BERKELEY LAB

LAWRENCE BERKELEY NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

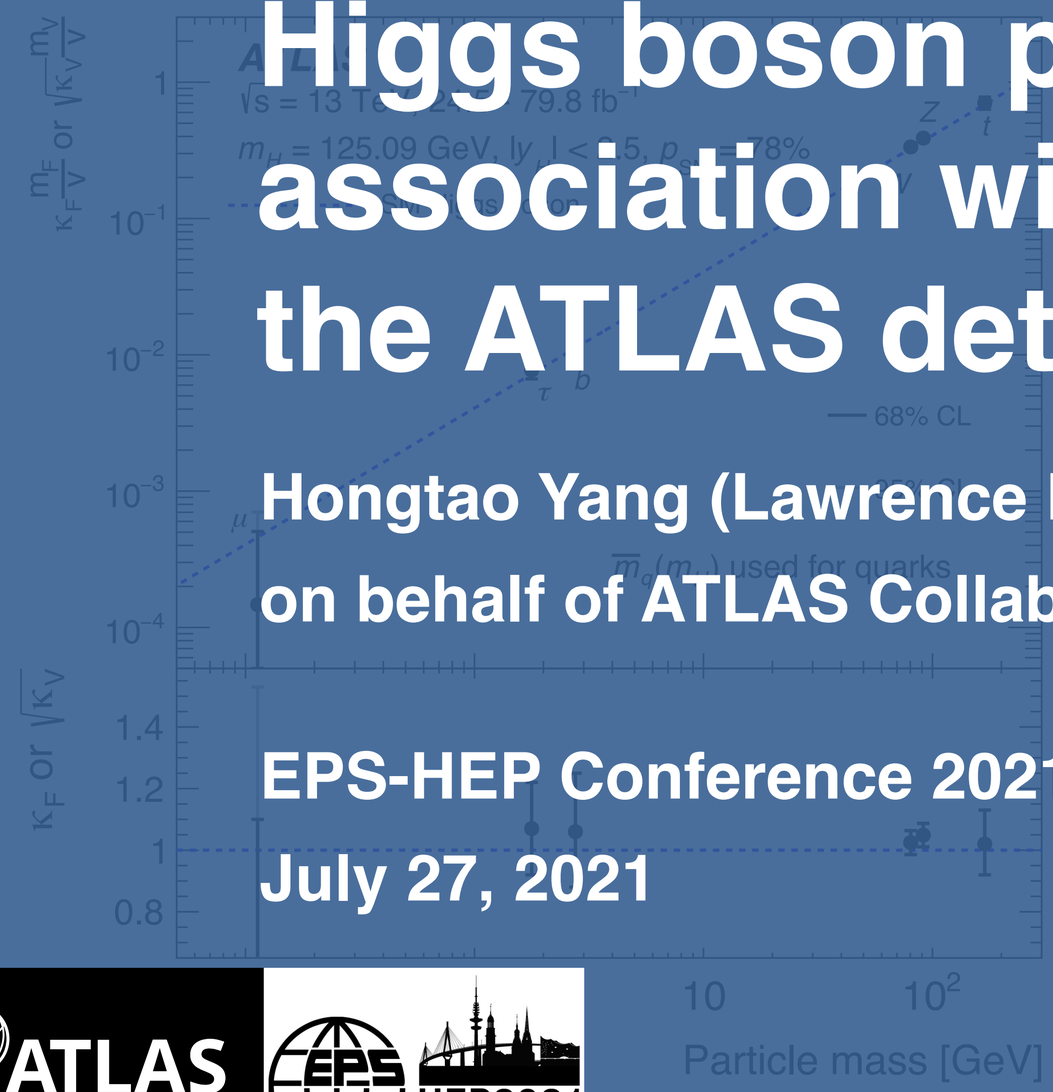
Higgs boson production in association with top quarks with the ATLAS detector

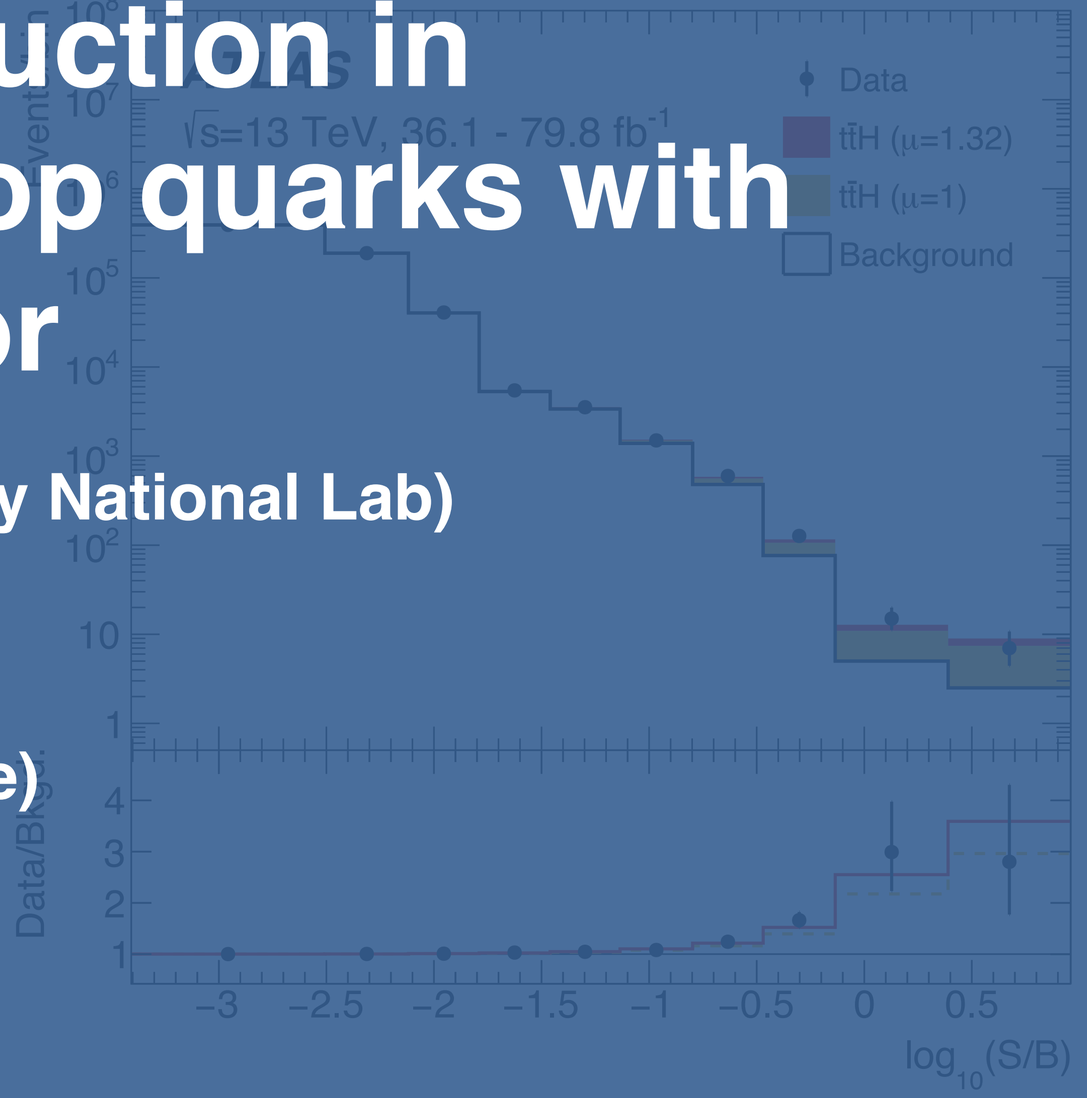
Hongtao Yang (Lawrence Berkeley National Lab)

on behalf of ATLAS Collaboration

EPS-HEP Conference 2021 (online)

July 27, 2021



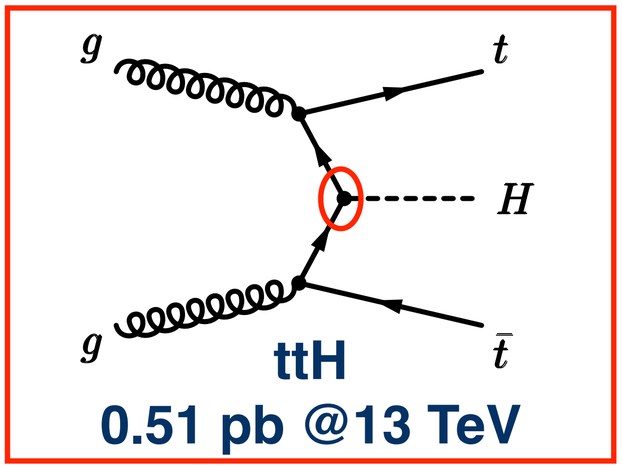


ATLAS EXPERIMENT

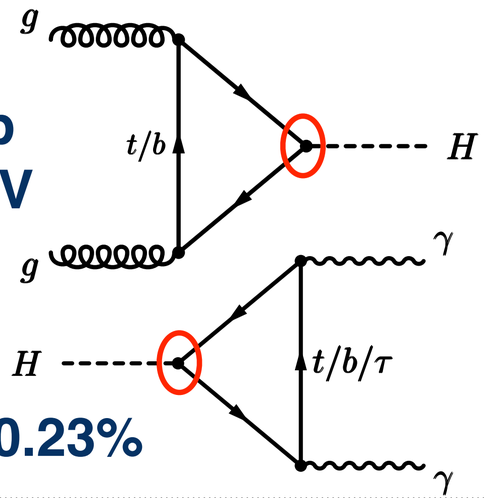
EPS HEP2021 VIRTUAL EDITION

Introduction

- Measurement of **top Yukawa coupling** at LHC is crucial for validating SM and searching for potential new physics
- **ttH/tH production mode** provides a **direct handle** to constrain top Yukawa coupling
  - Indirect constraints from ggF and H→γγ rely on model assumptions

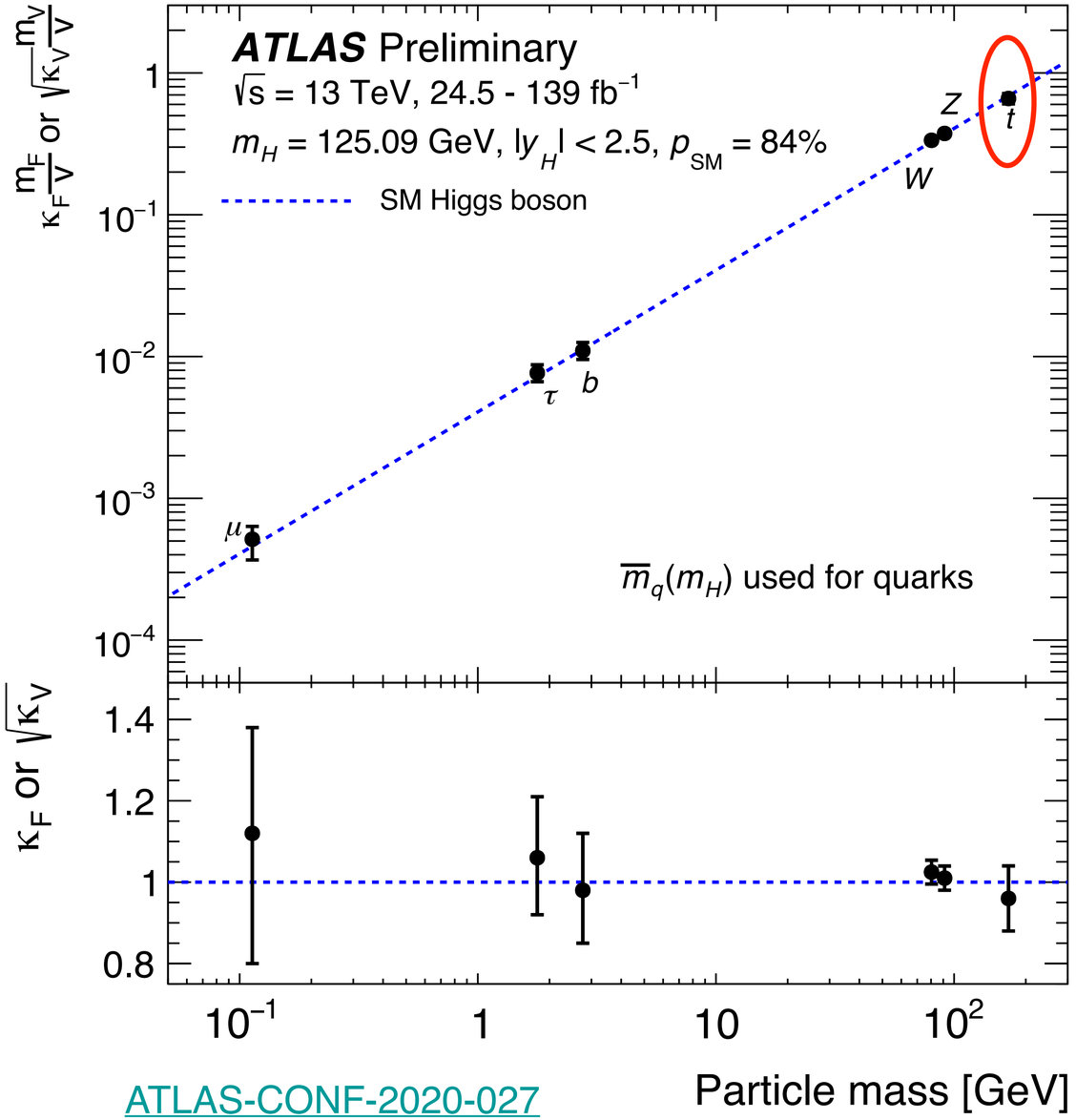


ttH
0.51 pb @13 TeV



ggF
48.5 pb
@13 TeV

H→γγ 0.23%

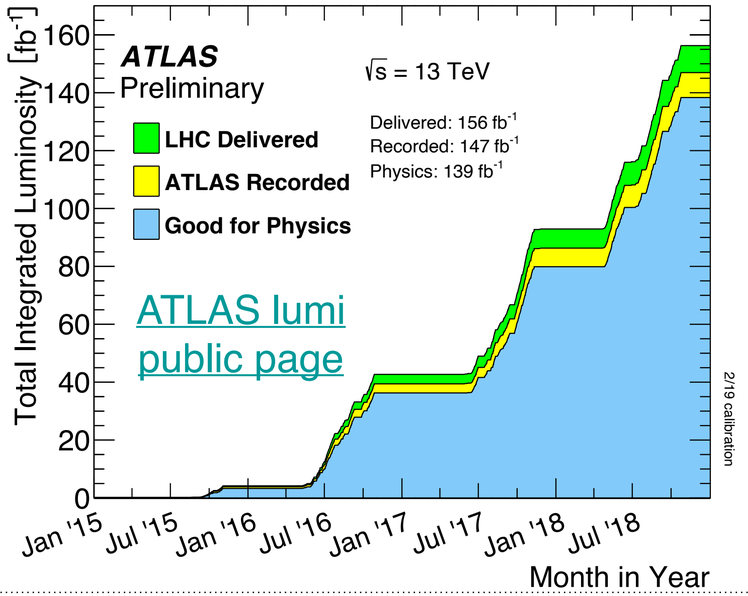


ATLAS-CONF-2020-027

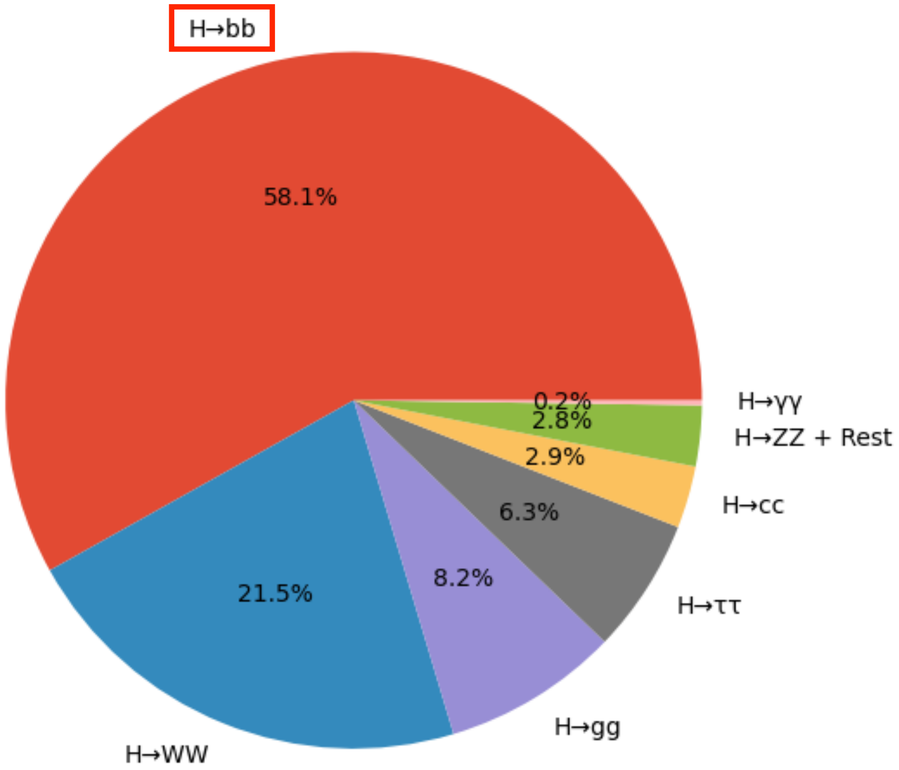
ttH channels

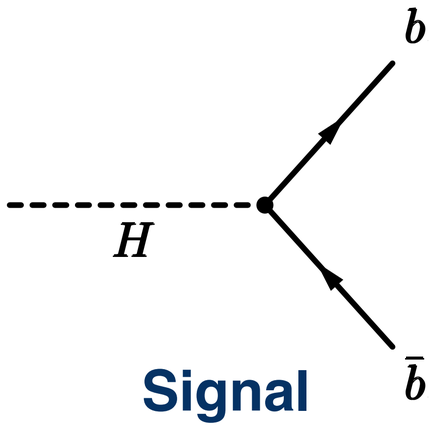
Channel	13 TeV pp dataset	Reference
H→bb	139 fb^{-1} (2015-2018)	ATLAS-CONF-2020-058
H→γγ	139 fb^{-1} (2015-2018)	ATLAS-CONF-2020-026
Combination	Up to 139 fb^{-1}	ATLAS-CONF-2020-027
H→ZZ→4l	139 fb^{-1} (2015-2018)	EPJC 80 (2020) 957
Multi-lepton (H→WW, ZZ, ττ)	80 fb^{-1} (2015-2017)	ATLAS-CONF-2019-045
H→μμ	139 fb^{-1} (2015-2018)	PLB 812 (2021) 135980
H→ττ	139 fb^{-1} (2015-2018)	M. Mlynarikova's talk

