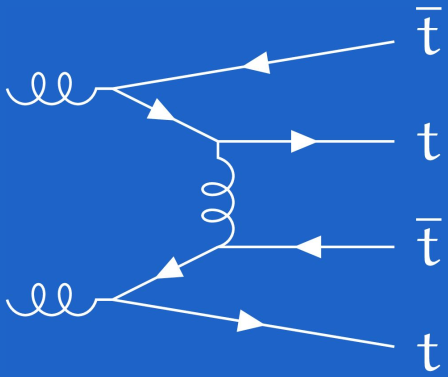
TOP2023 - 16th International Workshop on Top Quark Physics

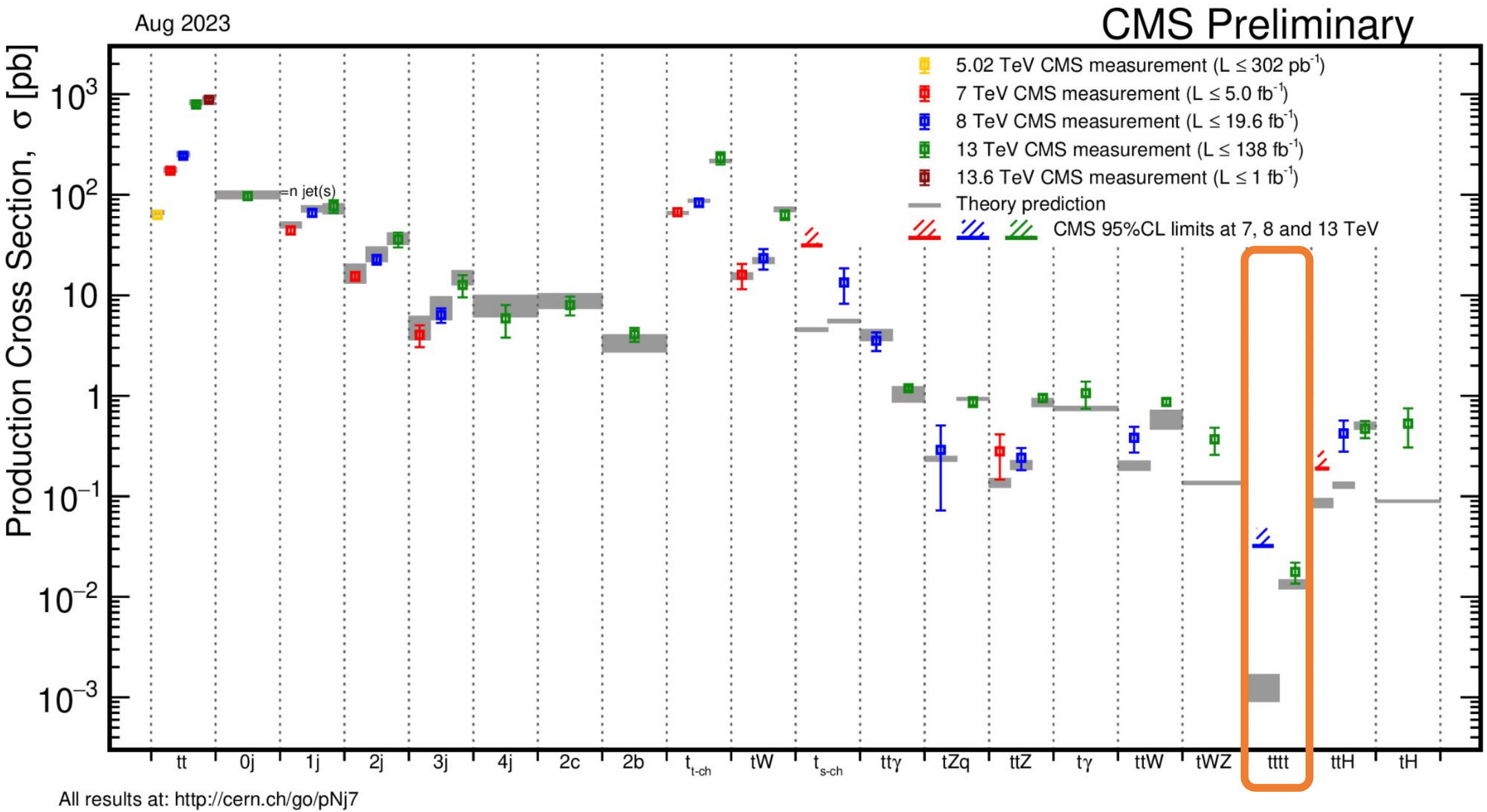


Four top searches and measurements

Niels Van den Bossche on behalf of the ATLAS and CMS Collaborations

GHENT UNIVERSITY

Four top production in the top sector



Introduction

Four top production ($t\bar{t}t\bar{t}$): a very rare standard model (SM) process

- $\sigma(t\bar{t}t\bar{t})_{NLO(QCD+EW)} = 12.0 \pm 2.4$ fb [JHEP 02 (2018) 031]
- $\sigma(t\bar{t}t\bar{t})_{NLO(QCD+EW)+NLL'} = 13.4^{+1.0}_{-1.8}$ fb [arXiv:2212.03259]

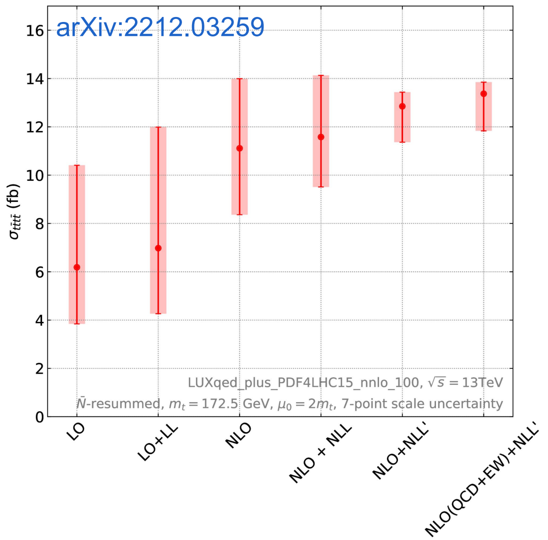
From a SM point of view:

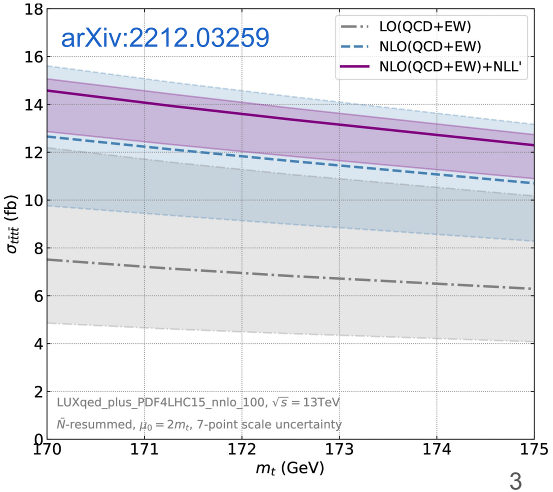
- Probe of top-Higgs Yukawa coupling
- Heaviest final state observed at LHC

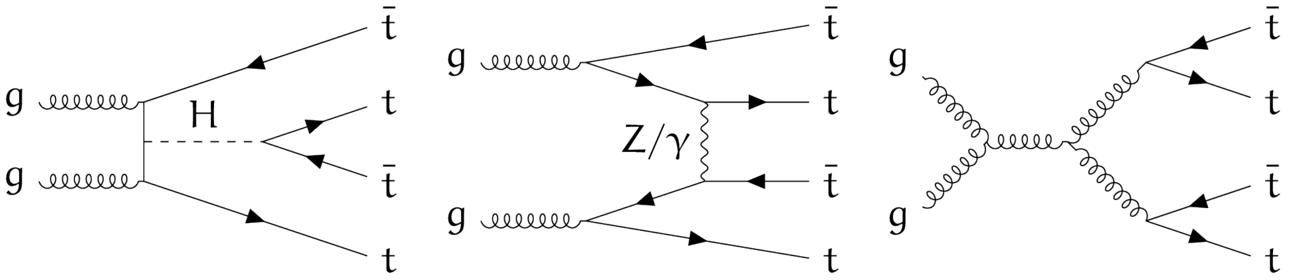
And for new physics:

- Sensitivity to wide range of new physics scenarios and effective field theory (EFT) operators

See next talk by A. Sharma!







