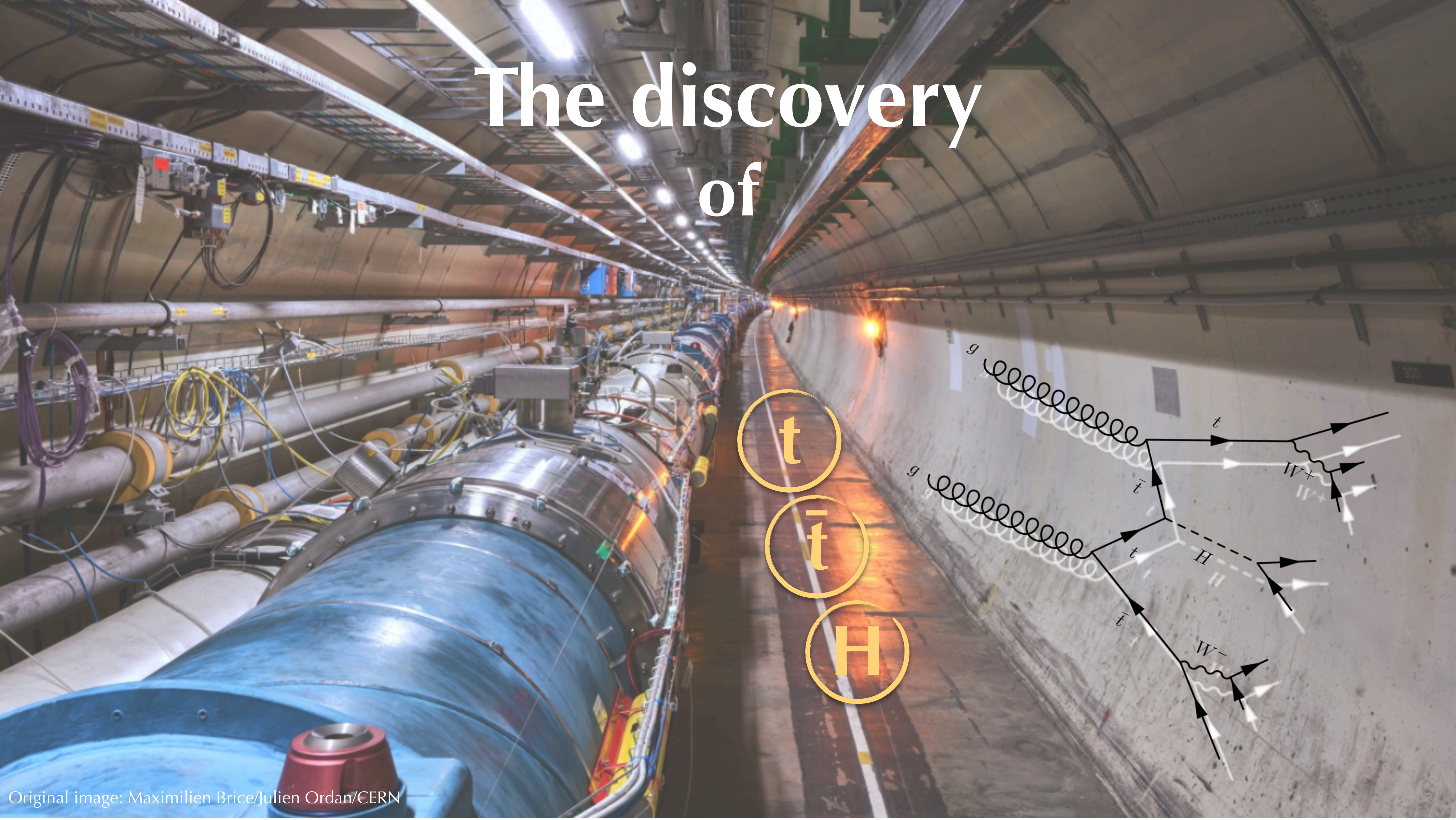
The discovery of $t\bar{t}H$



Original image: Maximilien Brice/Julien Ordan/CERN

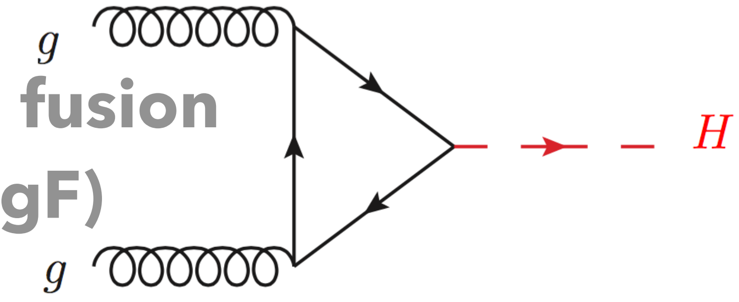
Tamara Vázquez Schröder (CERN)

Higgs@10 symposium, Birmingham/Virtual
01 July 2022

The Higgs boson: production

- At the LHC, the Higgs boson is dominantly produced via gluon fusion

gluon fusion (ggF)



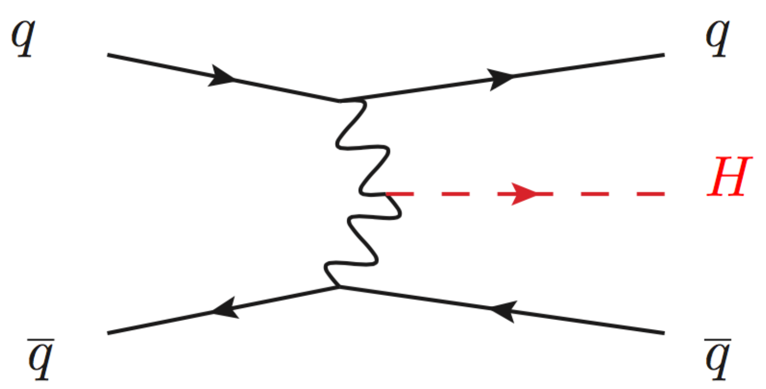
$\sigma_{H,ggF}$ ~ 49 pb at 13 TeV

6.9M events in Run-2

vector boson fusion (VBF)

σ_{VBF} ~ 3.8 pb

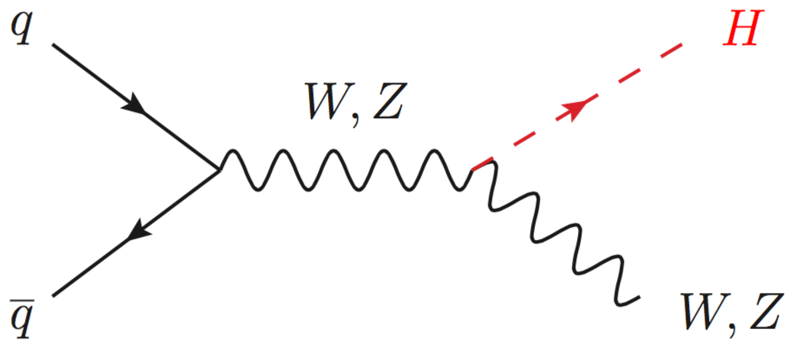
0.5M events in Run-2



W, Z associated production (VH)

$\sigma_{W/ZH}$ ~ 1.4-0.9 pb

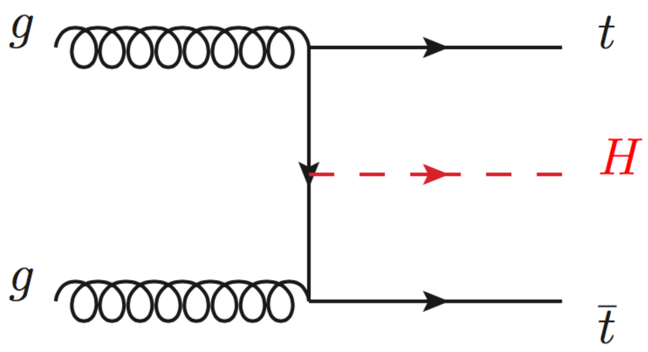
200-130k events in Run-2



top associated production ($t\bar{t}H$)

σ_{ttH} ~ 0.5 pb

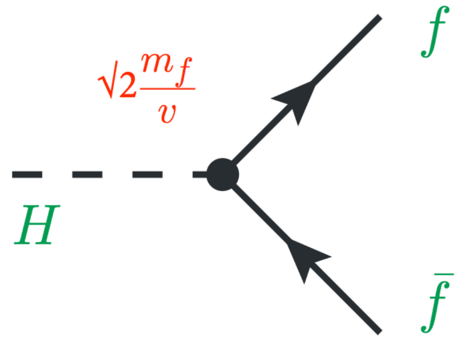
70k events in Run-2



y_{top} … why should we care?

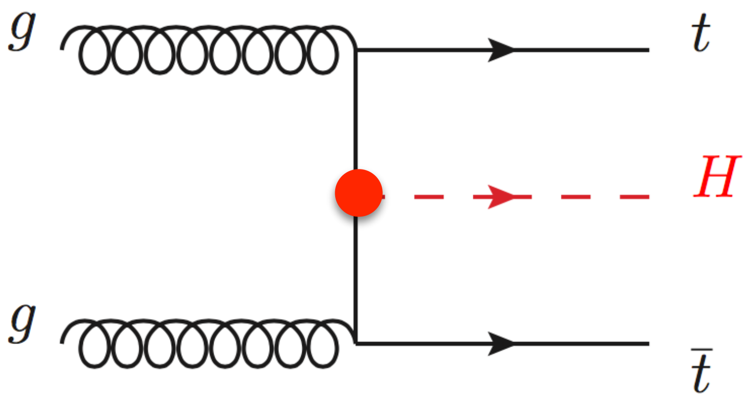
Top quark is the heaviest fermion in the SM ➜

Largest Yukawa coupling

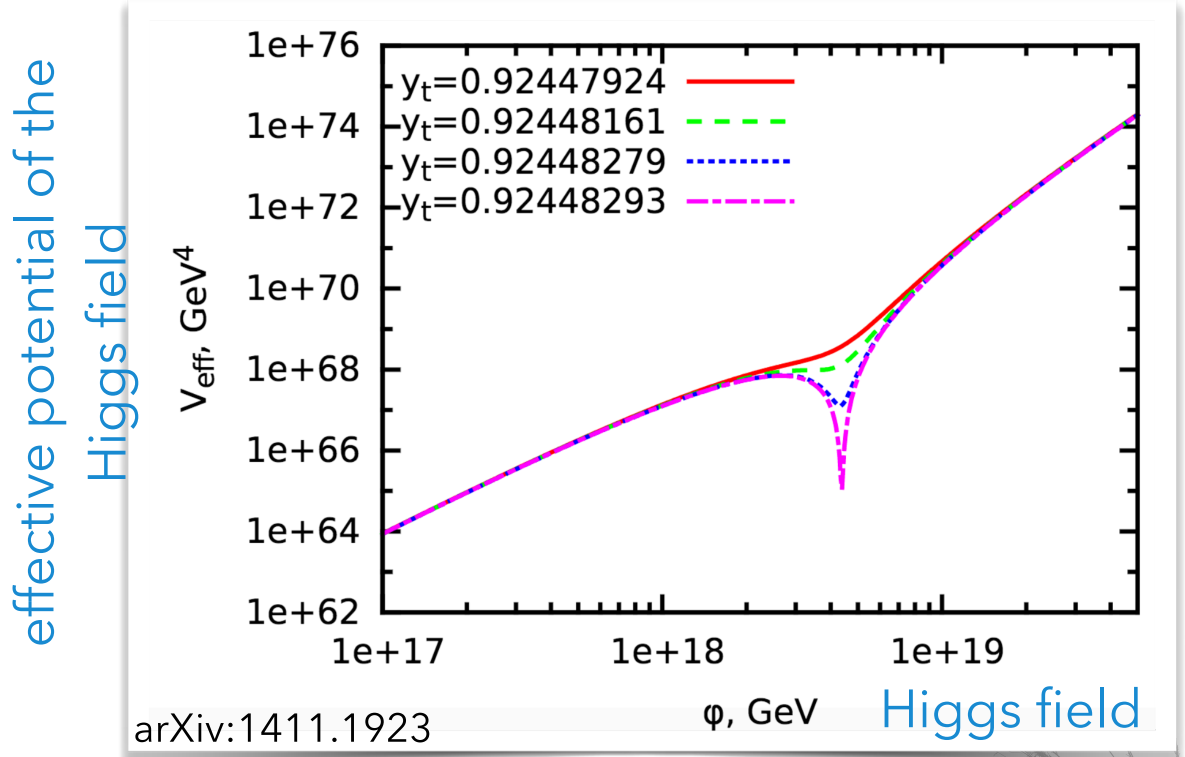


- The only fermion with predicted Yukawa coupling ~ 1
- Does this point to a special role in electroweak symmetry breaking or beyond the SM physics?
- Top quark Yukawa coupling is relevant for the stability of the Higgs potential and the required energy scale for new physics

direct top Yukawa coupling measurement only possible at the LHC via tt̄H and tH

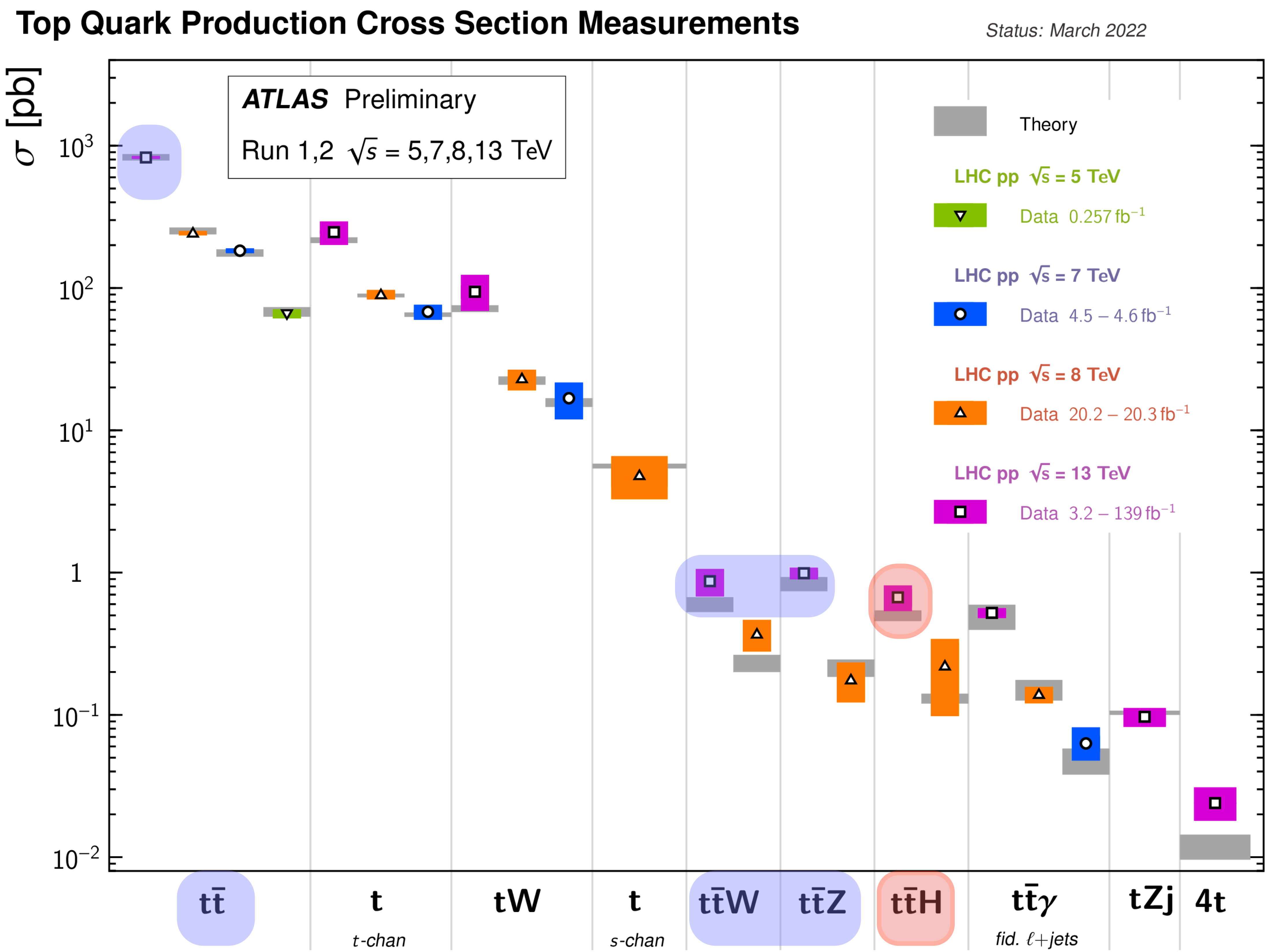


Is the Universe stable or only metastable?



arXiv:1411.1923

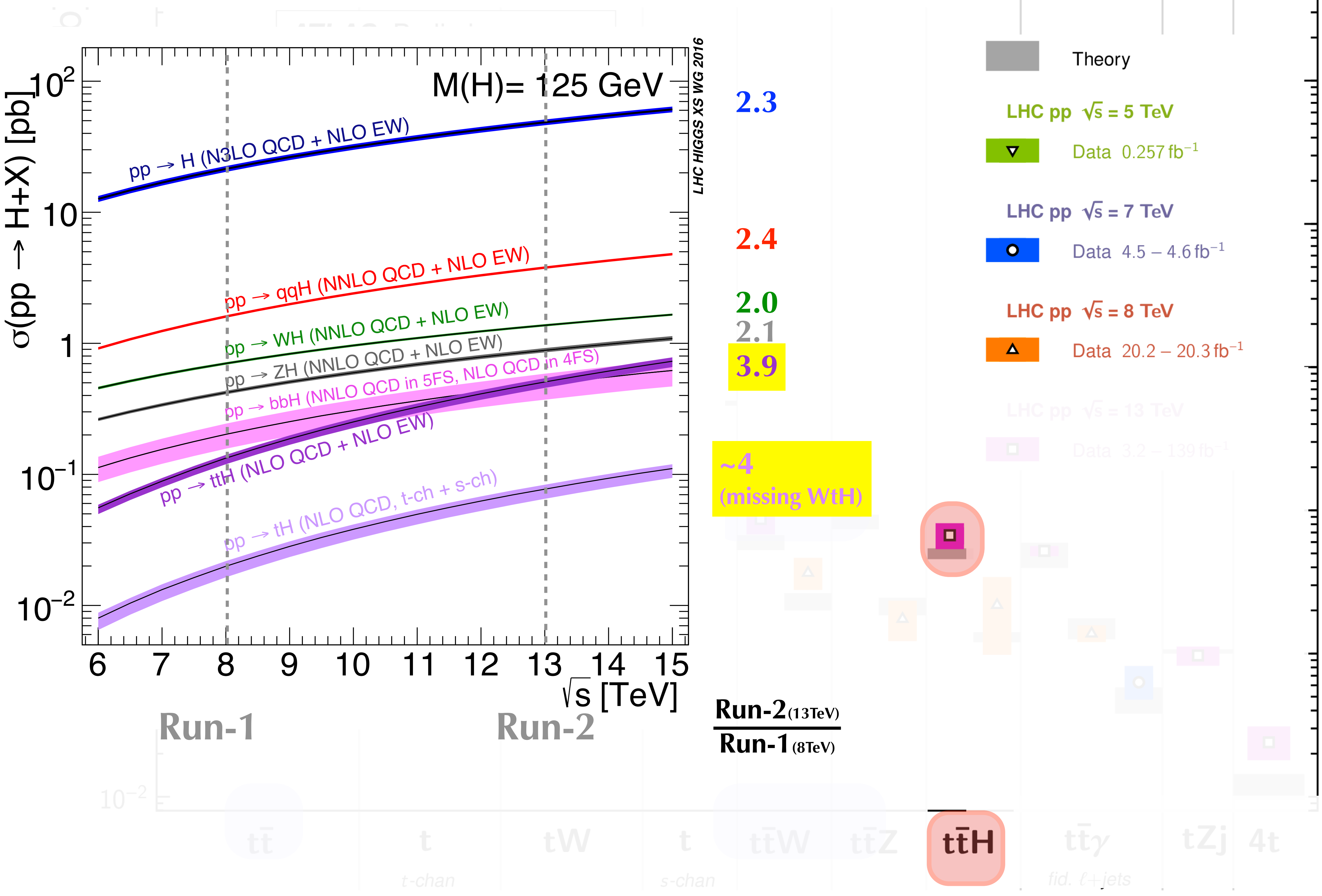
tt̄H: one of the tiniest rates!



ATL-PHYS-PUB-2021-014

$t\bar{t}H$: … but fastest growth!

- Large cross section increase from 8 to 13 TeV for $t\bar{t}H$ and tH



Where to look for $t\bar{t}H$ production?

- $t\bar{t}H$ production (~500 fb @ 13TeV) is:
  - **two orders** of magnitude smaller than ggF Higgs production
  - **three orders** of magnitude smaller than $t\bar{t}$ production

- Look for $t\bar{t}H$ in final states with distinctive signatures and features
  - Combination of top quark x Higgs boson decay modes

Top Pair Branching Fractions

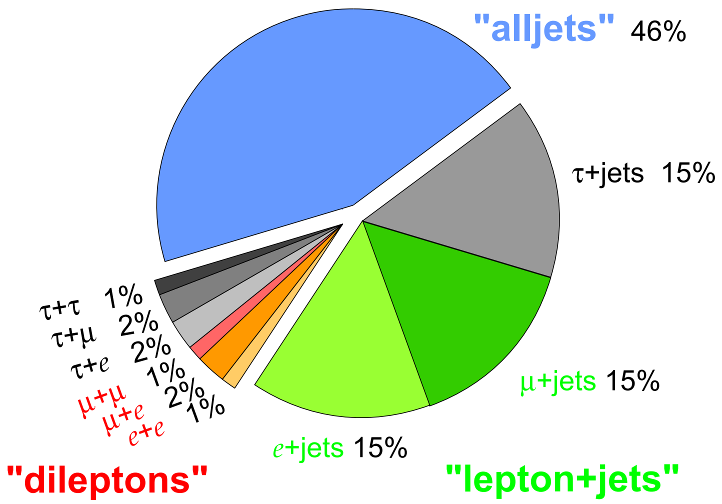
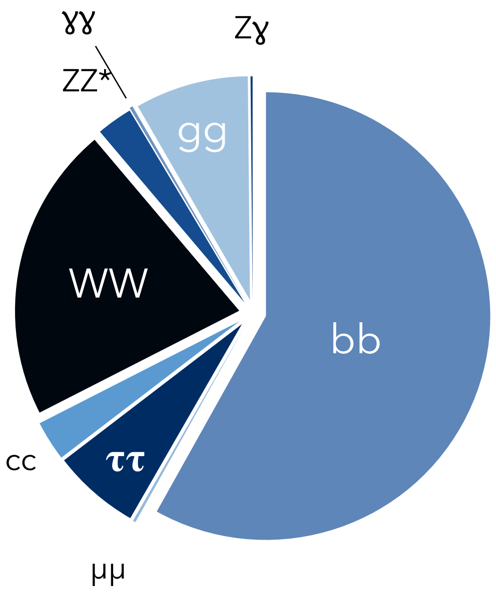


Diagram from D0

x

Higgs Branching Fractions



= …

$t\bar{t}H$ analysis channels

ttH
(H→bb)

ttH
(H→WW, ττ, ZZ)
'multilepton'

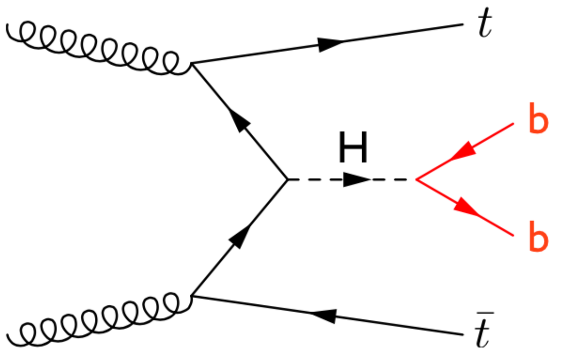
ttH
(H→γγ, ZZ(→4ℓ))

Low signal/background (need MVA) → **Clear peak** (clean bump hunt)

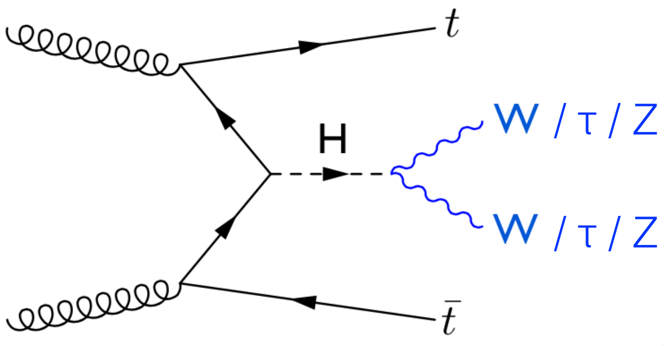
Large branching ratio (yields) ← **Small branching ratio**