- ttH production mode studied in all the major Higgs boson decay channels at LHC
- Today we will focus on H→bb, γγ, as well as multi-channel combination (including H→ZZ→4l) based on **full Run 2 pp dataset collected at 13 TeV!**



ttH, H→bb





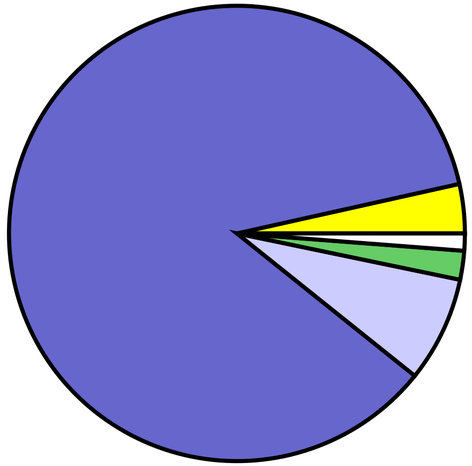
Signal

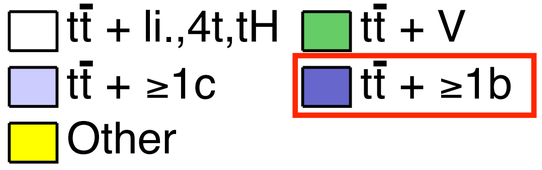
ATLAS Preliminary

$\sqrt{s}$ = 13 TeV

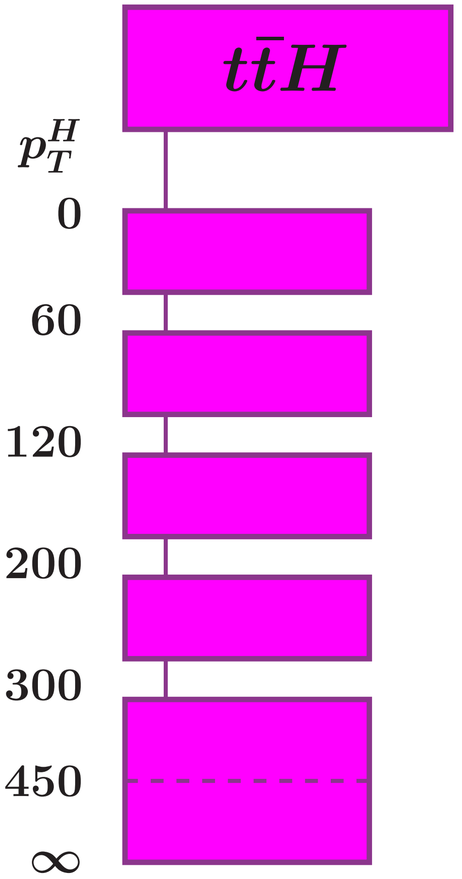
Dilepton ATLAS-CONF-2020-058

$SR^{\geq 4j}_{\geq 4b}$, $p_T^H \in [0,120)$ GeV





Background

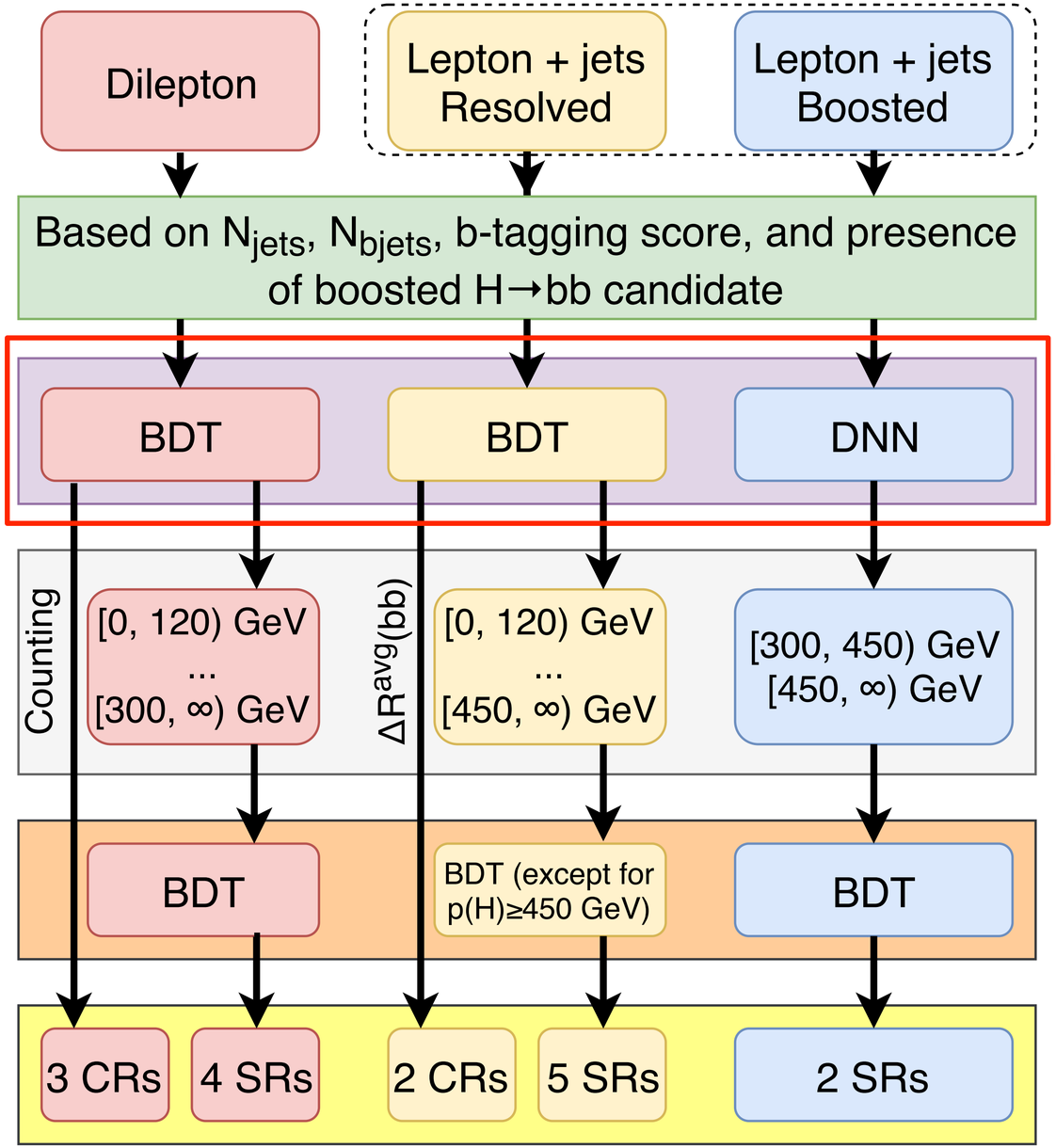


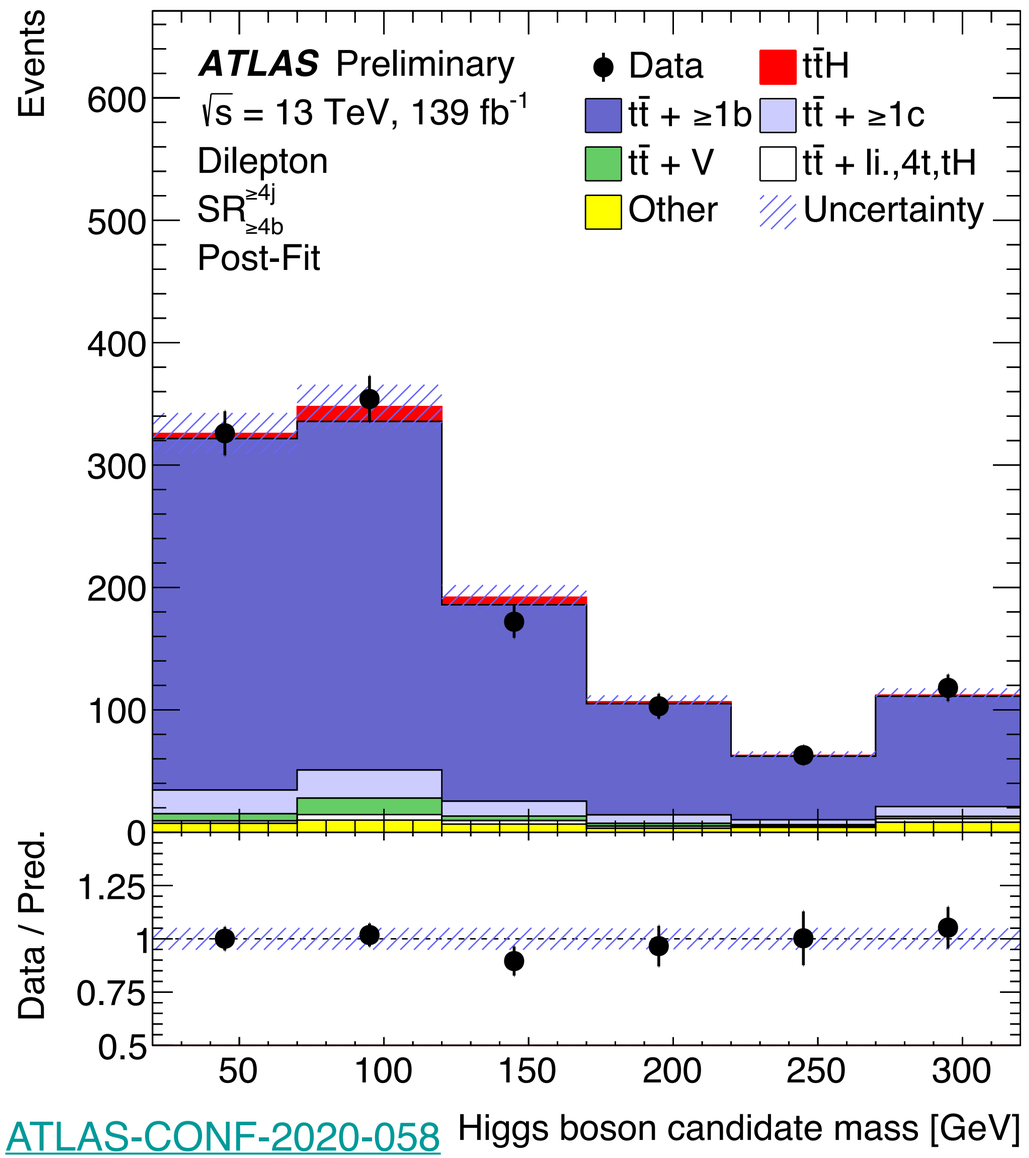
Simplified Template Cross-Section (STXS)

Les Houches 2019

[arXiv:2003.01700 [hep-ph]](arXiv:2003.01700 [hep-ph])

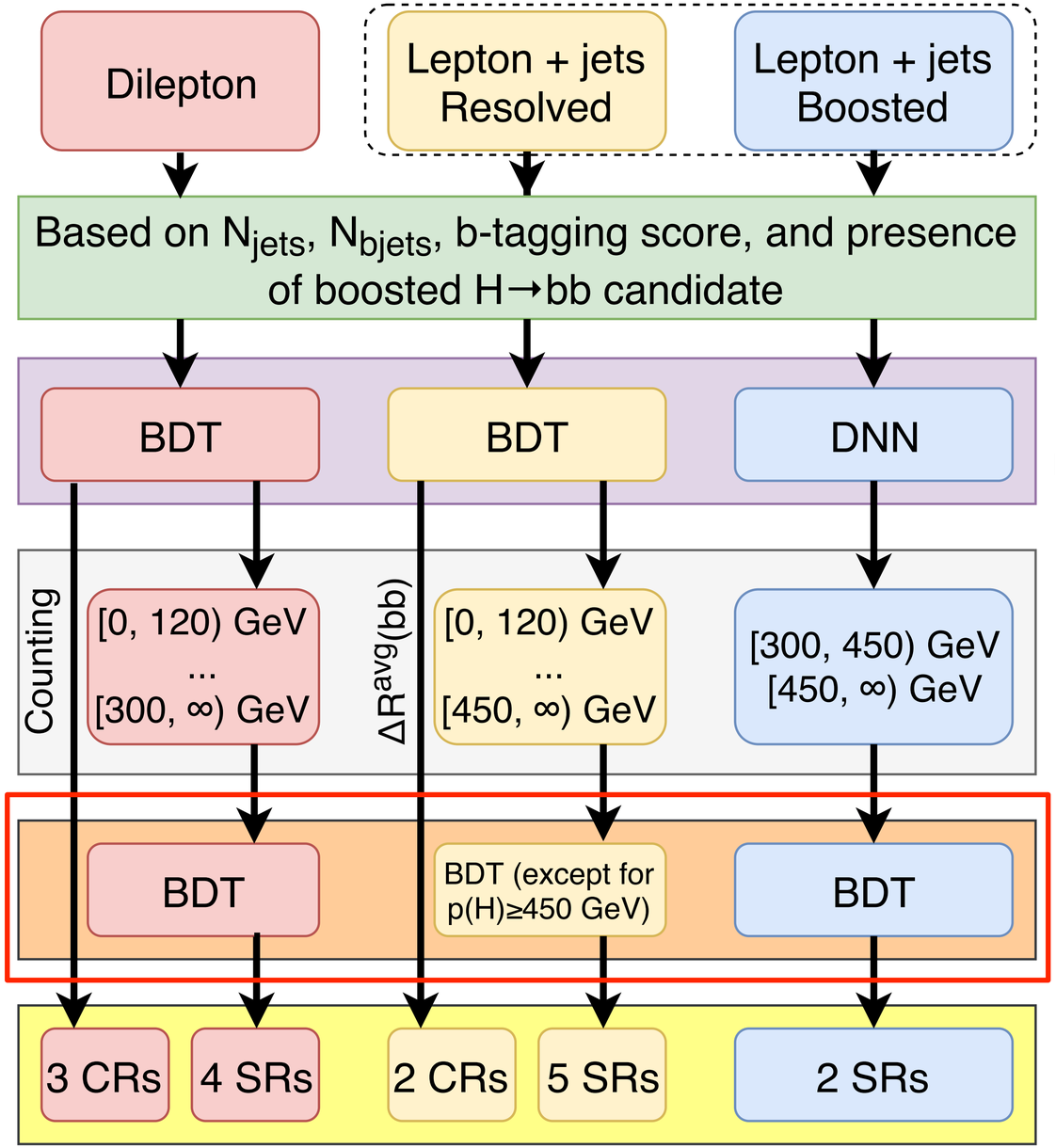
ttH(bb) analysis strategy





ATLAS-CONF-2020-058

ttH(bb) analysis strategy (cont'd)