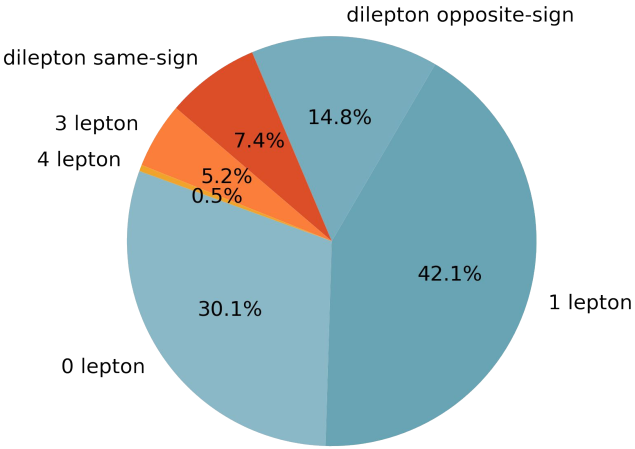
Four top final states

Four top production leads to large object multiplicity final state:

- 4 b-quarks leading to jets
- the decay products of 4 W bosons

Typically divided into three main analysis strategies:

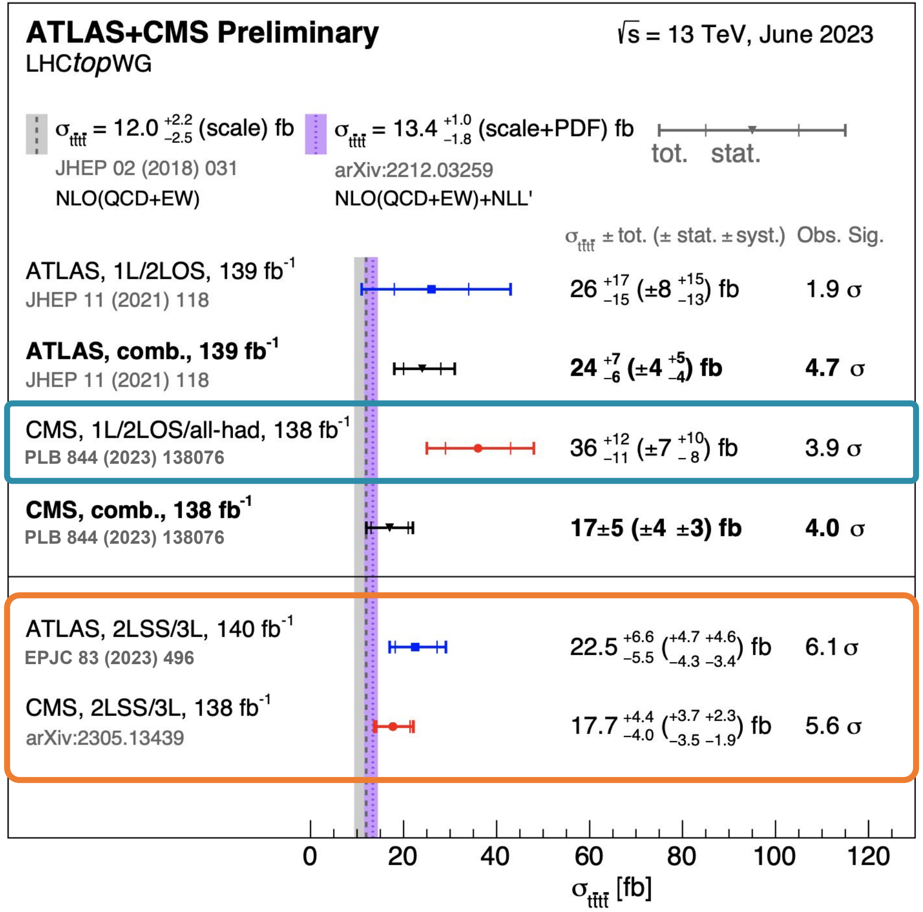
- All hadronic (0L)
- Single lepton and opposite sign dilepton (1L, OSDL)
  - Larger branching fraction and large irreducible background (from $t\bar{t}$)
- Same-sign dilepton and multilepton (SSDL, ML)
  - Smaller branching fraction and higher purity

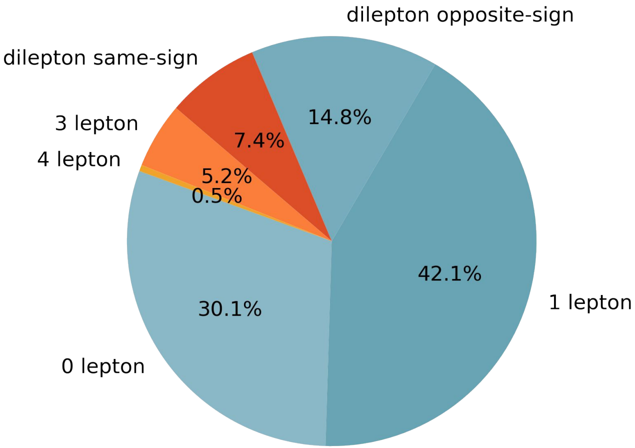


Heavy use of machine learning techniques for signal-to-background discrimination in all final states:

- Boosted decision trees (BDTs)
- Graph Neural Networks (GNNs)

Measurement landscape





CMS: all hadronic, 1L and OSDL final states

PLB 844 (2023) 138076

OSDL channel:

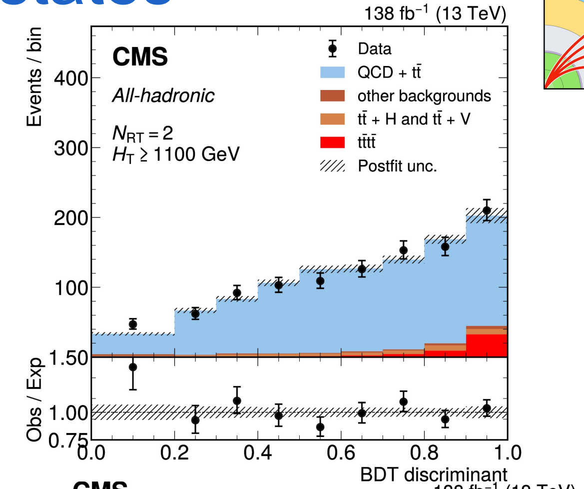
- Regions defined by leptons, jets, b-jets
- Low (b-)jet categories: control regions
- Fit HT distributions
  - HT: ∑ jet pT

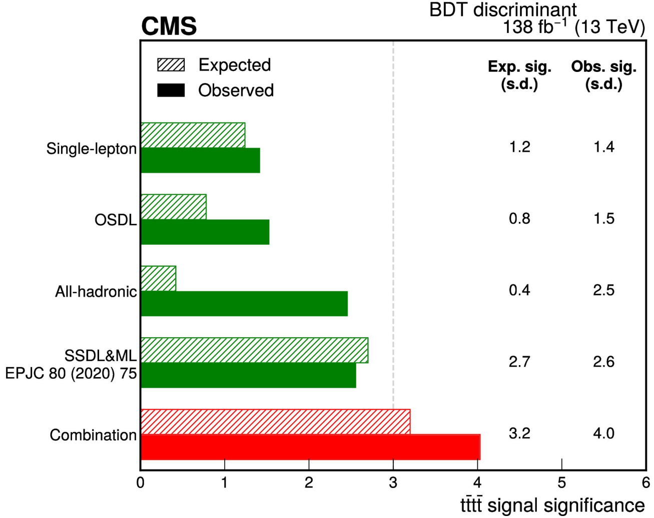
1L channel:

- Resolved top-tagger as BDT input
- Regions defined by leptons, jets, b-jets, top candidates
- BDT for signal/background separation

All-hadronic channel:

- First analysis to use it!
- Regions of resolved & boosted top candidates and HT
- Data-driven estimation of multijet and tt̄+jets background
- BDT for signal/background separation



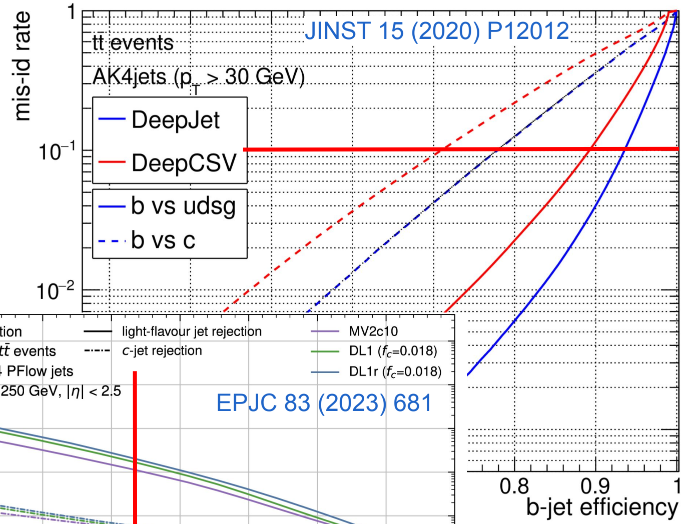


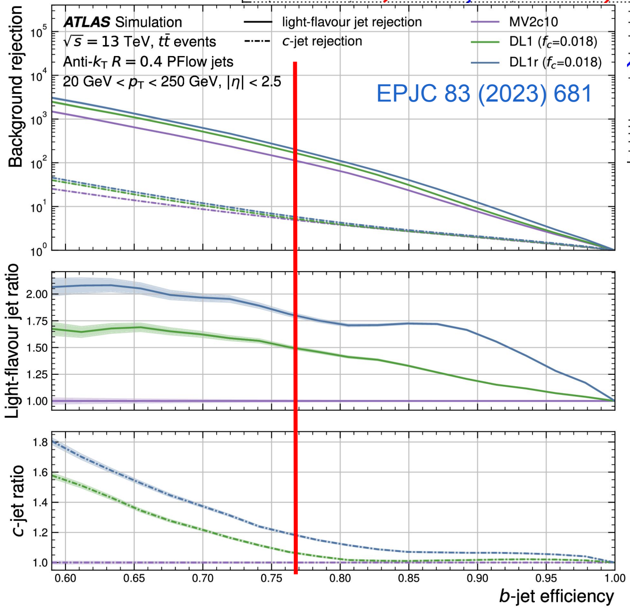
Observation in SSDL and ML channels

ATLAS: EPJC 83 (2023) 496
CMS: arXiv:2305.13439 (submitted to PLB)

