

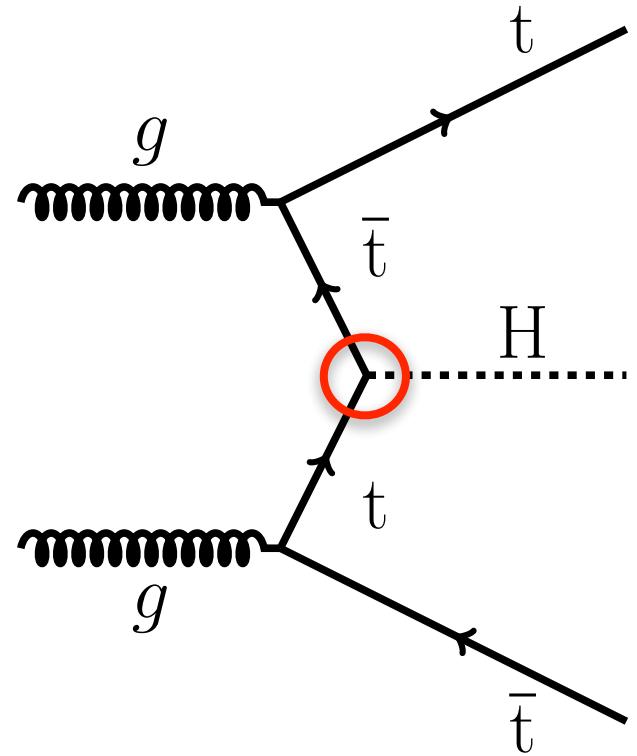
# Study of the top-Higgs coupling at LHC

Thomas Strebler

Centre de Physique  
des Particules de Marseille  
Aix-Marseille Université / CNRS-IN2P3