Complex background modelling → **Simpler background**

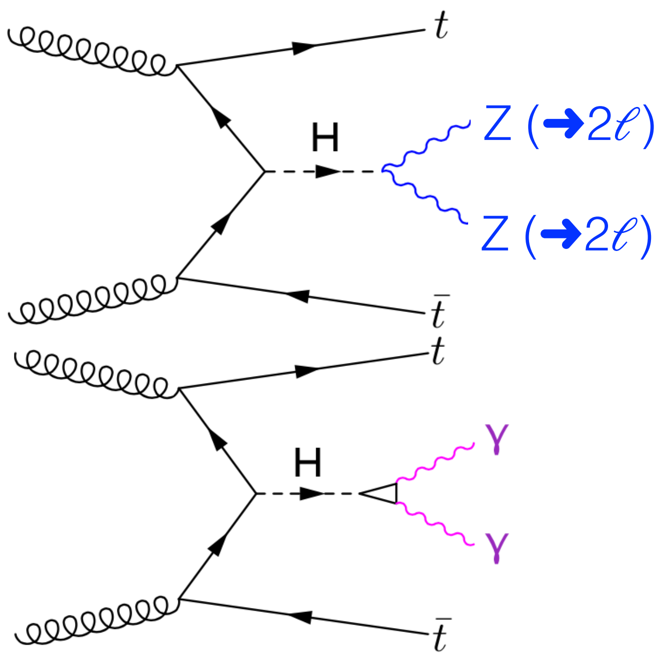
- large irreducible $t\bar{t}$+jets (HF) background
- final states with **multiple b-jets**



- leptonic decays of W / Z bosons and tau decays can give distinct **multilepton** signatures
- main background from $t\bar{t}$Z/W and non-prompt leptons



- **resonant** channels



motivation ← **challenge**

Towards $t\bar{t}H$ observation…

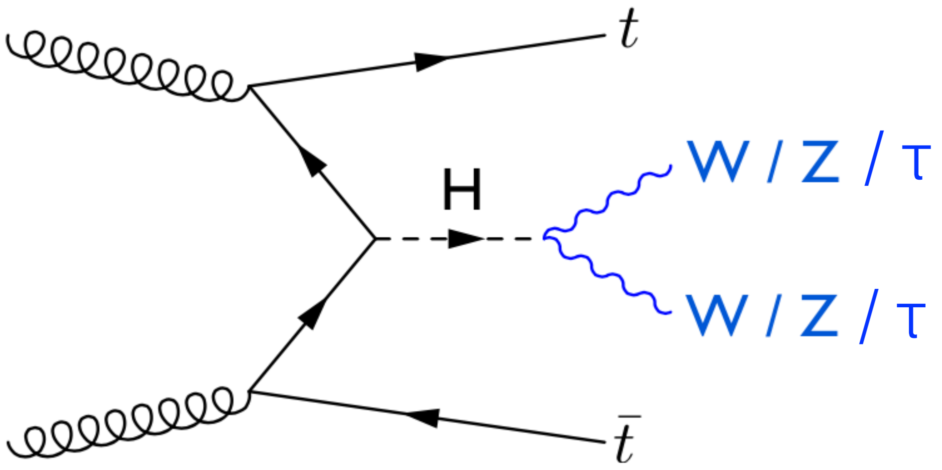
- In early 2018, the landscape was…

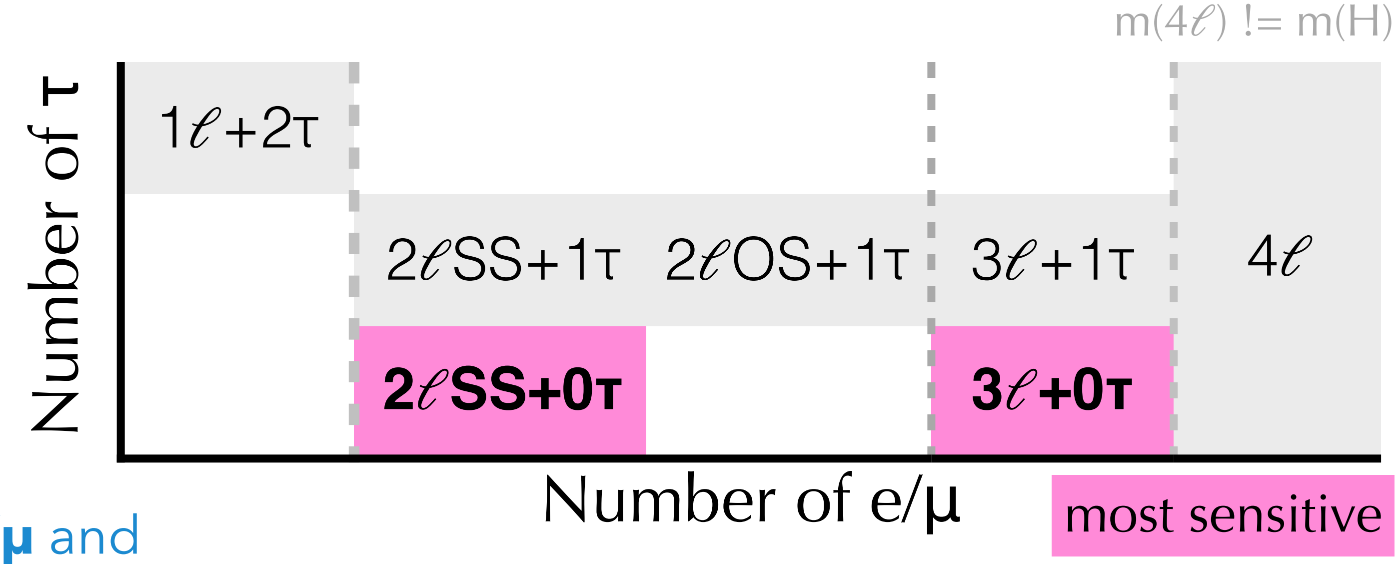
🟩 **2015+2016 data [~36 fb^{-1}]**

	ATLAS EXPERIMENT	CMS
ttH multilepton (H→WW/ττ/ZZ)	Phys. Rev. D 97 (2018) 072003 (including combination 36.1/fb → **evidence**)	CMS-HIG-17-018 $\mu_{ttH} = 1.23^{+0.45}_{-0.43}$
ttH(bb)	Phys. Rev. D 97 (2018) 072016 (leptonic)	CMS-HIG-17-026 (leptonic) $\mu_{ttH} = 0.72 \pm 0.45$ CMS-HIG-17-022 (all-hadronic) $\mu_{ttH} = 0.9 \pm 1.5$
ttH(ZZ→4ℓ)	arXiv:1712.02304 submitted to JHEP $\mu_{ttH} < 7.1$	arXiv:1706.09936 $\mu_{ttH} < 1.18$
ttH(γγ)	ATLAS-CONF-2017-045 1.0σ (exp: 1.8σ) $\mu_{ttH} = 0.5 \pm 0.6$	CMS-PAS-HIG-16-040 3.3σ (exp: 1.5σ) $\mu_{ttH} = 2.2^{+0.9}_{-0.8}$
ATLAS+CMS Run1 combination	JHEP 1608 (2016) 045 4.4σ (exp: 2.0σ) $\mu_{ttH} = 2.3^{+0.7}_{-0.6}$	

$t\bar{t}H$ (multileptons): analysis strategy

- **Target**: $t\bar{t}H$ with
  - $H \to WW/ZZ/\tau\tau \to \geq 1\ell$
  - $t\bar{t} \to (\ell$+jets, dilepton)
- **High multiplicity** final state
- **Rare in SM:** same-sign 2ℓ, 3ℓ, 4ℓ
  - Exploit presence of hadronically decaying τ
- **Analysis strategy**:
  - Split in categories based on **number of e/μ** and **number of τ**
  - Fit or cut on **BDTs (boosted decision tree)** to discriminate signal against the main background processes [except in $3\ell+1\tau$]
  - **2ℓSS0τ**: two BDTs combined ($t\bar{t}H$ vs. $t\bar{t}$; $t\bar{t}H$ vs. $t\bar{t}V$)
  - **3ℓ0τ**: 5D-classification BDT ($t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}$, VV)




m(4ℓ) != m(H)

Number of τ

$1\ell+2\tau$

2ℓSS+1τ

2ℓOS+1τ

$3\ell+1\tau$

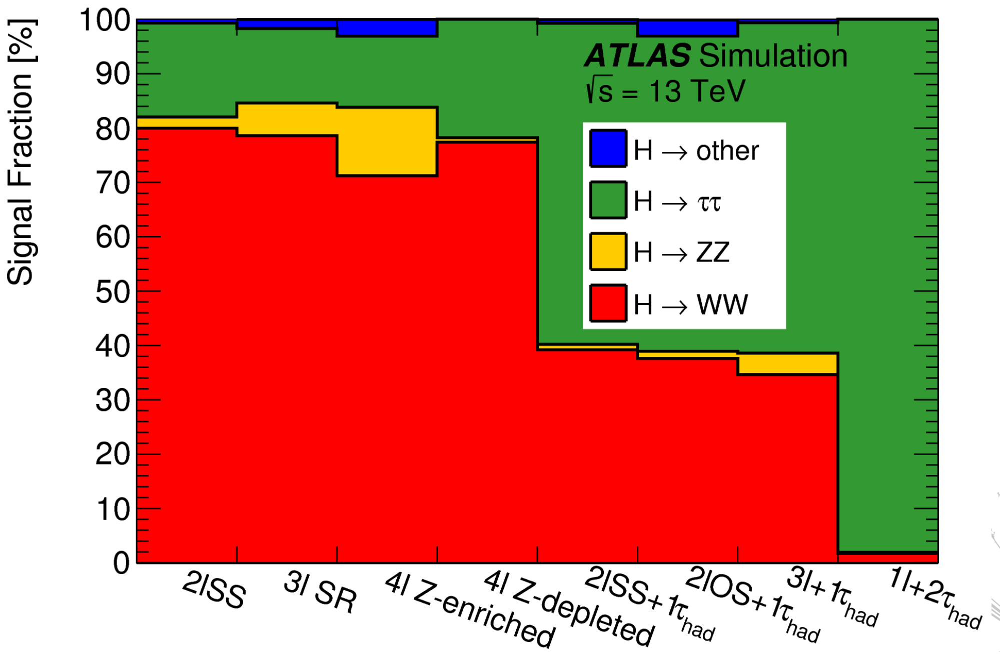
4ℓ

2ℓSS+0τ

3ℓ+0τ

Number of e/μ

most sensitive



$t\bar{t}H$ (multileptons): **backgrounds**

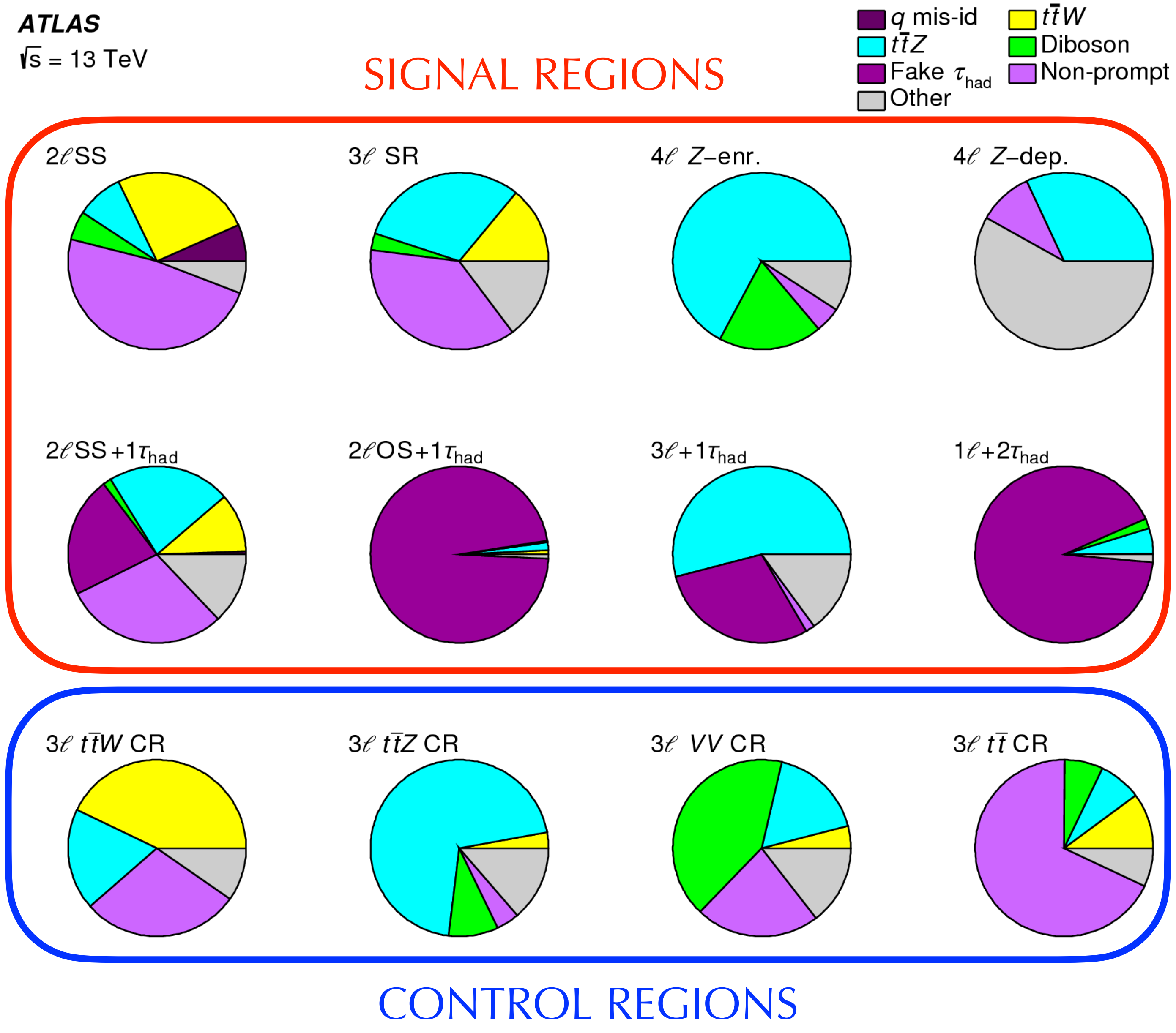
- **Non-prompt** lepton in $t\bar{t}$
  - semileptonic b-decay
  - ɣ conversions
- **Fake τ** from light/b-jets

DATA-DRIVEN (DD):
MATRIX METHOD (MM), FAKE FACTOR (FF)

FF ~ matrix method except prompt background is taken from MC

- **Misidentified charge** lepton
  - e.g. trident electrons (Bremsstrahlung)
  - using **3D likelihood method** [p_T, η, Tight/Loose]

DATA-DRIVEN (DD):
LIKELIHOOD FIT



- **Irreducible backgrounds** with prompt-leptons

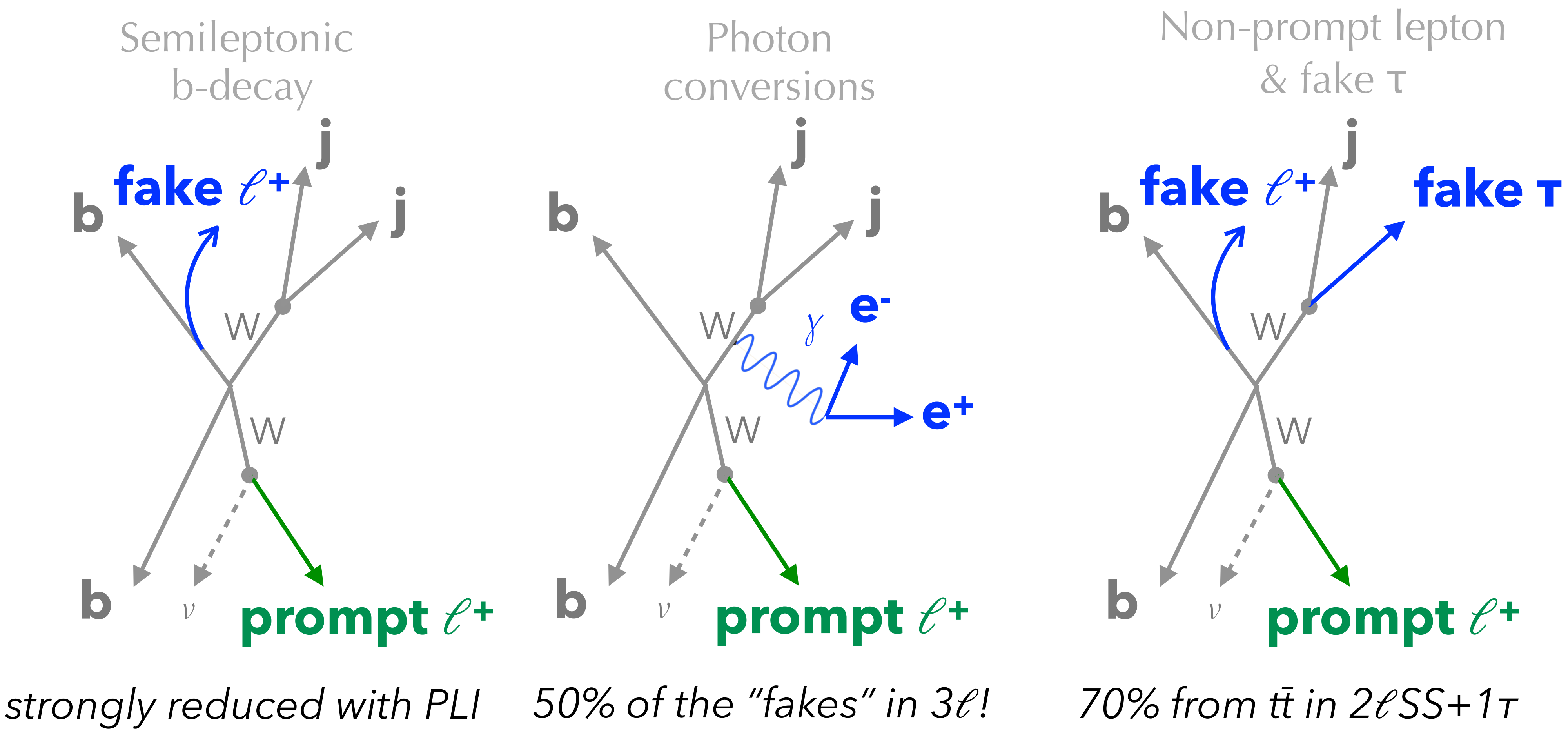
$t\bar{t}Z$ $t\bar{t}W$ VV

MC
(cross check: fit to data)

"Other": 4tops, $t\bar{t}WW$, tH, tZ

$t\bar{t}H$ (multileptons): non-prompt light ℓ

- **Common main/important background**: non-prompt leptons from semileptonic b-decay
- **New MVA lepton isolation (PromptLeptonIso=PLI)** to reject non-prompt ℓ based on:
  - lepton and overlapping **track jets** properties
  - lepton track/calorimeter **isolation** variables
  - Factor $\mathcal{O}(20)$ rejection for leptons originating from b-hadrons
- **New MVA cut to reduce QMIsID** for 2ℓSS and $3\ell+0\tau$
  - Factor $\mathcal{O}(17)$ background rejection for a 95% signal efficiency

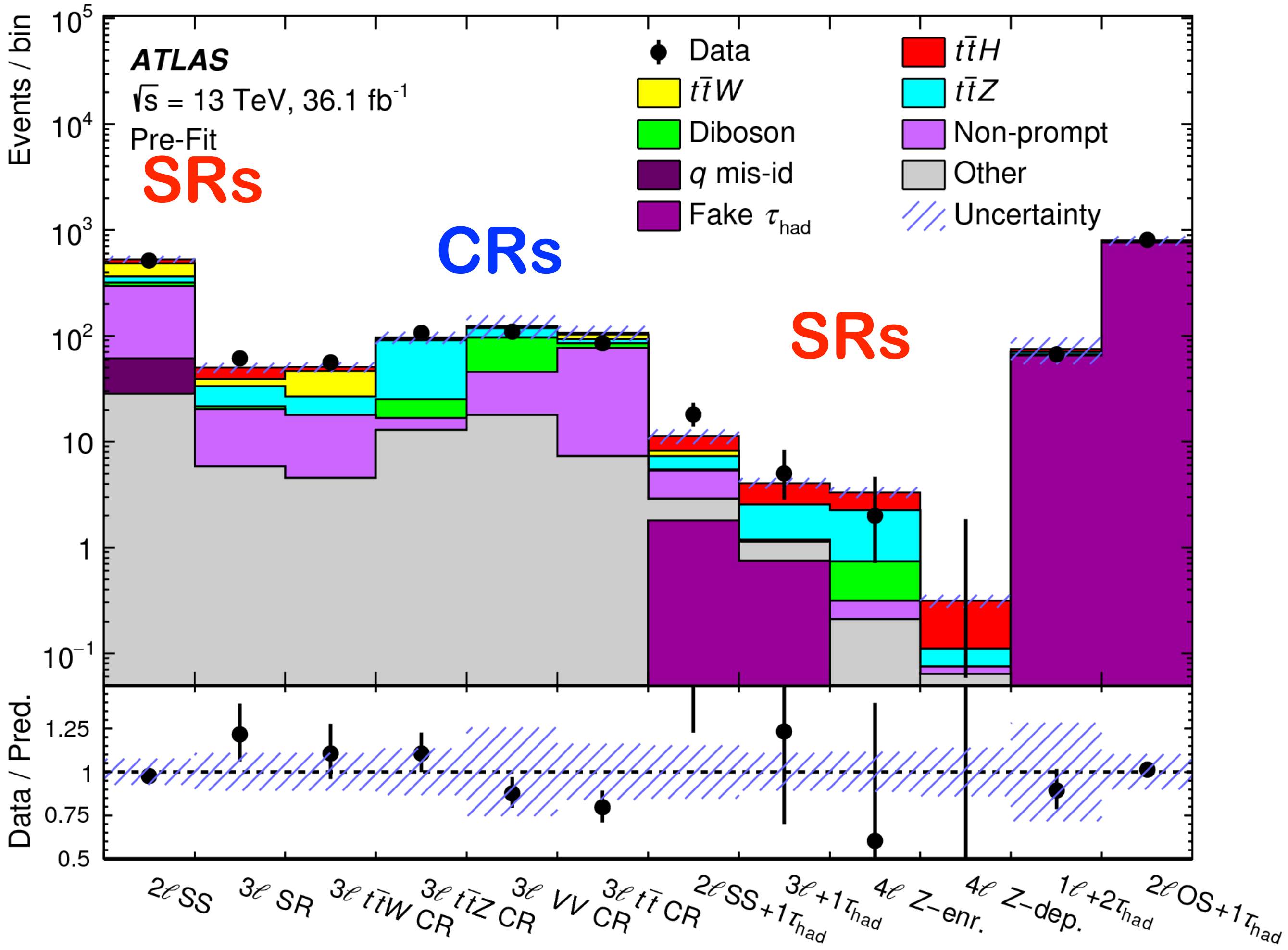


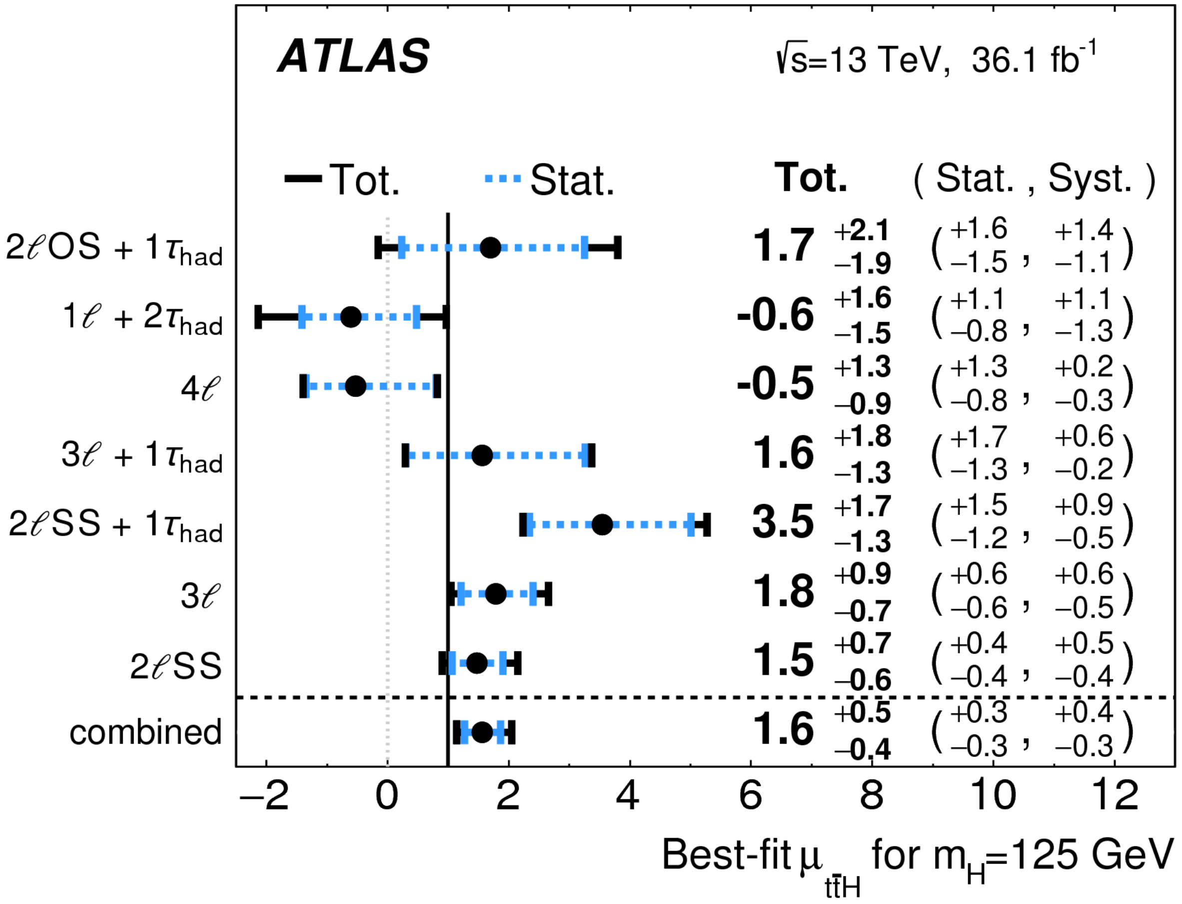
strongly reduced with PLI *50% of the "fakes" in 3ℓ!* *70% from $t\bar{t}$ in 2ℓSS+1τ*