First observation of four top production at both ATLAS and CMS

- Re-analysis of Run 2 datasets
  - Supersede previous results
- Profit significantly from general improvements in lepton and jet selection:
  - Better reconstruction methods
  - **Improved b-tagging**
  - Better lepton identification methods
- Major improvements in analysis methods
  - Stronger machine learning discriminants: GNNs (ATLAS) or multiclass BDTs (CMS)
  - Better handles on $t\bar{t}X$ backgrounds

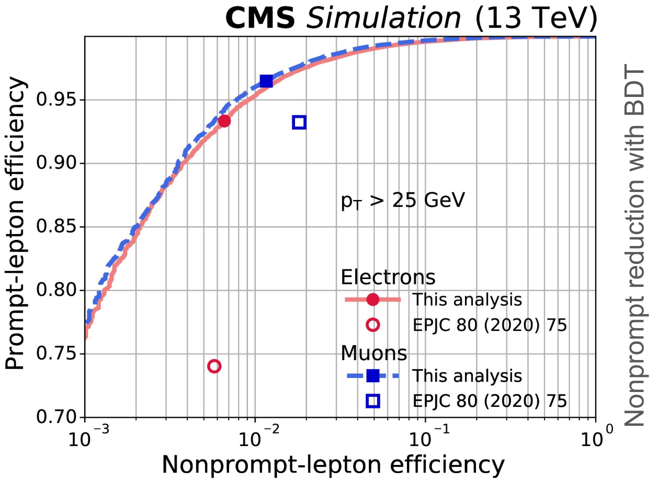


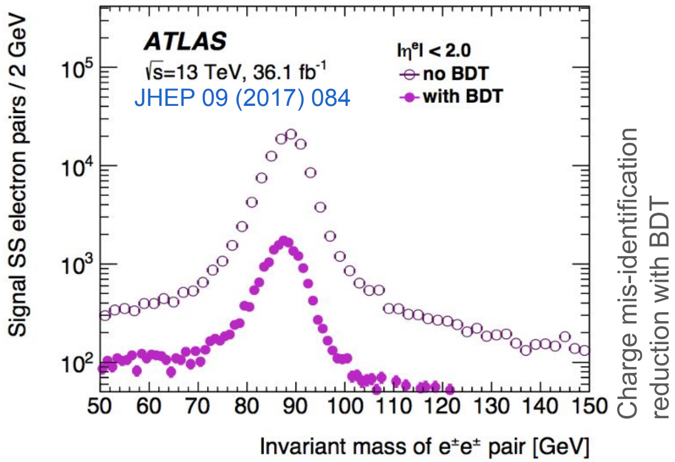


SSDL and ML channels: reducible backgrounds

tt̄ with additional nonprompt leptons

- e.g. from semileptonic b-decays
- Reduced with per-lepton BDTs using b-tagging and isolation information
- Standard datadriven methods for prediction
  - CMS: tight-to-loose ratio
  - ATLAS: MC shapes, normalization from fit



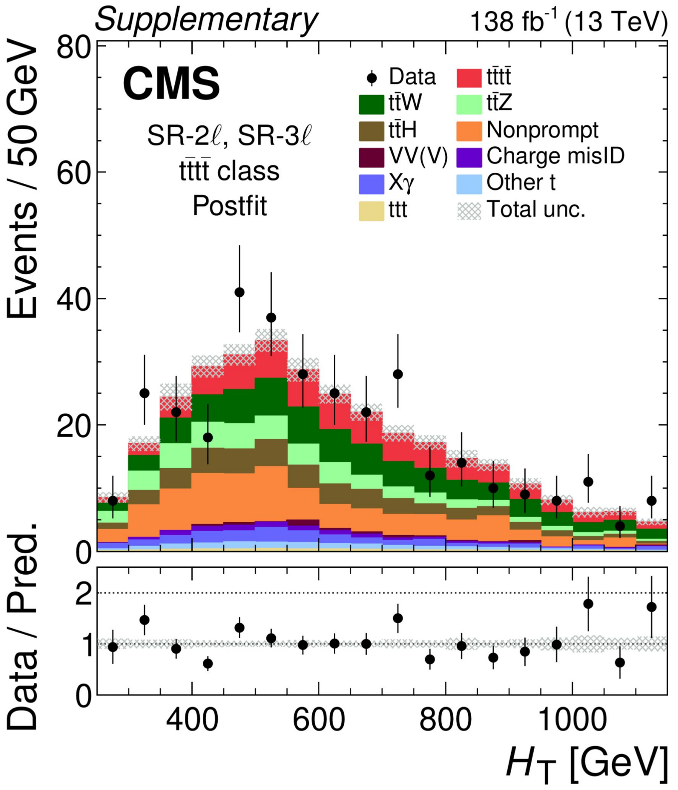


Events with charge mis-identified electrons

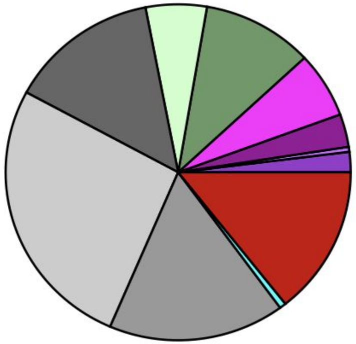
- e.g. due to Bremsstrahlung
- Mainly Z+jets and tt̄+jets events
- Datadriven prediction using DY events

SSDL and ML channels: irreducible backgrounds

- $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}H$
  - Large systematic uncertainty on $t\bar{t}X$ + b(b) component
  - $t\bar{t}W$ dominant in SSDL final state, $t\bar{t}Z$ in 3L
  - $t\bar{t}H$ subdominant contribution in all considered final states
- Diboson processes
  - mainly WZ, same-sign WW
- Single top processes with associated vector bosons



Background mix in ATLAS signal region



ATLAS: $t\bar{t}W$ modelling

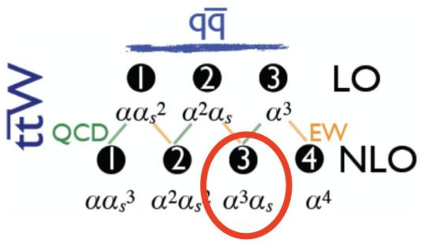
$t\bar{t}W$ modelling at large jet multiplicities corrected using data

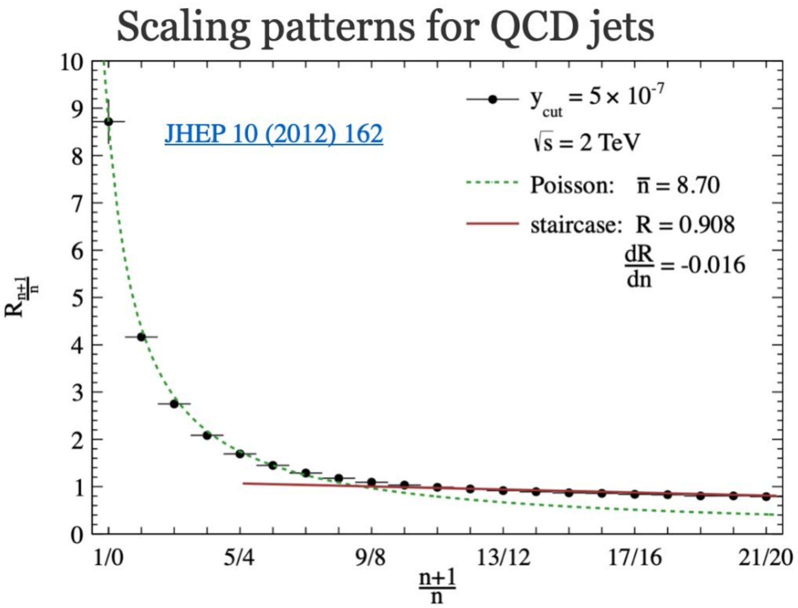
- Main sherpa sample: 1j@NLO, 2j@LO QCD
  - Additive weight considering LO3 and NLO4
- Additional $t\bar{t}W$ EW sample: NLO3 term
  - Known to be sizeable (JHEP 02 (2018) 031)

N_{Jets} distribution corrected using jet scaling regimes:

- R(j) := N(j+1)/N(j), j is the jet multiplicity
- Staircase: R(j) = a_0, valid at high jet multiplicities
- Poisson: R(j) = a_1 / (1+n), n the number of additional jets

