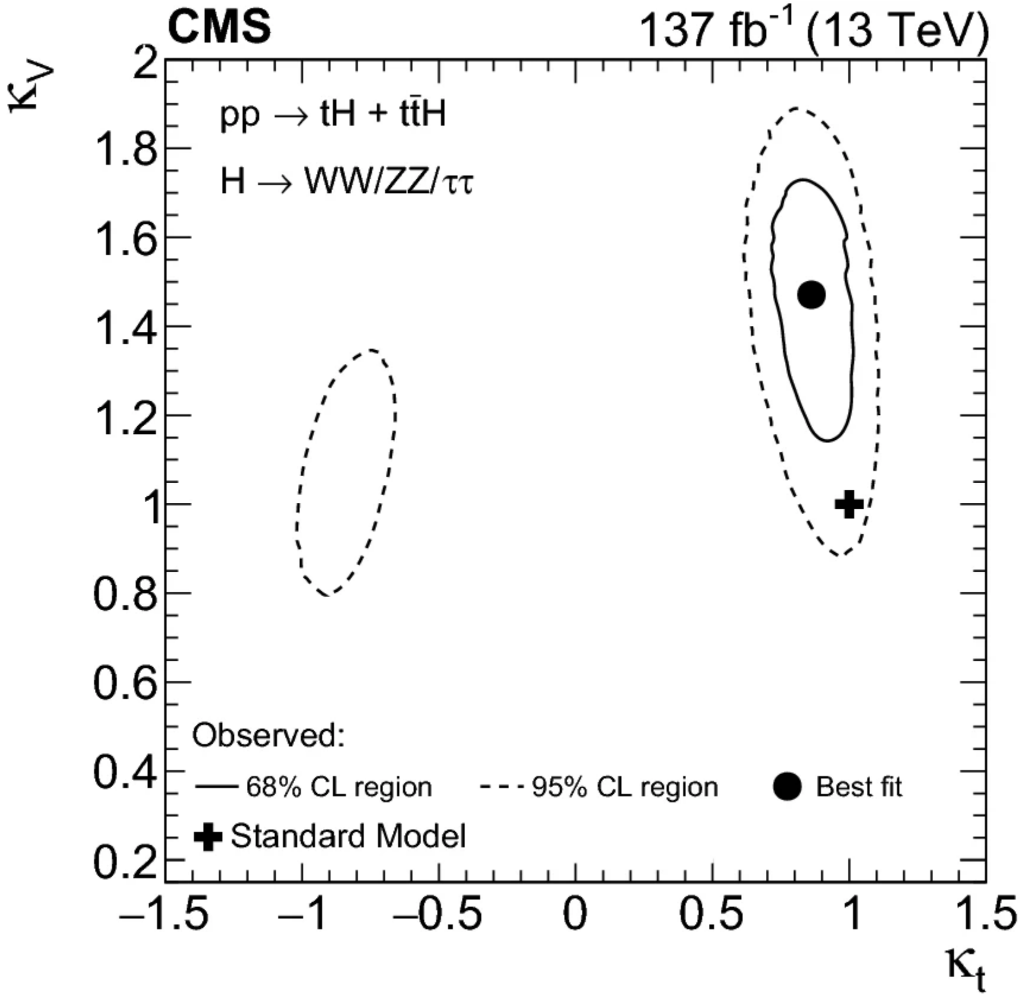
ttW syst

Source	Uncertainty [%]
Experimental uncertainties	
Integrated luminosity	1.9
b tagging efficiency	1.6
Trigger efficiency	1.2
Pileup reweighting	1.0
L1 inefficiency	0.7
Jet energy scale	0.6
Jet energy resolution	0.4
Lepton selection efficiency	0.4
Background uncertainties	
$t\bar{t}H$ normalization	2.6
Charge misidentification	1.6
Nonprompt leptons	1.3
VVV normalization	1.2
$t\bar{t}VV$ normalization	1.2
Conversions normalization	0.7
$t\bar{t}\gamma$ normalization	0.6
ZZ normalization	0.6
Other normalizations	0.5
$t\bar{t}Z$ normalization	0.3
WZ normalization	0.2
tZq normalization	0.2
tHq normalization	0.2
Modeling uncertainties	
$t\bar{t}W$ scale	1.8
$t\bar{t}W$ color reconnection	1.0
ISR & FSR scale for $t\bar{t}W$	0.8
$t\bar{t}\gamma$ scale	0.4
VVV scale	0.3
$t\bar{t}H$ scale	0.2
Conversions	0.2
Simulation statistical uncertainty	1.8
Total systematic uncertainty	5.8