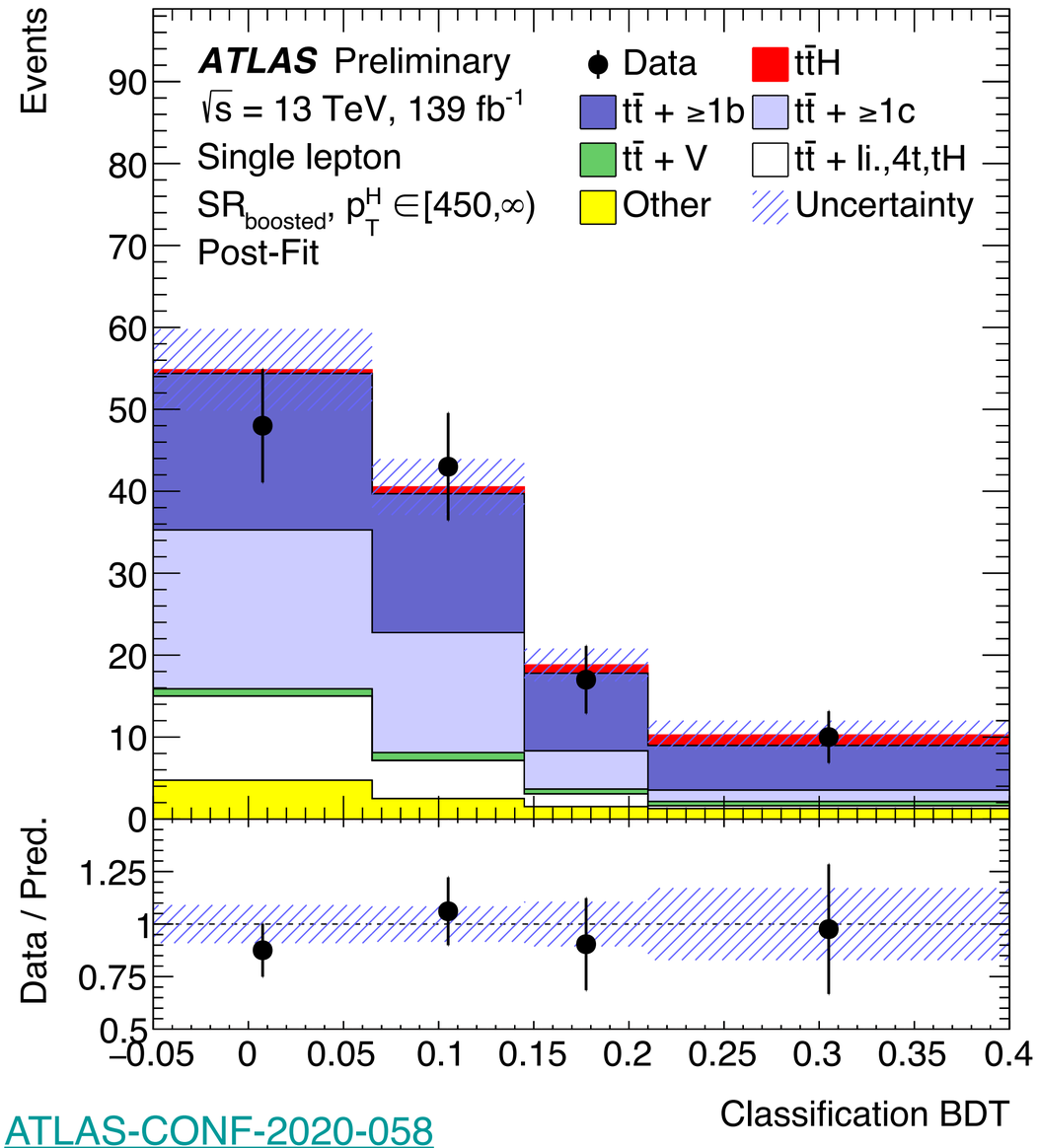
Define signal & control regions

Event reconstruction

Divide in $p_T(H)$ bins according to STXS framework

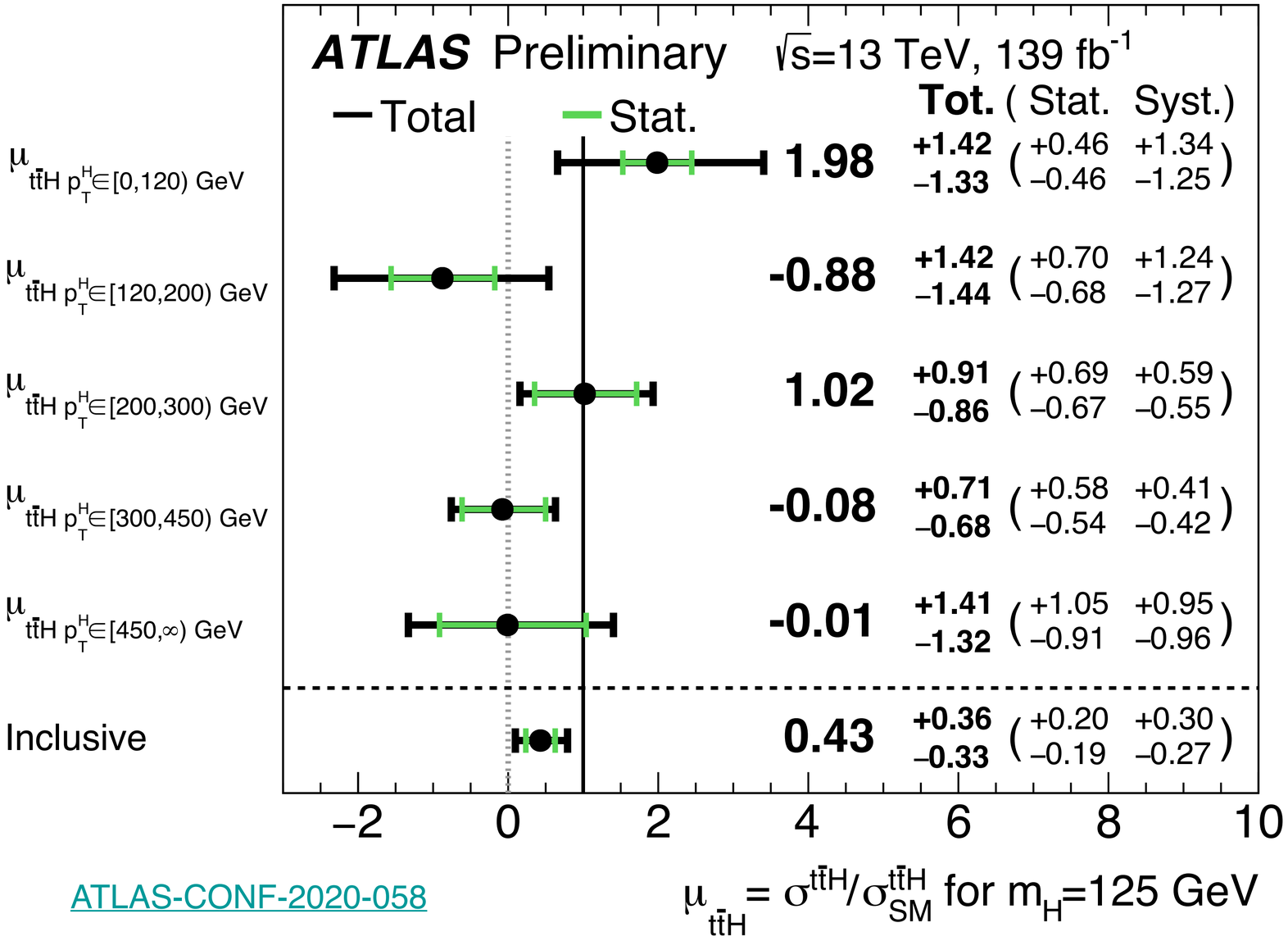
Event classification

Final categories



ATLAS-CONF-2020-058

ttH(bb) results

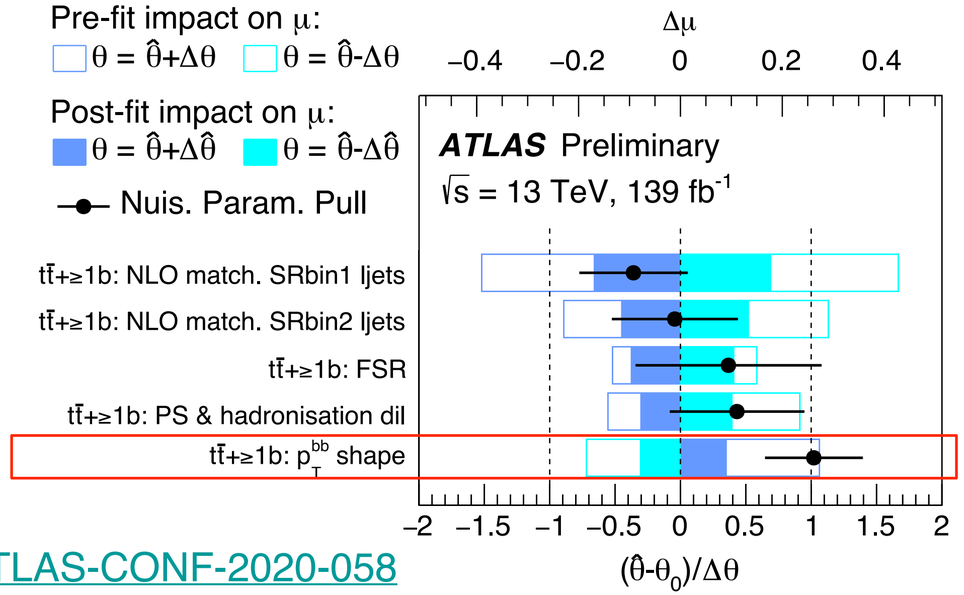


ATLAS-CONF-2020-058

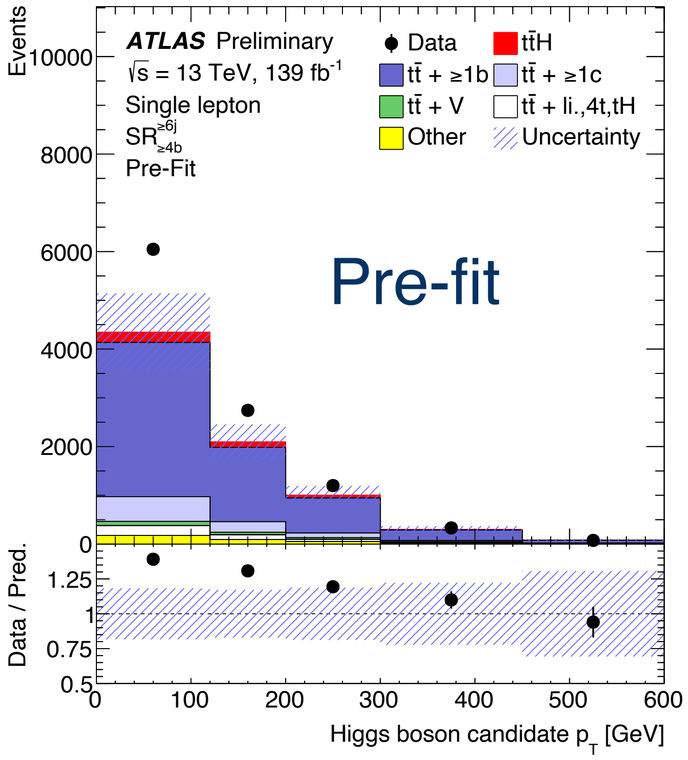
- Observed significance 1.3σ (expected 3.0σ)
- **First STXS measurement in ttH(bb) channel!**
- Inclusive signal strength measurement dominated by systematic uncertainty
  - Uncertainty reduced by almost factor of two compared with analysis based on 36 fb^{-1} partial Run 2 dataset (PRD 97 (2018) 072016)

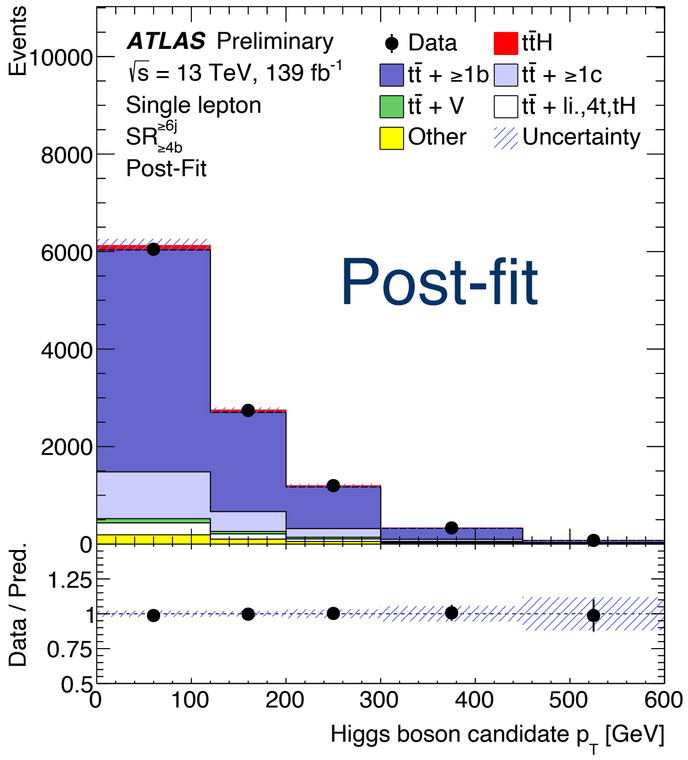
ttH(bb) systematic uncertainties

Uncertainty source	$\Delta\mu$	
$t\bar{t}+\geq 1b$ modelling	+0.25	−0.24
$t\bar{t}H$ modelling	+0.14	−0.06
tW modelling	+0.08	−0.08
b-tagging efficiency and mis-tag rates	+0.05	−0.05
Background-model statistical uncertainty	+0.05	−0.05
Jet energy scale and resolution	+0.03	−0.03
$t\bar{t}+\geq 1c$ modelling	+0.03	−0.03
$t\bar{t}+$ light modelling	+0.02	−0.02
Luminosity	+0.01	−0.00
Other sources	+0.03	−0.03



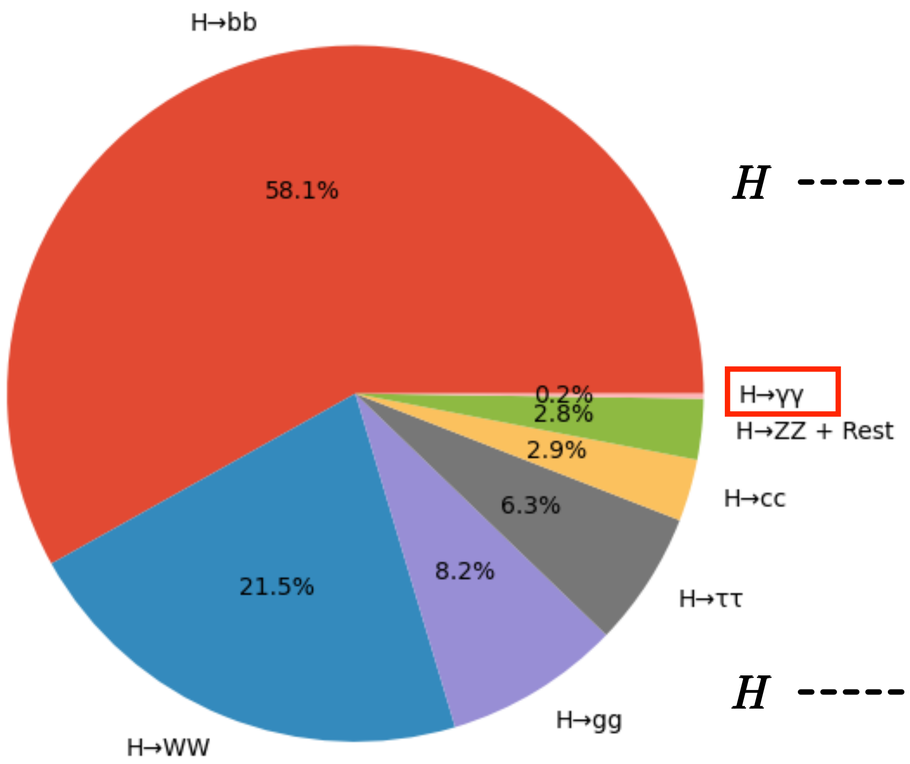
ATLAS-CONF-2020-058

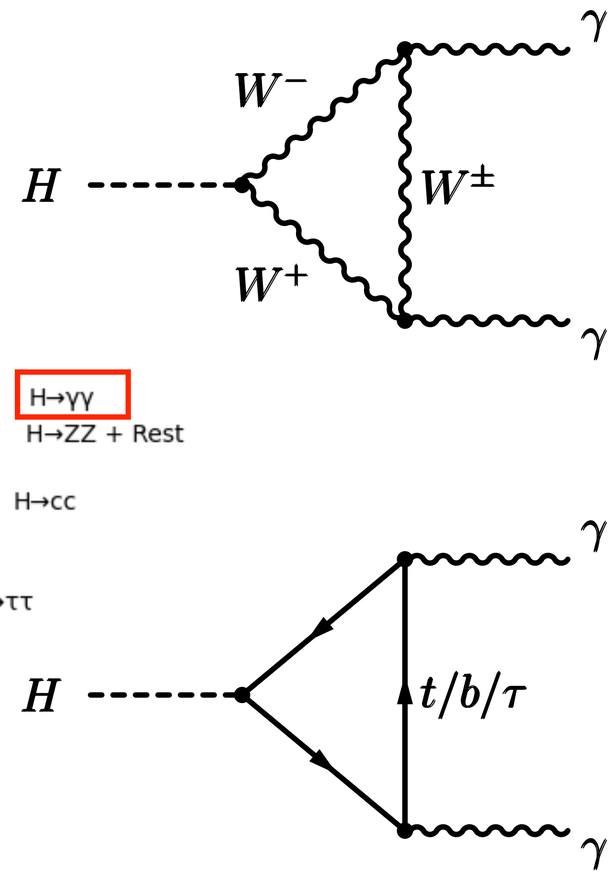




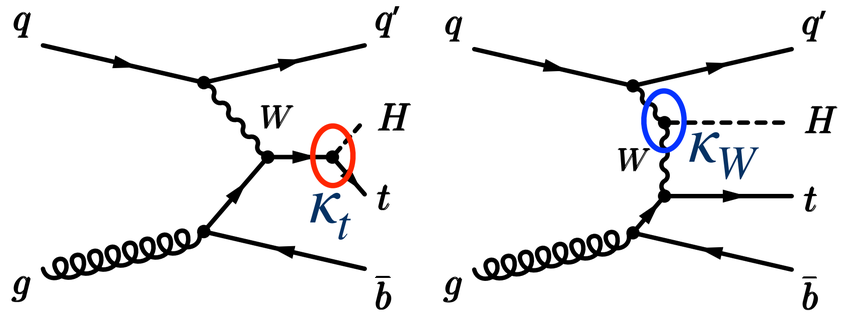
- $t\bar{t}+\geq 1b$ process modeled with **PowhegBoxRes & OpenLoops + Pythia8** in **4-flavor scheme**: $\mu_R = \sqrt[4]{m_T(t)\cdot m_T(\bar{t})\cdot m_T(b)\cdot m_T(\bar{b})}$ and $\mu_F = 0.5\times\sum_{i=t,\bar{t},b,\bar{b},j} m_T(i)$
  - Normalization decided as a free parameter by data: $k(t\bar{t}+\geq 1b) = 1.26\pm 0.09$
- Leading uncertainties from two-point $t\bar{t}+\geq 1b$ PowhegBox + Pythia8 vs. MG5_aMC + Pythia8 NLO matching (uncorrelated between p_T(H) bins)
- Additional uncertainty to cover p_T(H) mis-modeling