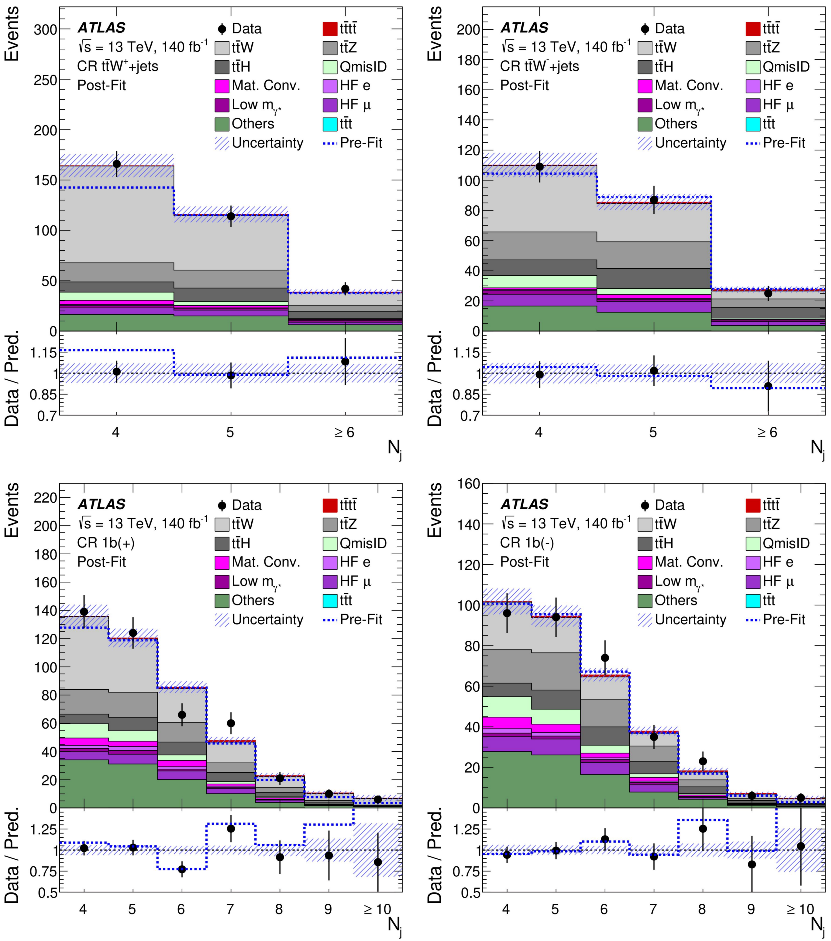
In addition: separate normalization factors (NFs) for $t\bar{t}W^+$ and $t\bar{t}W^-$

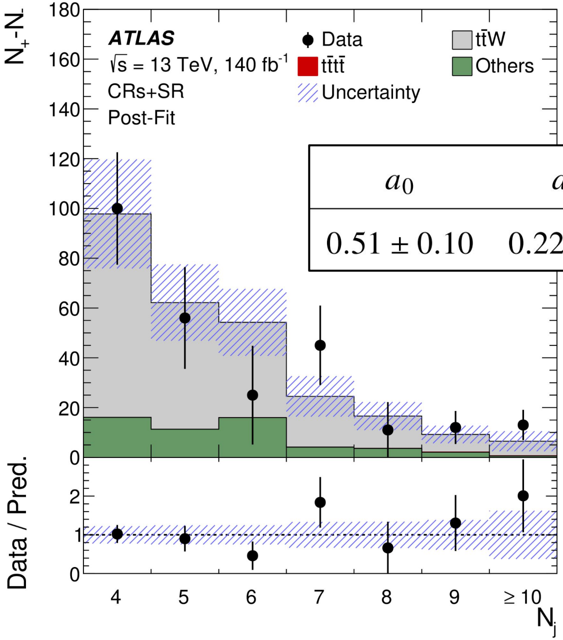




ATLAS: $t\bar{t}W$ modelling

- 4 dedicated control regions to determine a_0, a_1 and 2 NFs
- N_+-N_- to examine $t\bar{t}W$ modelling
  - Good agreement between data and prediction



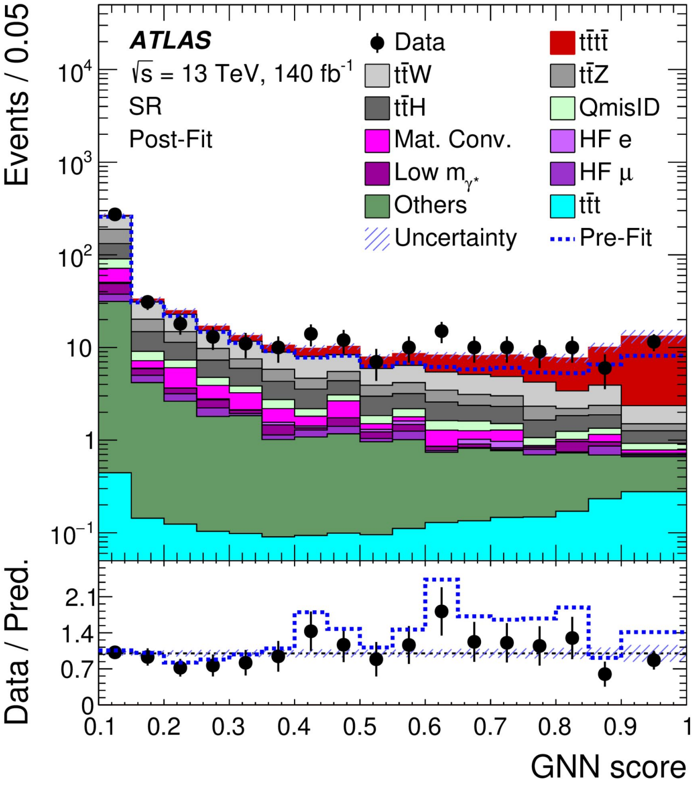


a_0	a_1	$NF_{t\bar{t}W^+(4jet)}$	$NF_{t\bar{t}W^-(4jet)}$
0.51 ± 0.10	$0.22^{+0.25}_{-0.22}$	$1.27^{+0.25}_{-0.22}$	$1.11^{+0.31}_{-0.28}$

ATLAS: summary

Signal extraction:

- 4 control regions for tt̄W modelling
- 4 control regions for nonprompt and conversions background
- Single signal region (combines SSDL and ML events)
  - ≥ 6 jets, ≥ 2 b-jets, HT > 500 GeV
  - HT: ∑ jet and lepton pT
- Graph neural network to separate signal/background



Sensitivity: 6.1σ observed (4.7σ expected)

Measured cross section: $22.5^{+4.7}_{-4.3}(\text{stat})^{+4.6}_{-3.4}(\text{syst})$ fb

SM expectation: σ(tt̄tt̄) = 13.4 fb (arXiv:2212.03259)

Bunch of interpretations: see next talk by A. Sharma

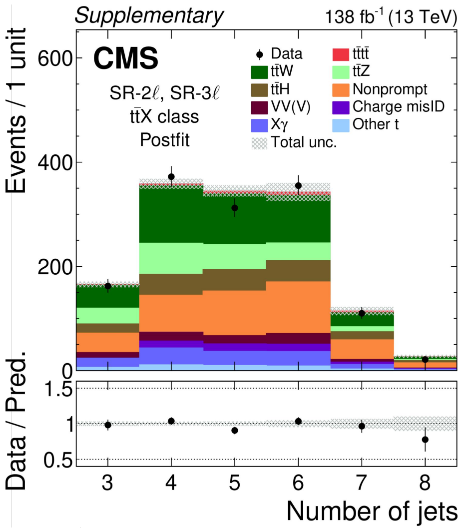
CMS: background prediction

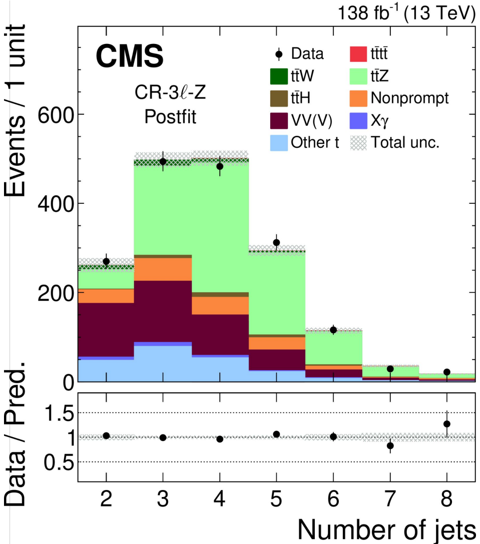
ttW modelling: NLO QCD MC

- Additional large uncertainty on tt̄W + jets
- Free-floating normalization in fit
  - Postfit normalization: 990 ± 98 fb
  - Compatible with inclusive CMS measurement (868 ± 65 fb) (JHEP 07 (2023) 219)
  - Constrained by 2 control regions and multiclass BDT

Two on-Z control regions (3 and 4 lepton channels)

- $|m_{ll} - m_Z| < 15$ GeV
- Allows for free-floating tt̄Z normalization in fit
  - Postfit normalization: 945 ± 81 fb
  - Compatible (and competitive) with 2016+2017 CMS measurement (JHEP 03 (2020) 056)
- Control over WZ & ZZ with additional (b)-jets





CMS: signal extraction

3 signal regions defined (**SSDL**, **3L** and **4L**)