Htt coupling

- $\kappa_t = y_t / y_t^{SM}$ in good agreement with SM
- tH sensitive to the relative sign of kv kt
  - BSM yt =-1→ constructive interference → σ_{tH} ~ 0.8 pb (10 times greater than in the SM)

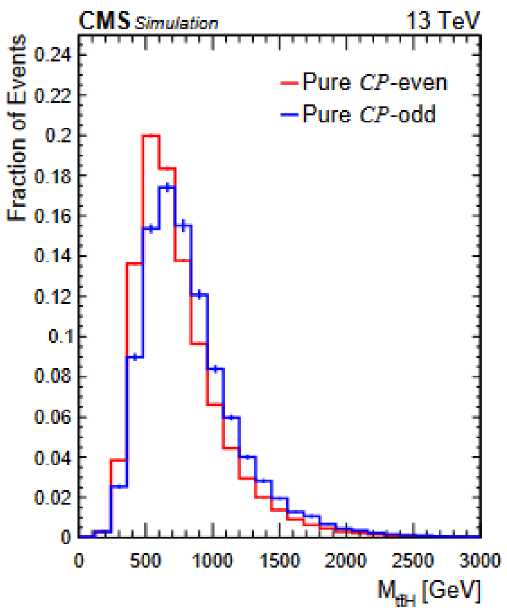


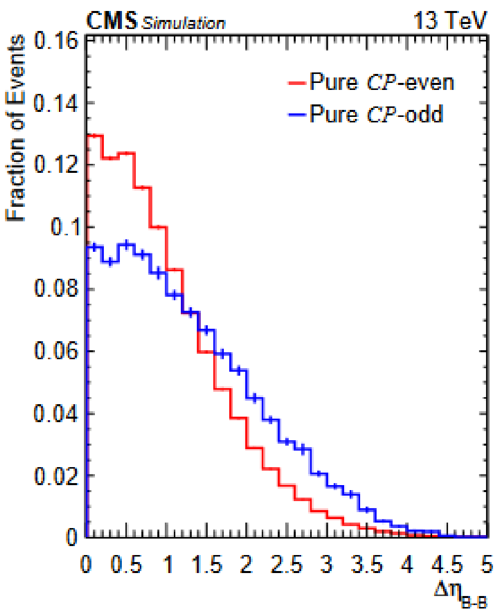


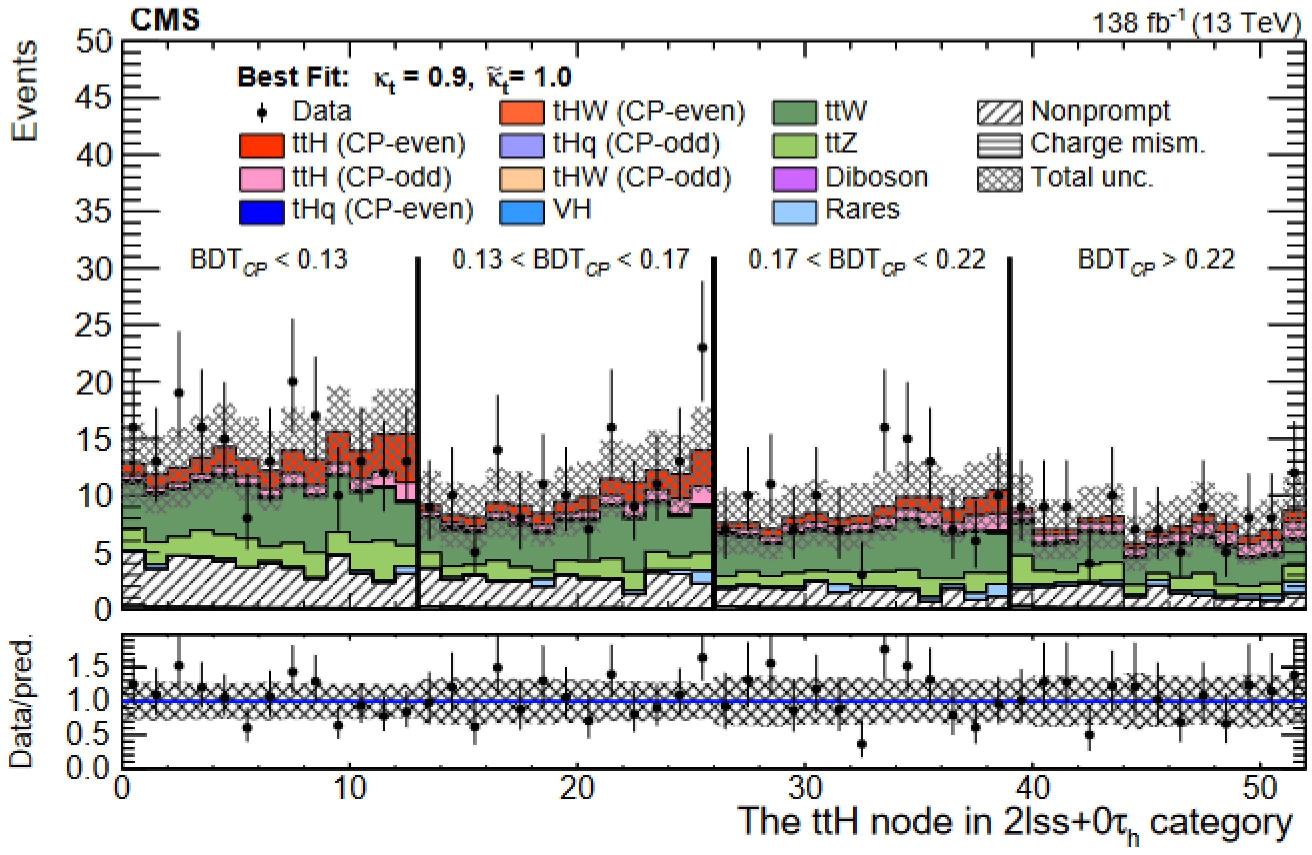
CP interpretation