*ATLAS collaboration*

CPPM seminar  
March 4th, 2019

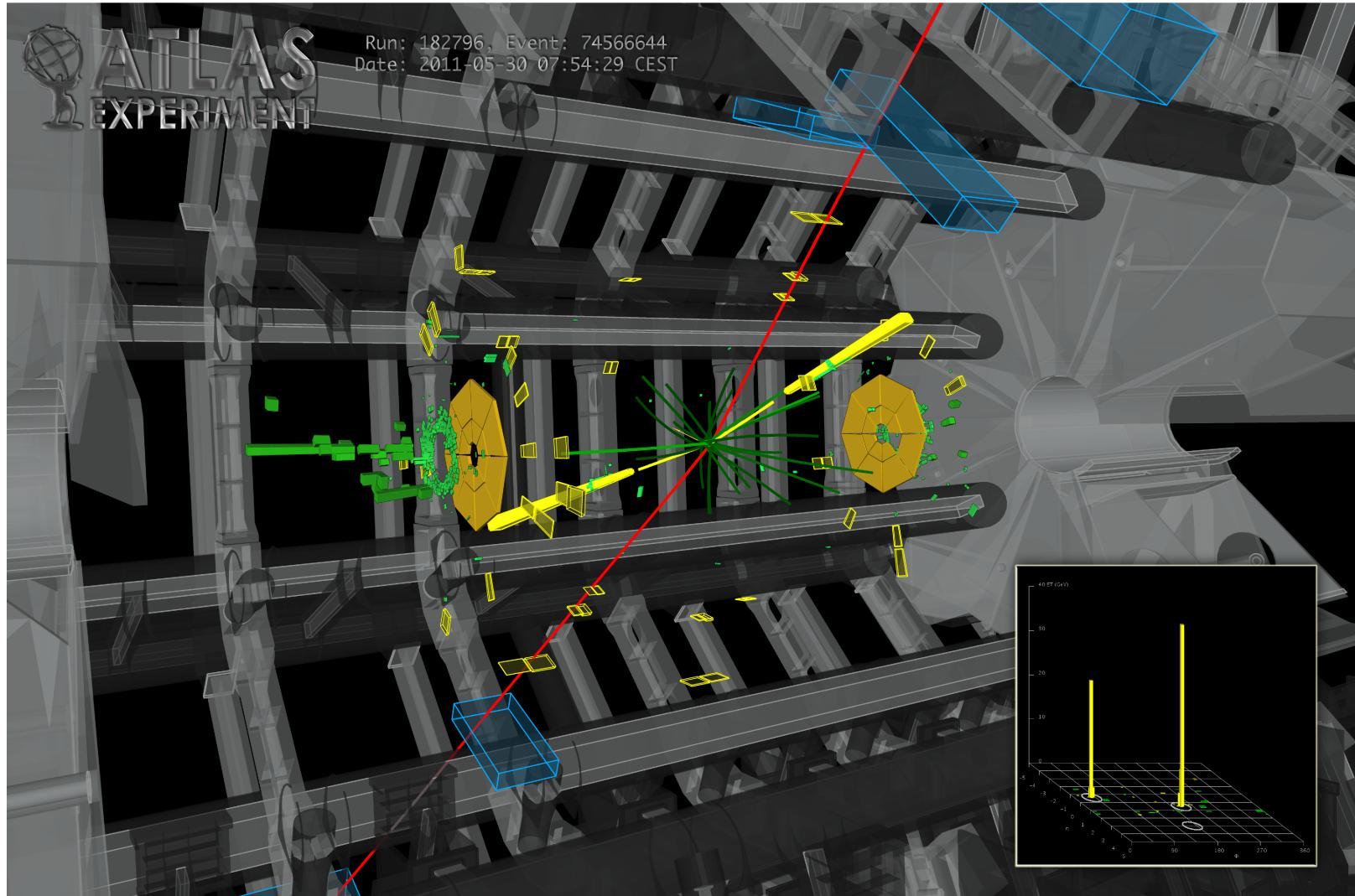


IN2P3  
Les deux infinis



# Introduction

Run I of the LHC brought evidence for the last missing piece of the Standard Model:  
**the Higgs boson!**



# Introduction

- After observation and mass measurement, all Higgs boson properties predicted within the SM
- Almost **all accessible decay modes have been observed** by both ATLAS and CMS collaborations during Run 1

Channel	References for individual publications		Signal strength [ $\mu$ ]		Signal significance [ $\sigma$ ]	
	ATLAS	CMS	ATLAS	CMS	ATLAS	CMS
$H \rightarrow \gamma\gamma$	[51]	[52]	$1.15^{+0.27}_{-0.25}$ ( $+0.26$ ) ( $-0.24$ )	$1.12^{+0.25}_{-0.23}$ ( $+0.24$ ) ( $-0.22$ )	5.0 (4.6)	5.6 (5.1)
$H \rightarrow ZZ \rightarrow 4\ell$	[53]	[54]	$1.51^{+0.39}_{-0.34}$ ( $+0.33$ ) ( $-0.27$ )	$1.05^{+0.32}_{-0.27}$ ( $+0.31$ ) ( $-0.26$ )	6.6 (5.5)	7.0 (6.8)
$H \rightarrow WW$	[55, 56]	[57]	$1.23^{+0.23}_{-0.21}$ ( $+0.21$ ) ( $-0.20$ )	$0.91^{+0.24}_{-0.21}$ ( $+0.23$ ) ( $-0.20$ )	6.8 (5.8)	4.8 (5.6)
$H \rightarrow \tau\tau$	[58]	[59]	$1.41^{+0.40}_{-0.35}$ ( $+0.37$ ) ( $-0.33$ )	$0.89^{+0.31}_{-0.28}$ ( $+0.31$ ) ( $-0.29$ )	4.4 (3.3)	3.4 (3.7)
$H \rightarrow bb$	[38]	[39]	$0.62^{+0.37}_{-0.36}$ ( $+0.39$ ) ( $-0.37$ )	$0.81^{+0.45}_{-0.42}$ ( $+0.45$ ) ( $-0.43$ )	1.7 (2.7)	2.0 (2.5)