$t\bar{t}H$ (multileptons): fit results

- Main systematic uncertainties: signal modelling, JES and JER, and the non-prompt light ℓ estimates
- **Alternative fit**: $t\bar{t}Z$ and $t\bar{t}W$ normalisation free-floating
  - 15% loss in sensitivity: $\mu(t\bar{t}H) = 1.57^{+0.57}_{-0.50}$
  - $\mu(t\bar{t}Z/W)$ in agreement with SM:
    - $\mu_{t\bar{t}W} = 0.92 \pm 0.32$; $\mu_{t\bar{t}Z} = 1.17^{+0.25}_{-0.22}$

Channel	Significance	
	Observed	Expected
2ℓOS+$1\tau_{\text{had}}$	0.9σ	0.5σ
1ℓ+$2\tau_{\text{had}}$	-	0.6σ
4ℓ	-	0.8σ
3ℓ+$1\tau_{\text{had}}$	1.3σ	0.9σ
2ℓSS+$1\tau_{\text{had}}$	3.4σ	1.1σ
3ℓ	2.4σ	1.5σ
2ℓSS	2.7σ	1.9σ
Combined	4.1σ	2.8σ

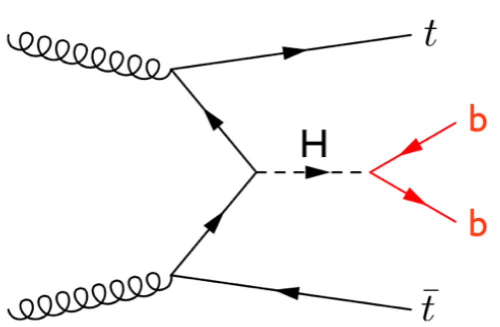




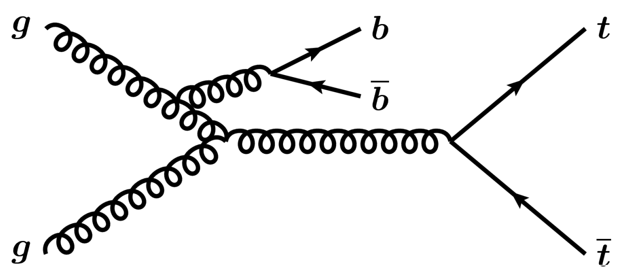
$t\bar{t}H(b\bar{b})$: analysis strategy

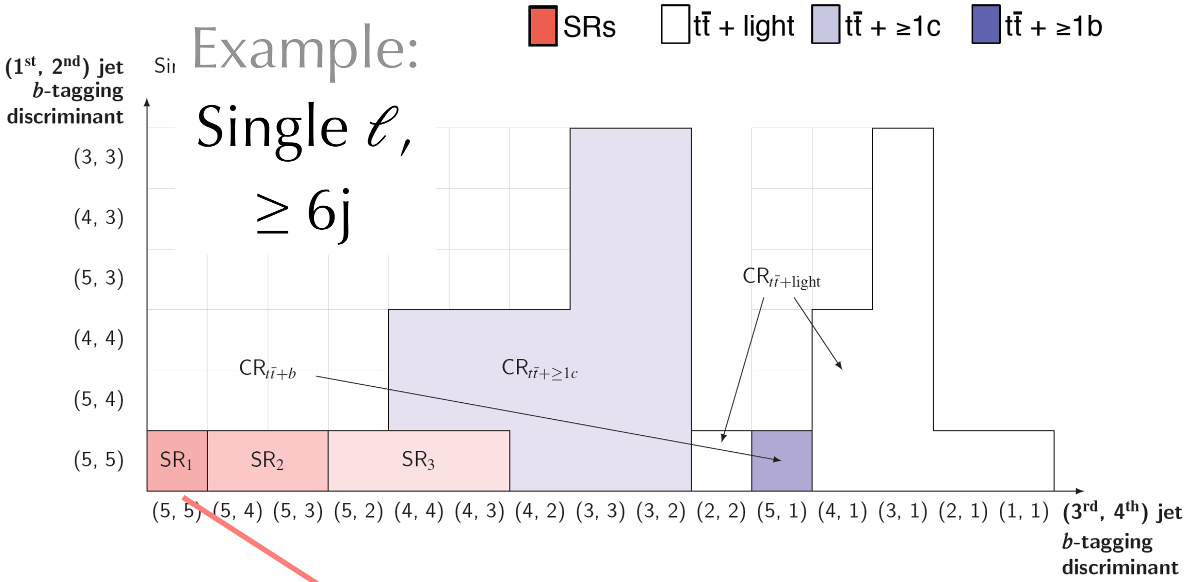
- **Biggest challenge**: good modelling of the $t\bar{t}$+HF (≥1b, ≥1c) background
  - Nominal sample: 5-flavour scheme
  - Relative contribution of $t\bar{t}$+≥1b sub-components reweighted to $t\bar{t}$+$b\bar{b}$ predictions by Sherpa+OpenLoops (4-flavour scheme)
- **Channel categorisation** based on
  - Number of ℓ (1 or 2 opposite-sign)
  - Number of jets
  - Requirements on the b-tagging discriminant (4 calibrated working points)
  - Resolved or boosted, for single lepton channel
- **MVA analysis** needed to discriminate signal from the overwhelming background
  - The '**classification BDT**' includes as input variables: kinematic variables, reconstruction BDTs (resolved), likelihood and matrix element method discriminants (where available), discrete btagging discriminant

signal (ttH)

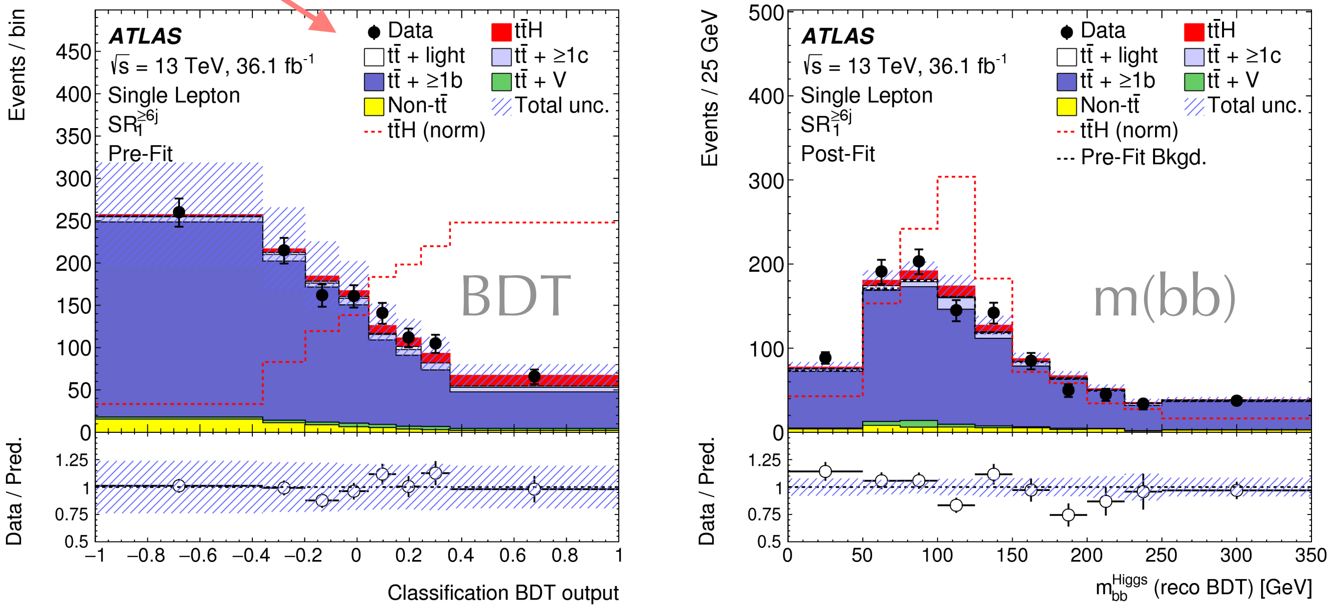


ttbb

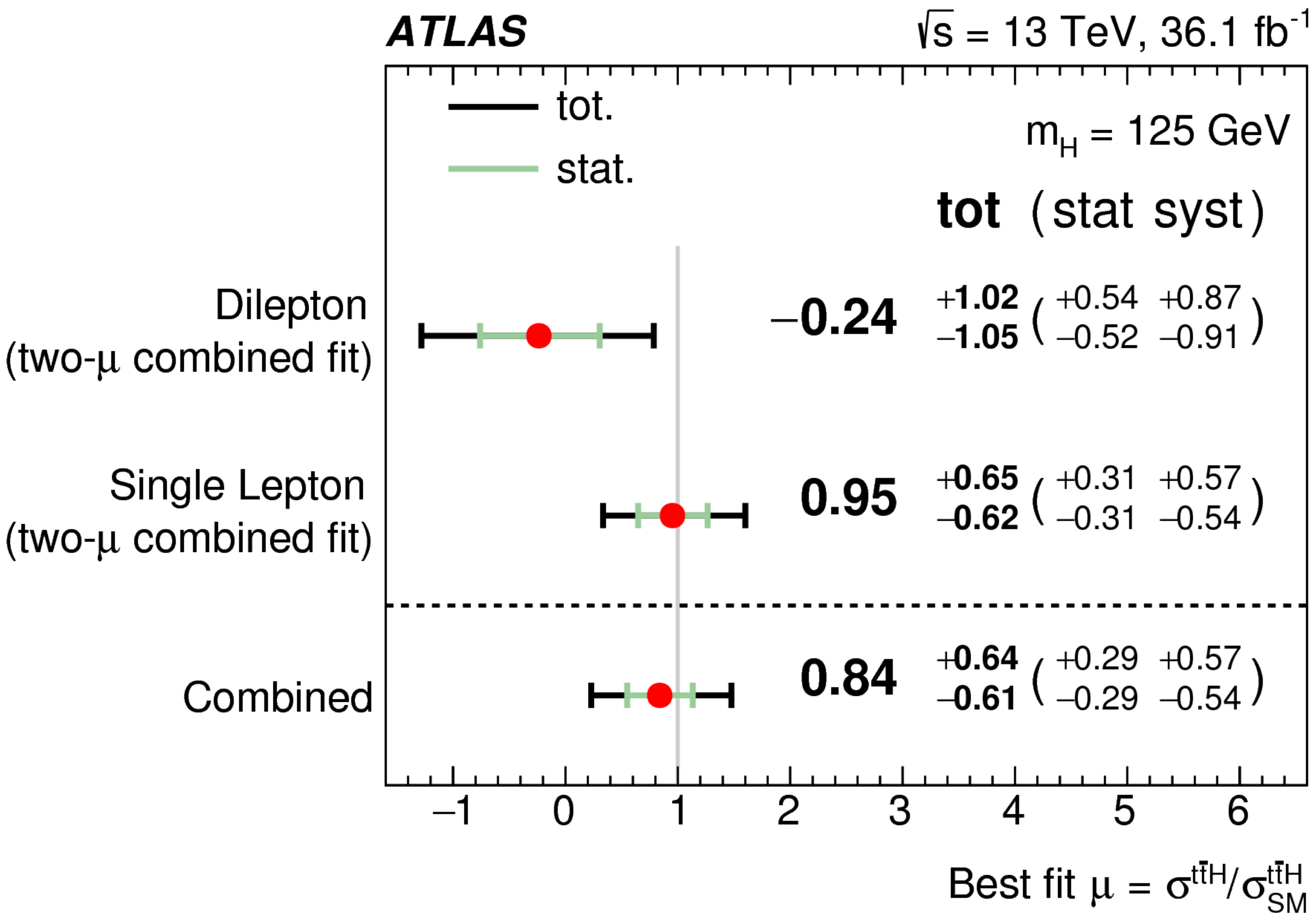




Example:
Single ℓ,
≥ 6j



$t\bar{t}H(b\bar{b})$: results



Uncertainty source	$\Delta\mu$	
$t\bar{t}+\geq 1b$ modeling	$+0.46$	-0.46
Background-model stat. unc.	$+0.29$	-0.31
b-tagging efficiency and mis-tag rates	$+0.16$	-0.16
Jet energy scale and resolution	$+0.14$	-0.14
$t\bar{t}H$ modeling	$+0.22$	-0.05
$t\bar{t}+\geq 1c$ modeling	$+0.09$	-0.11
JVT, pileup modeling	$+0.03$	-0.05
Other background modeling	$+0.08$	-0.08
$t\bar{t}$ + light modeling	$+0.06$	-0.03
Luminosity	$+0.03$	-0.02
Light lepton (e, μ) id., isolation, trigger	$+0.03$	-0.04
Total systematic uncertainty	$+0.57$	-0.54
$t\bar{t}+\geq 1b$ normalization	$+0.09$	-0.10
$t\bar{t}+\geq 1c$ normalization	$+0.02$	-0.03
Intrinsic statistical uncertainty	$+0.21$	-0.20
Total statistical uncertainty	$+0.29$	-0.29
Total uncertainty	$+0.64$	-0.61

- Normalisation factors for $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ left free-floating in the fit:
  - NF($t\bar{t}+\geq 1b$) = 1.24 ± 0.10
  - NF($t\bar{t}+\geq 1c$) = 1.63 ± 0.23
- Most relevant uncertainties related to $t\bar{t}+\geq 1b$ background modelling
- Analysis is **dominated by systematic** uncertainties
- Significance w.r.t background-only hypothesis: **1.4σ (1.6σ) obs (exp)**

New $t\bar{t}H(\gamma\gamma)$ and $t\bar{t}H(ZZ\rightarrow4\ell)$ results!

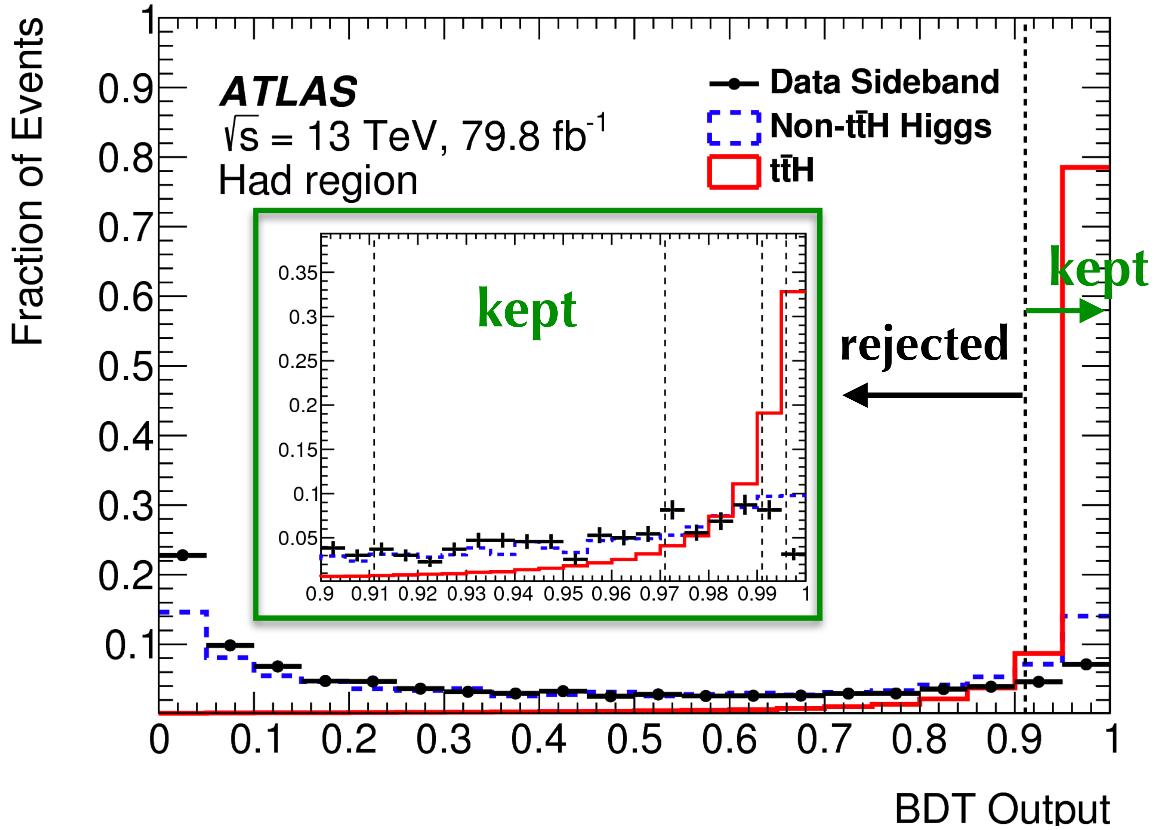
🟩 **2015-2016 data [~36 fb^{-1}]**

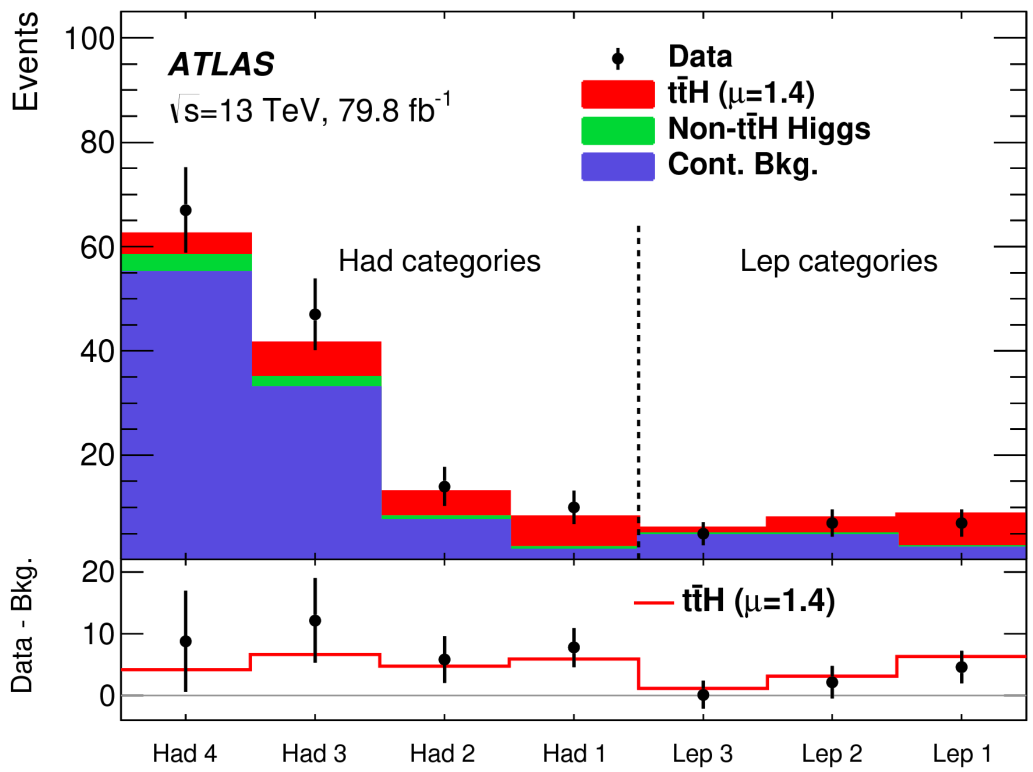
🟦 **2015-2017 data [~80 fb^{-1}]**

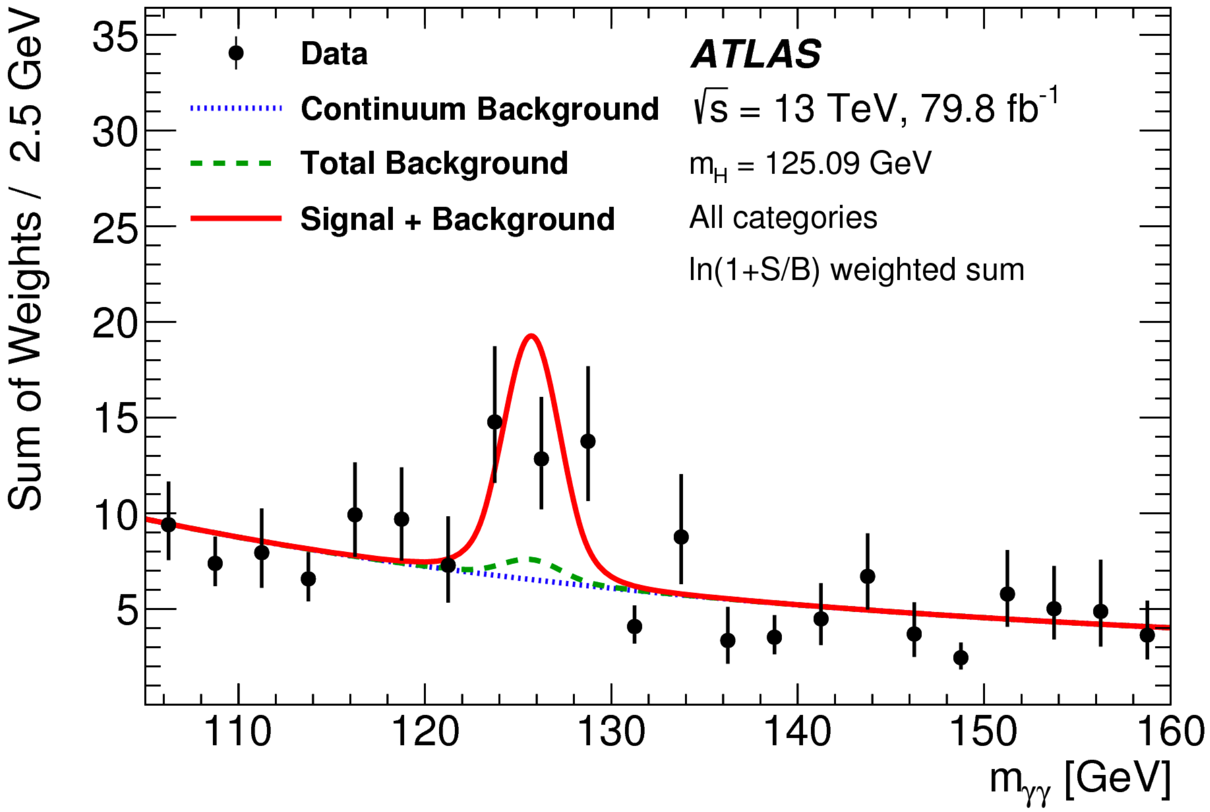
	ATLAS EXPERIMENT	CMS Compact Muon Solenoid
ttH multilepton ($H\rightarrow WW/\tau\tau/ZZ$)	Phys. Rev. D 97 (2018) 072003 (including ttH combination 36.1/fb)	JHEP 08 (2018) 066 $\mu_{ttH} = 1.23\ ^{+0.45}_{-0.43}$
ttH(bb)	Phys. Rev. D 97 (2018) 072016 (leptonic)	arXiv:1804.03682 (leptonic) $\mu_{ttH} = 0.72 \pm 0.45$ JHEP 06 (2018) 101 (all-hadronic) $\mu_{ttH} = 0.9 \pm 1.5$
ttH($ZZ\rightarrow4\ell$)	Phys. Lett. B 784 (2018) 173 (including ttH combination 80/fb + 36.1/fb ➜ Observation)	JHEP 11 (2017) 047 $\mu_{ttH} < 1.19$
ttH(γγ)		arXiv:1804.02716 3.3σ (exp: 1.5σ) $\mu_{ttH} = 2.2\ ^{+0.9}_{-0.8}$
Combination		*Phys. Rev. Lett. 120 (2018) 231801* ➜ Observation
ATLAS+CMS Run1 combination	JHEP 1608 (2016) 045 4.4σ (exp: 2.0σ) $\mu_{ttH} = 2.3\ ^{+0.7}_{-0.6}$	

$t\bar{t}H(H\rightarrow\gamma\gamma)$

- **Seven categories** optimised for ttH, with 0-≥1ℓ from $t\bar{t}$ decays
  - Based on cuts on a **BDT per channel** to discriminate against non-resonant diphoton production and non-ttH Higgs production
  - Input variables: photon, jets, MET and leptons (Lep channel) observables
- **Leptonic channel:** ≥1 b-tagged jets
- **Hadronic channel:** ≥2 jets, ≥1 b-tagged jets