ttH/tH, H→γγ

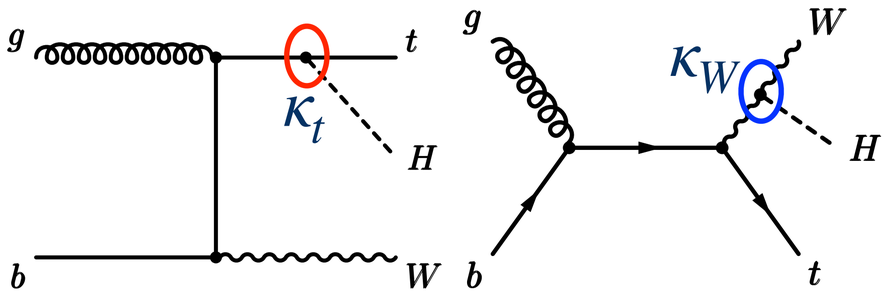




tHqb



tHW



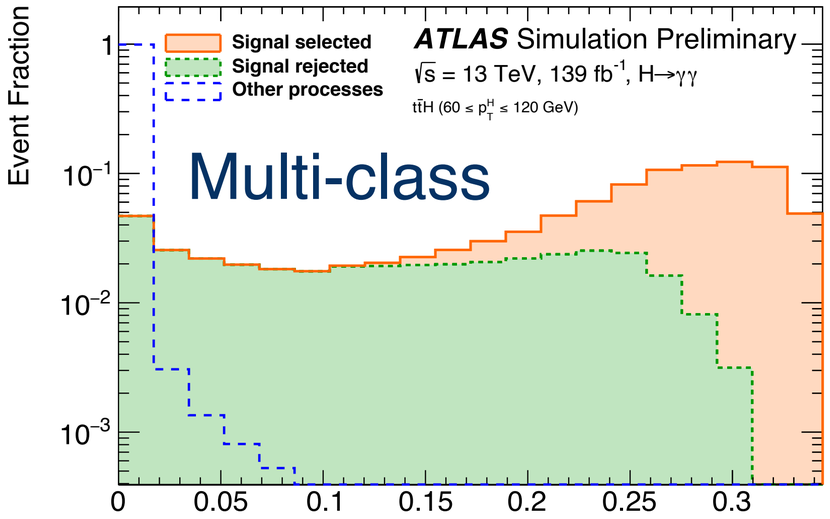
tH

0.09 pb @13 TeV

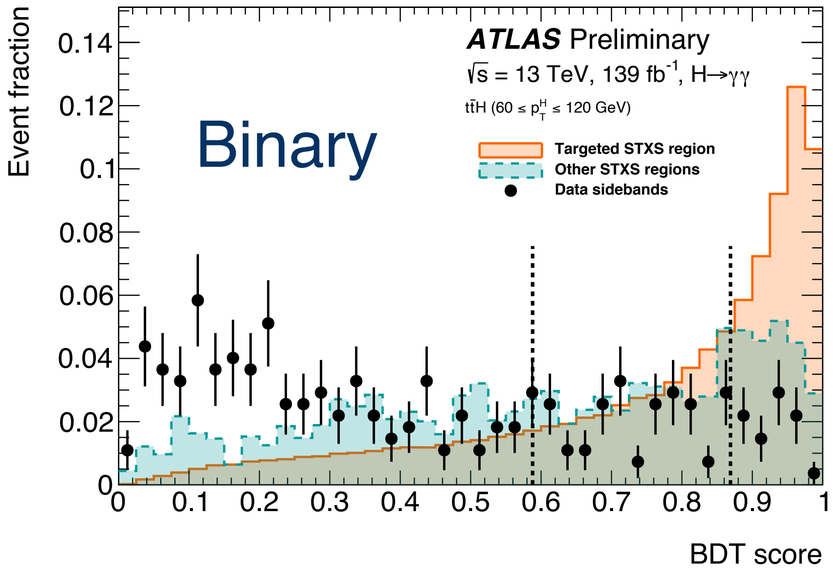
- Excellent sensitivity thanks to good photon energy resolution
- No combinatorics issue in H reconstruction

- Distinguish relative sign of top Yukawa coupling (w.r.t. W-H coupling) using **tH production process**

ttH/tH(γγ) analysis strategy



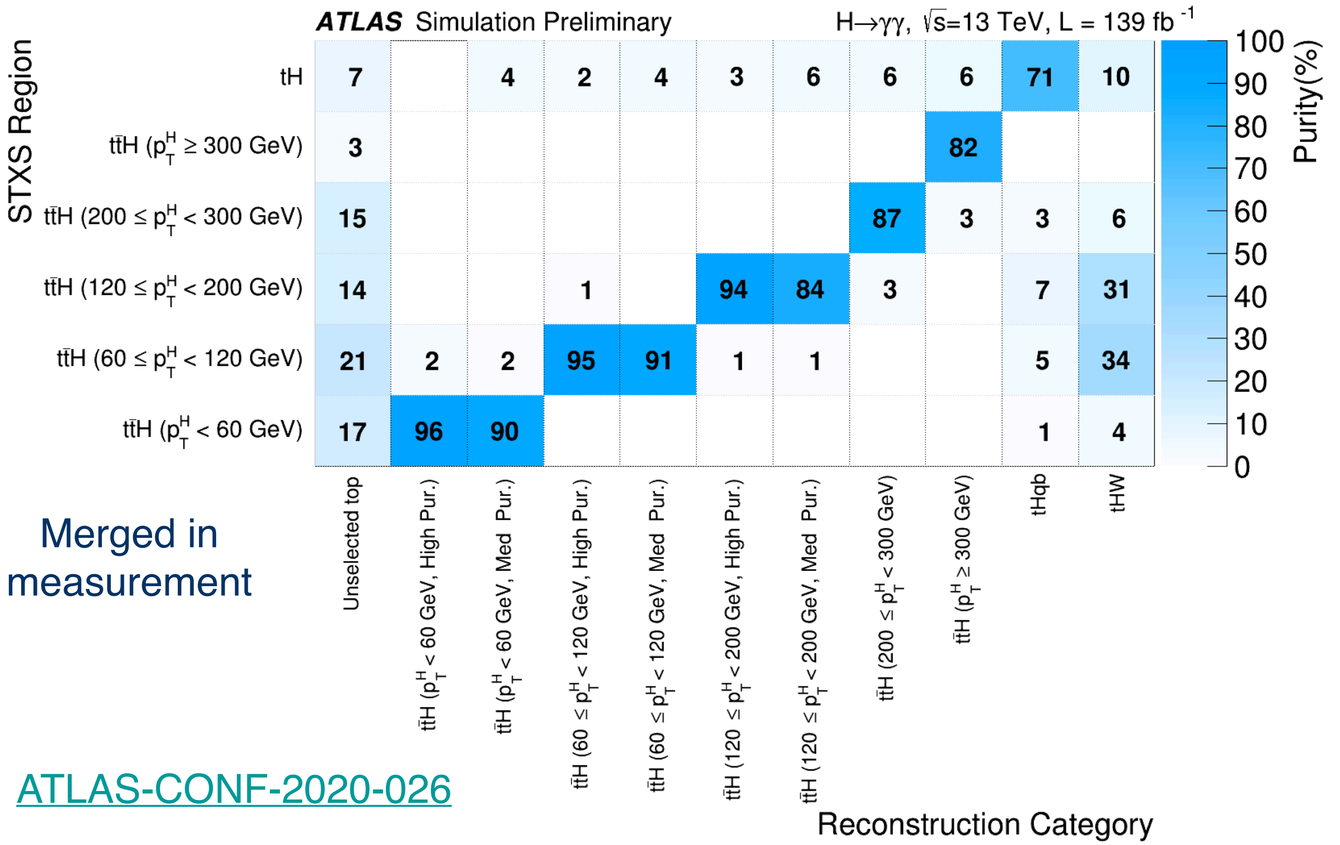
ATLAS-CONF-2020-026



- **Multi-class BDT** trained to divide Higgs boson signal into different STXS bins
  - Use kinematics of diphoton system, N_{jet}, N_{bjet}, N_{lepton}, N_{top}, jet kinematics…
- **Binary BDT** trained to isolate out targeted signal from continuum background
  - Use four-vector of photons/jets/leptons, b-tagging scores of jets, top quark reconstruction BDT etc.
- Define final analysis categories by cutting on Binary BDT. Fit $m_{\gamma\gamma}$ spectrum to extract signal

Categorization

Category	S_{90}	B_{90}	S/(S+B)	Z
$p_T(H) < 60$ GeV High	3.2	5.0	0.39	1.3
$p_T(H) < 60$ GeV Med	3.5	15	0.18	0.8
$60 \le p_T(H) < 120$ GeV High	5.1	4.3	0.54	2.1
$60 \le p_T(H) < 120$ GeV Med	3.7	10	0.26	1.1
$120 \le p_T(H) < 200$ GeV High	6.1	3.8	0.62	2.6
$120 \le p_T(H) < 200$ GeV Med	3.1	8.1	0.28	1.0
$200 \le p_T(H) < 300$ GeV	4.6	1.7	0.73	2.7
$p_T(H) \ge 300$ GeV	3.6	1.0	0.78	2.6
tHW	0.80	2.4	0.25	0.5
tHqb	0.88	2.7	0.24	0.5
Other	11	120	0.08	1.0

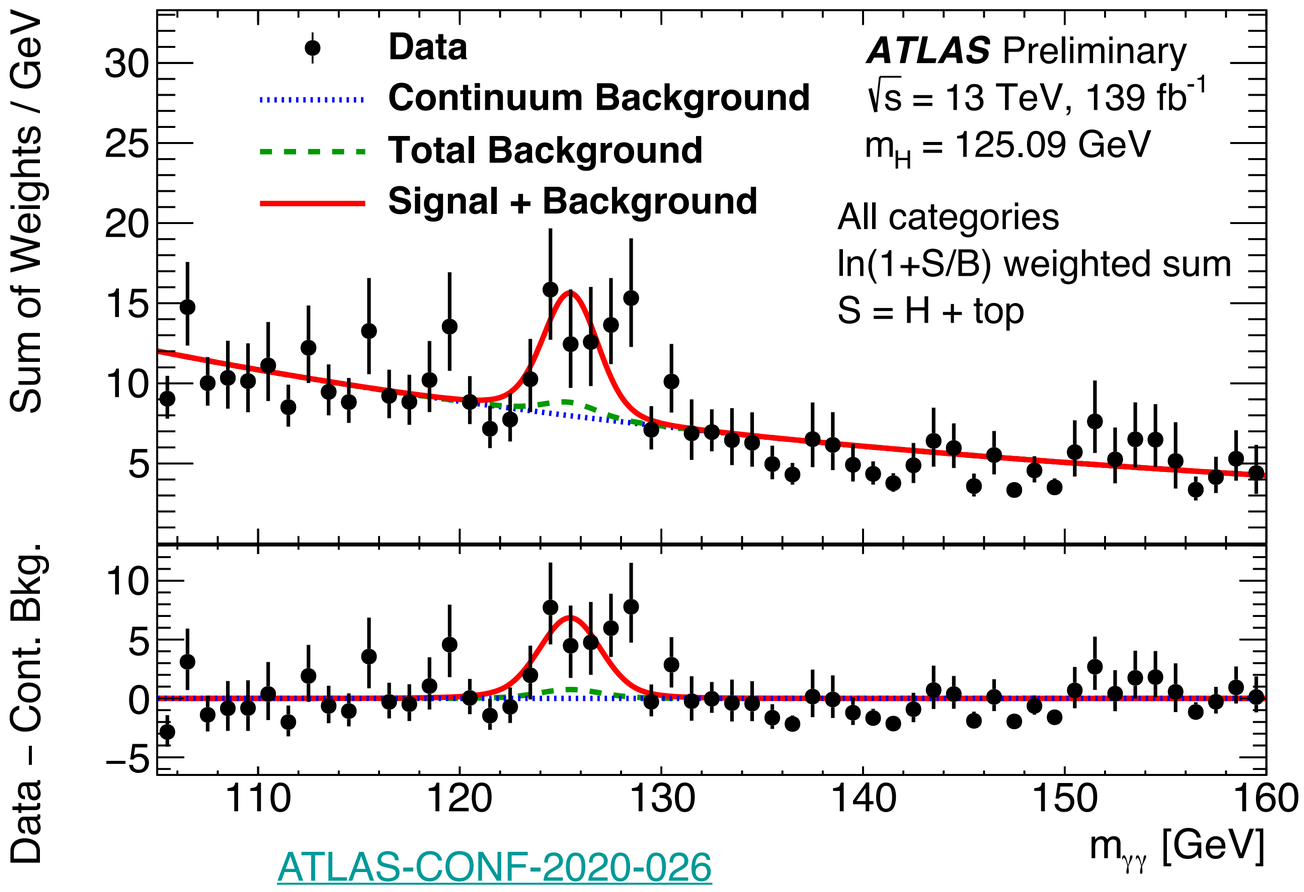


Merged in measurement

ATLAS-CONF-2020-026

- High S/(S+B) and good purity of targeted ttH/tH STXS regions
  - Higgs boson candidate p_T can be reconstructed with high resolution in the diphoton channel
- Sensitivity mostly contributed by high binary BDT categories

ttH/tH(γγ) results



ATLAS-CONF-2020-026

Uncertainty source	Δσ(ttH+tH) [%]
Photon energy resolution	±4.9
Underlying event and parton shower	±3.5
Jet/missing ET	±3.0
Photon efficiency	±2.4
Background modeling	±2.4
Luminosity and trigger	±2.3
Photon energy scale	±2.2
Higgs boson mass	±1.9
Modeling of heavy flavor jets in non-ttH process	±1.3
Pileup	±1.1

- ttH+tH observed significance 4.7σ (exp. 5.0σ)
- $\sigma(ttH+tH)/\sigma_{SM} = 0.92^{+0.25}_{-0.23}\,(\text{Stat.})^{+0.09}_{-0.07}\,(\text{Syst.})$

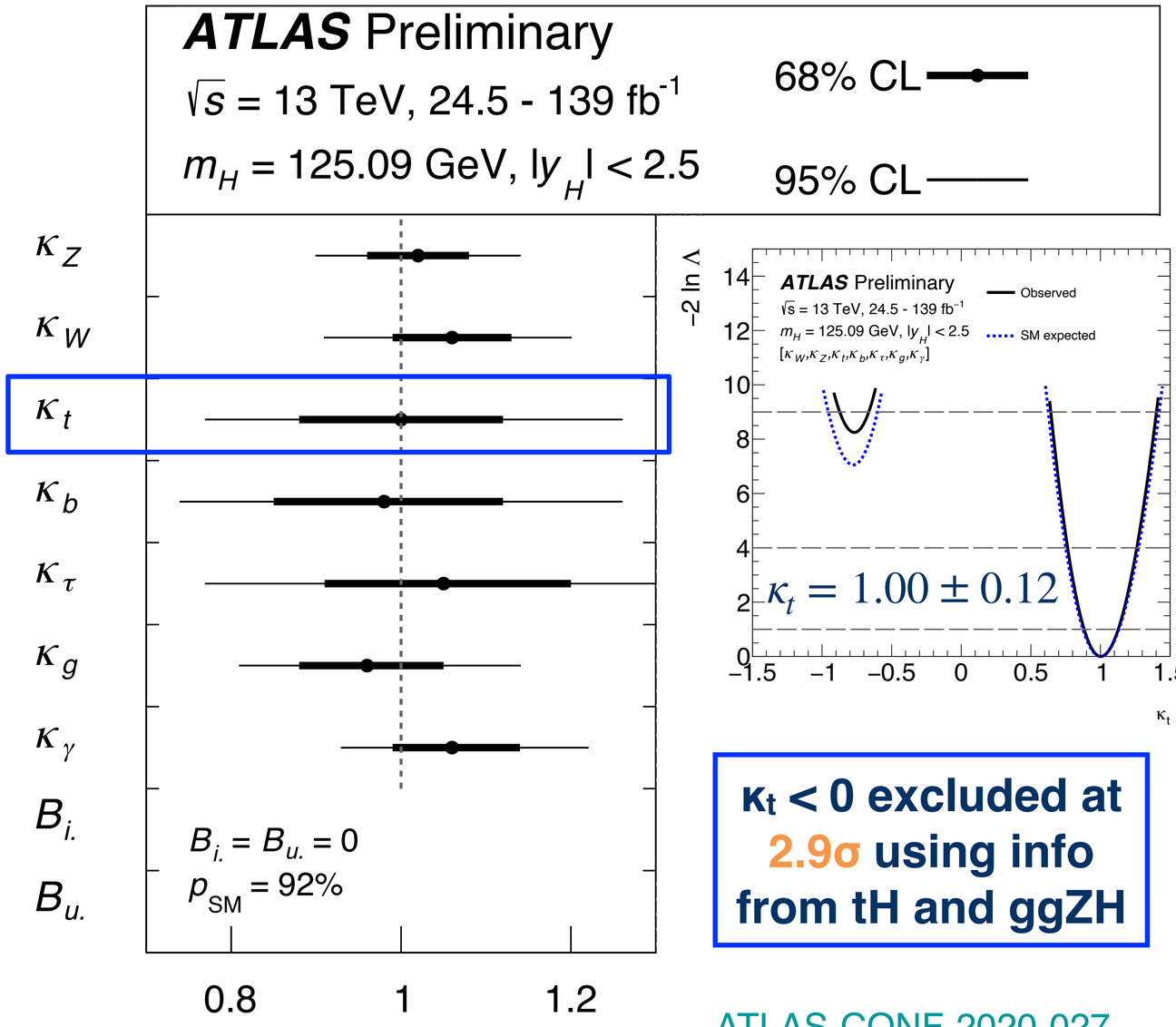
- $\sigma(tH) < 8 \times \sigma_{SM}$ @95% confidence level
- Cross-sections of other production modes decided by data in other analysis categories