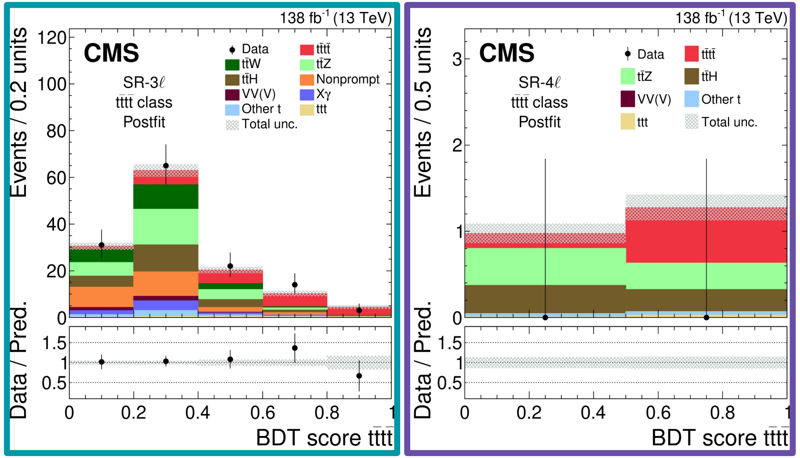
- Varying cuts on number of (b-)jets, HT

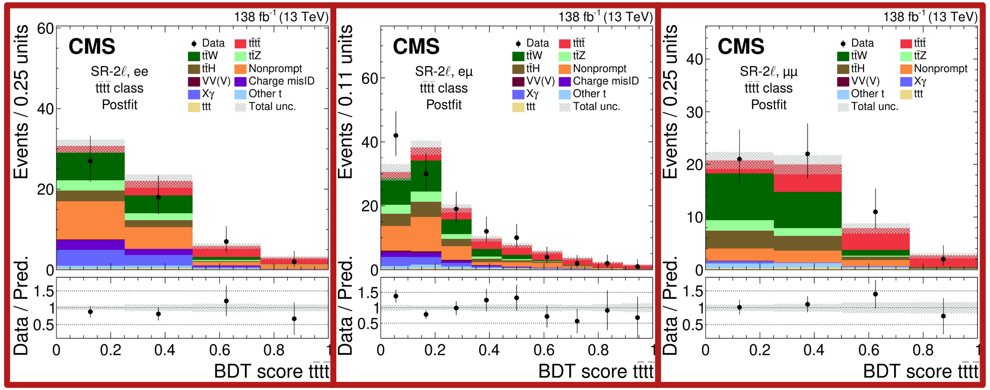
Two multiclass BDTs

- One for SSDL signal region, one for 3L+4L signal regions
- Trained on 3 classes:
  - $t\bar{t}t\bar{t}$
  - $t\bar{t}X$: $t\bar{t}W$, $t\bar{t}Z$ and $t\bar{t}H$
  - $t\bar{t}$: nonprompt and charge misID

Fit optimization:

- Signal regions split in 3 BDT categories
- **SSDL signal region** split in lepton flavors





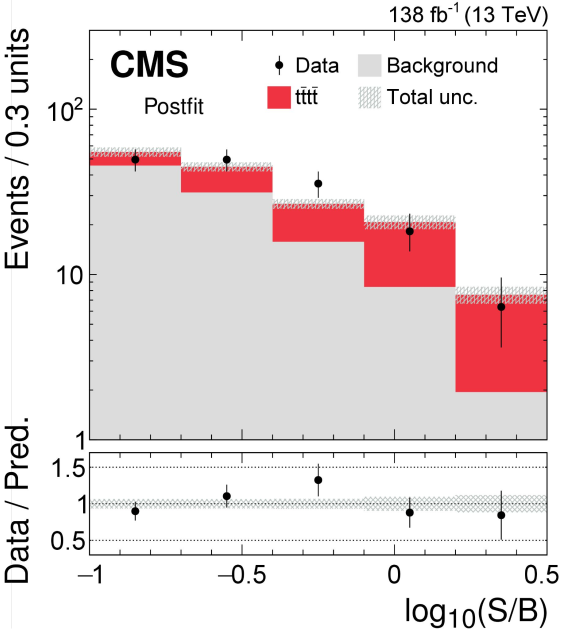
CMS: summary

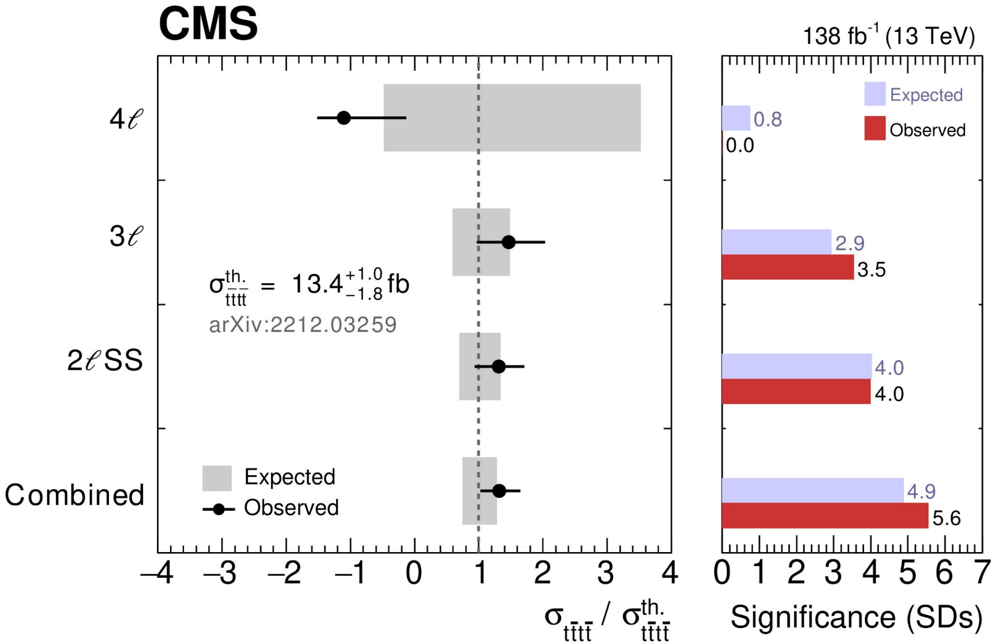
Sensitivity: 5.6σ observed (4.9σ expected)

Measured cross section: $17.7^{+3.7}_{-3.5}(\text{stat})^{+2.3}_{-1.9}(\text{syst})$ fb

SM expectation: $\sigma(t\bar{t}t\bar{t})$ = 13.4 fb (arXiv:2212.03259)

Interpretation with triple top production: see next talk by A. Sharma





Summary

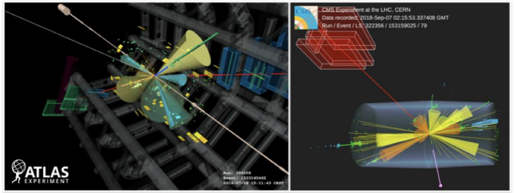
Many new four top results:

- Evidence at CMS using the all-hadronic, 1L and OSDL channels
  - PLB 844 (2023) 138076
- Observation at both CMS and ATLAS using SSDL, 3L and 4L channels
  - ATLAS: EPJC 83 (2023) 496
  - CMS: arXiv:2305.13439 (submitted to PLB)

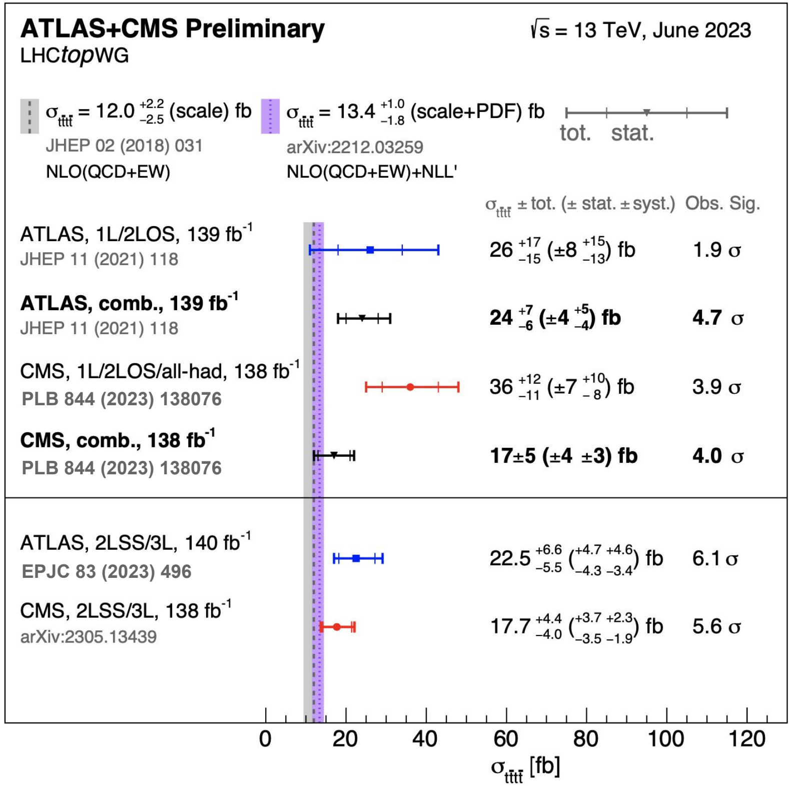
ATLAS and CMS observe simultaneous production of four top quarks

The ATLAS and CMS collaborations have both observed the simultaneous production of four top quarks, a rare phenomenon that could hold the key to physics beyond the Standard Model

24 MARCH, 2023 | By Naomi Dinmore



Event displays of four-top-quark production from ATLAS (left) and CMS (right).