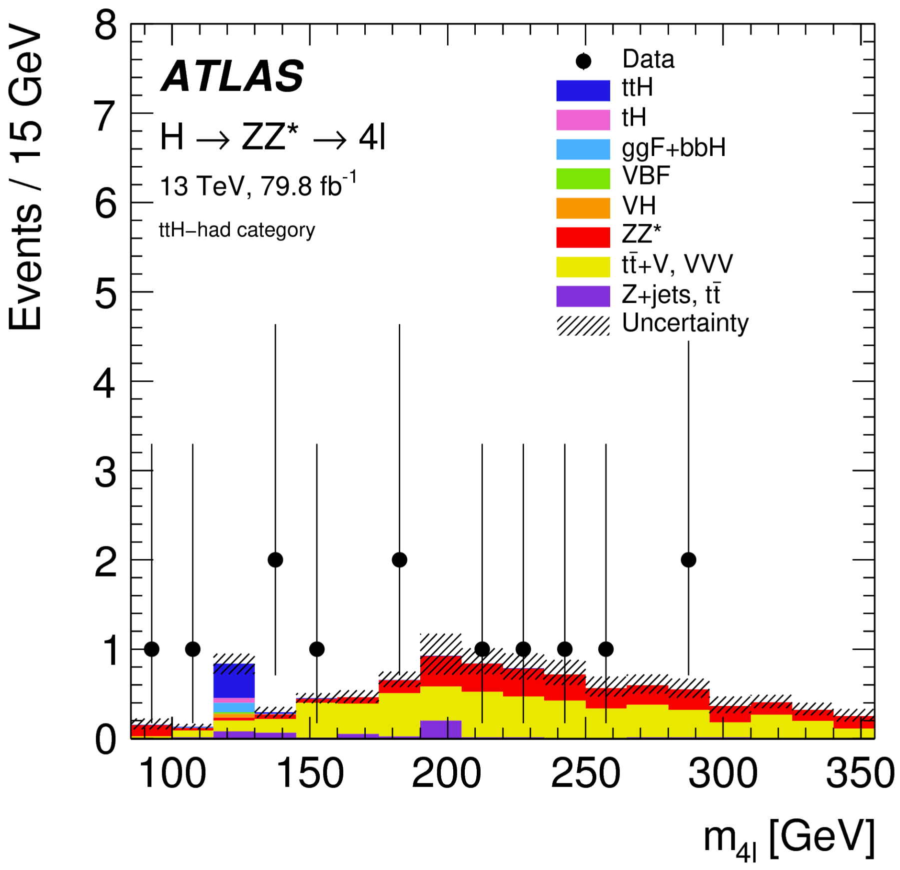






$t\bar{t}H(H\to ZZ\to 4\ell)$ resonant

- Higgs boson candidates with 115 < $m(4\ell)$ < 130 GeV
- $t\bar{t}H$ enriched category:
  - ≥1 b-tagged jet
  - 0 additional ℓ + ≥3 jets **[Had]** or 1 additional ℓ + ≥1 jets **[Lep]**
  - **BDT in Had channel** with jet, MET and lepton observables, as well as LO Matrix-Element value of Higgs boson decay, as input variables

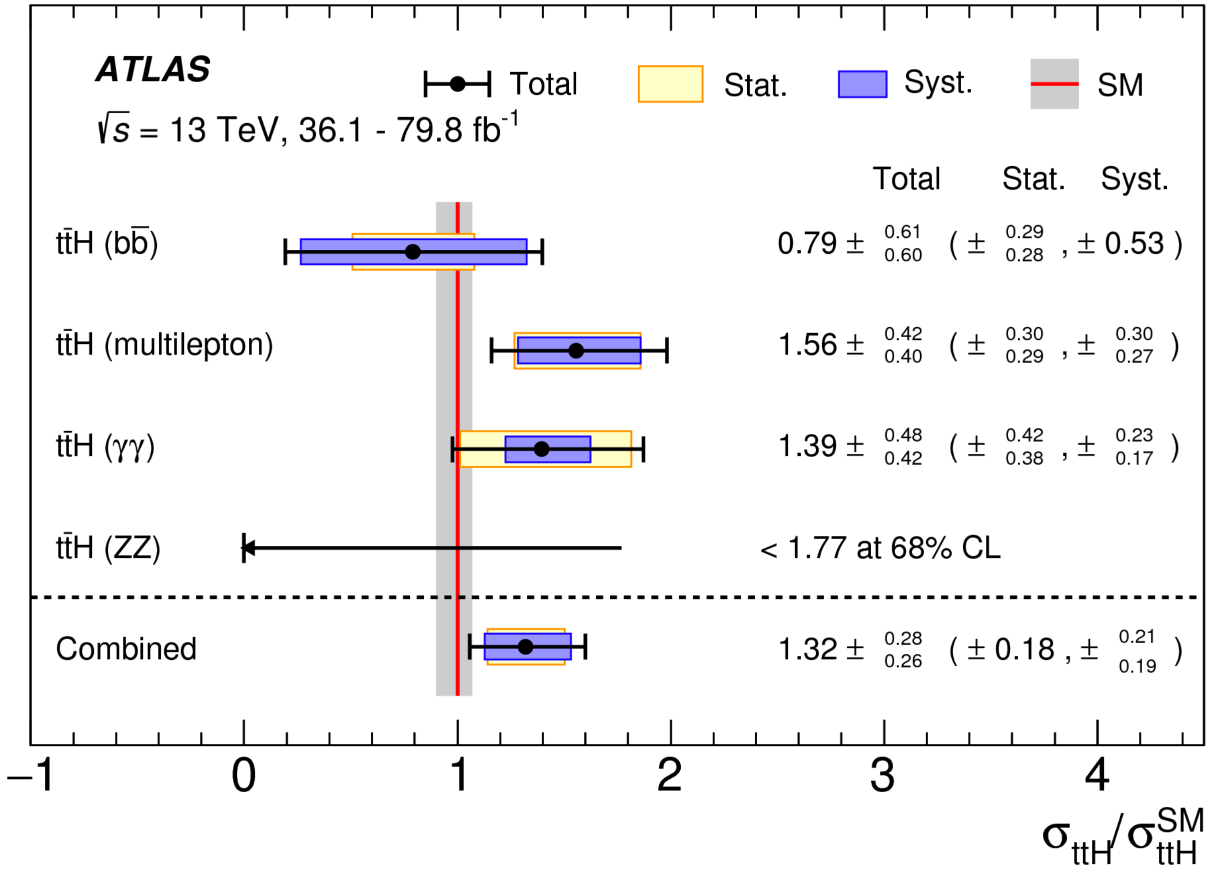


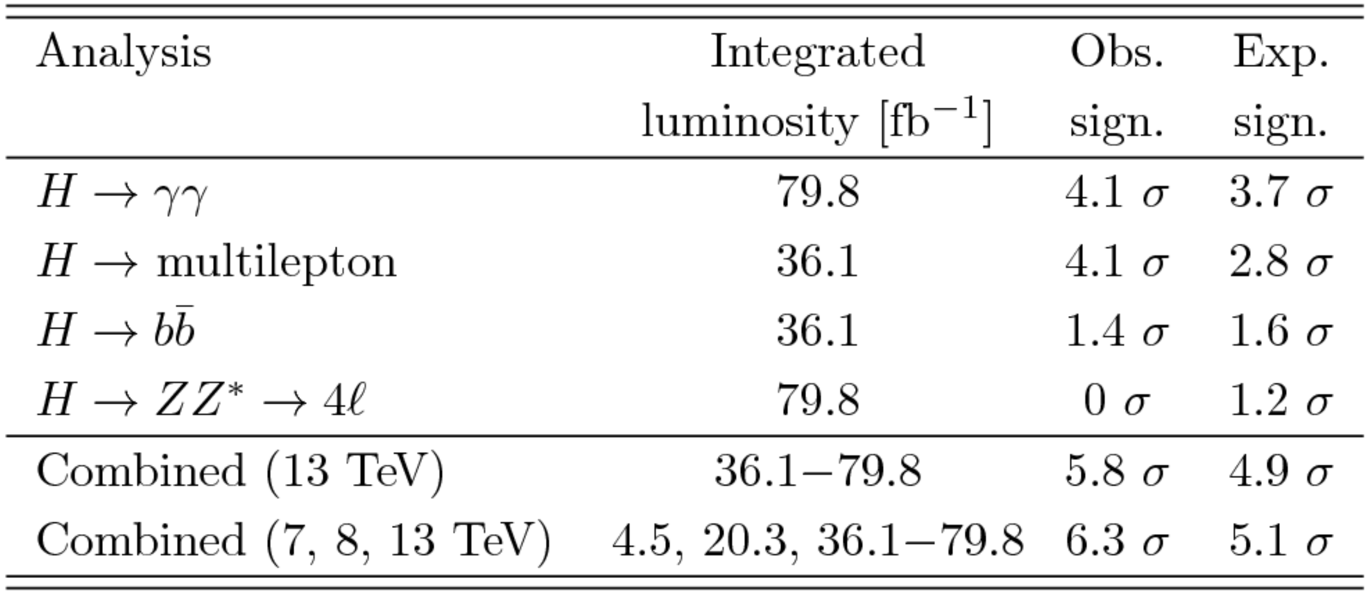
Bin	Expected				Observed
	$t\bar{t}H$ (signal)	Non-$t\bar{t}H$ Higgs	Non-Higgs	Total	Total
	$H \to \gamma\gamma$				
Had 1	4.2(11)	0.49(33)	1.76(55)	6.4(13)	10
Had 2	3.41(74)	0.69(56)	7.5(11)	11.6(15)	14
Had 3	4.70(88)	2.0(17)	32.9(22)	39.6(32)	47
Had 4	3.00(55)	3.2(31)	55.0(28)	61.3(47)	67
Lep 1	4.5(10)	0.25(9)	2.19(59)	6.9(12)	7
Lep 2	2.23(39)	0.27(10)	4.59(91)	7.1(10)	7
Lep 3	0.82(18)	0.30(13)	4.58(91)	5.70(88)	5
	$H \to ZZ^* \to 4\ell$				
Had 1	0.169(31)	0.021(7)	0.008(8)	0.198(33)	0
Had 2	0.216(32)	0.20(9)	0.22(12)	0.63(16)	0
Lep	0.212(31)	0.0256(23)	0.015(13)	0.253(34)	0

- **Purity** of Had 1 (signal-enriched BDT bin) and Lep **> 80%**
- No events observed → Upper limits on $t\bar{t}H$

$t\bar{t}H$ observation: *in combination*

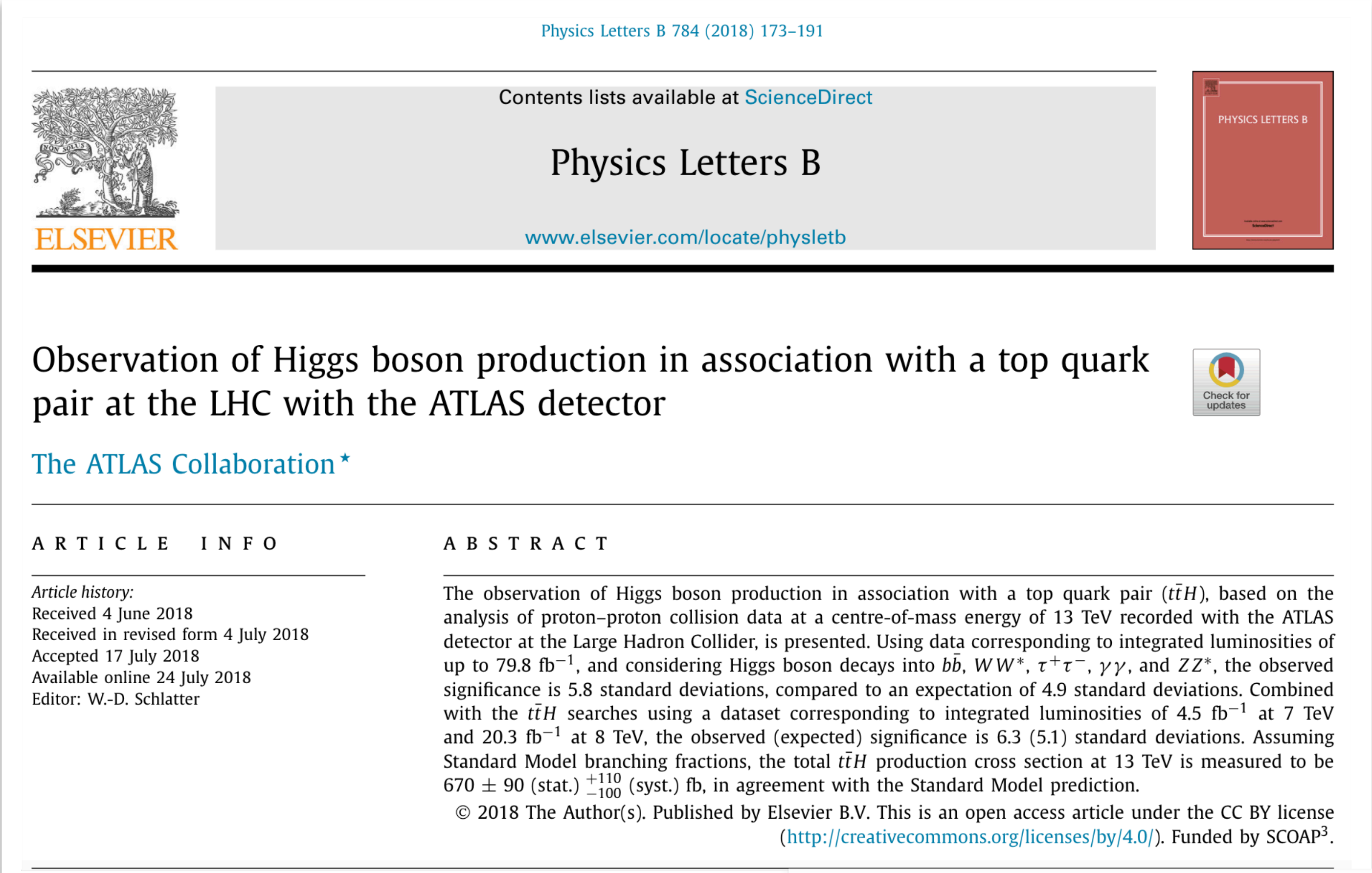
- Combination of multilepton, $b\bar{b}$, $\gamma\gamma$, and $ZZ\rightarrow 4\ell$ $t\bar{t}H$ analyses
- Results in agreement with the SM predictions
  - $\sigma(ttH) = 670^{+142}_{-135}$ fb
  - $\sigma_{SM}(ttH) = 507^{+35}_{-50}$ fb
- Significance w.r.t background-only hypothesis when combining with Run 1:
  - **6.3σ (5.1σ) obs (exp)**
  - Observation of $t\bar{t}H$ production!





Analysis	Integrated luminosity [fb^{-1}]	Obs. sign.	Exp. sign.
$H\rightarrow\gamma\gamma$	79.8	4.1 σ	3.7 σ
$H\rightarrow$ multilepton	36.1	4.1 σ	2.8 σ
$H\rightarrow b\bar{b}$	36.1	1.4 σ	1.6 σ
$H\rightarrow ZZ^*\rightarrow 4\ell$	79.8	0 σ	1.2 σ
Combined (13 TeV)	36.1−79.8	5.8 σ	4.9 σ
Combined (7, 8, 13 TeV)	4.5, 20.3, 36.1−79.8	6.3 σ	5.1 σ

Observation of $t\bar{t}H$ by ATLAS and CMS



Physics Letters B 784 (2018) 173–191

Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector

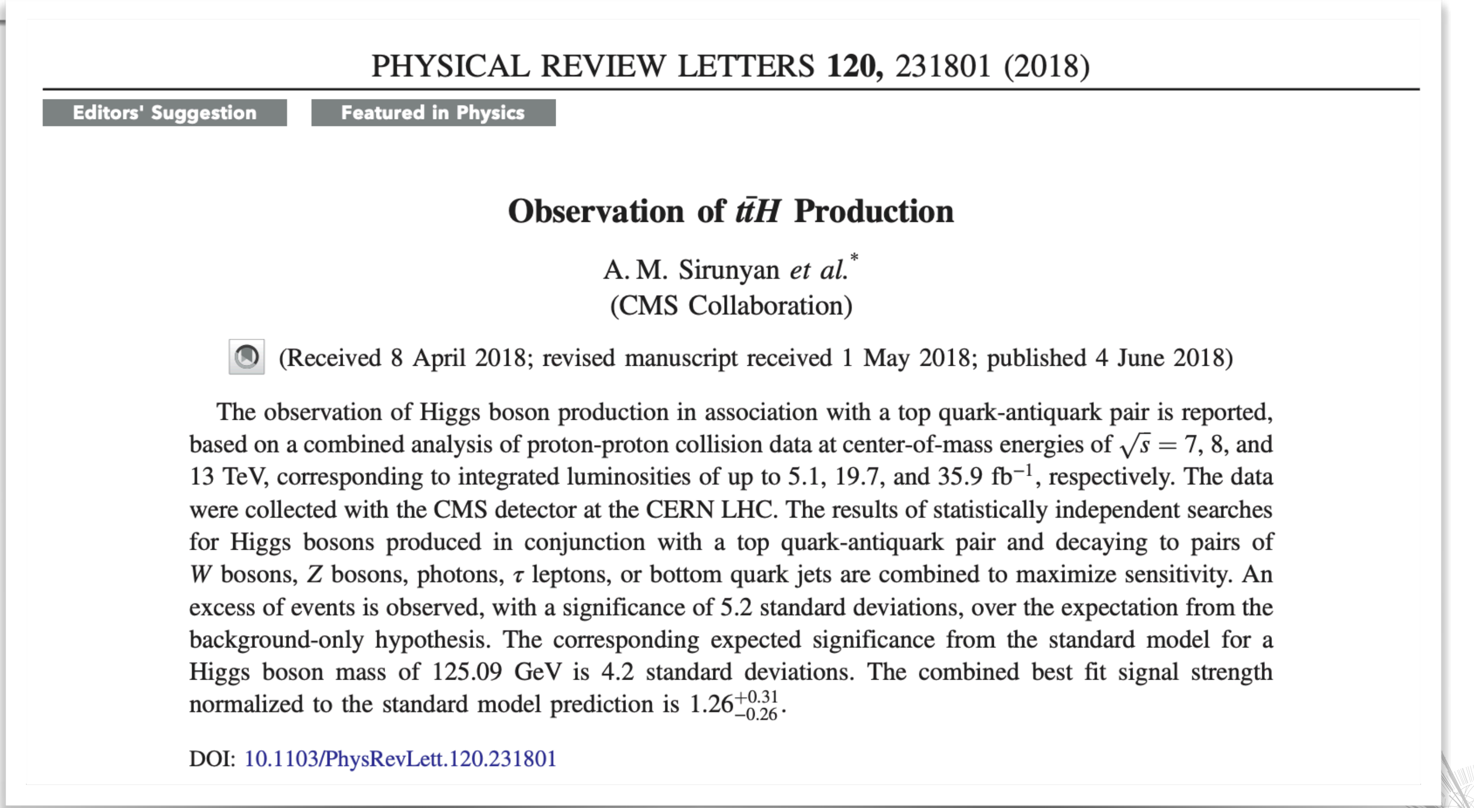
The ATLAS Collaboration *

A R T I C L E I N F O

Article history:
Received 4 June 2018
Received in revised form 4 July 2018
Accepted 17 July 2018
Available online 24 July 2018
Editor: W.-D. Schlatter

A B S T R A C T

The observation of Higgs boson production in association with a top quark pair ($t\bar{t}H$), based on the analysis of proton–proton collision data at a centre-of-mass energy of 13 TeV recorded with the ATLAS detector at the Large Hadron Collider, is presented. Using data corresponding to integrated luminosities of up to 79.8 fb^{-1}, and considering Higgs boson decays into $b\bar{b}$, WW^*, $\tau^+\tau^-$, $\gamma\gamma$, and ZZ^*, the observed significance is 5.8 standard deviations, compared to an expectation of 4.9 standard deviations. Combined with the $t\bar{t}H$ searches using a dataset corresponding to integrated luminosities of 4.5 fb^{-1} at 7 TeV and 20.3 fb^{-1} at 8 TeV, the observed (expected) significance is 6.3 (5.1) standard deviations. Assuming Standard Model branching fractions, the total $t\bar{t}H$ production cross section at 13 TeV is measured to be 670 ± 90 (stat.) $^{+110}_{-100}$ (syst.) fb, in agreement with the Standard Model prediction.



PHYSICAL REVIEW LETTERS **120,** 231801 (2018)

Editors' Suggestion   Featured in Physics

Observation of $t\bar{t}H$ Production

A. M. Sirunyan *et al.*[*]
(CMS Collaboration)

(Received 8 April 2018; revised manuscript received 1 May 2018; published 4 June 2018)

The observation of Higgs boson production in association with a top quark-antiquark pair is reported, based on a combined analysis of proton-proton collision data at center-of-mass energies of $\sqrt{s} = 7$, 8, and 13 TeV, corresponding to integrated luminosities of up to 5.1, 19.7, and 35.9 fb^{-1}, respectively. The data were collected with the CMS detector at the CERN LHC. The results of statistically independent searches for Higgs bosons produced in conjunction with a top quark-antiquark pair and decaying to pairs of W bosons, Z bosons, photons, τ leptons, or bottom quark jets are combined to maximize sensitivity. An excess of events is observed, with a significance of 5.2 standard deviations, over the expectation from the background-only hypothesis. The corresponding expected significance from the standard model for a Higgs boson mass of 125.09 GeV is 4.2 standard deviations. The combined best fit signal strength normalized to the standard model prediction is $1.26^{+0.31}_{-0.26}$.