Combination of multiple decay channels

Channel	ggF	VBF	VH	ttH
H→γγ (139 fb^{-1})	✔	✔	✔	✔
H→ZZ (139 fb^{-1})	✔	✔	✔	✔
H→WW (36 fb^{-1})	✔	✔		✔
H→ττ (36 fb^{-1})	✔	✔		✔
H→bb (VH 139 fb^{-1}, others 36 fb^{-1})		✔	✔	✔
H→μμ (139 fb^{-1})	✔	✔	✔	✔
H→inv. (139 fb^{-1})		✔		

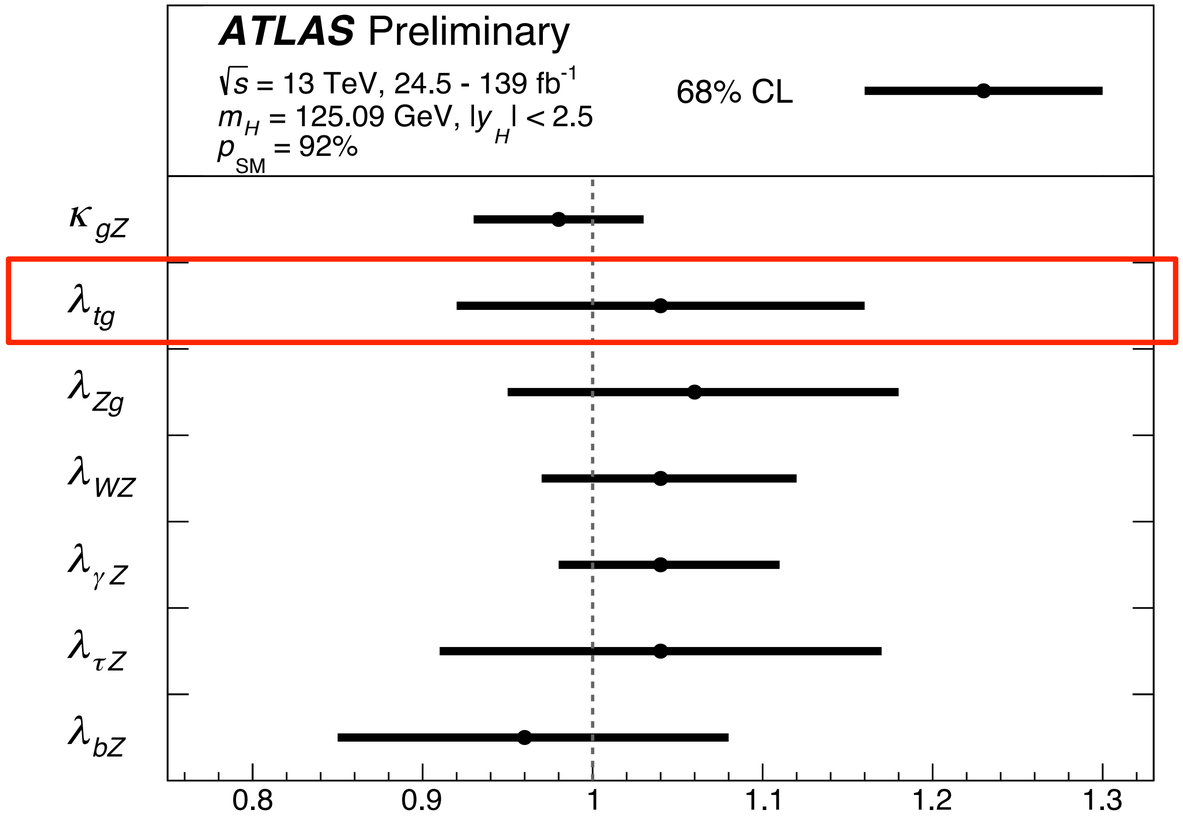
- **ttH, H→γγ** (ATLAS-CONF-2020-026) and **H→ZZ→4l** (EPJC 80, 957 (2020)) measurements based on **139 fb^{-1} full Run 2 dataset** and provide **STXS measurements**
- **ttH, H→bb** (PRD 97 (2018) 072016) and **multi-lepton** (PRD 97 (2018) 072003) measurements in this combination use only 36 fb^{-1} partial Run 2 dataset, and provide only inclusive measurements

Test of top Yukawa coupling strength

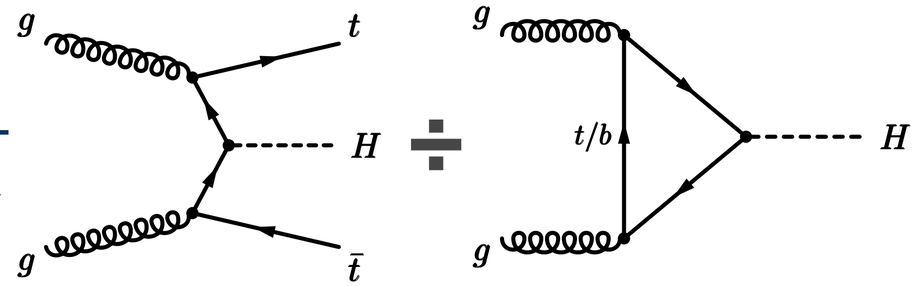


$\kappa_t < 0$ excluded at 2.9σ using info from tH and ggZH

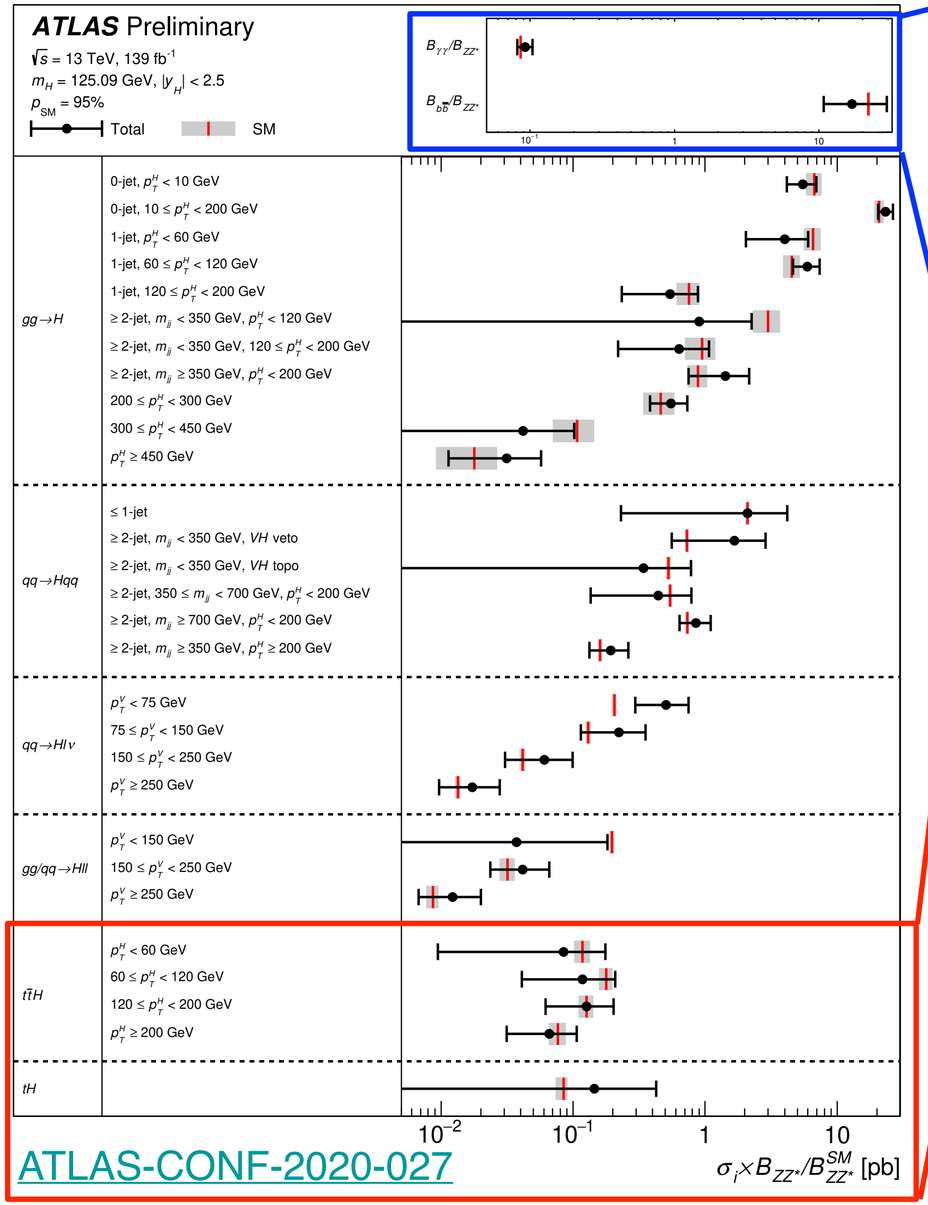
ATLAS-CONF-2020-027



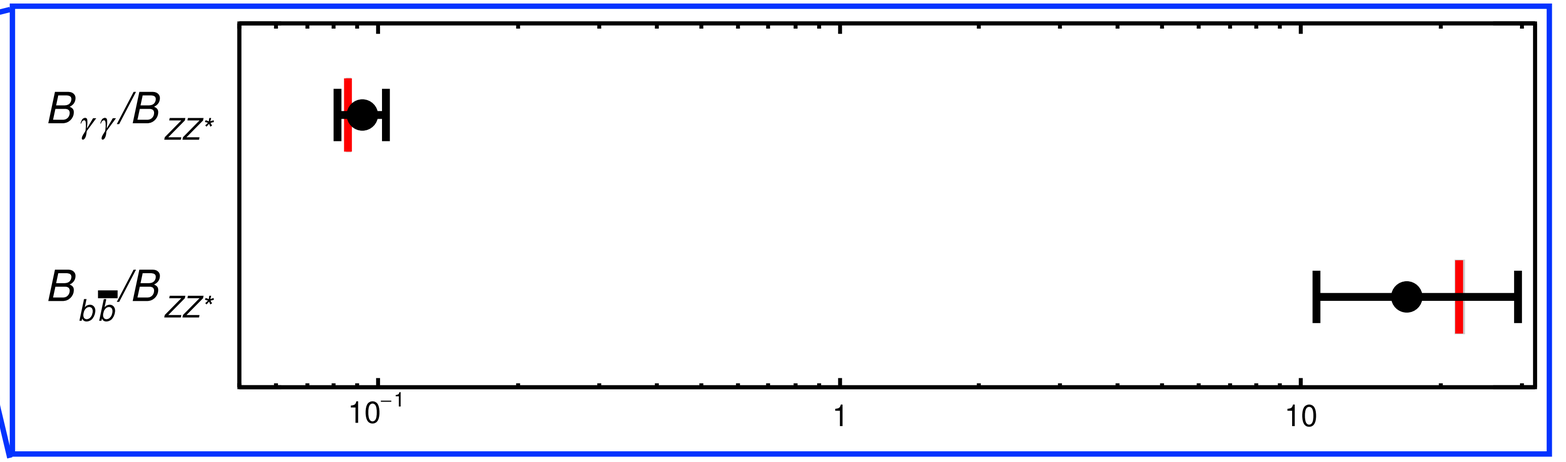
$$\lambda_{tg} = \frac{\kappa_t}{\kappa_g}$$

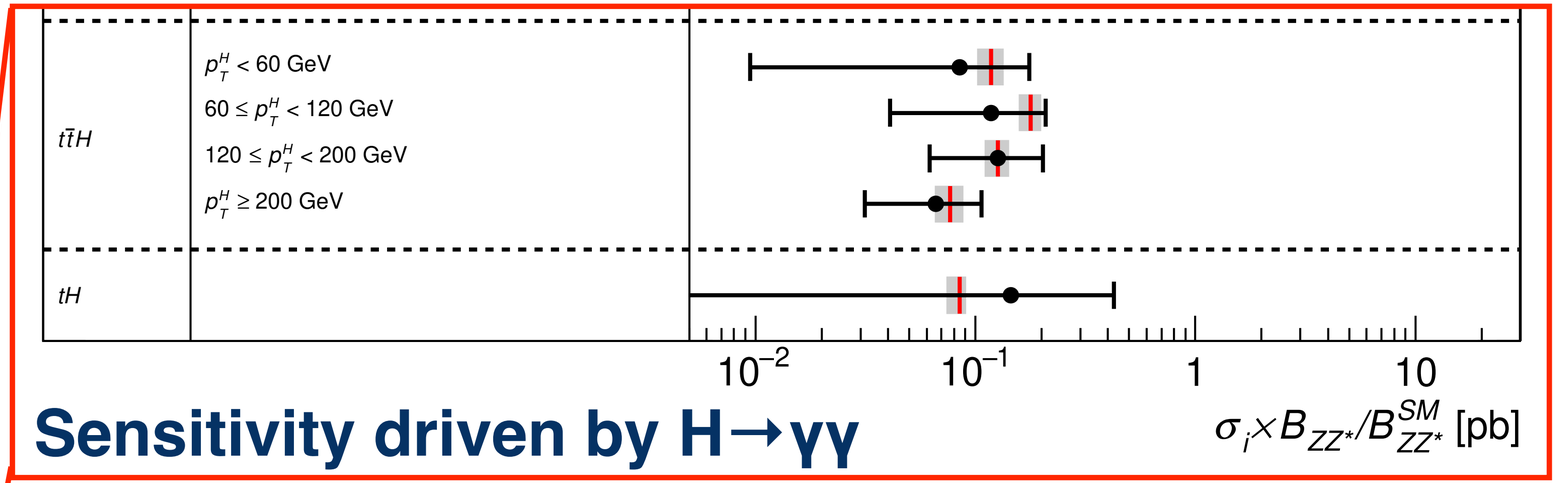


STXS measurements



ATLAS-CONF-2020-027



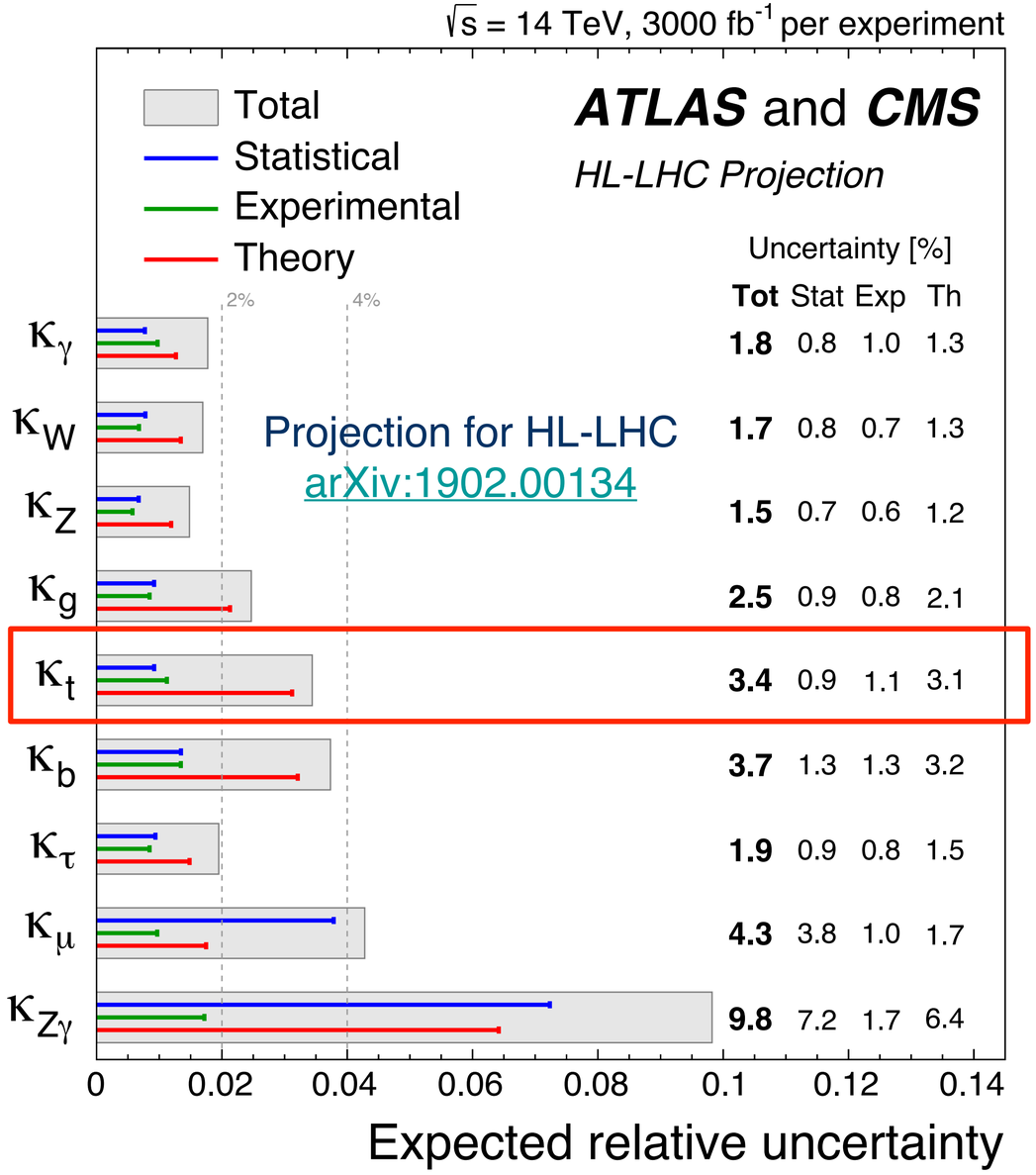


Sensitivity driven by H→γγ

- Combination helps reduce model dependence and improve precision
- Once included, **ttH(bb) will extend p_T(H) measurement to >450 GeV**