BACKUP

Next experimental steps

- Final states with hadronic taus

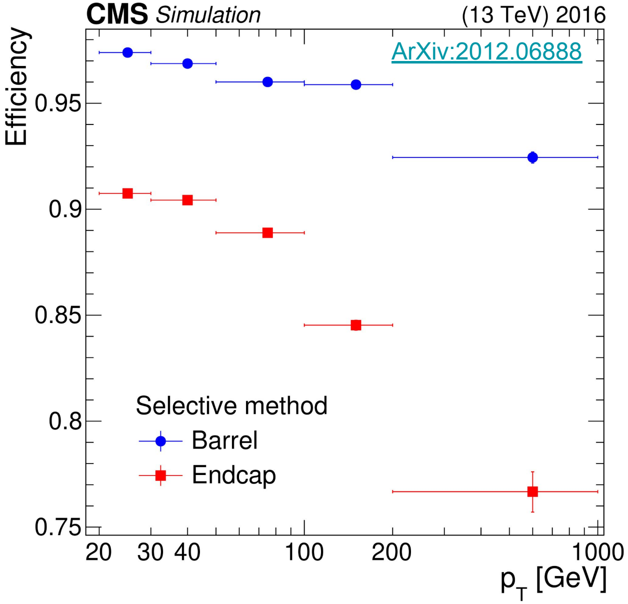
Background prediction:

- ttW modelling
- NLO MC for ttVV and triple top backgrounds

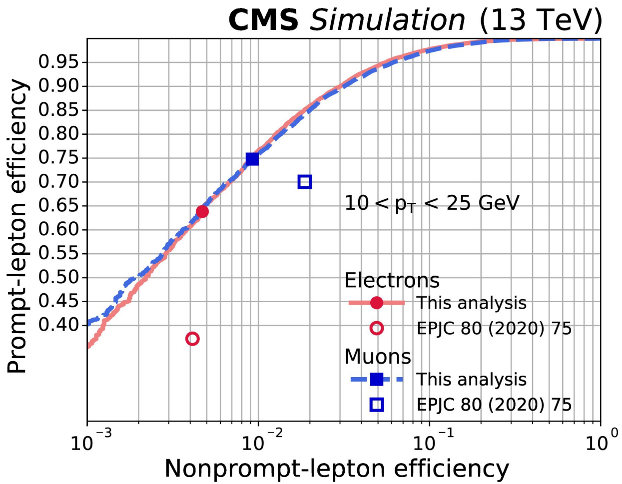
Signal/background selection:

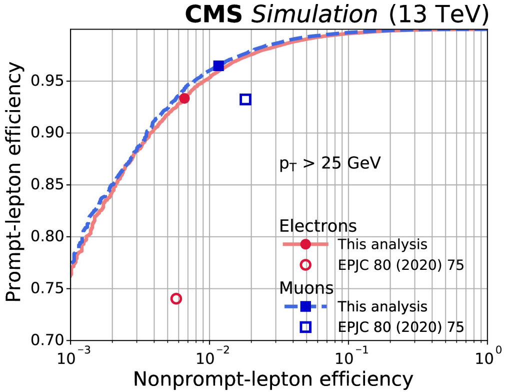
- Constantly improving b-tagging
- Advanced ML techniques (e.g. GNN used by ATLAS)

SSDL & ML: lepton selection



Efficiency of the charge misID reduction method

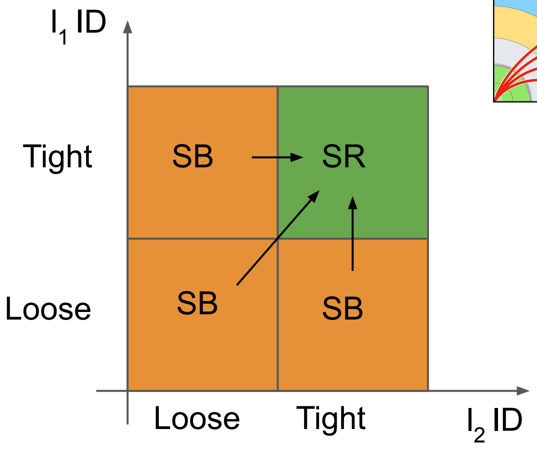


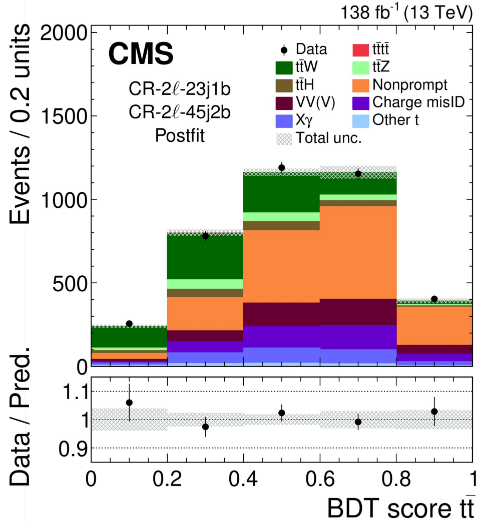


Per-lepton BDT at low (above) and high (below) pT for prompt lepton identification

Background prediction: nonprompt

- **Tight-to-loose ratio method (data driven method)**
  - Define data sideband based on one or more leptons failing tight ID
  - Fake rates (FR) measured in QCD multijet events in data
  - Measured as function of pT and eta
- Validation of FR (from QCD MC) in ttbar and DY MC
- Uncertainties:
  - Shapes (statistical variation of FR)
  - Flat (20% uncorr. ⊕ 20% corr.)
  - Individual nuisances per lepton flavor



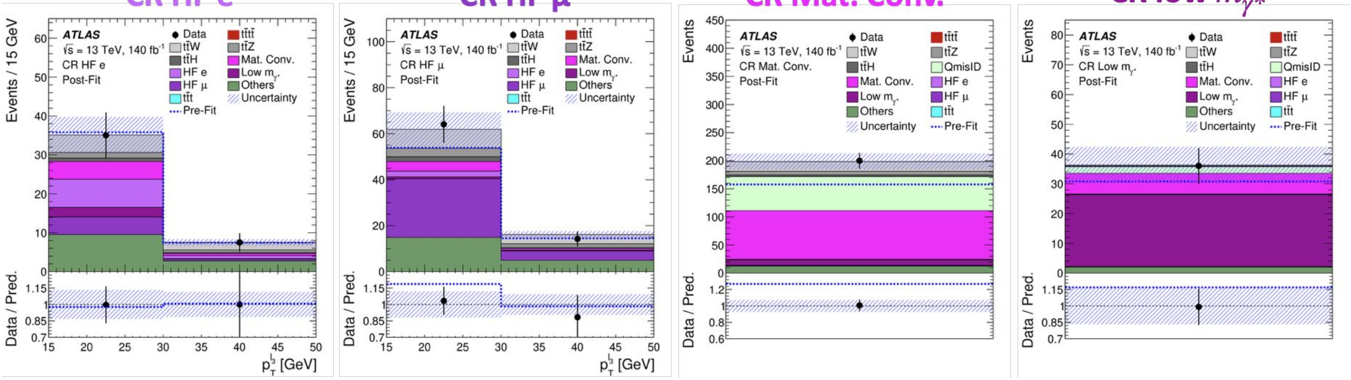


Non-prompt lepton modeling

ATLAS fake with **Template fit method: (semi-data method)**

- Rely on Monte Carlo simulation for shapes for different components
- 4 normalization factors are allowed to float in the fit:
  - **fake electrons from semi-leptonic b-decay (HF e)**
  - **fake muons from semi-leptonic b-decay (HF μ)**
  - **material conversion (Mat. Conv.)**
  - **virtual photon conversion (low m_{γ^*})**
- Systematics evaluated in the isolation loosed region

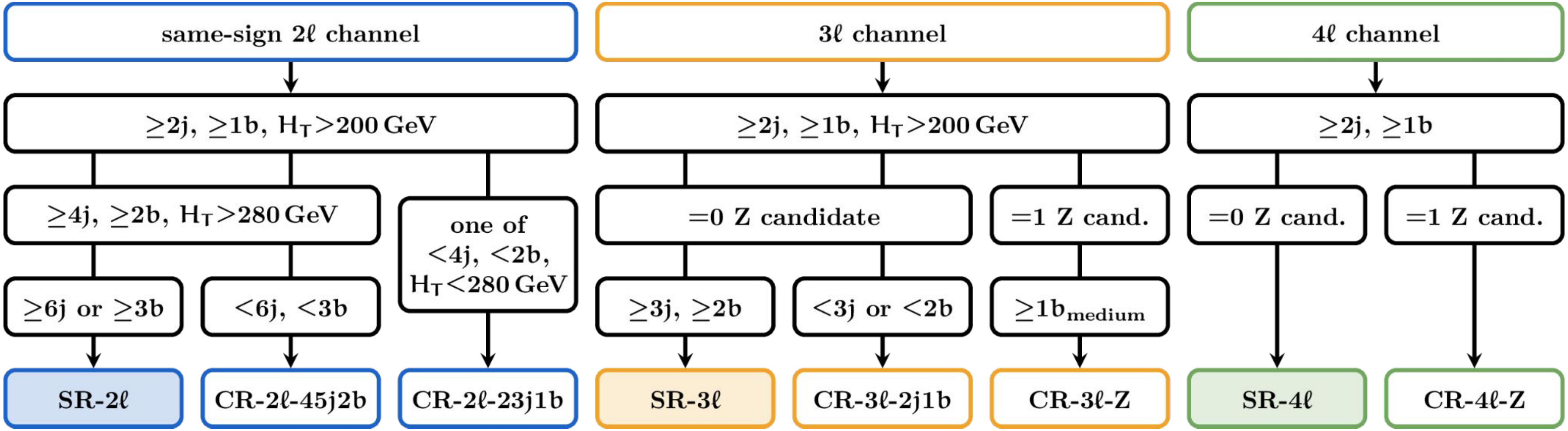
CR HF e | CR HF μ | CR Mat. Conv. | CR low m_{γ^*}