DOI: 10.1103/PhysRevLett.120.231801

Current $t\bar{t}H$ landscape

- What have we been working on since then?
  - **New measurements**
    - $t\bar{t}H$ STXS `[simplified template cross section]` measurements (p_T^{Higgs})
    - $t\bar{t}H$ CP-odd contribution searches
  - **Addressing long-standing / recent issues:**
    - Improve estimation of non-prompt lepton background ($t\bar{t}H$ multiℓ)
    - Understand observed tension with theory prediction in $t\bar{t}W$-like regions ($t\bar{t}H$ multiℓ)
    - Improve estimation of $t\bar{t}b\bar{b}$ background ($t\bar{t}Hbb$)
    - Improve evaluation of background modelling uncertainties (all)

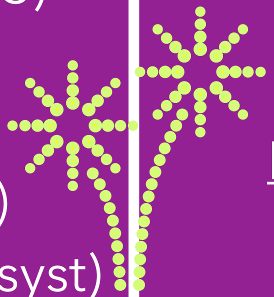
$t\bar{t}H$ state of the art Run 2

🟩 2015-2016 [~36 fb^{-1}]

🟦 2015-2017 [~80 fb^{-1}]

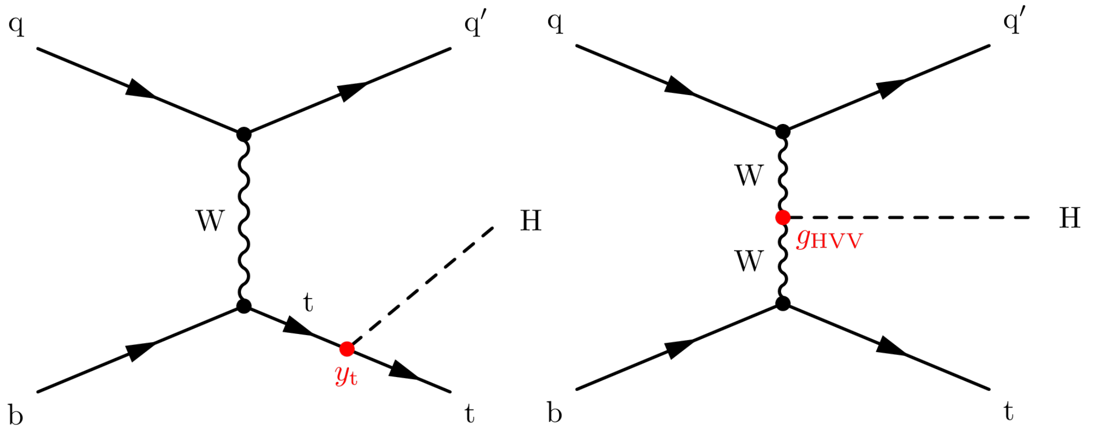
🟪 2015-2018 [~140 fb^{-1}]

	ATLAS EXPERIMENT	CMS
ttH multilepton (H→WW/ττ/ZZ)	🟦 ATLAS-CONF-2019-045 $\mu_{ttH} = 0.58\ ^{+0.26}_{-0.25}$	🟪 arXiv:2011.03652 $\mu_{ttH} = 0.92 \pm 0.19$ (stat) $^{+0.17}_{-0.13}$ (syst)
ttH(bb)	🟪 arXiv:2111.06712 $\mu_{ttH} = 0.35\ ^{+0.36}_{-0.34}$	🟦 CMS-PAS-HIG-18-030 $\mu_{ttH} = 1.15\ ^{+0.15}_{-0.15}$ (stat) $^{+0.28}_{-0.25}$ (syst)
ttH(ZZ→4ℓ)	🟪 Eur. Phys. J. C 80 (2020) 957 (+STXS) $\mu_{ttH} = 1.7\ ^{+1.7}_{-1.2} \pm 0.2 \pm 0.2$	🟪 arXiv:2103.04956 (+STXS) $\mu_{ttH} = 0.13\ ^{+0.92}_{-0.13}$ (stat) $^{+0.11}_{-0.00}$ (syst)
ttH(γγ) **Observation in a single channel!**	🟪 ATLAS-CONF-2020-026 (+ STXS) $\mu_{ttH+tH} = 0.92\ ^{+0.27}_{-0.24}$ **4.7 (5.0) σ** obs (exp) PRL 125 (2020) 061802 (+CP) $\mu_{ttH} = 1.43\ ^{+0.33}_{-0.31}$ (stat) $^{+0.21}_{-0.15}$ (syst) **5.2 (4.4) σ** obs (exp)	🟪 JHEP07(2021)027 (+STXS) $\mu_{ttH} = 1.35\ ^{+0.34}_{-0.28}$ PRL 125 (2020) 061801 (+CP) $\mu_{ttH} = 1.38\ ^{+0.36}_{-0.29}$ **6.6 (4.7) σ** obs (exp)
Combination	🟩🟦 Phys. Lett. B 784 (2018) 173 (80/fb + 36.1/fb ➜ **Observation**)	🟩 Phys. Rev. Lett. 120 (2018) 231801 ➜ **Observation**



tH state of the art

- Central top and Higgs, back-to-back
- **Destructive interference** in SM (top Yukawa coupling competing against g_{HVV})
- Very challenging due to low SM cross section and larger background than tt̄H



🟩 2015-2016 [~36 fb^{-1}]

🟪 2015-2018 [~140 fb^{-1}]

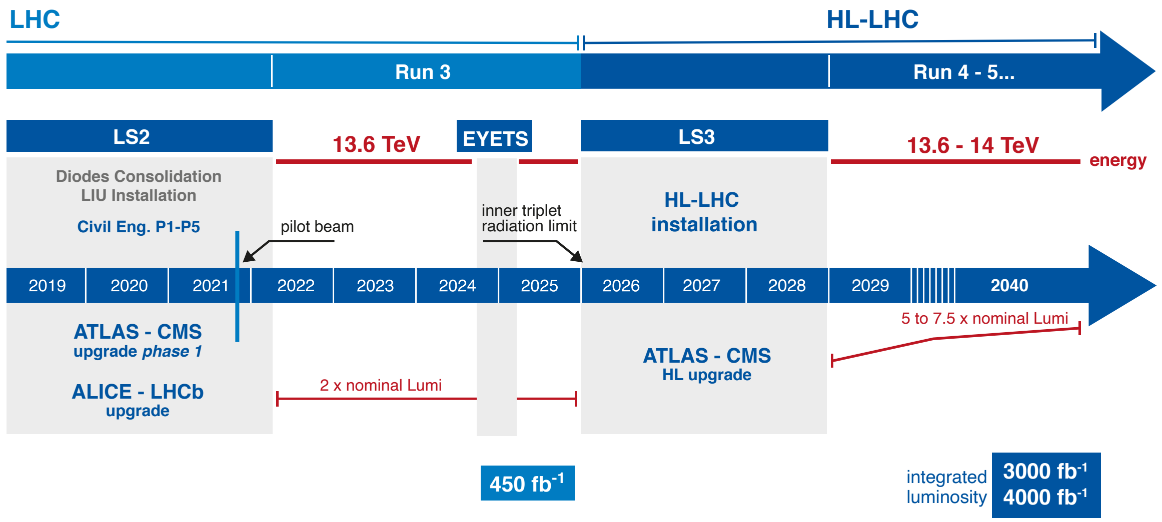
	ATLAS EXPERIMENT	CMS
tH multilepton (H→WW/ττ/ZZ)	*ongoing*	🟪 Eur. Phys. J. C 81, 378 (2021) μ_{tH} = 5.7 ± 2.7 (stat) ± 3.0 (syst)
tH(bb)	*ongoing*	🟩 *(see below)*
tH(γγ)	🟪 ATLAS-CONF-2020-026 μ_{tH} < 8 x SM @95% CL	🟪 JHEP07(2021)027 μ_{tH} < 14 x SM @95% CL
Combination	*ongoing*	🟩 PRD 99 (2019) 092005 Expected and observed 95% CL upper limits on the tH XS x BR

Scenario	Channel	Observed	Expected
$\kappa_t = -1$	$b\bar{b}$	4.98 pb	$2.52^{+1.29}_{-0.81}$ pb
	$\gamma\gamma$	0.84 pb	$0.88^{+0.46}_{-0.28}$ pb
	$\mu^\pm\mu^\pm + e^\pm\mu^\pm + \ell\ell\ell$	0.85 pb	$0.77^{+0.36}_{-0.24}$ pb
	Combined	0.74 pb	$0.53^{+0.24}_{-0.16}$ pb
$\kappa_t = 1$ (SM-like)	$b\bar{b}$	6.88 pb	$3.19^{+1.64}_{-1.02}$ pb
	$\gamma\gamma$	3.68 pb	$2.03^{+1.05}_{-0.67}$ pb
	$\mu^\pm\mu^\pm + e^\pm\mu^\pm + \ell\ell\ell$	1.36 pb	$1.18^{+0.53}_{-0.35}$ pb
	Combined	1.94 pb	$0.92^{+0.40}_{-0.27}$ pb

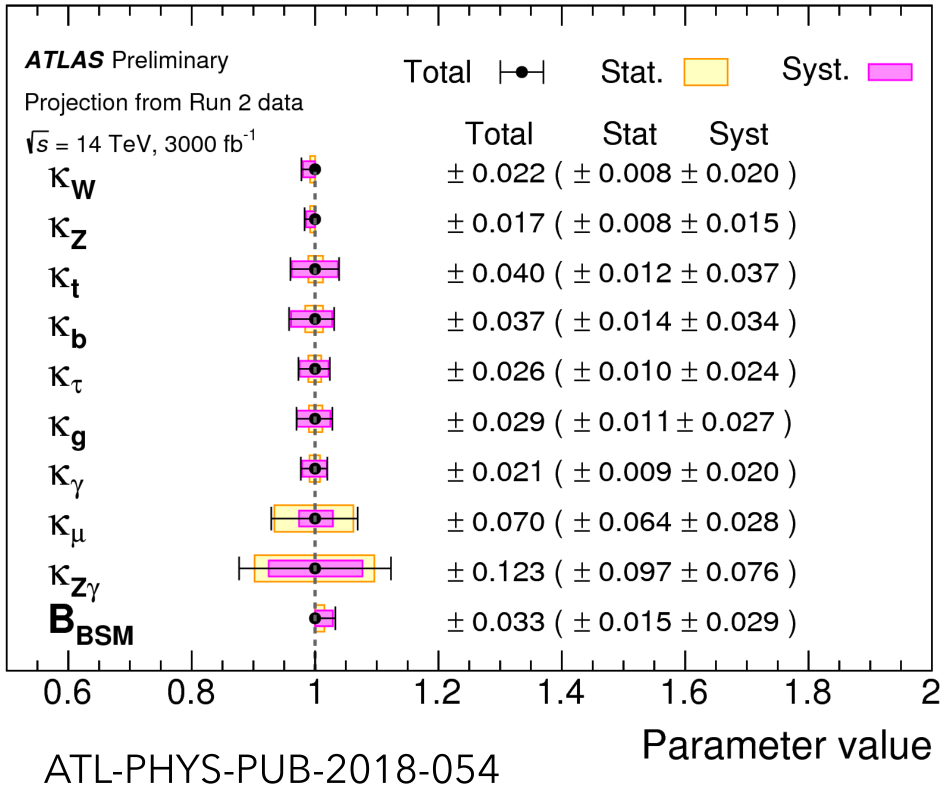
SM tH (XS x BR) = 0.077 pb

The Future

- **At the LHC:**
  - Still analysing full Run-2 dataset!
  - **Run 3 13.6 TeV stable beams imminently**: triple integrated luminosity!
  - **HL-LHC**: 10x more luminosity, explore less accessible processes such as di-Higgs (self-coupling of Higgs boson)



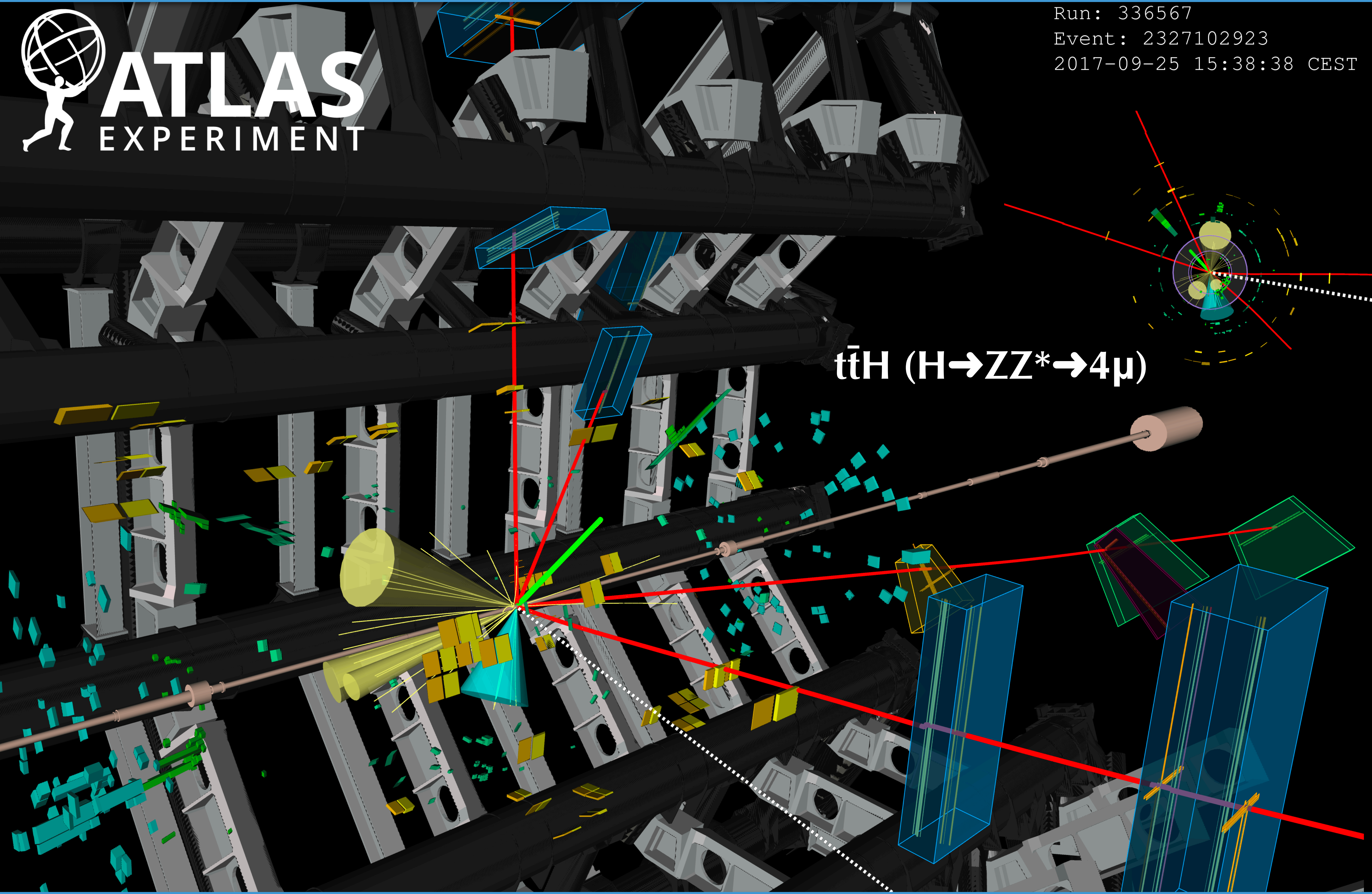
LHC/ HL-LHC Plan (last update February 2022)



ATL-PHYS-PUB-2018-054

- Expect to measure the top Yukawa coupling (modifier) $\boldsymbol{\kappa}_t$ at **4% level** at the end of HL-LHC
  - **Systematics-limited!**

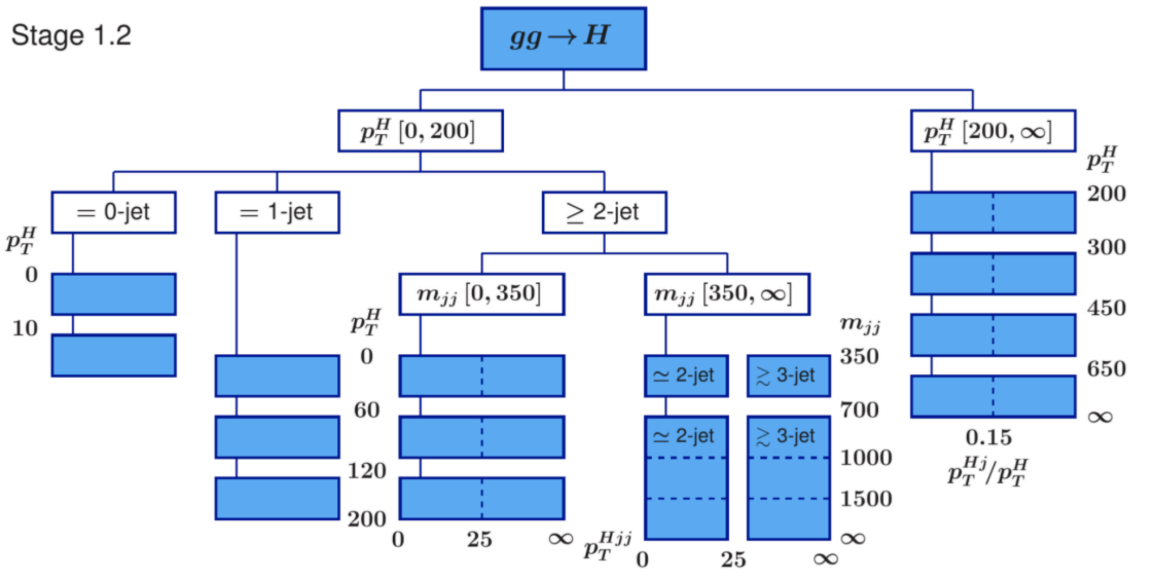
Stay tuned for upcoming results!

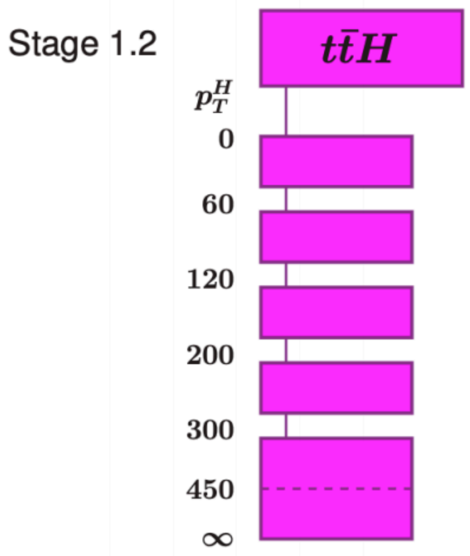


Thanks for your attention!

Simplified Template Cross Section (STXS)

- Measure **production modes** separately, categorising each into bins of key (truth) quantities (p_T^H, Njets, m_{jj}, …)
  - Chosen as most sensitive variables to theory predictions / signal sensitivity / new physics
  - Different stages (e.g. stage 0, stage 1, stage 1.2) with varying degrees of granularity
  - Decay mode agnostic: well-suited for combinations
- **How to design an STXS analysis?**
  - How are events categorised?
    - Reconstructed quantities as proxy for truth quantities or multivariate classifier
  - How many / which bins to target?
    - Driven by analysis sensitivity





David Shope, Higgs STXS, Moriond'21

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

- Probe the charge conjugation and parity (CP) properties of the Yukawa coupling of the Higgs boson to the top quark

- Any measured CP-odd contribution would be a sign of physics beyond the SM
  - explain observed baryon asymmetry of the universe?

$$\begin{array}{lcr} |f_{\mathrm{CP}}^{\mathrm{Htt}}| = \frac{|\tilde{\kappa}_t|^2}{|\tilde{\kappa}_t|^2+|\kappa_t|^2} & \Leftrightarrow & \sin^2\alpha \\ \mu_{\mathrm{ttH}} & \Leftrightarrow & \kappa_t^2 \end{array}$$

CP-structure ttH parametrisation:

$$\mathcal{A}(\mathrm{Htt}) = -\frac{m_{\mathrm{t}}}{v}\overline{\psi}_{\mathrm{t}}\left(\kappa_{\mathrm{t}} + \mathrm{i}\tilde{\kappa}_{\mathrm{t}}\gamma_5\right)\psi_{\mathrm{t}},$$

$$f_{\mathrm{CP}}^{\mathrm{Htt}} = \frac{|\tilde{\kappa}_{\mathrm{t}}|^2}{|\kappa_{\mathrm{t}}|^2 + |\tilde{\kappa}_{\mathrm{t}}|^2}\,\mathrm{sign}(\tilde{\kappa}_{\mathrm{t}}/\kappa_{\mathrm{t}})$$

SM (CP-even): $\boldsymbol{\kappa}_{\mathrm{t}} = 1$; $\tilde{\boldsymbol{\kappa}}_{\mathrm{t}} = 0$

CP-odd: $\boldsymbol{\kappa}_{\mathrm{t}} = 0$; $\tilde{\boldsymbol{\kappa}}_{\mathrm{t}} = 1$

or

$$\mathcal{L} = -\frac{m_t}{v}\left\{\bar{\psi}_t\kappa_t\left[\cos(\alpha) + \mathrm{i}\sin(\alpha)\gamma_5\right]\psi_t\right\}H$$

SM (CP-even): $\boldsymbol{\alpha} = 0°$

CP-odd: $\boldsymbol{\alpha} = 90°$

$t\bar{t}H(b\bar{b})$: analysis strategy

- **Biggest challenge**: good modelling of the $t\bar{t}$+HF (≥1b, ≥1c) background

Nominal sample: @NLO 4-flavour scheme $t\bar{t}b\bar{b}$

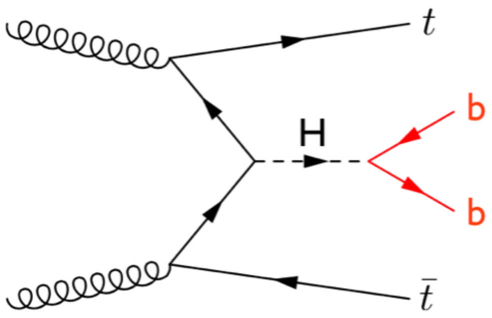
$t\bar{t}$+≥1c and $t\bar{t}$+light modelled by $t\bar{t}$ @NLO

Nominal sample: @NLO 5-flavour scheme

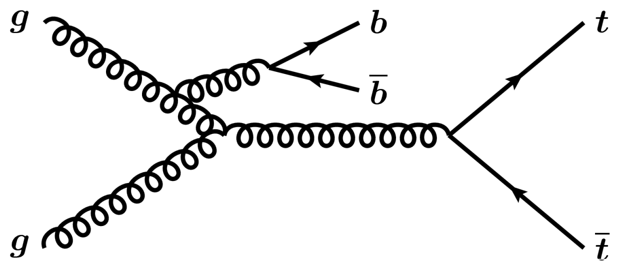
Split in further sub-components: $t\bar{t}+b\bar{b}$, $t\bar{t}$+2b (unresolved), $t\bar{t}$+b (extra b missed)

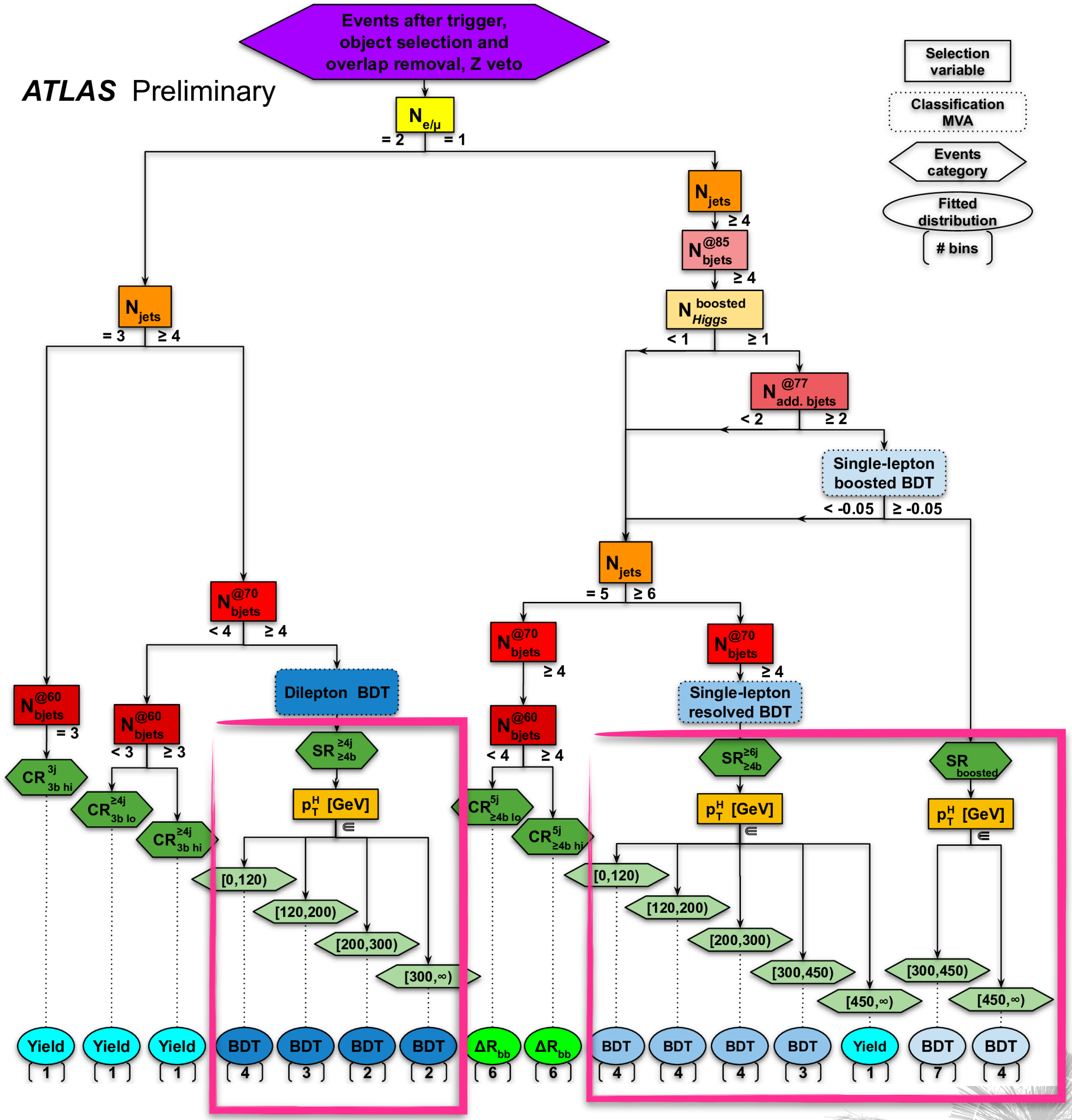
- **Channel categorisation** based on
  - Number of ℓ (0, 1 or 2 opposite-sign)
  - Number of jets
  - Requirements on the b-tagging discriminant (based on **4 or 1** calibrated working points)
  - Resolved or boosted, for single lepton channel
  - Multi-classification ANN decisions for single lepton channel
  - Reconstructed p_T^{Higgs} categories

signal (ttH)