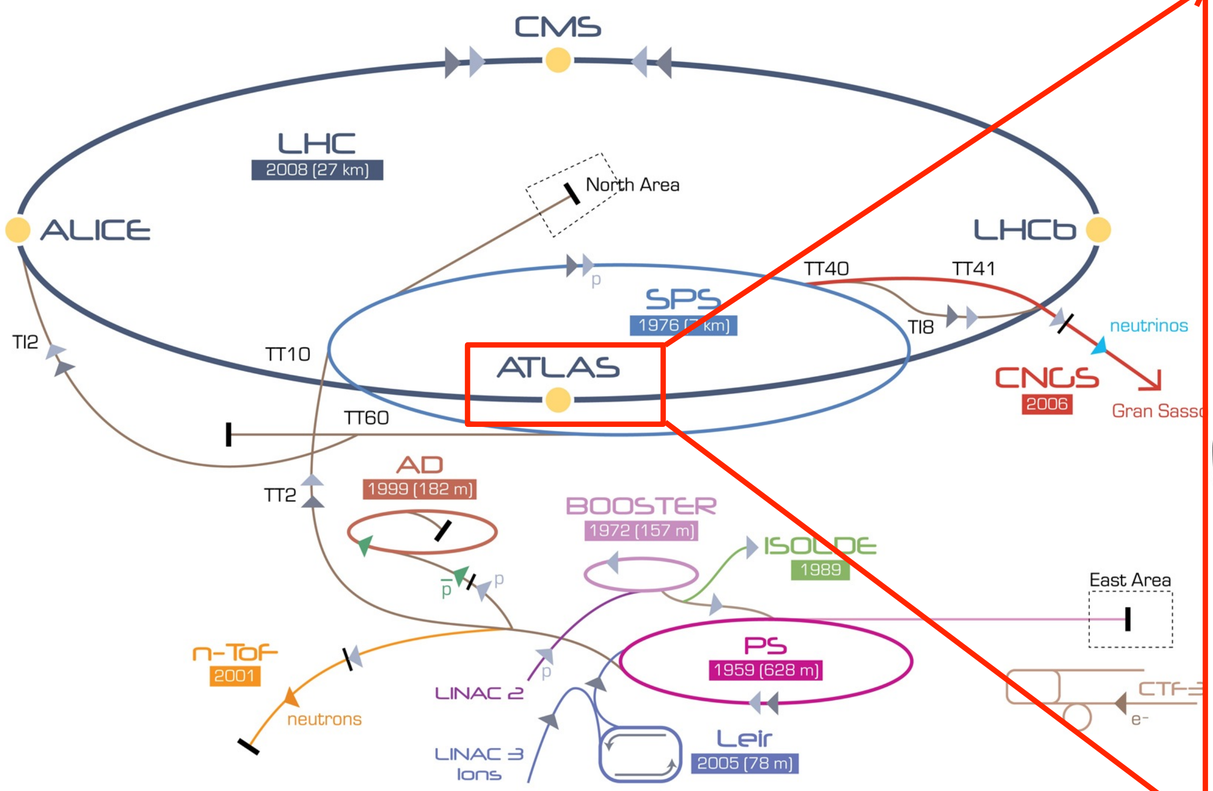
Conclusions

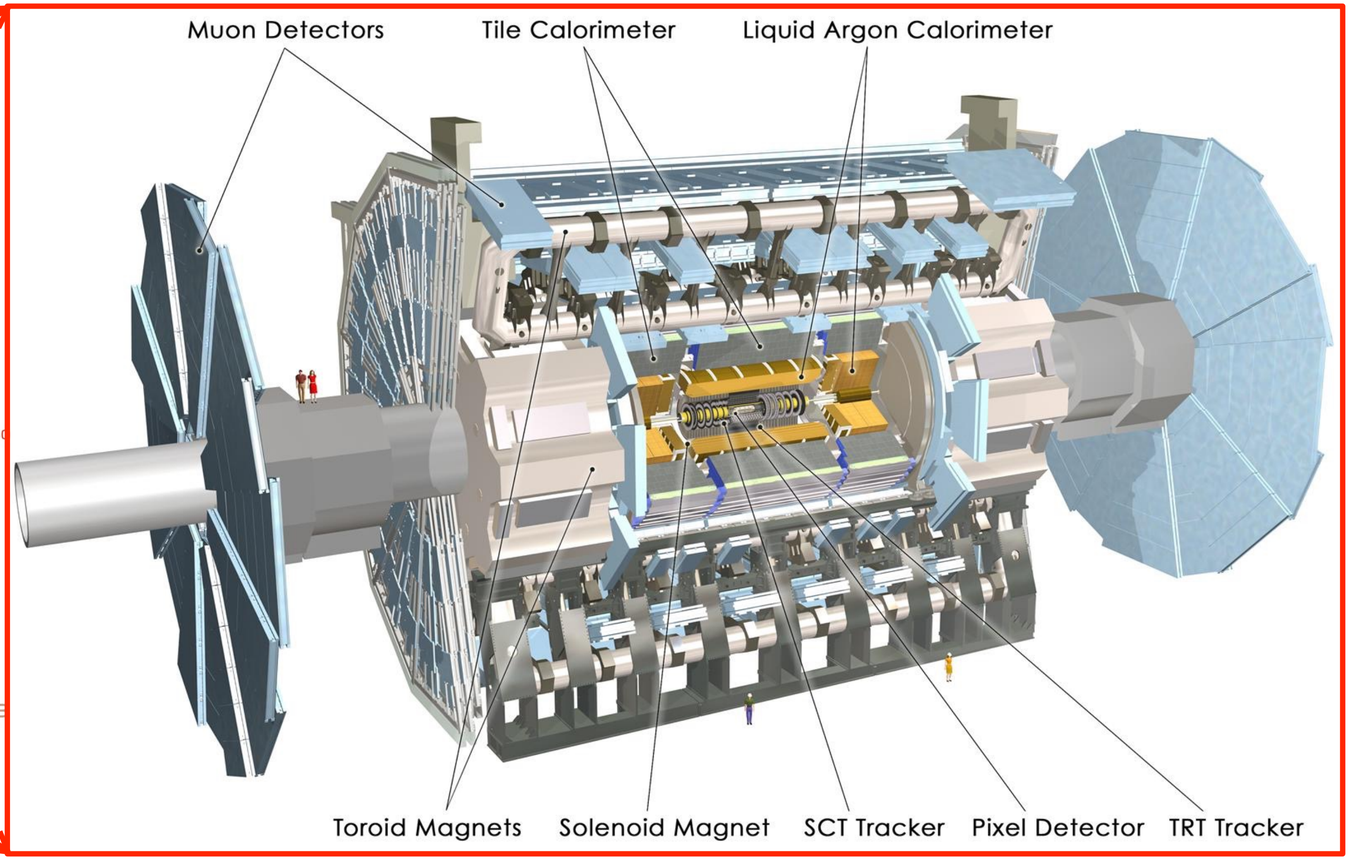
- ttH/tH remains a hot topic in LHC Higgs physics
  - Run 2 analyses being finalized, highlighting ttH differential measurements in $p_T(H)$ bins based on STXS framework
  - ATLAS standalone as well as ATLAS + CMS combinations will follow
- Factor of ~3 improvement in precision expected at High-Luminosity LHC
  - While waiting, work hard to resolve bkg. modeling challenges (e.g. $t\bar{t} + \geq 1b$ for ttH(bb)) & brainstorm new ideas to explore rich physics in ttH/tH using Run 2 & Run 3 data!



Backup

LHC and ATLAS detector





- Successful Run 2 data-taking at $\sqrt{s}$ = **13 TeV**. **139 fb^{-1}** proton-proton collision data for physics thanks to the excellent LHC performance!

Efficiency of Higgs boson reconstruction in ttH(bb)

p_{T}^{H} [GeV]	Dilepton $\mathrm{SR}^{\geq 4j}_{\geq 4b}$	Single-lepton $\mathrm{SR}^{\geq 6j}_{\geq 4b}$	Single-lepton $\mathrm{SR}_{\mathrm{boosted}}$
Inclusive	51%	43%	91%
$[0, 120)$	43%	35%	—
$[120, 200)$	50%	45%	—
$[200, 300)$	64%	57%	—
$[300, 450)$	78%	59%	90%
$[450, \infty)$			93%

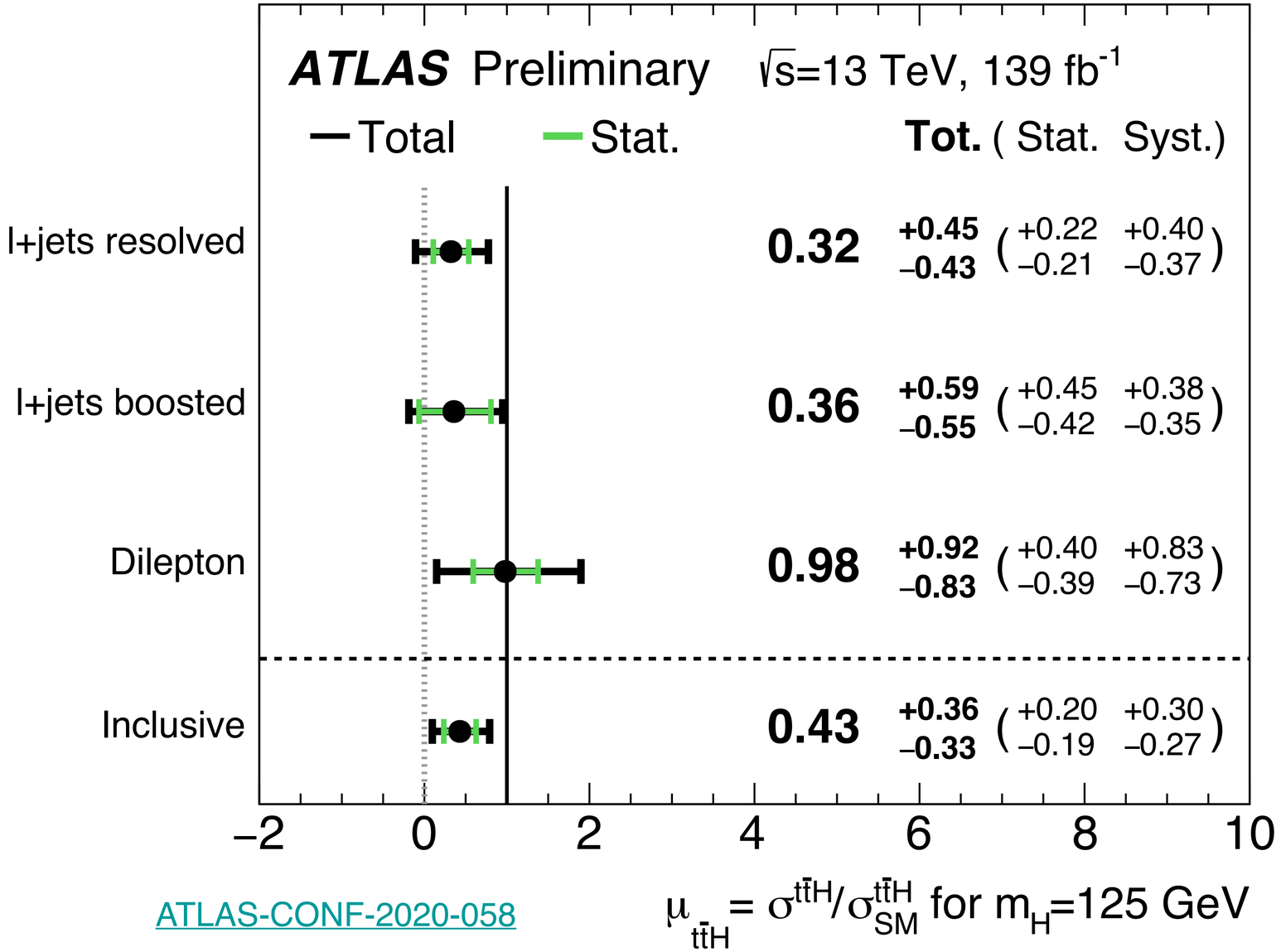
ATLAS-CONF-2020-058

Systematic sources for $t\bar{t}$ + jets modeling in ttH(bb)

Uncertainty source	Description		Components
$t\bar{t}$ cross-section	$\pm 6\%$		$t\bar{t}+$ light
$t\bar{t}+\geq 1b$ normalisation	Free-floating		$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1c$ normalisation	$\pm 100\%$		$t\bar{t}+\geq 1c$
NLO matching	MadGraph5_aMC@NLO+Pythia8 vs. PowhegBox+Pythia8		All
PS & hadronisation	PowhegBox+Herwig7 vs. PowhegBox+Pythia8		All
ISR	Varying α_S^{ISR} (PS), μ_{R}&μ_{F} (ME)	in PowhegBoxRes+Pythia8	$t\bar{t}+\geq 1b$
		in PowhegBox+Pythia8	$t\bar{t}+\geq 1c$, $t\bar{t}+$ light
FSR	Varying α_S^{FSR} (PS)	in PowhegBoxRes+Pythia8	$t\bar{t}+\geq 1b$
		in PowhegBox+Pythia8	$t\bar{t}+\geq 1c$, $t\bar{t}+$ light
$t\bar{t}+\geq 1b$ fractions	PowhegBox+Herwig7 vs. PowhegBox+Pythia8		$t\bar{t}+1b/1B$, $t\bar{t}+\geq 2b$
p_{T}^{bb} shape	Shape mismodelling measured from data		$t\bar{t}+\geq 1b$

ATLAS-CONF-2020-058

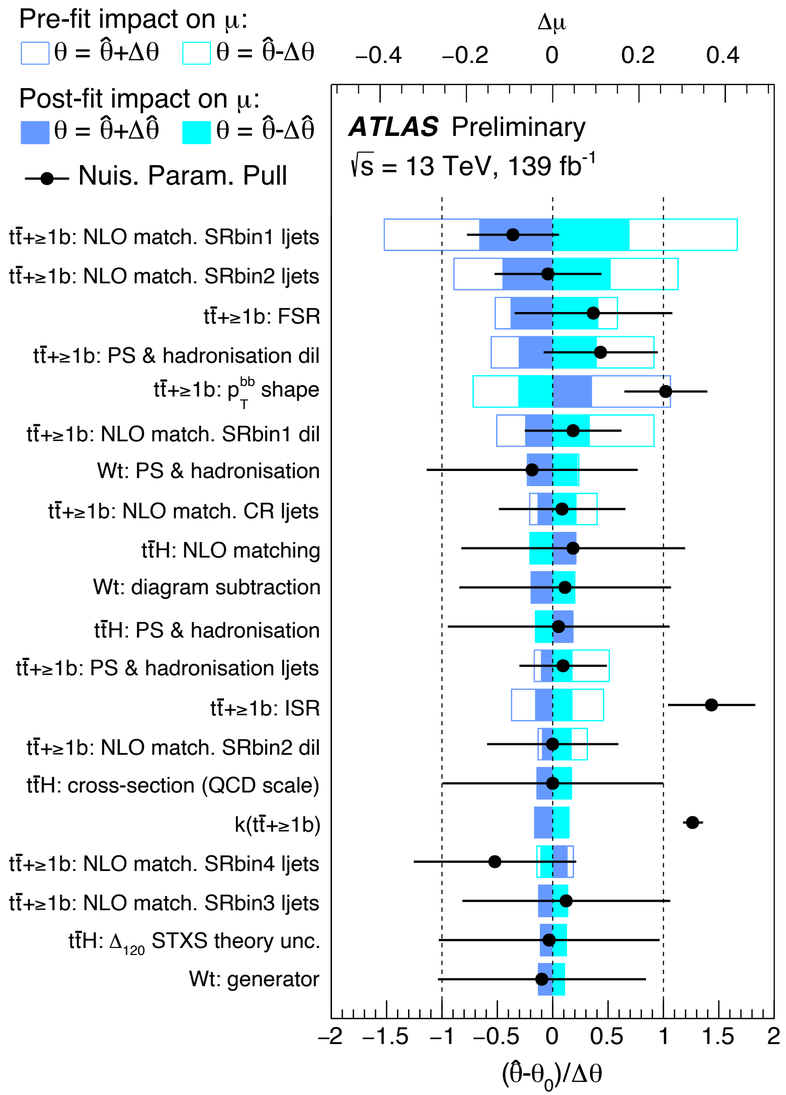
ttH(bb) signal strength in different regions