ATLAS: event selection

Region	Channel	N_j	N_b	Other selection	Fitted variable
CR Low m_{γ^*}	SS, ee or eμ	$4 \leq N_j < 6$	≥ 1	ℓ_1 or ℓ_2 is from virtual photon (γ^*) decay ℓ_1 and ℓ_2 are not from photon conversion	counting
CR Mat. Conv.	SS, ee or eμ	$4 \leq N_j < 6$	≥ 1	ℓ_1 or ℓ_2 is from photon conversion	counting
CR HF μ	e$\mu\mu$ or $\mu\mu\mu$	≥ 1	$= 1$	$100 < H_T < 300$ GeV $E_T^{miss} > 50$ GeV total charge $= \pm 1$	$p_T^{\ell_3}$
CR HF e	eee or eeμ	≥ 1	$= 1$	$100 < H_T < 275$ GeV $E_T^{miss} > 35$ GeV total charge $= \pm 1$	$p_T^{\ell_3}$
CR $t\bar{t}W^+$+jets	SS, eμ or $\mu\mu$	≥ 4	≥ 2	$\lvert\eta(e)\rvert < 1.5$ when $N_b = 2$: $H_T < 500$ GeV or $N_j < 6$ when $N_b \geq 3$: $H_T < 500$ GeV total charge > 0	N_j
CR $t\bar{t}W^-$+jets	SS, eμ or $\mu\mu$	≥ 4	≥ 2	$\lvert\eta(e)\rvert < 1.5$ when $N_b = 2$: $H_T < 500$ GeV or $N_j < 6$ when $N_b \geq 3$: $H_T < 500$ GeV total charge < 0	N_j
CR 1b(+)	2LSS+3L	≥ 4	$= 1$	ℓ_1 and ℓ_2 are not from photon conversion $H_T > 500$ GeV total charge > 0	N_j
CR 1b(-)	2LSS+3L	≥ 4	$= 1$	ℓ_1 and ℓ_2 are not from photon conversion $H_T > 500$ GeV total charge < 0	N_j
SR	2LSS+3L	≥ 6	≥ 2	$H_T > 500$ GeV	GNN score

CMS: event selection



Results: overview

ATLAS EXPERIMENT

Central result

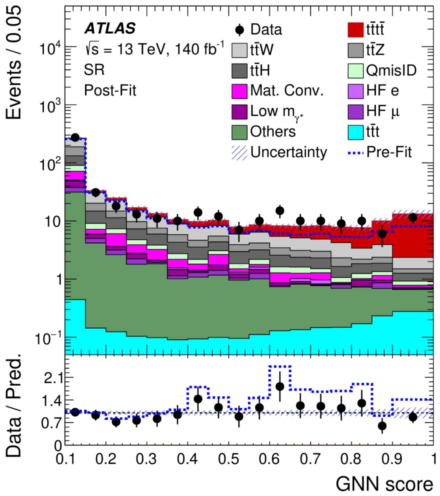
$$\sigma_{t\bar{t}t\bar{t}} = 22.5^{+4.7}_{-4.3}(\text{stat})^{+4.7}_{3.4}(\text{syst})\ \text{fb} = 22.5^{+6.6}_{-5.6}\ \text{fb}$$

Relative uncertainty

Total uncertainty: +29%, -25%
Systematic: +21%, -15%
Statistical: +21%, -19%

Obs. (exp.) significance (@ 13.4 fb - NLO+NLL')

6.1 (4.7)σ



CMS

Central result

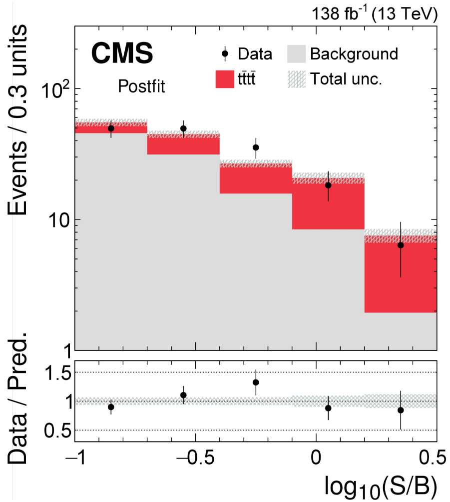
$$\sigma(t\bar{t}t\bar{t}) = 17.7^{+3.7}_{-3.5}(\text{stat})^{+2.3}_{-1.9}(\text{syst})\,\text{fb} = 17.7^{+4.4}_{-4.0}\,\text{fb}$$

Relative uncertainty

Total uncertainty: +25%, -23%
Systematic: +13%, -11%
Statistical: +21%, -20%

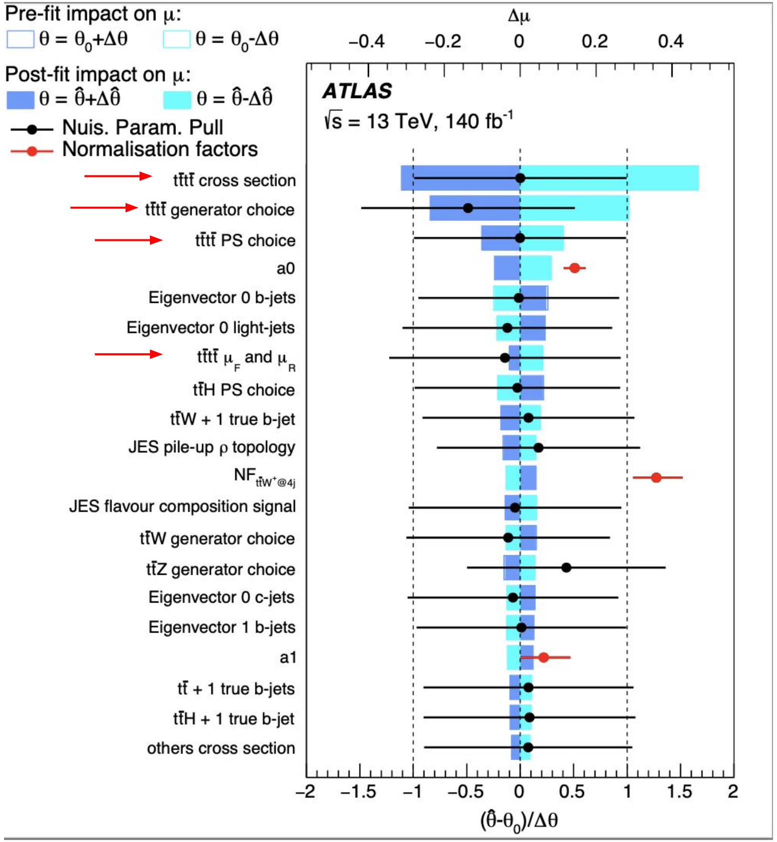
Obs. (exp.) significance (@ 13.4 fb - NLO+NLL')