ttbb

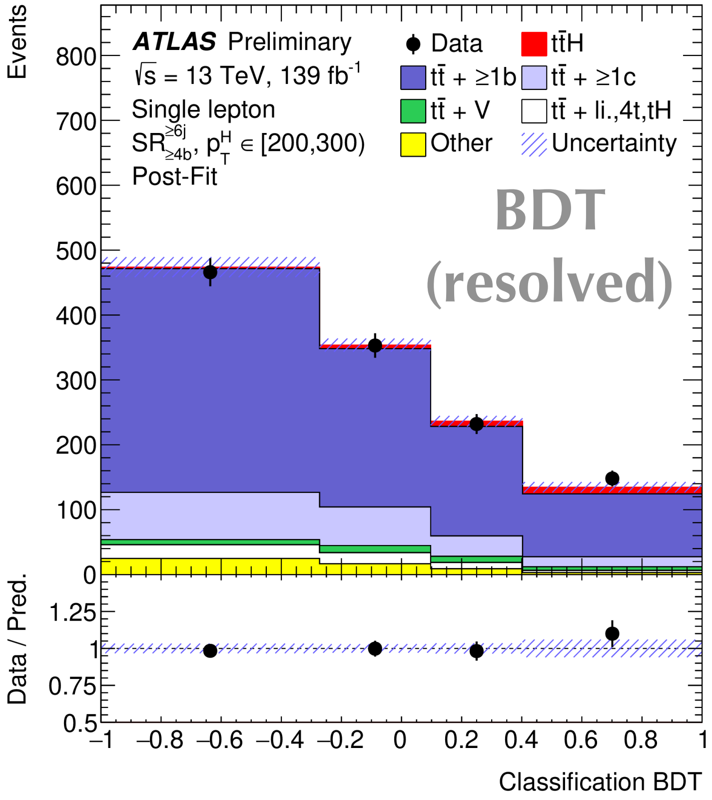


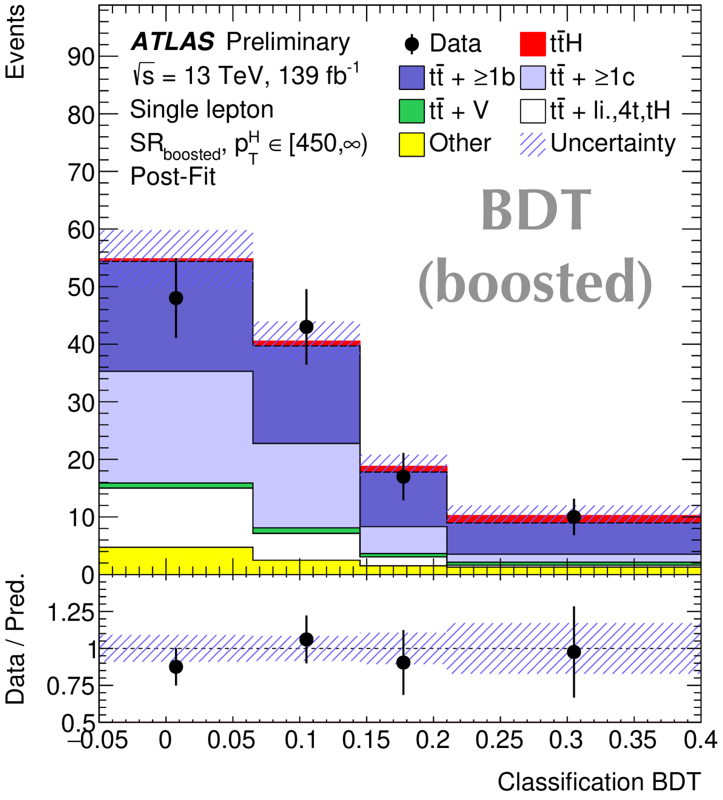


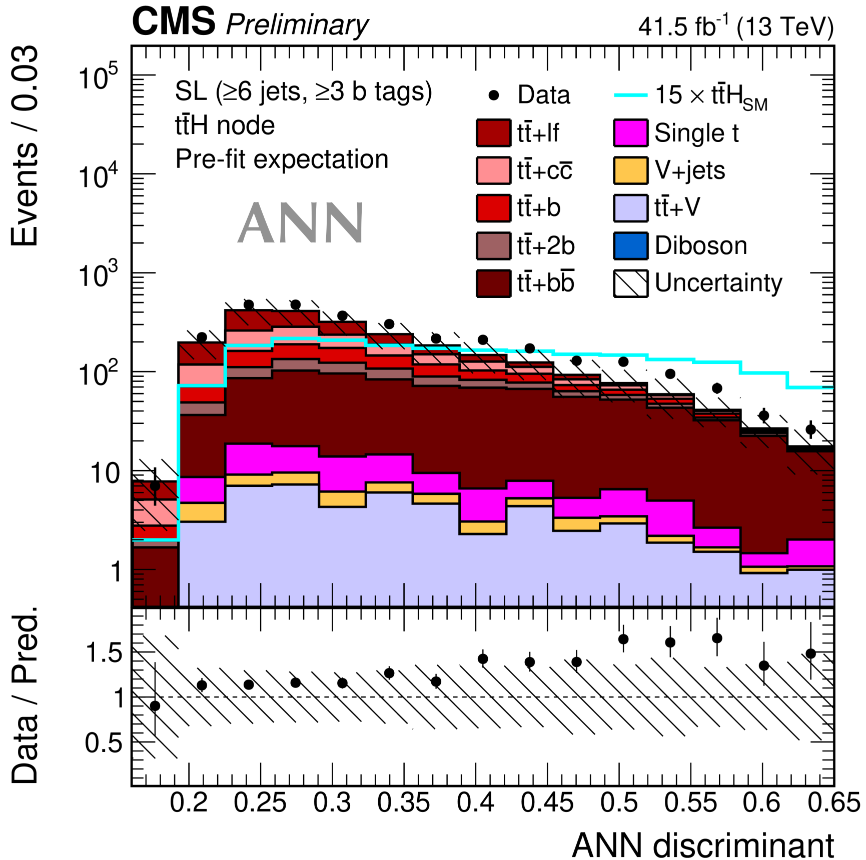
STXS analysis design

$t\bar{t}H(b\bar{b})$: MVA discriminants

- **MVA analysis** needed to discriminate signal from the overwhelming background
  - Input variables of **classification BDT**: kinematic variables, reconstruction BDTs (resolved), likelihood, and **discrete** btagging discriminants
  - **MEM** in 0ℓ (ttH, tt+bb), **ANN** in SL and **BDT** in DL (ttH, tt+jets) with MEM input, as well the **continuous** b-tagging score



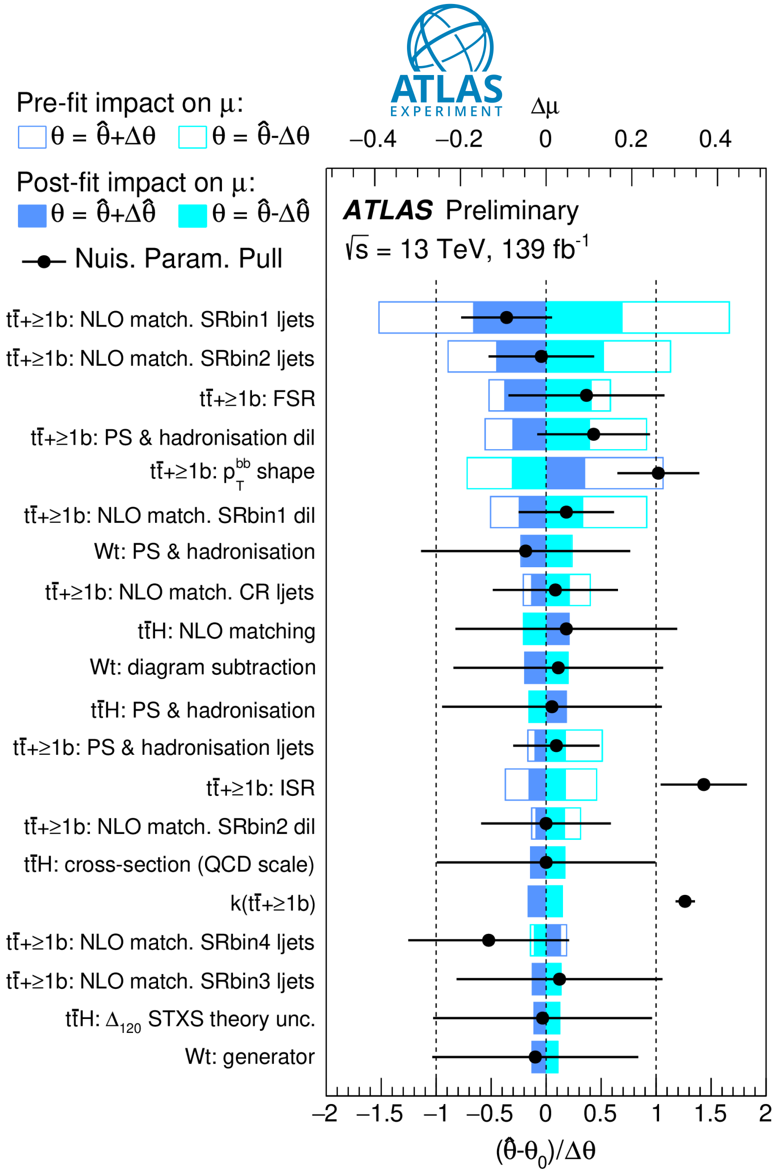


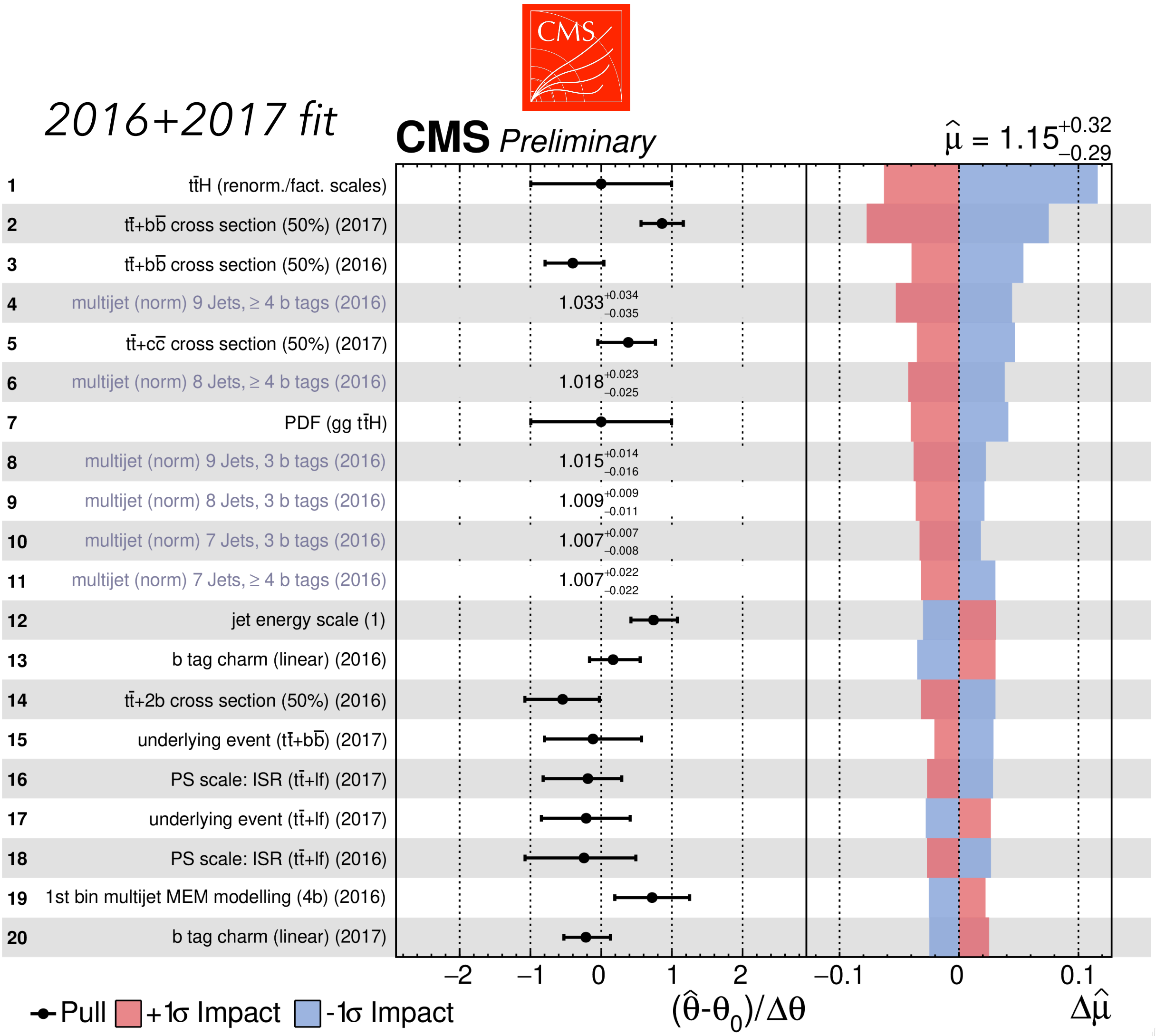


$t\bar{t}H(b\bar{b})$: modelling uncertainties

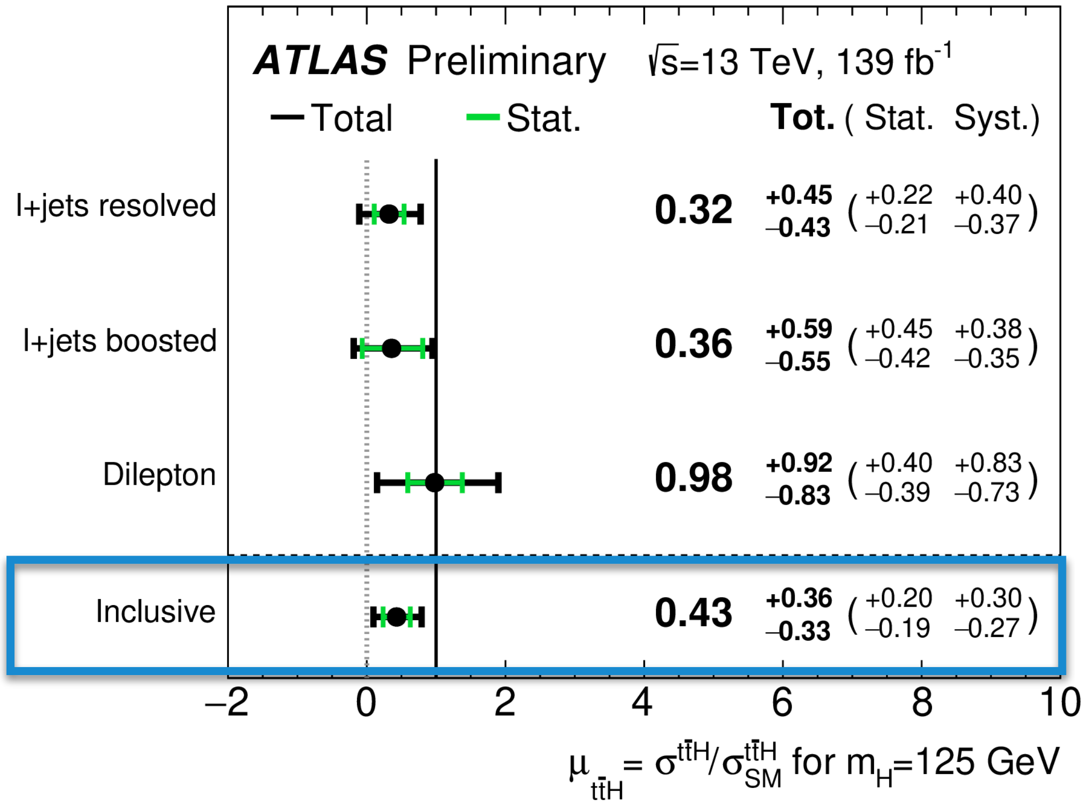
- Generator: Powheg+Pythia8 vs aMC@NLO+Pythia8 (5FS)
- Parton shower: Powheg+Pythia8 vs Powheg+Herwig7
- ISR (+scale), FSR, $t\bar{t}$+1b vs $t\bar{t}$+≥2b fraction uncertainties
- **p_T^{bb} shape uncertainty (ad-hoc)**
- **Free-floating** normalisation $t\bar{t}$+≥1b
- Nuisance parameter (100% prior) $t\bar{t}$+≥1c normalisation

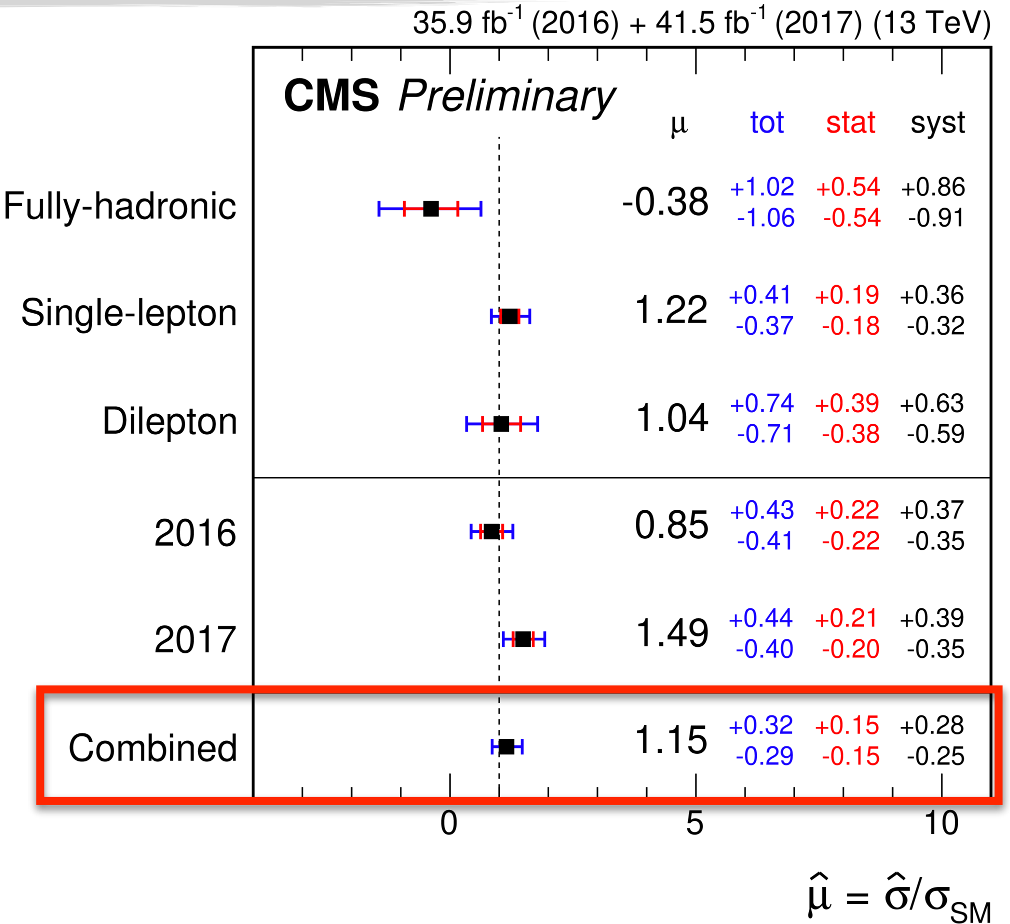
- Parton shower: ISR/FSR
- $t\bar{t}$ underlying event
- $t\bar{t}$ hdamp
- Scale variations
- **Nuisance parameters** for normalisation of tt+bb, tt+2b, tt+b, and $t\bar{t}$+≥1c (50% prior) and **decorrelated** between years

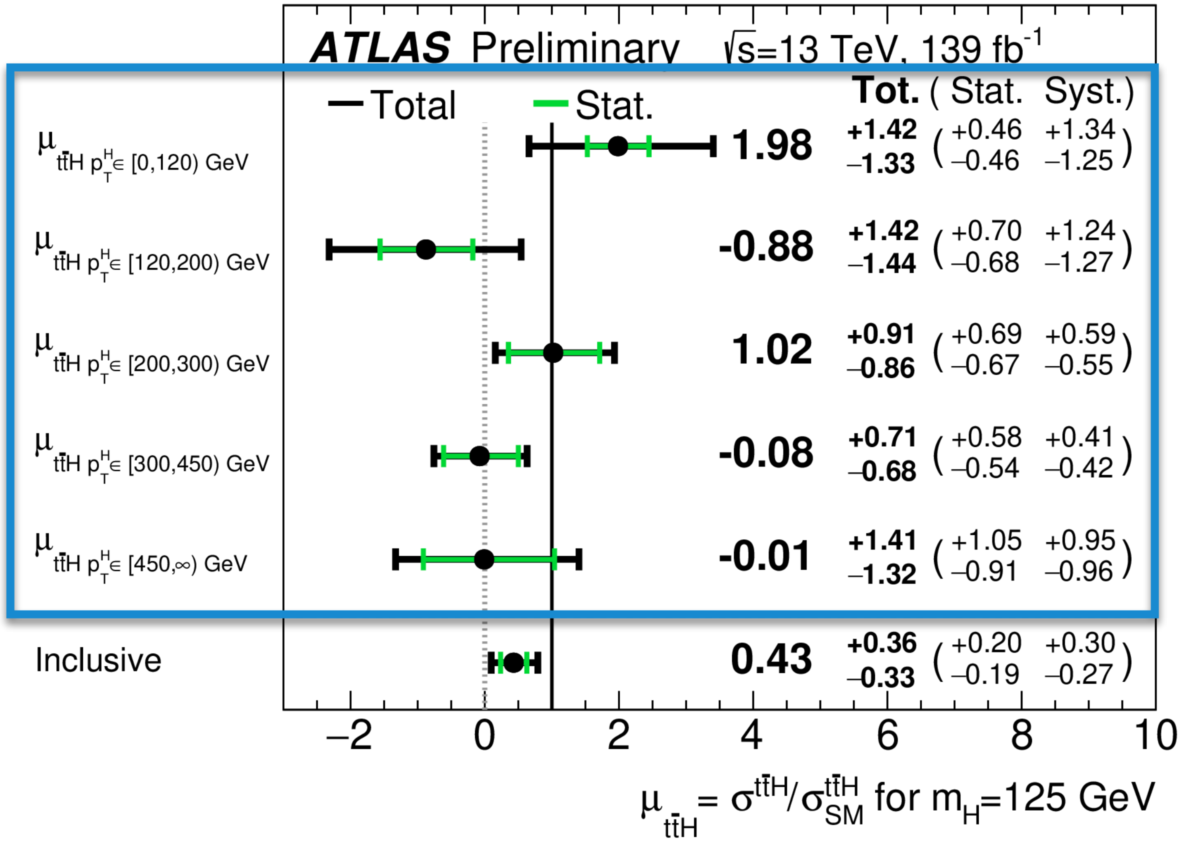




$t\bar{t}H(b\bar{b})$: results





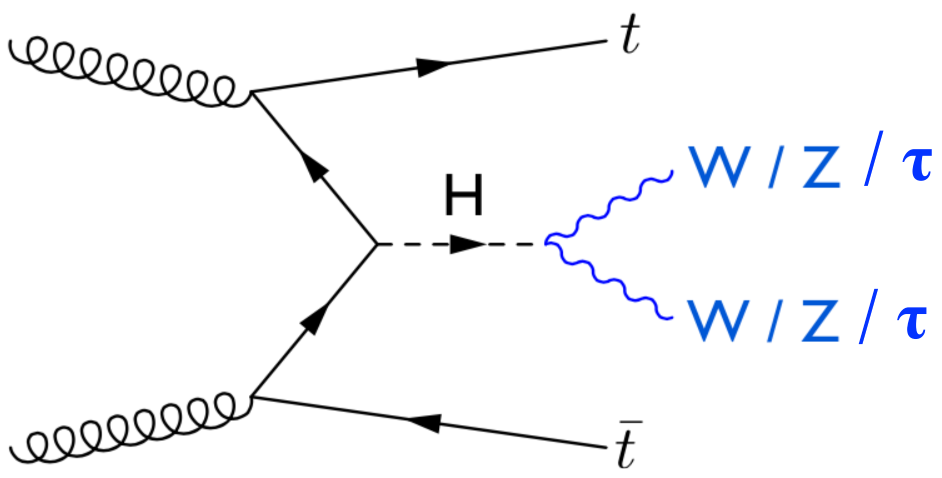


- NF($t\bar{t}$+≥1b) = 1.26 ± 0.09
- **Dominated by systematic** uncertainties
- Most relevant uncertainties related to $t\bar{t}$+≥1b background modelling (Δμ/μ = 60% and 15%)
- Significance w.r.t background-only hypothesis: **1.3 (3.0σ)** and **3.9σ (3.5σ)** obs (exp)
  - **Evidence** for $t\bar{t}H$ in H➔bb channel
- **First $t\bar{t}H$(bb) STXS measurement**
  - Complements $t\bar{t}H(\gamma\gamma)$ STXS measurements **at high p_T^H**

$t\bar{t}$H (multiℓ): analysis strategy

- **Target**: $t\bar{t}$H with
  - H→WW/ZZ/ττ→≥1ℓ
  - $t\bar{t}$→(ℓ+jets, dilepton)

→ **“multilepton” final state**



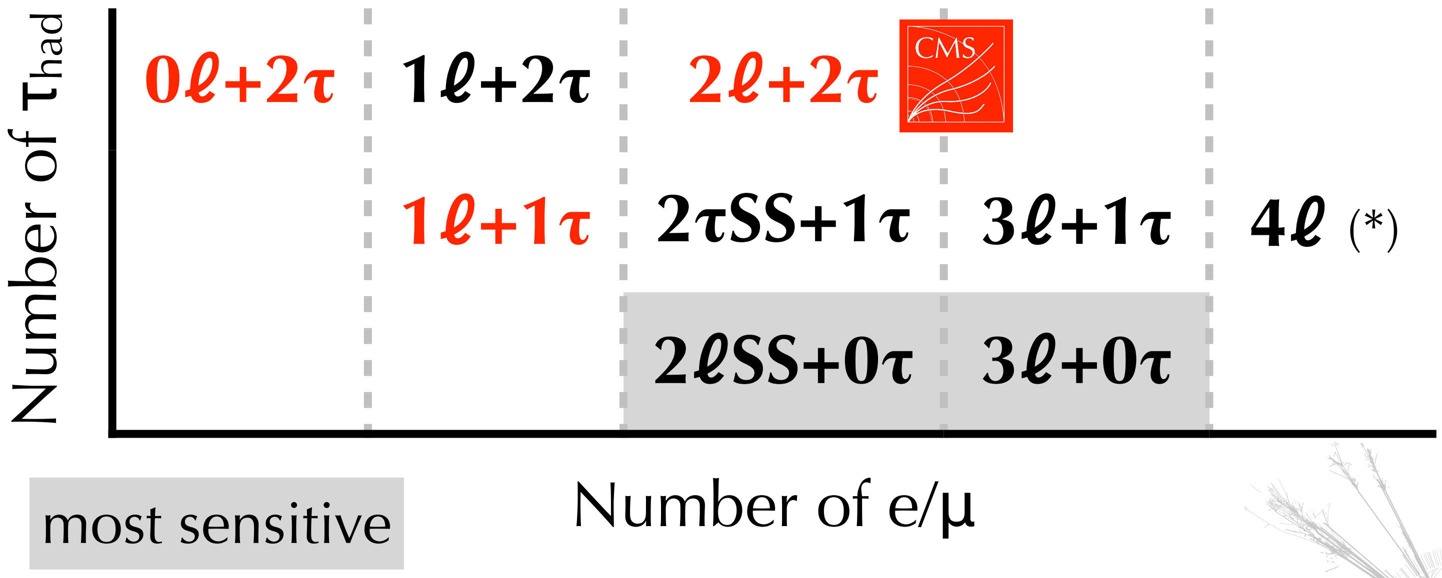
- **High multiplicity** final state
- **Rare in SM:** same-sign 2ℓ, 3ℓ, 4ℓ