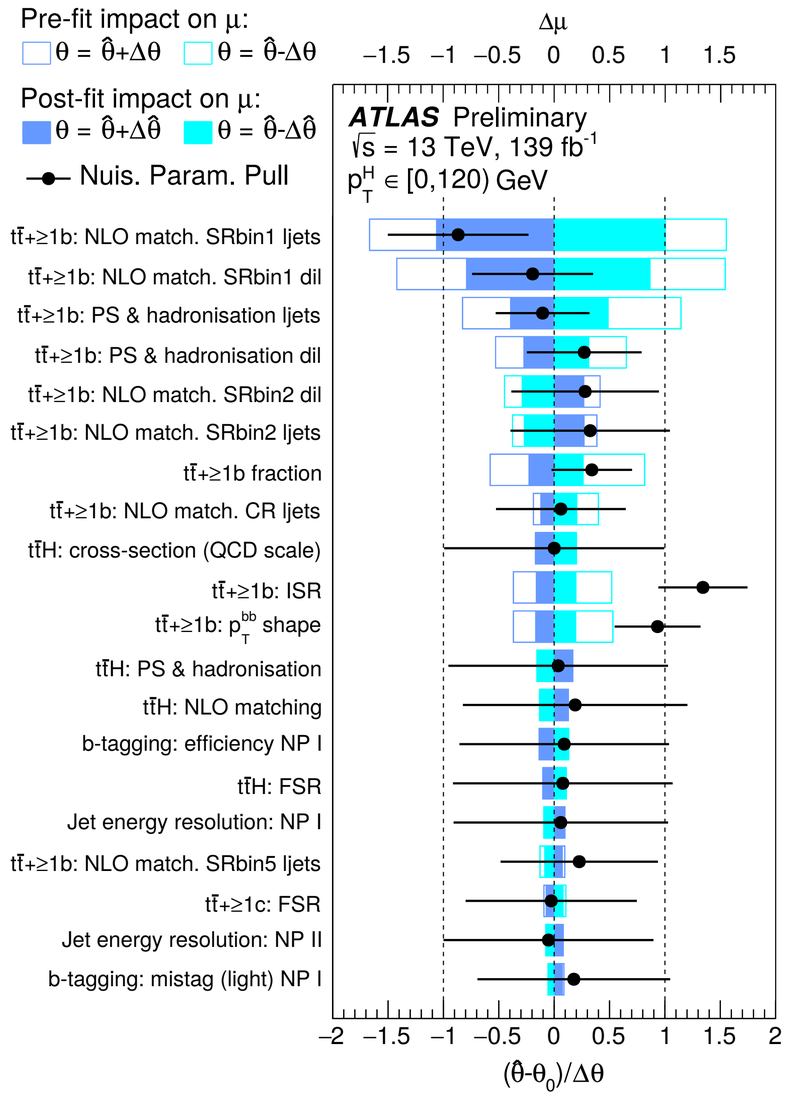
ATLAS-CONF-2020-058

ttH(bb) ranking results

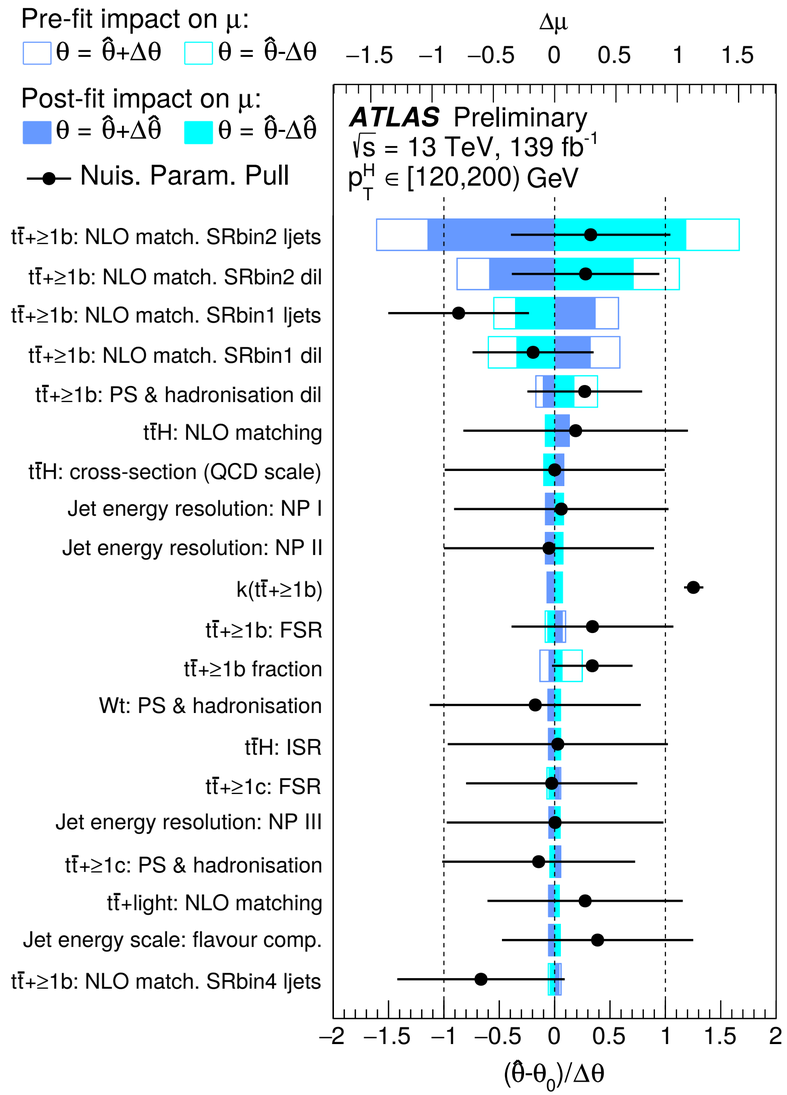
ATLAS-CONF-2020-058



Inclusive



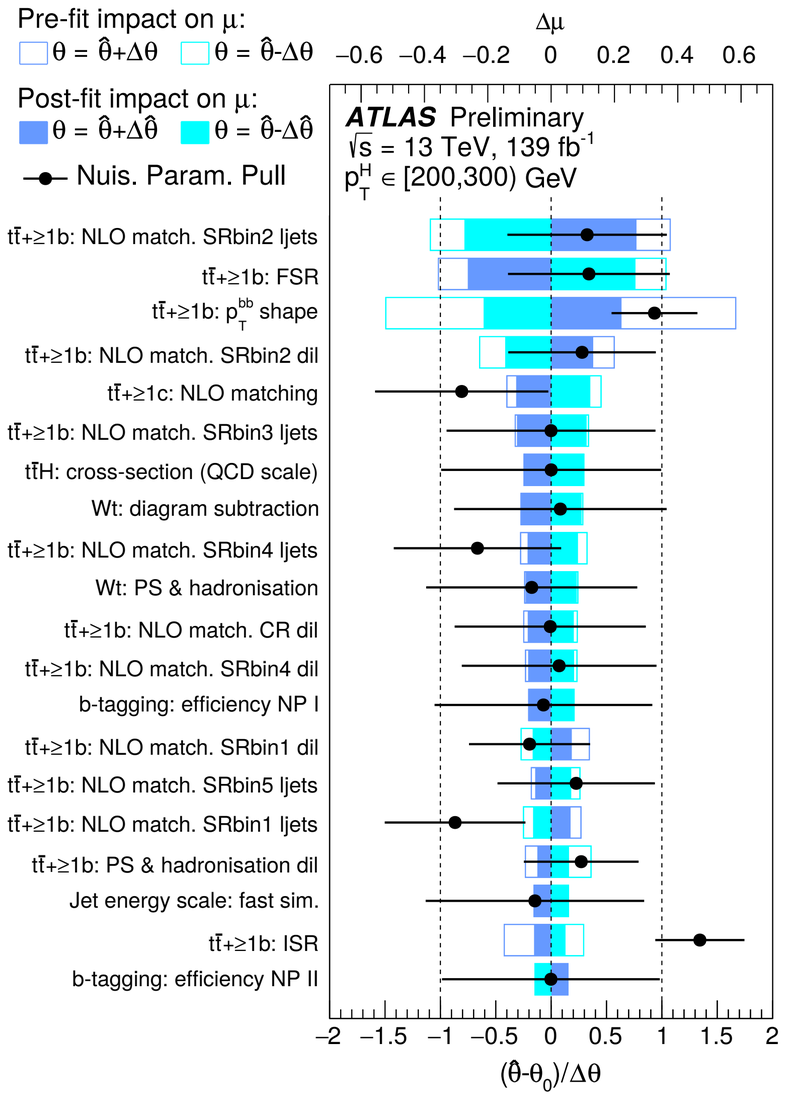
$p_T(H) < 120$ GeV

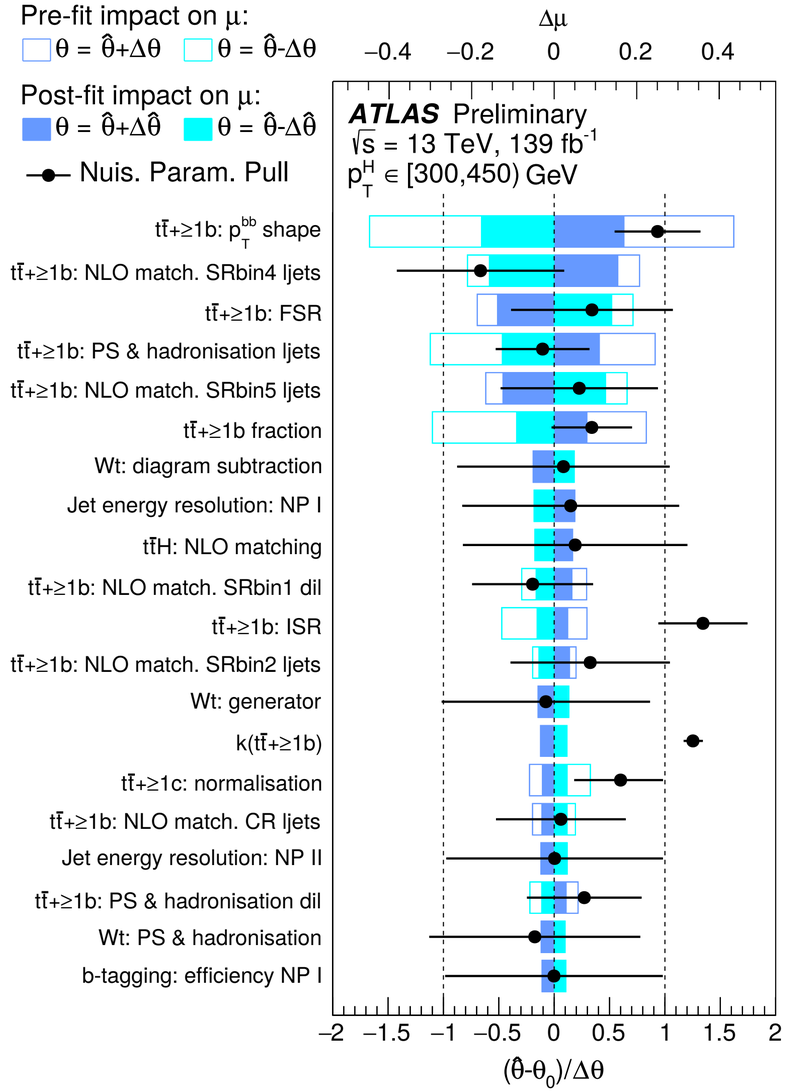


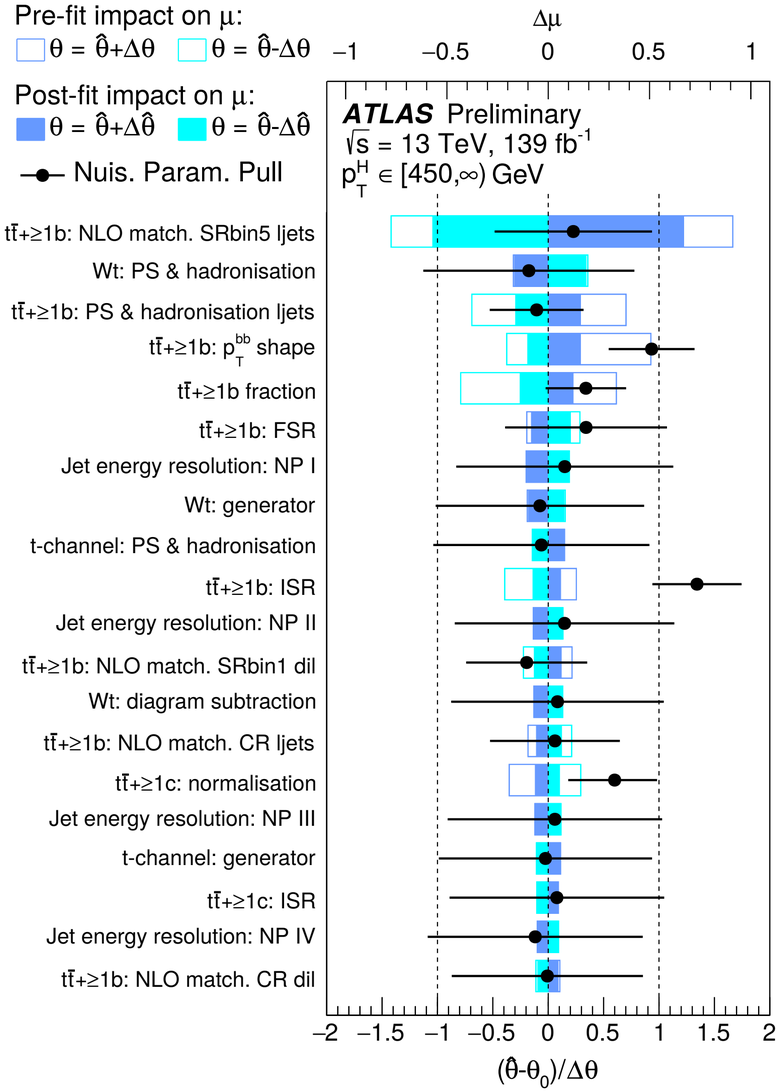
$120 \leq p_T(H) < 200$ GeV

ttH(bb) ranking results

ATLAS-CONF-2020-058







200 ≤ p_T(H) < 300 GeV

300 ≤ p_T(H) < 450 GeV

p_T(H) ≥ 450 GeV

Post-fit event yields in ttH(bb) di-lepton channel

	$\mathrm{SR}^{\geq 4j}_{\geq 4b}$				$\mathrm{CR}^{3j}_{3b\ \mathrm{hi}}$	$\mathrm{CR}^{\geq 4j}_{3b\ \mathrm{hi}}$	$\mathrm{CR}^{\geq 4j}_{3b\ \mathrm{lo}}$
p_{T}^{H} range [GeV]	[0,120)	[120,200)	[200,300)	[300,∞)			
$t\bar{t}H$	14 ± 12	6.7 ± 5.3	3.3 ± 2.6	1.6 ± 1.2	10.5 ± 8.4	51 ± 41	33 ± 27
$t\bar{t} + \geq 1b$	557 ± 28	265 ± 17	117.6 ± 9.6	37.4 ± 5.6	2030 ± 130	4080 ± 210	2540 ± 170
$t\bar{t} + \geq 1c$	48.7 ± 9.5	14.4 ± 4.4	6.2 ± 1.4	3.9 ± 1.0	523 ± 130	1190 ± 260	2550 ± 500
$t\bar{t}$ + light, $4t$, tH	7.9 ± 5.8	4.2 ± 2.8	2.1 ± 1.5	1.4 ± 1.3	123 ± 66	221 ± 120	923 ± 360
$t\bar{t} + Z$	12.5 ± 2.0	7.6 ± 1.6	4.15 ± 0.71	2.03 ± 0.44	10.7 ± 1.7	57.4 ± 7.3	52.5 ± 6.8
$t\bar{t} + W$	0.75 ± 0.31	0.41 ± 0.12	0.27 ± 0.11	0.128 ± 0.069	1.83 ± 0.55	10.9 ± 1.6	22.0 ± 3.5
Other top sources	19.0 ± 6.7	7.7 ± 4.2	4.4 ± 4.0	2.0 ± 1.5	126 ± 34	208 ± 60	254 ± 71
Fakes	3.6 ± 1.1	1.32 ± 0.51	0.40 ± 0.23	0.57 ± 0.30	6.31 ± 1.8	46.3 ± 12	55.7 ± 14
Total	664 ± 24	307 ± 16	138.5 ± 8.9	48.9 ± 5.1	2830 ± 54	5860 ± 79	6430 ± 82
Data	647	306	135	48	2827	5865	6429