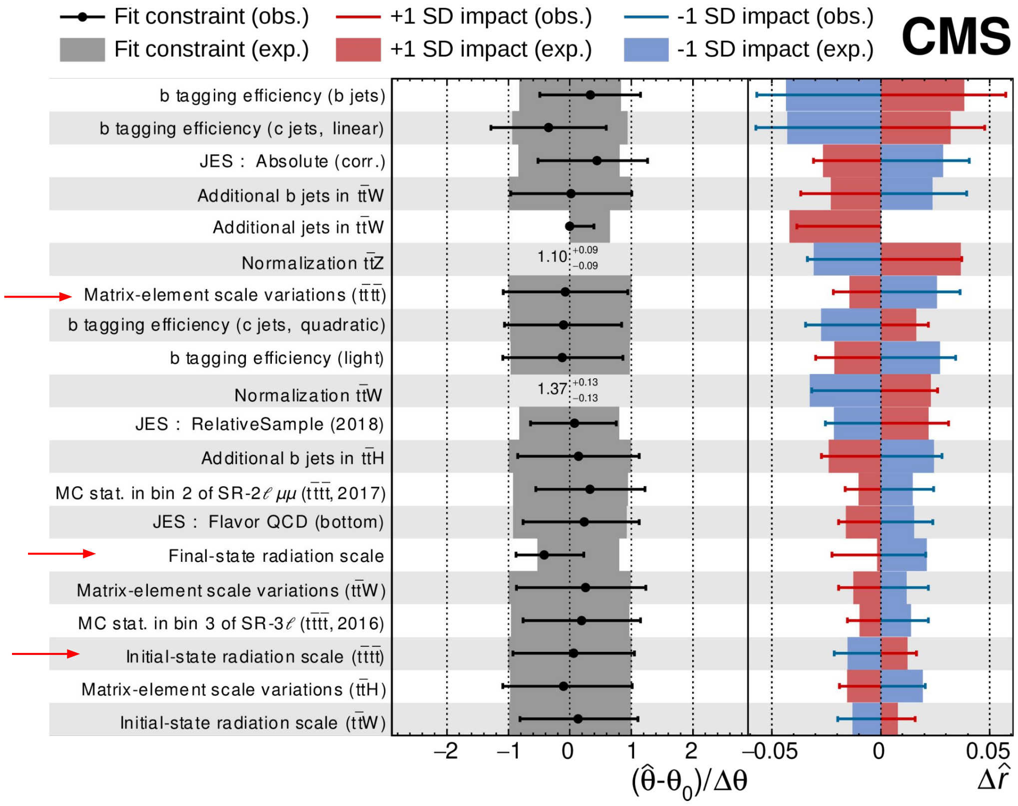
5.6 (4.9)σ



Ranking and pulls on uncertainties

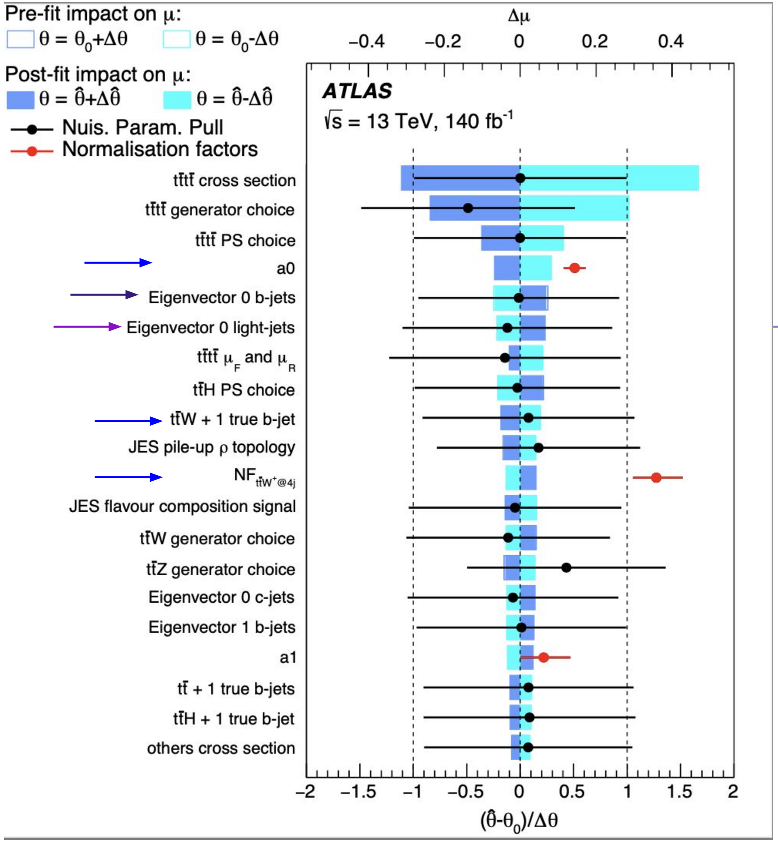
In ATLAS, signal modeling has large impact in the cross section measurement compared with CMS

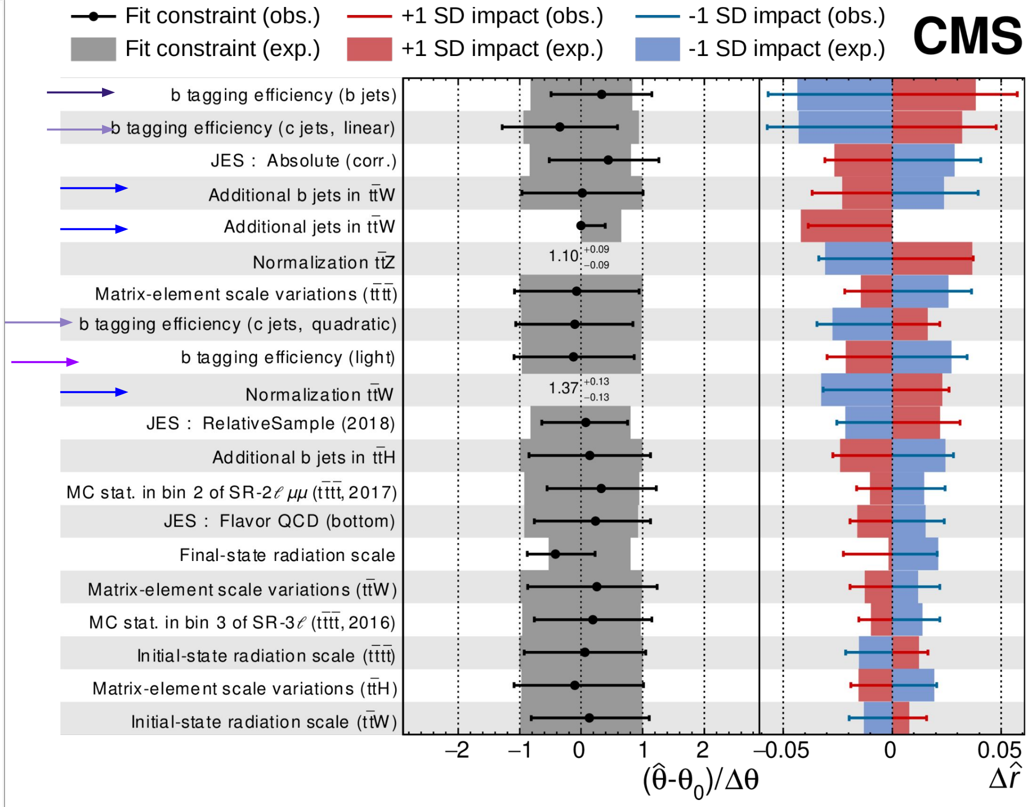




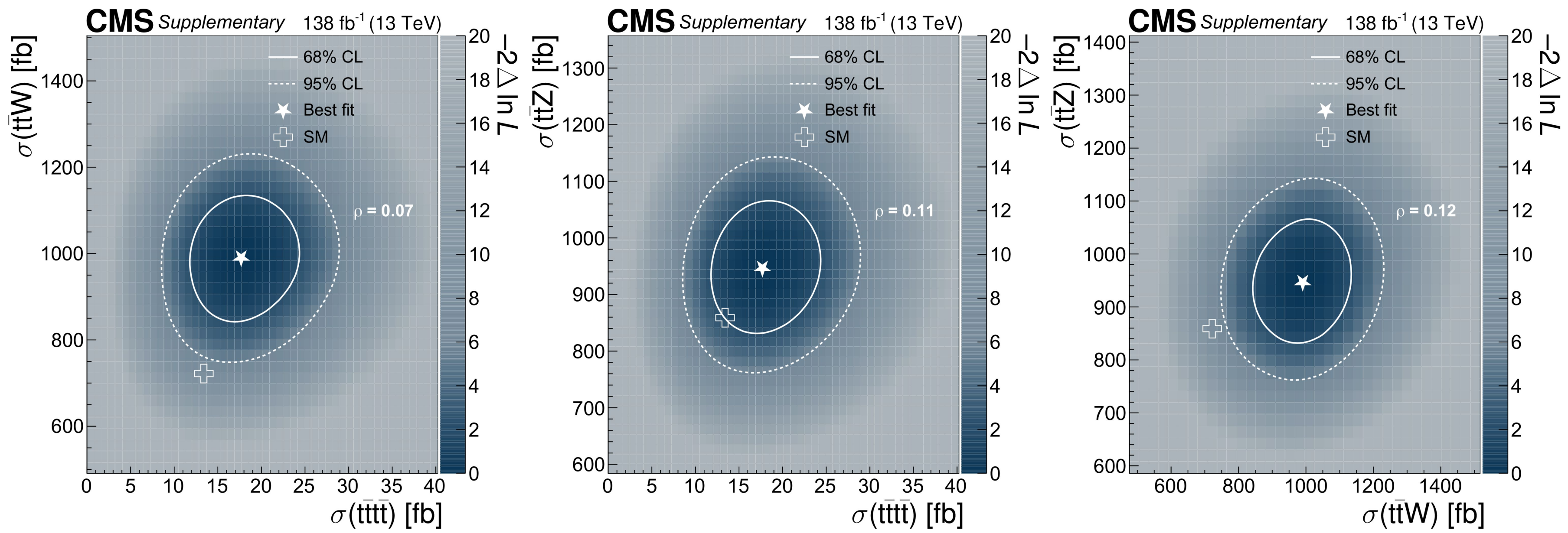
Ranking and pulls on uncertainties

Leading impact comes from b-jets and ttW modelling in both CMS and ATLAS





CMS: 2D scans



CMS Supplementary 138 fb-1 (13 TeV)
68% CL
95% CL
Best fit
SM
ρ = 0.07
ρ = 0.11
ρ = 0.12
σ(tt̄W) [fb]
σ(tt̄Z) [fb]
σ(tt̄tt̄) [fb]
−2 Δ ln L