- Main reducible backgrounds are: non-prompt ℓ, charge misID electrons, and electrons from photon conversions
  - Specific **lepton BDT isolation** suppressing ℓ from semi-leptonic b-decays, **BDT to reject charge misID**, material and internal (γ*➜ℓ±ℓ∓) electron **conversion** (CO) candidates further suppressed with track invariant masses and conversion radius

- Main irreducible backgrounds are: $t\bar{t}$Z, $t\bar{t}$W, VV

- Several categorisation stages

- **#1 categorisation:** split in categories based on **number of e/μ** and **number of τ**



0ℓ+2τ	1ℓ+2τ	2ℓ+2τ		
	1ℓ+1τ	2τSS+1τ	3ℓ+1τ	4ℓ (*)
		2ℓSS+0τ	3ℓ+0τ	

(*) m(4ℓ) != Higgs mass window

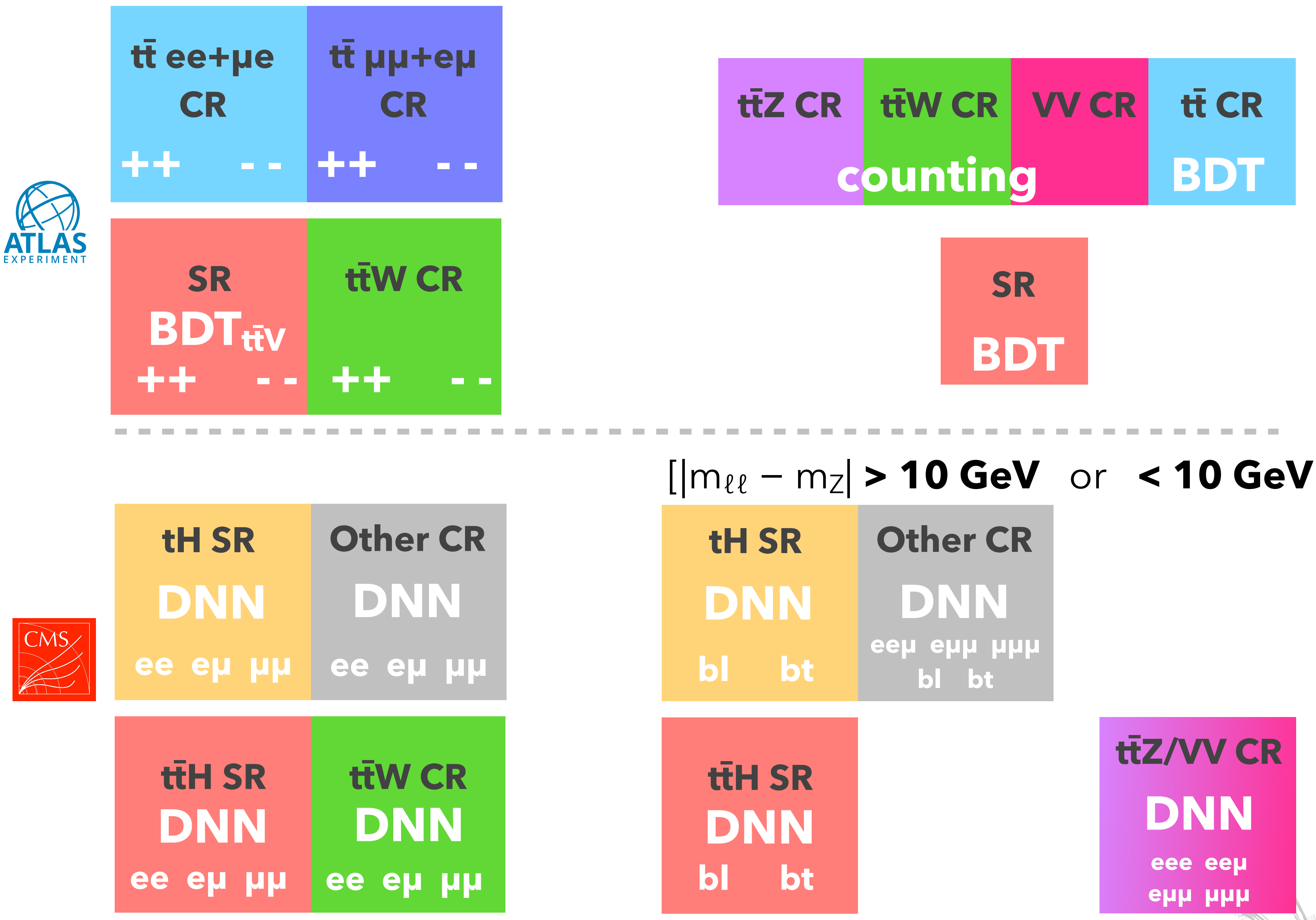
tt̄H (multiℓ): categories

- **#2 categorisation** (“high NJets”):
  - **2ℓSS0τ:** a combination of 2 **BDTs** (vs. tt̄V, vs. fakes/tt̄) in a **2D space**, or
  - **3ℓ0τ:** a **multi-dimensional BDT** (vs. tt̄W, vs. fakes/tt̄, vs. tt̄Z, vs. VV)
  - **2ℓSS0τ, 3ℓ0τ and 2ℓSS1τ: DNN** (vs tH vs other backgrounds)**; BDT** in the other channels

Categories based on
*n***D BDT/DNN space**,
ℓ±ℓ± charge *(input to DNN)*,
btagging discriminants,
and
lepton flavour

2ℓSS0τ [≥4j, ≥1bj]

3ℓ0τ [≥2j, ≥1bj]



tt̄ ee+μe CR
++ - -

tt̄ μμ+eμ CR
++ - -

SR
BDT$_{t\bar{t}V}$
++ - -

tt̄W CR
++ - -

tt̄Z CR **tt̄W CR** **VV CR** **tt̄ CR**
counting **BDT**

SR
BDT

[$|m_{\ell\ell} - m_Z|$ **> 10 GeV** or **< 10 GeV**]

tH SR
DNN
ee eμ μμ

Other CR
DNN
ee eμ μμ

tt̄H SR
DNN
ee eμ μμ

tt̄W CR
DNN
ee eμ μμ

tH SR
DNN
bl bt

Other CR
DNN
eeμ eμμ μμμ
bl bt

tt̄H SR
DNN
bl bt

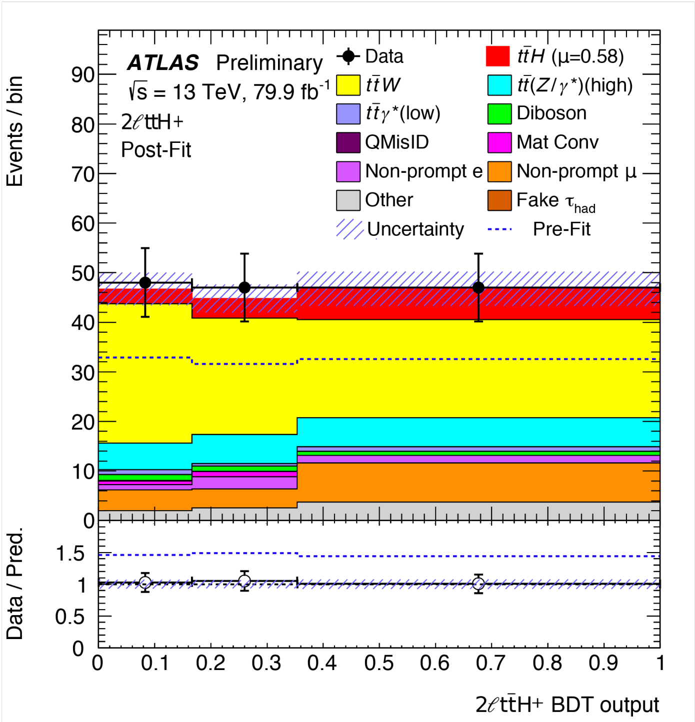
tt̄Z/VV CR
DNN
eee eeμ
eμμ μμμ

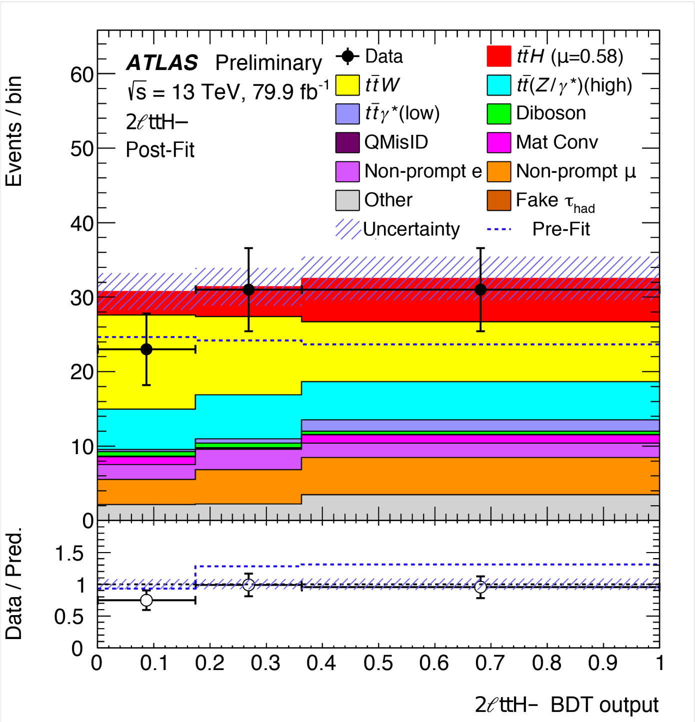
bl = loose btag WP; bt = tight btag WP

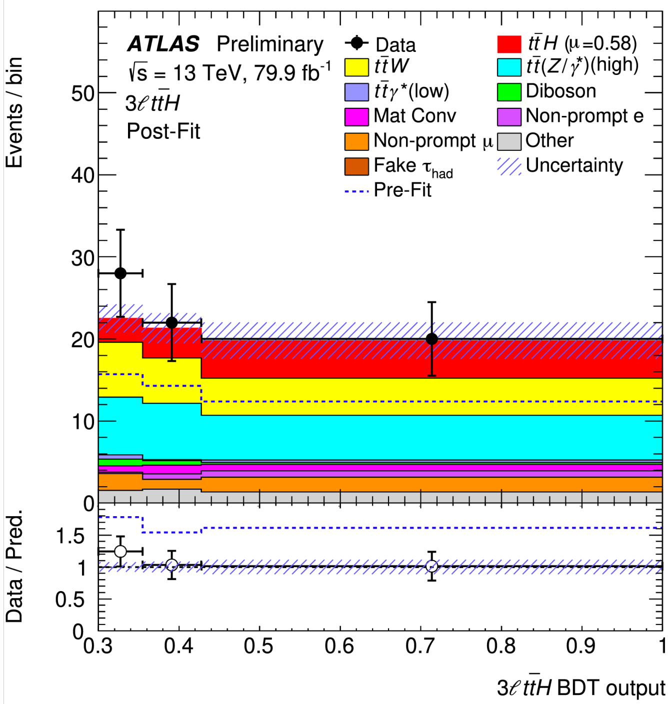
$t\bar{t}$H (multiℓ): signal regions

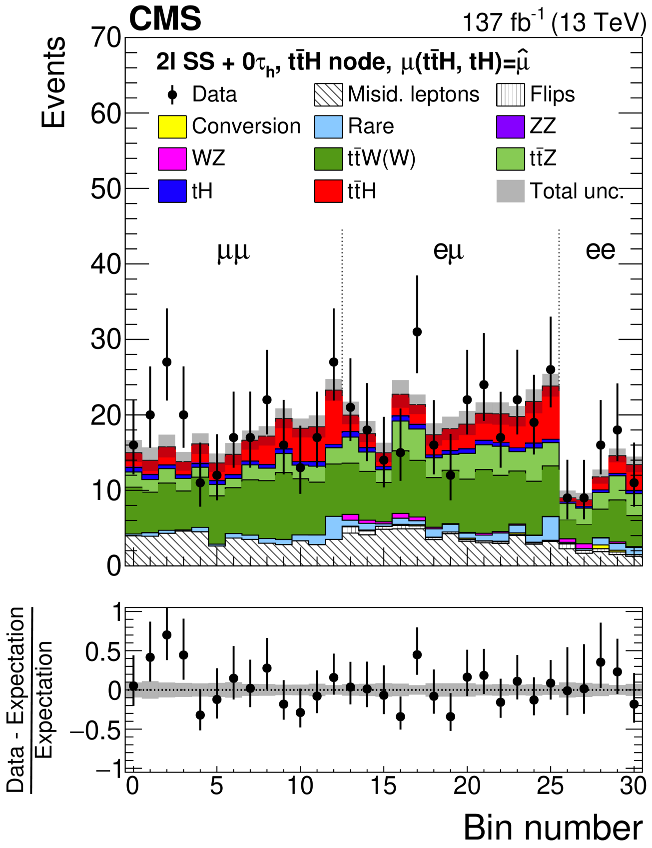
2ℓSS0τ [≥4j, ≥1bj] SRs

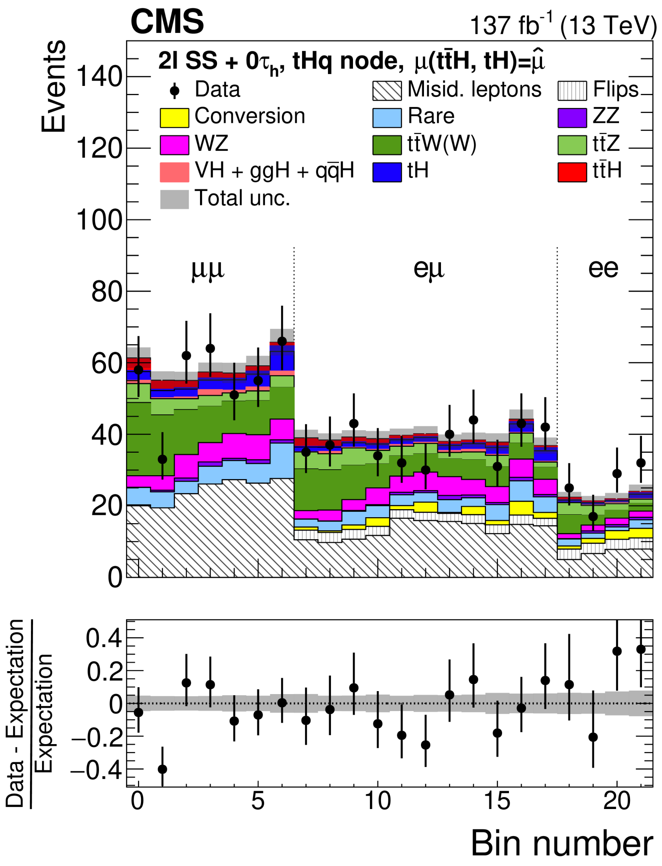
3ℓ0τ [≥2j, ≥1bj] SR

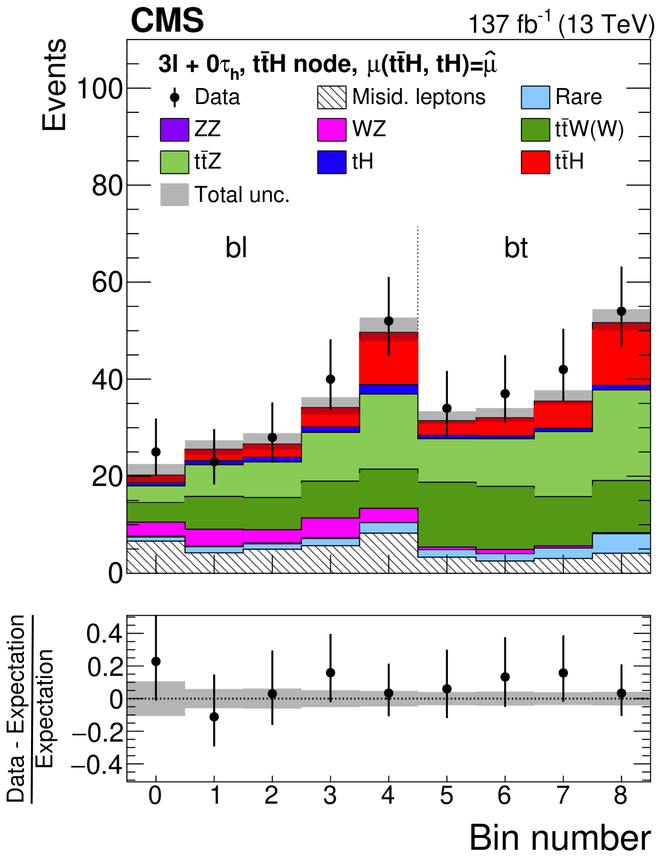






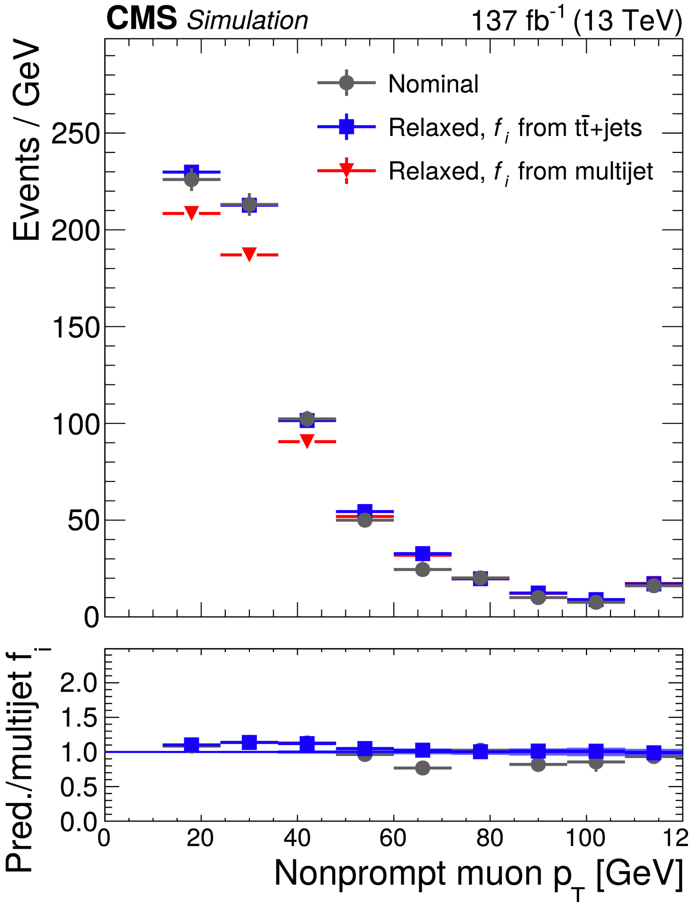






$t\bar{t}H$ (multiℓ): fakes (and more) estimate

- **Fakes estimated from data in a QCD CR with relaxed object ID**
- **#3 categorisation: add CR categories to the fit model** ("low NJets" and conversion CRs)
  - 2ℓSS0τ/3ℓ0τ: ≥1 electron passing **material / internal conversion** selection
  - 2ℓSS0τ: **2-3 jets**, enriched in **non-prompt leptons** and **$t\bar{t}W$**
- Normalisation of **non-prompt leptons** (electrons and muons), electrons from **material CO**, electron from **internal CO** [low mass], **$t\bar{t}W$** (decorrelated between 2ℓSS0τ low NJets, 2ℓSS0τ high NJets, and 3ℓ0τ), and **$t\bar{t}Z$** are measured *simultaneously* in the fit to data
  - Shapes from MC simulation, extensive set of systematic uncertainties included



2ℓSS0τ [2-3j, ≥1bj]

ee+μe 1b — **dR(ℓℓ)**

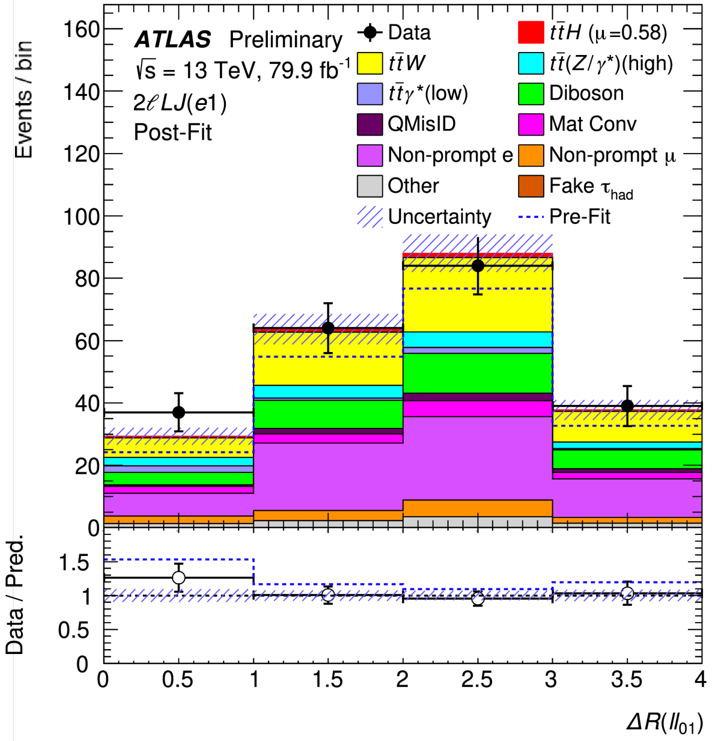
ee+μe 2b — **HT_{lep}**

μμ+eμ — **HT_{lep}**

2ℓSS0τ + 3ℓ0τ [≥2j, ≥1bj]