Kinematic differences between ttH CP-even and CP-odd components are exploited → dedicated BDT in each of the 3 most sensitive ttH enriched categories

Inputs: momentum of leptons and jets, angular variables, mases, object multiplicities and a specific tagger targeting hadronic top quark decays.







CP interpretation

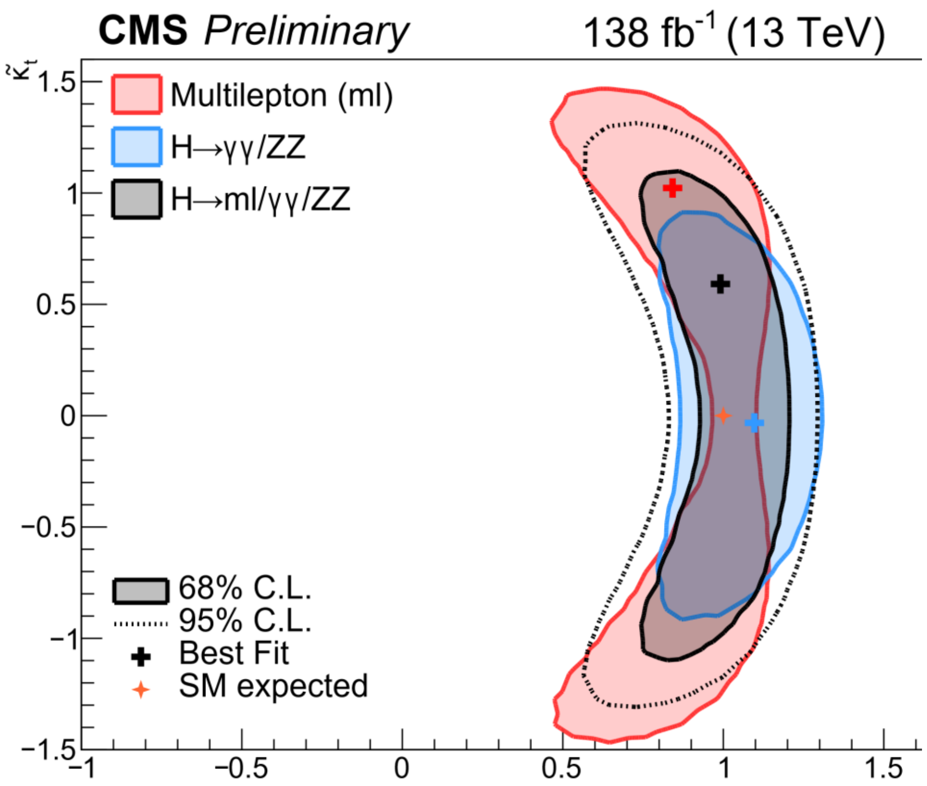
Yields are parametrized using:

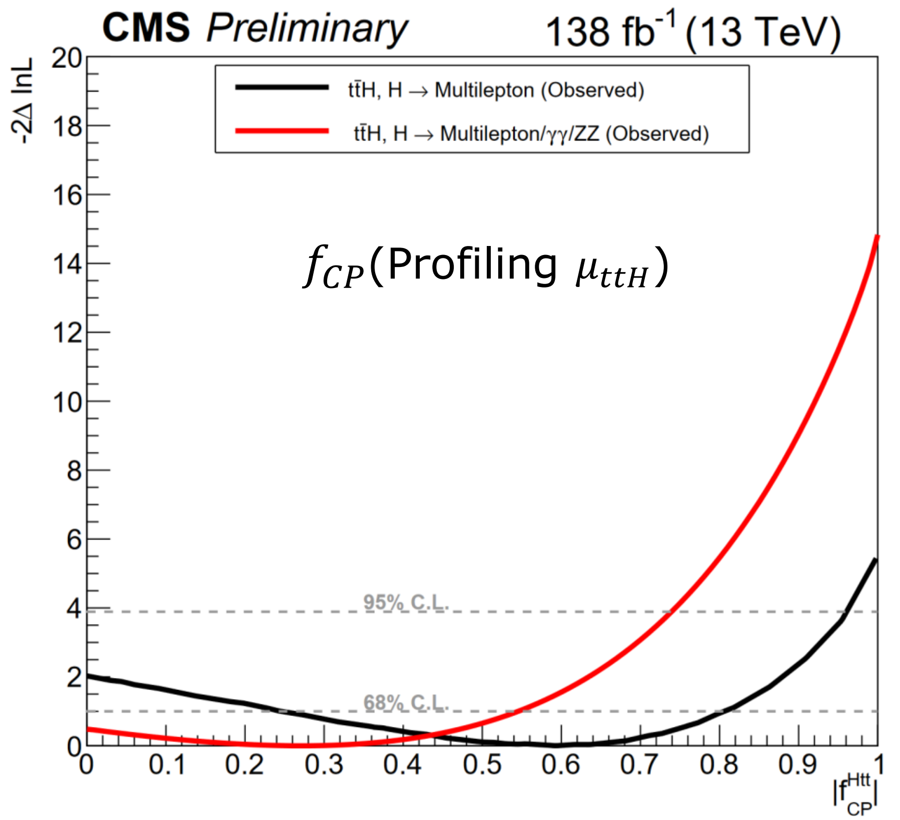
- κ_t and $\tilde{\kappa}_t$ (ratio of the CP-even and CP-odd terms to SM expectation, respectively)

$$f_{CP} = \frac{|\tilde{\kappa}_t|^2}{|\tilde{\kappa}_t|^2 + |\kappa_t|^2}$$

The result is combined with already published ttH measurements:

- ZZ, Phys. Rev. D 104, 052004
- $\gamma\gamma$, Phys. Rev. Lett. 125, 061801





$|f^{Htt}_{CP}| = 1$ **excluded with 3.7σ**

ttH diff/EFT

Neural Network for p_T^H regression input variables:

Input	Number of Variables	Which Channels
$l_1(p_T, \eta, \phi)$	3	$2\ell ss$ & 3ℓ
$l_2(p_T, \eta, \phi)$	3	$2\ell ss$ & 3ℓ
$l_3(p_T, \eta, \phi)$	3	3ℓ
$t_{had}(p_T, \eta, \phi)$	3	$2\ell ss$ & 3ℓ
t_{had} BDT Score	1	$2\ell ss$ & 3ℓ
E_T^{miss}	1	$2\ell ss$ & 3ℓ
$\phi_{E_T^{miss}}$	1	$2\ell ss$ & 3ℓ
$\sum_{n=1}^{5} j_n(p_T, \eta, \phi)$	3	$2\ell ss$ & 3ℓ
$\sum_{n>5} j_n(p_T, \eta, \phi)$	3	$2\ell ss$ & 3ℓ
$\sum_n j_n + \sum_n l_n(p_T, \eta, \phi)$	3	$2\ell ss$ & 3ℓ
Total	21	-

Table 19: Input variables to DNN used for p_T regression.

Neural Network for p_T^H regression input variables:

Name	Operator	Comments
ctp	$\bar{q}_i u_j \tilde{\phi} \left(\phi^\dagger \phi\right)$	Effects on tHq
cpt	$\left(\phi^\dagger i\overleftrightarrow{D}_\mu \phi\right)\left(\bar{u}_i \gamma^\mu u_j\right)$	Effects on $t\bar{t}$H, $t\bar{t}$W, $t\bar{t}$Z, and tZq
cptb	$\left(\tilde{\phi}^\dagger iD_\mu \phi\right)\left(\bar{u}_i \gamma^\mu d_j\right)$	Effects on tHqand tZq
ctG	$\left(\bar{q}_i \sigma^{\mu\nu} T^A u_j\right)\tilde{\phi} G^A_{\mu\nu}$	Effects on every process with a top quark
cpG	$\left(\phi^\dagger \phi\right) G^A_{\mu\nu} G^{A\mu\nu}$	Effects on every QCD process

ttW diff/EFT

We are exploring several observables

- Number of jets
- HT (scalar sum of jet pt)
- Number of b-jets
- Leading b-jet pt
- Leading lepton pt
- Minimum ΔR(leading lepton, jet)
- ΔR(leptons)
- Maximum | η(lepton) |
- Leading lepton pt and eta
- Leading jet pt
- Leading b-jet pt
- Number of jets
- Δη(ll)

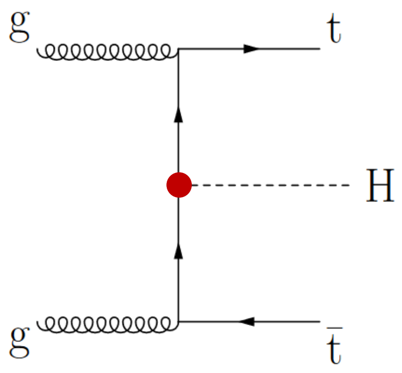
Introduction

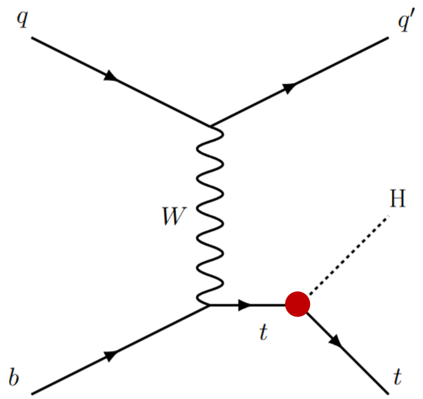
Aim: Study the production of a top antitop quark pair produced in association with a H or W boson using full run 2 dataset (13 TeV,138 fb^{-1})

$\mathbf{t\bar{t}H}$ allows to study directly the Yukawa coupling

- $y_t \sim 1$
- **BSM physics** could introduce **modified couplings,** in particular **CP-violating coupling**

Will summarize the studies from EPJC 81, 378 (2021) and JHEP 07 (2023) 092





$\mathbf{t\bar{t}W}$ production is one of the **main background for $\mathbf{t\bar{t}H}$, $\mathbf{t\bar{t}t\bar{t}}$** …

- Measured $\boldsymbol{\sigma}(\mathbf{t\bar{t}W})$ **consistently above theory** value (both in ATLAS and CMS)
- Active discussion about the modelling of this background

Will summarized the studies from JHEP 07 (2023) 219