ATLAS-CONF-2020-058

Post-fit event yields in ttH(bb) lepton+jet channel

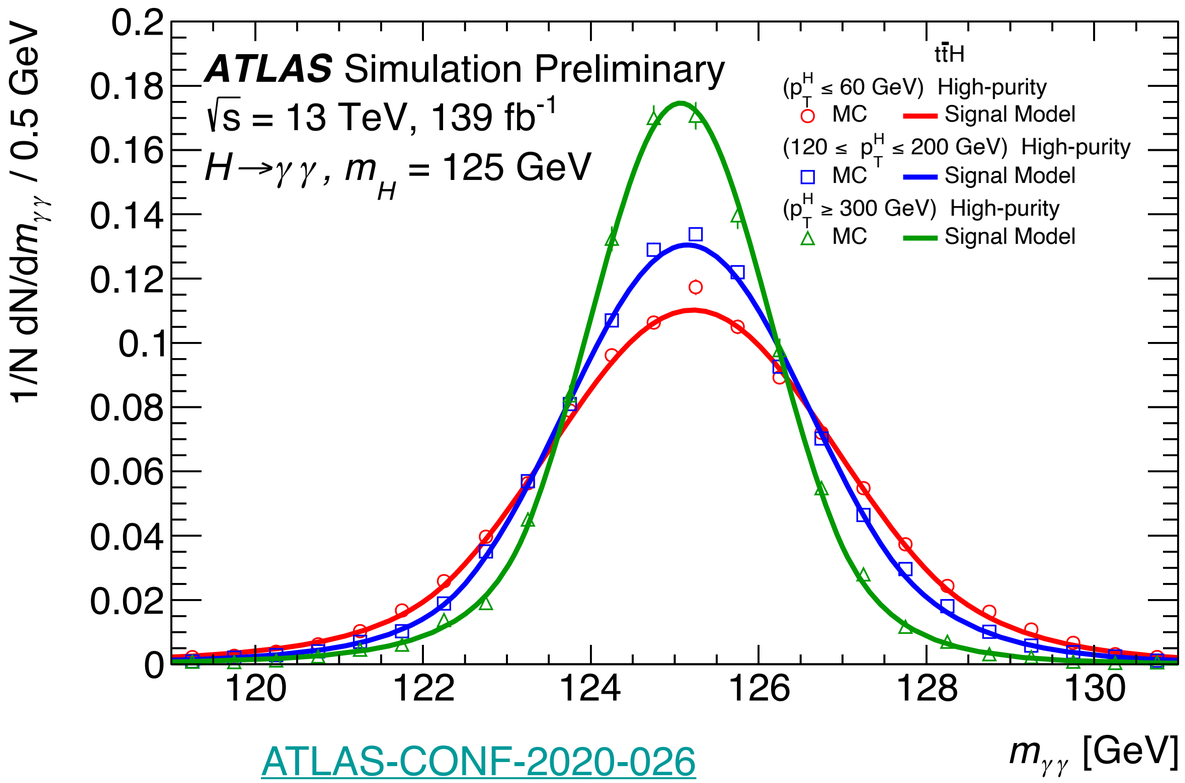
	$\mathrm{SR}^{\geq 6j}_{\geq 4b}$					$\mathrm{SR}_{\mathrm{boosted}}$		$\mathrm{CR}^{5j}_{\geq 4b\ \mathrm{lo}}$	$\mathrm{CR}^{5j}_{\geq 4b\ \mathrm{hi}}$
p_{T}^{H} range [GeV]	[0,120)	[120,200)	[200,300)	[300,450)	[450,∞)	[300,450)	[450,∞)		
$t\bar{t}H$	93 ± 74	49 ± 39	26 ± 21	5.9 ± 4.6	1.26 ± 1.00	15 ± 12	3.6 ± 2.8	26 ± 20	26 ± 21
$t\bar{t} + \geq 1b$	4450 ± 160	2040 ± 85	855 ± 43	234 ± 20	43.4 ± 8.2	297 ± 27	51.0 ± 9.8	1595 ± 80	1102 ± 51
$t\bar{t} + \geq 1c$	960 ± 210	404 ± 87	179 ± 38	46 ± 11	12.9 ± 3.3	157 ± 37	40 ± 11	630 ± 140	90 ± 23
$t\bar{t}$ + light, $4t$, tH	250 ± 140	105 ± 57	52 ± 26	15.4 ± 8.8	3.5 ± 2.2	62 ± 25	16.9 ± 7.6	270 ± 100	26 ± 16
$t\bar{t} + W$	7.3 ± 1.1	4.46 ± 0.87	2.54 ± 0.48	1.09 ± 0.31	0.48 ± 0.14	1.89 ± 0.36	0.57 ± 0.17	2.62 ± 0.46	0.53 ± 0.12
$t\bar{t} + Z$	79 ± 10	46.0 ± 6.4	31.1 ± 4.9	11.8 ± 2.3	2.12 ± 0.64	11.0 ± 2.1	2.34 ± 0.60	25.9 ± 3.5	22.8 ± 3.1
Single top Wt	80 ± 43	44 ± 27	18.7 ± 7.8	9.5 ± 9.0	6.1 ± 5.4	14.0 ± 8.3	4.9 ± 4.3	60 ± 32	28 ± 20
Other top sources	48 ± 25	24 ± 16	14 ± 10	4.5 ± 2.7	1.09 ± 0.54	4.4 ± 3.0	0.88 ± 0.78	41 ± 16	28 ± 11
V & VV + jets	63 ± 24	30 ± 11	20.6 ± 8.2	8.1 ± 3.4	1.92 ± 0.84	13.1 ± 5.6	4.2 ± 2.0	43 ± 15	24.9 ± 8.8
Total	6026 ± 84	2747 ± 52	1198 ± 31	336 ± 15	72.8 ± 7.0	575 ± 23	124.4 ± 9.7	2700 ± 52	1348 ± 38
Data	6047	2742	1199	331	75	581	118	2696	1362

ATLAS-CONF-2020-058

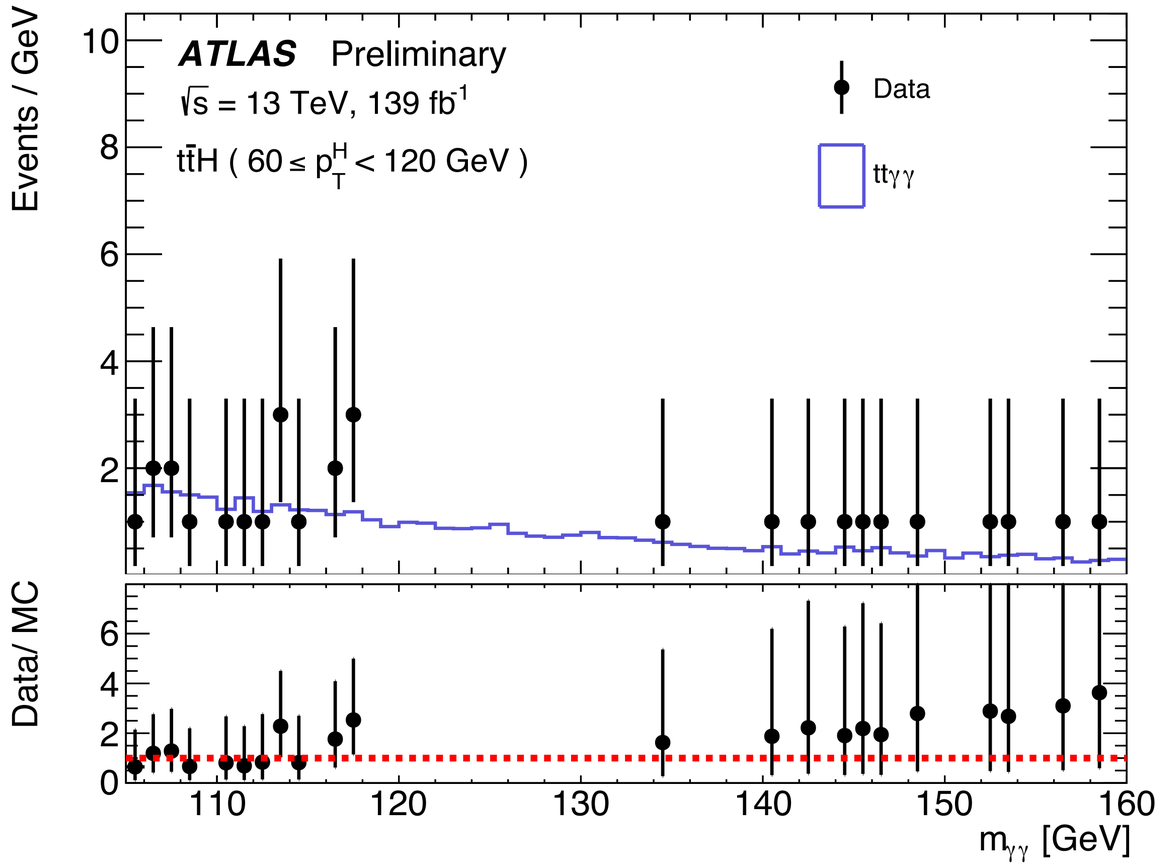
Summary of training variables in H→γγ

STXS regions	Multi-class BDT	STXS regions	Binary BDT
$gg \to H$	di-photon p_{T} and absolute rapidity; di-jet p_{T}, mass, Δy, $\Delta\phi$, $\Delta\eta$ between the 2 jets; p_{T}, mass of $\gamma\gamma + j$ and $\gamma\gamma + jj$, Δy, $\Delta\phi$ between $\gamma\gamma$ and jj, minimum ΔR between jets and photons, mass of the sum of all jets; di-lepton p_{T}, di-e or di-μ mass, E_{T}^{miss}, p_{T} of lepton + E_{T}^{miss}; p_T, η, ϕ, mass of top candidates; Number of jets, barrel jets ($\lvert\eta\rvert < 2.5$), b-jets and leptons; leading jet p_T, sum p_{T} of all jets $\sum E^{\mathrm{T}}$, E_{T}^{miss} significance; Average interaction per crossing, number of primary vertices	individual STXS regions from $gg \to H$ or $qq' \to Hqq'$	Multi-class BDT variables, and $\Delta\phi$, $\Delta\eta$ between the 2 photons ($\Delta\phi_{\gamma\gamma}$, $\Delta\eta_{\gamma\gamma}$); Number of electrons and muons; E_{T}^{miss}, $\sum E^{\mathrm{T}}$, E_{T}^{miss} significance, and E_{T}^{miss} azimuthal angle computed from hardest vertex; $\gamma\gamma$ $\vec{p}_{\mathrm{T}}$ projected to its thrust axis ($p_{\mathrm{Tt}}^{\gamma\gamma}$); Half difference between di-photon η and sum η of leading 2 jets (η^{Zepp}); $\phi^*_{\gamma\gamma} = \tan(\frac{\pi - \lvert\Delta\phi_{\gamma\gamma}\rvert}{2})\sqrt{1-\tanh^2(\frac{\Delta\eta_{\gamma\gamma}}{2})}$ $\cos\theta^*_{\gamma\gamma} = \lvert\frac{(E^{\gamma_1}+p_z^{\gamma_1})\cdot(E^{\gamma_2}-p_z^{\gamma_2}) - (E^{\gamma_1}-p_z^{\gamma_1})\cdot(E^{\gamma_2}+p_z^{\gamma_2})}{m_{\gamma\gamma}+\sqrt{(m_{\gamma\gamma}^2+(p_{\mathrm{T}}^{\gamma\gamma})^2)}}\rvert$
$qq' \to Hqq'$			
$qq \to H\ell\nu$		WH STXS regions combined	$p_{\mathrm{T}}/m_{\gamma\gamma}$, η, ϕ of 2 leading photons; p_{T}, η, ϕ of 2 leading leptons; E_{T}^{miss}, E_{T}^{miss} significance, E_{T}^{miss} azimuthal angle; Whether or not the E_{T}^{miss} built from di-photon vertex is larger than that built from the hardest vertex by more than 30 GeV; di-lepton mass, and transverse mass of lepton + E_{T}^{miss}
$qq \to H\ell\ell$		ZH STXS regions combined	
$t\bar{t}H$	ATLAS-CONF-2020-026	$t\bar{t}H$ STXS regions combined	p_{T}, η, ϕ of 2 leading photons; p_{T}, η, ϕ and B-tagging scores of 6 leading jets; E_{T}^{miss}, E_{T}^{miss} significance, E_{T}^{miss} azimuthal angle; Top reconstruction BDT scores
tH		tWH, $tHqb$	

H→γγ Signal and background modeling

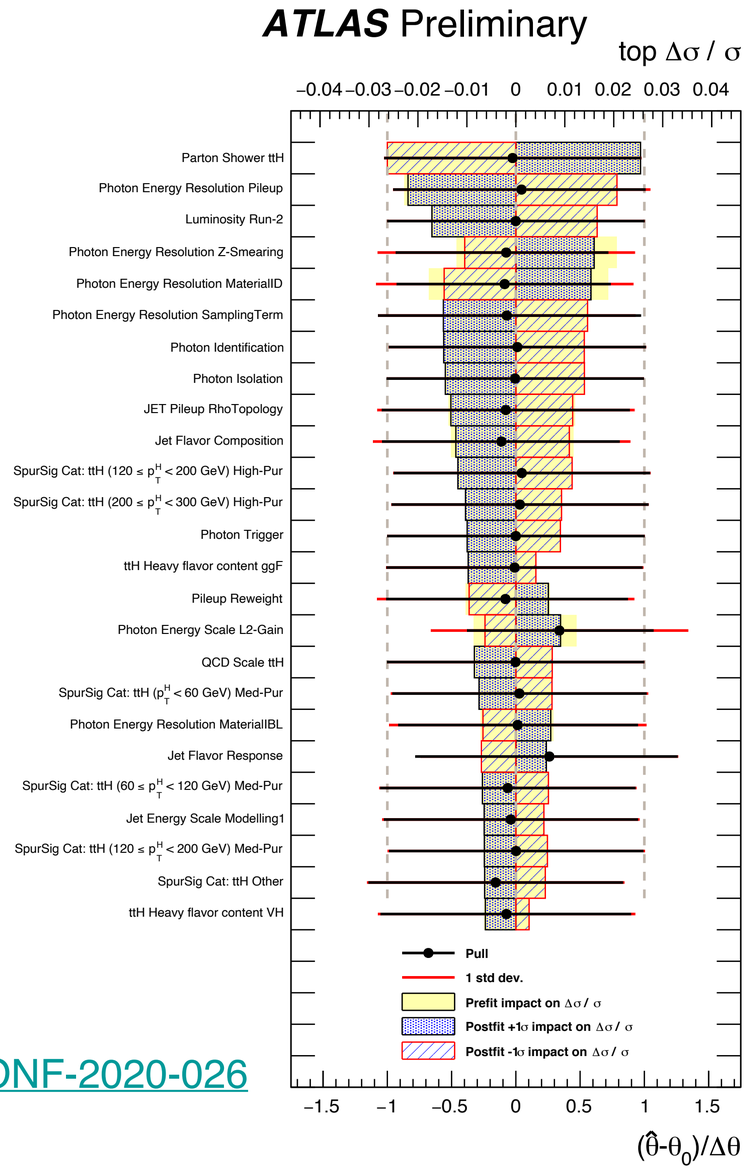


ATLAS-CONF-2020-026



- Parameterize signal with double-sided crystal ball function
- Select background model on MC template using spurious signal procedure.
  Exponential function chosen for all ttH/tH categories

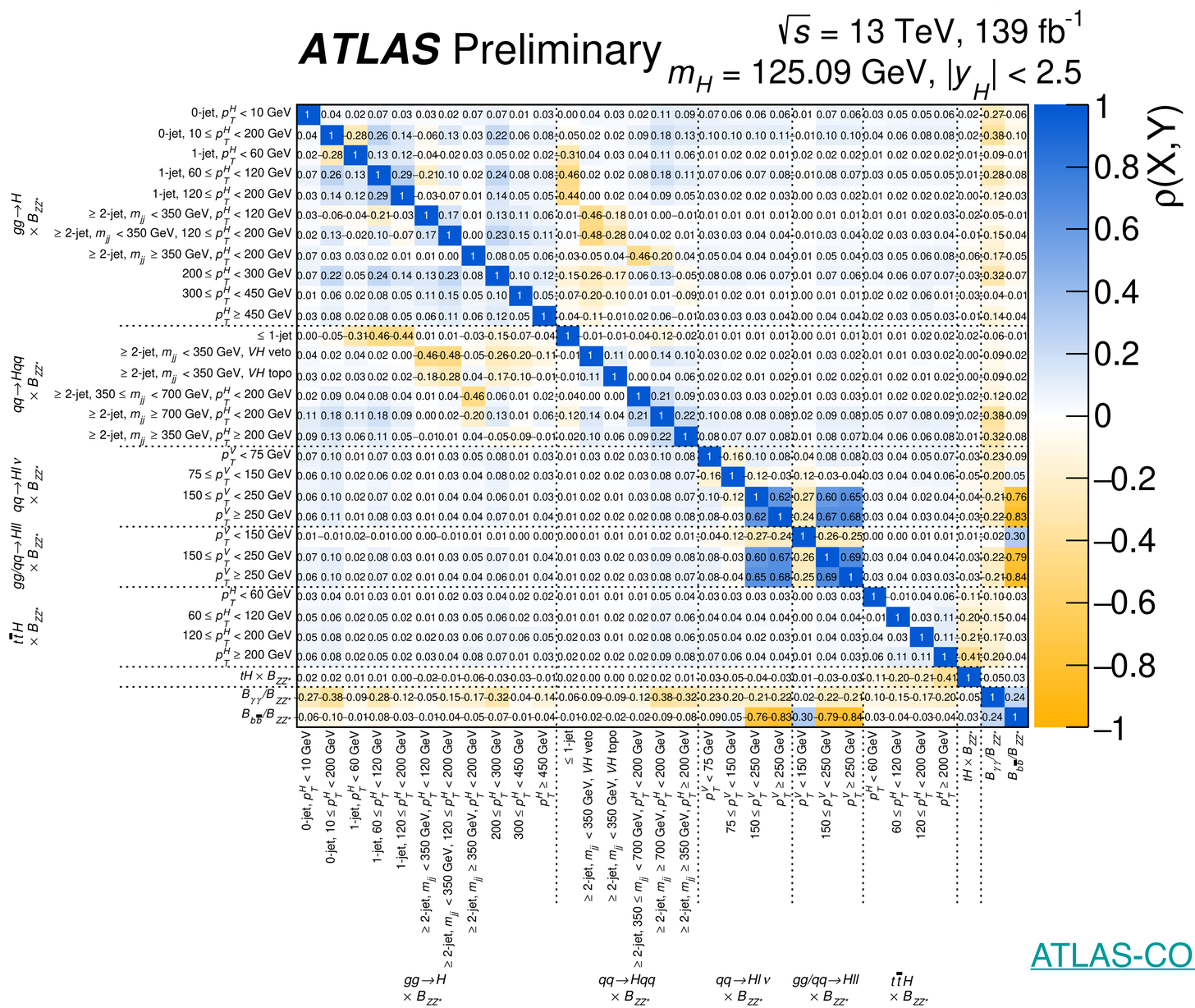
Nuisance parameter ranking from H→γγ



ATLAS-CONF-2020-026

- No significant pull or over-constraint of nuisance parameter for systematic uncertainties

STXS correlation matrix



ATLAS-CONF-2020-027