Mat CO CR — **counting**

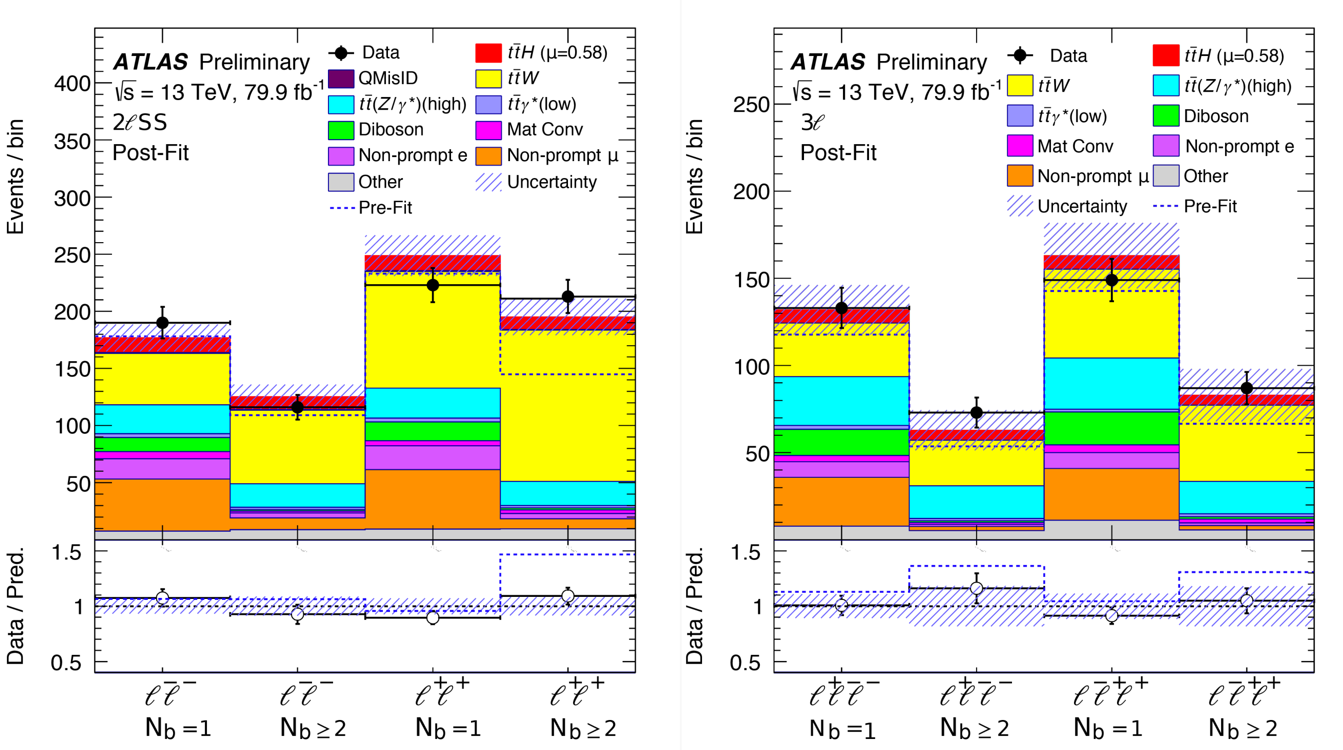
Int CO CR — **counting**



$t\bar{t}H$ (multiℓ): systematics

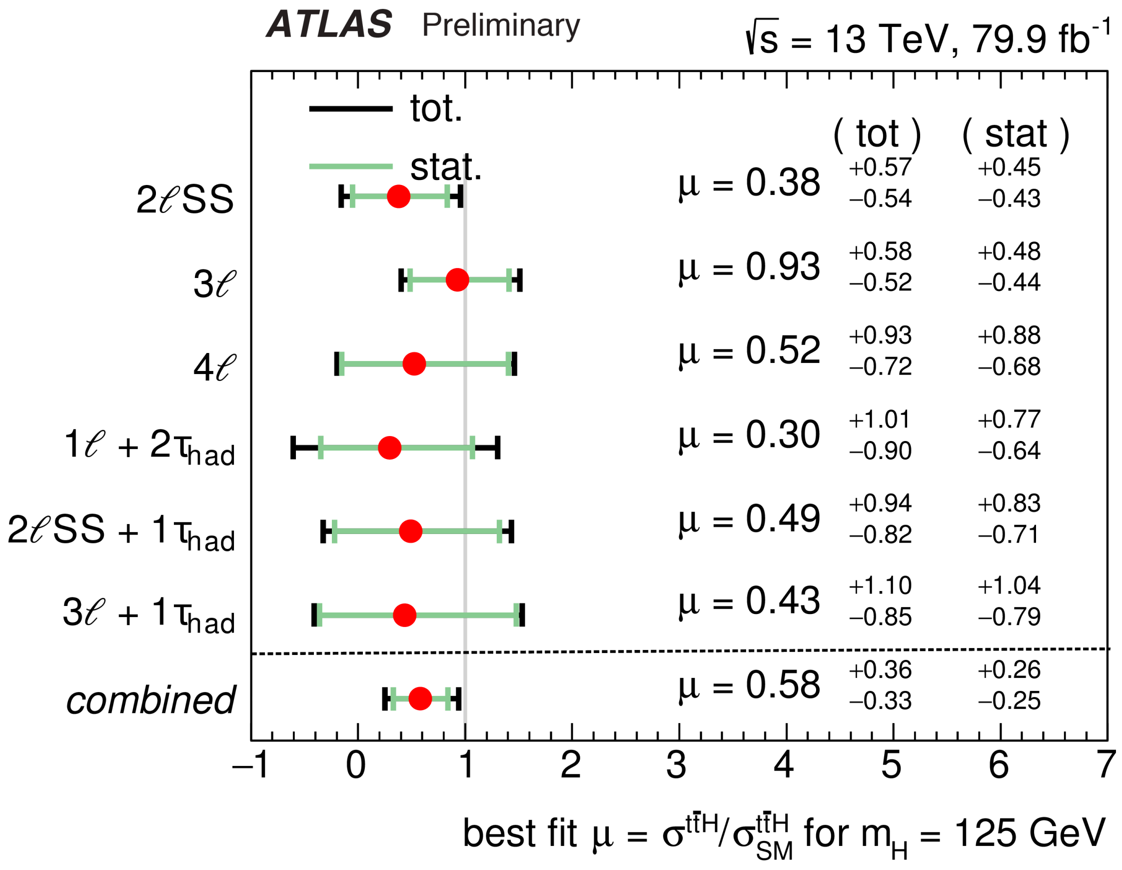
Uncertainty source	$\Delta\hat{\mu}$	
Jet energy scale and resolution	+0.13	−0.13
$t\bar{t}(Z/\gamma^*)$ (high mass) modelling	+0.09	−0.09
$t\bar{t}W$ modelling (radiation, generator, PDF)	+0.08	−0.08
Fake τ_{had} background estimate	+0.07	−0.07
$t\bar{t}W$ modelling (extrapolation)	+0.05	−0.05
$t\bar{t}H$ cross section	+0.05	−0.05
Simulation sample size	+0.05	−0.05
$t\bar{t}H$ modelling	+0.04	−0.04
Other background modelling	+0.04	−0.04
Jet flavour tagging and τ_{had} identification	+0.04	−0.04
Other experimental uncertainties	+0.03	−0.03
Luminosity	+0.03	−0.03
Diboson modelling	+0.01	−0.01
$t\bar{t}\gamma^*$ (low mass) modelling	+0.01	−0.01
Charge misassignment	+0.01	−0.01
Template fit (non-prompt leptons)	+0.01	−0.01
Total systematic uncertainty	+0.25	−0.22
Intrinsic statistical uncertainty	+0.23	−0.22
$t\bar{t}W$ normalisation factors	+0.10	−0.10
Non-prompt leptons normalisation factors (HF, material conversions)	+0.05	−0.05
Total statistical uncertainty	+0.26	−0.25
Total uncertainty	+0.36	−0.33

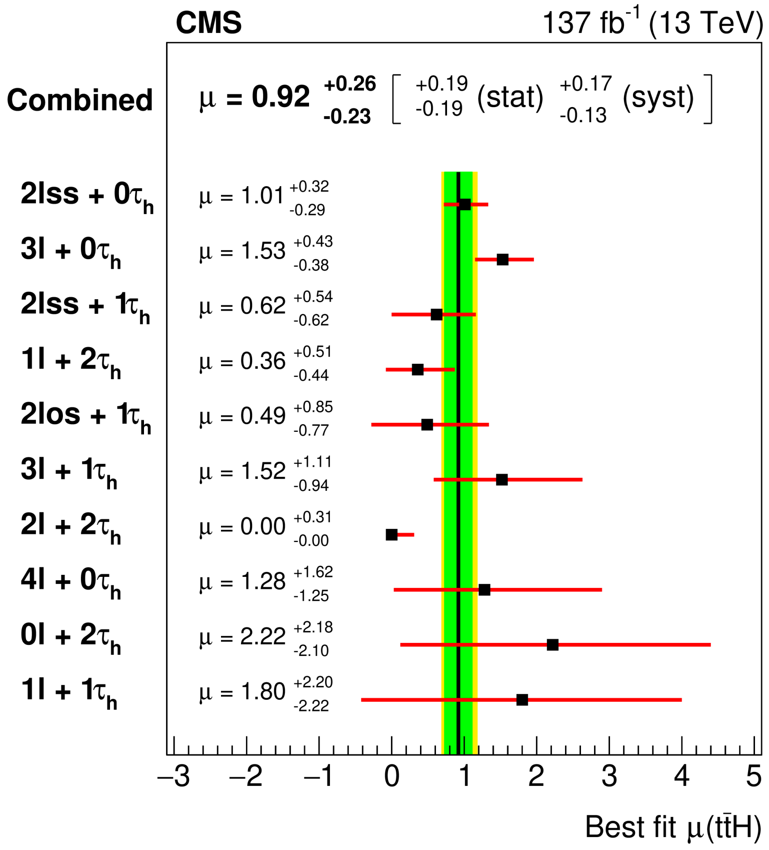
Source	$\Delta\mu_{t\bar{t}H}/\mu_{t\bar{t}H}$ [%]	$\Delta\mu_{tH}/\mu_{tH}$ [%]
Trigger efficiency	2.3	8.1
e, μreconstruction and identification efficiency	2.9	7.1
τ_h identification efficiency	4.6	9.1
b tagging efficiency and mistag rate	3.6	13.6
Misidentified leptons and flips	6.0	36.8
Jet energy scale and resolution	3.4	8.3
MC sample and sideband statistical uncertainty	7.1	27.2
Theory-related sources	4.6	18.2
Normalization of MC-estimated processes	13.3	12.3
Integrated luminosity	2.2	4.6
Statistical uncertainty	20.9	48.0



- Largest **systematic uncertainties** come from $t\bar{t}W$ and $t\bar{t}\ell\ell$ modelling
  - Additional uncertainties to cover data/MC disagreements as a function of NBjets and Lepton charge for $t\bar{t}W$
- Fakes impact is reducing its size with more statistics!
- Non-prompt leptons + QMisID uncertainties large impact on tH

$t\bar{t}H$ (multiℓ): fit results





- Combined $\mu_{t\bar{t}H}$ measured with **2015, 2016, 2017 and 2018** dataset:

	Main result ATLAS EXPERIMENT	**Alternative fit**	CMS
$\mu_{t\bar{t}H}$	**0.58 $^{+0.36}_{-0.33}$**	**0.70 $^{+0.36}_{-0.33}$**	**0.92 $^{+0.26}_{-0.23}$**
NF($t\bar{t}W$) (to compare with CMS take ~1.1xATLAS)	1.56 $^{+0.30}_{-0.28}$ **(2ℓ LNJ)** 1.26 $^{+0.19}_{-0.18}$ **(2ℓ HNJ)** 1.68 $^{+0.30}_{-0.28}$ **(3ℓ)**	1.39 $^{+0.17}_{-0.16}$ *[SM ref: 727 fb]*	1.43 ±0.21 *[SM ref: 650 fb]*
NF($t\bar{t}Z$)	-		1.03 ±0.14

Significance with respect to background-only hypothesis = **1.8 (3.1σ)** and **4.7σ (5.2σ)** obs (exp)

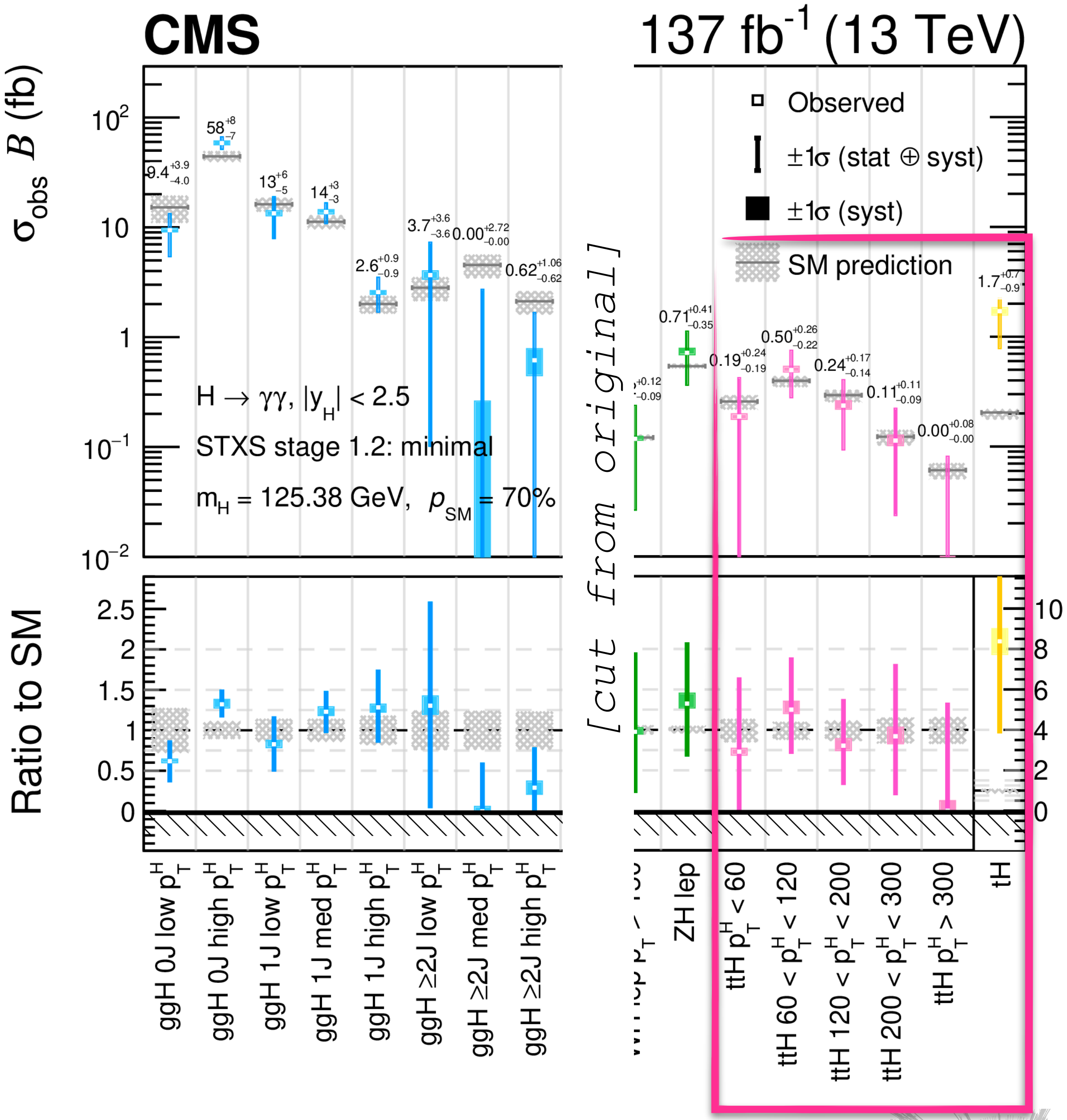
Compatibility between main and alternative fit = 0.59 σ

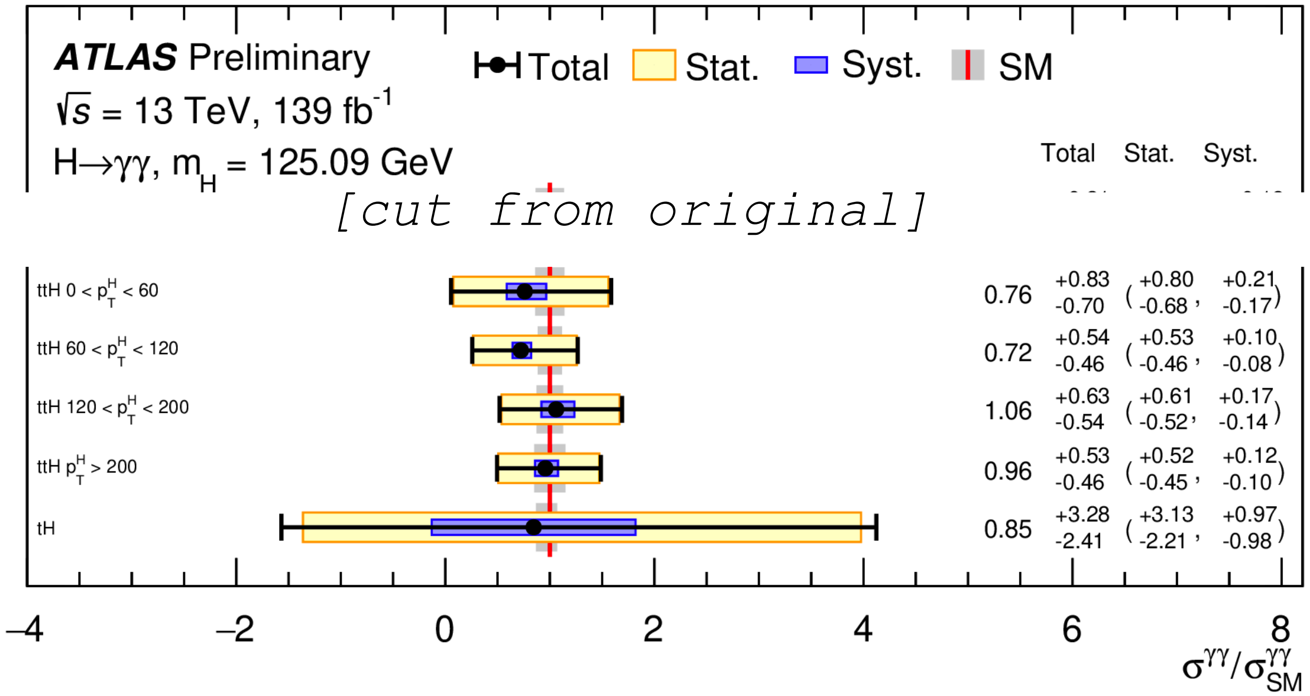
$t\bar{t}W$ measured consistently higher than SM in both experiments!

tt̄H(H→γγ): STXS

- First channel to perform tt̄H measurement differentially
- Leptonic (tt̄H & tH) and hadronic channels (tt̄H & tH)
- Mixture of **multiclass BDT** (STXS signal vs other signals) and **binary BDTs** (STXS signal vs background)
- Mixture of **Top DNN** (tt̄H vs tH) and **BDT** (STXS signal vs non-Higgs SM background), and final classification based on **reco $p_T(\gamma\gamma)$**

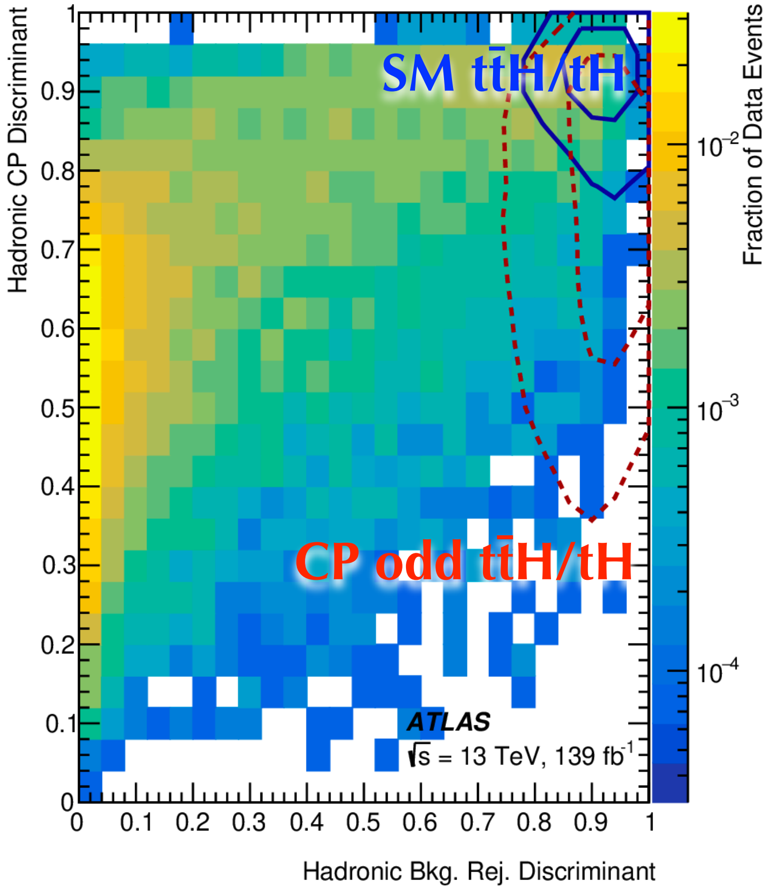
- Dominated by stat uncertainty but overall compatible with SM predictions

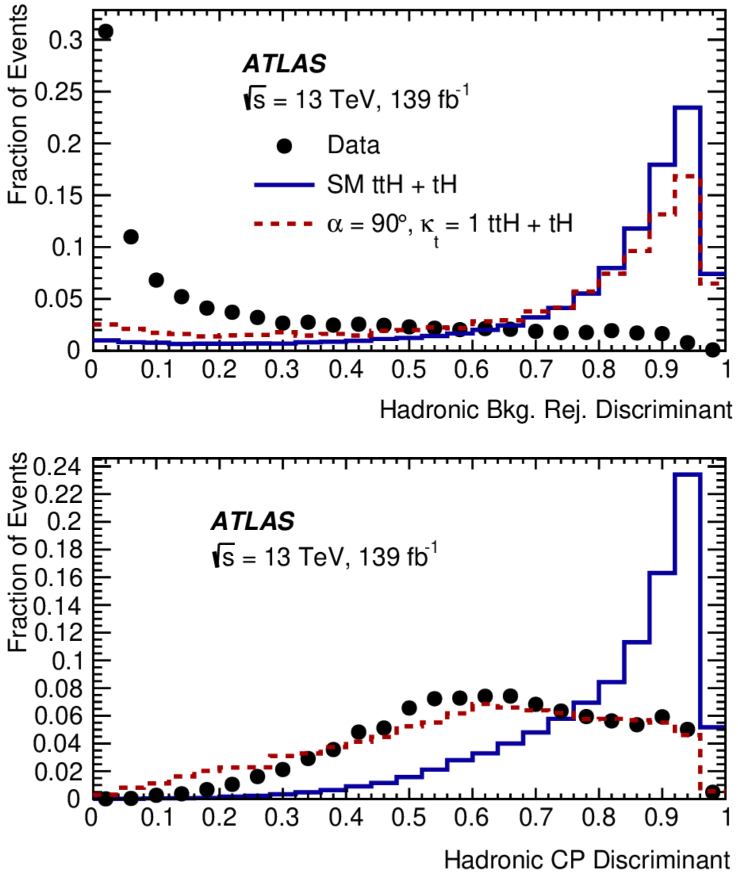


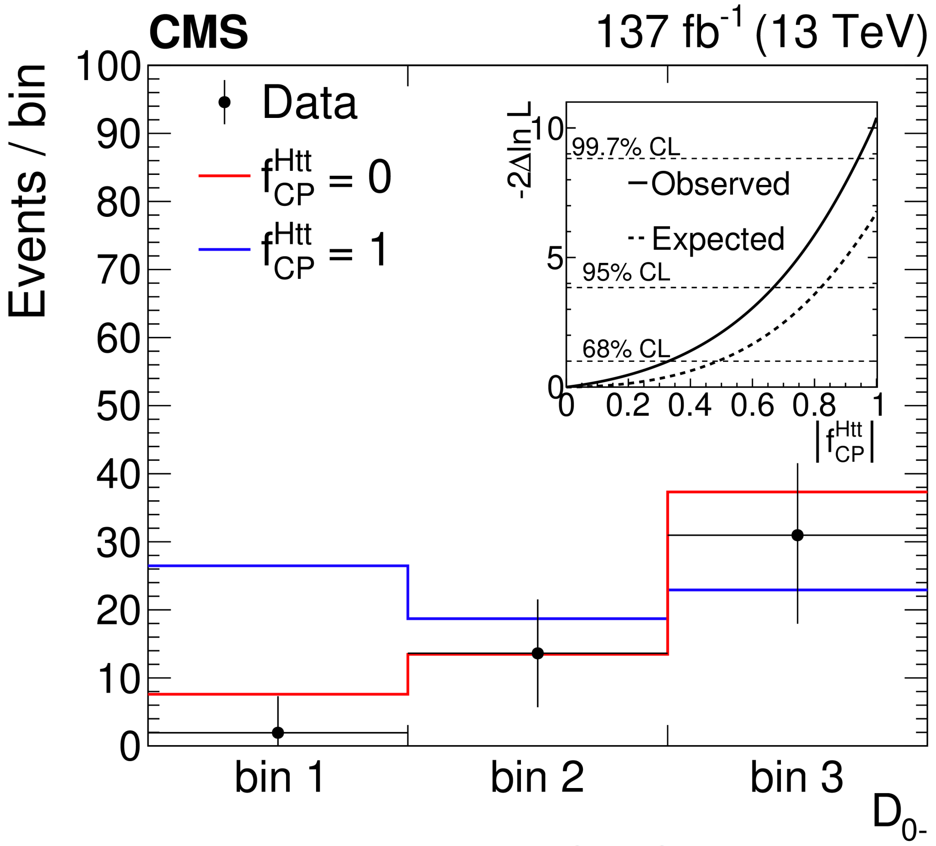


tt̄H(H➔γγ): CP analysis

- 2D partitioning /categorisation using BDT-bkg and BDT-CP (D_{0-})
  - 20 (12had + 8lep) vs 12 (6had + 6lep) categories
- Constrains: observed (expected under CP-even hypothesis)
  - $|\alpha^{CP}| < 43°$ (63°) @ 95 CL; $\alpha=90°$ excluded at 3.9σ
  - $|f_{CP}| < 0.67 \Rightarrow |\alpha^{CP}| < 55°$ (66°) @ 95 CL; $\alpha=90°$ excluded at 3.2σ







$$D_{0-} = \frac{P(odd)}{P(odd)+P(even)}$$

$t\bar{t}H(H\rightarrow ZZ\rightarrow 4\ell)$ resonant

- Higgs boson candidates: $115 < m(4\ell) < 130$ GeV
- Both analyses use NN-based categorisation either to define the categories or as observable to fit
- **$\mu_{t\bar{t}H} = 1.7^{+1.7}_{-1.2}$ (stat) ±0.2 (exp) ±0.2 (th)** and **$\mu_{t\bar{t}H} = 0.17^{+0.88}_{-0.17}$ (stat) $^{+0.42}_{-0.00}$ (syst)**
- Also computed the Stage 0/1.1/1.2(merged) STXS cross-sections (1 bin for $t\bar{t}H$)
  - Largely **statistically** limited
- Additionally, performed SMEFT fit