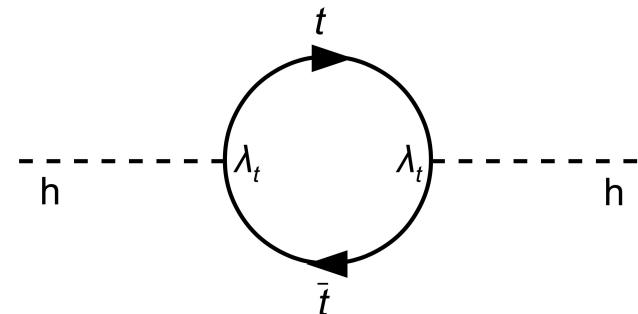
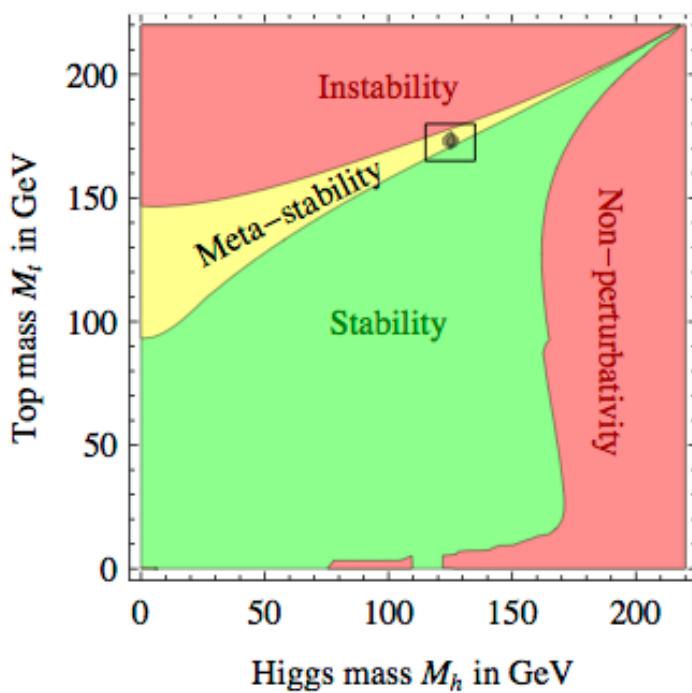
JHEP08(2016)045

[arXiv:1606.02266 \[hep-ex\]](https://arxiv.org/abs/1606.02266)

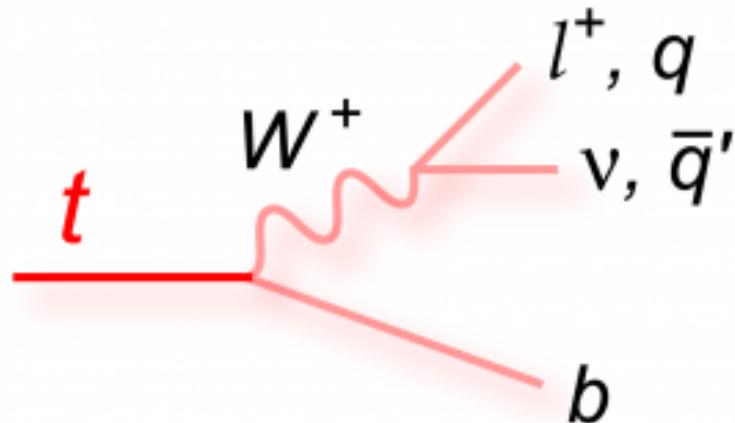
- Unambiguous evidence of **Higgs boson coupling to gauge bosons**

# Top-Higgs Yukawa coupling

- Unlike W/Z bosons, **masses of fermions** generated within SM with **ad hoc Yukawa couplings**
- **Top quark** only fermion with **Yukawa coupling of order 1**: very special role in **naturalness problem and vacuum stability**

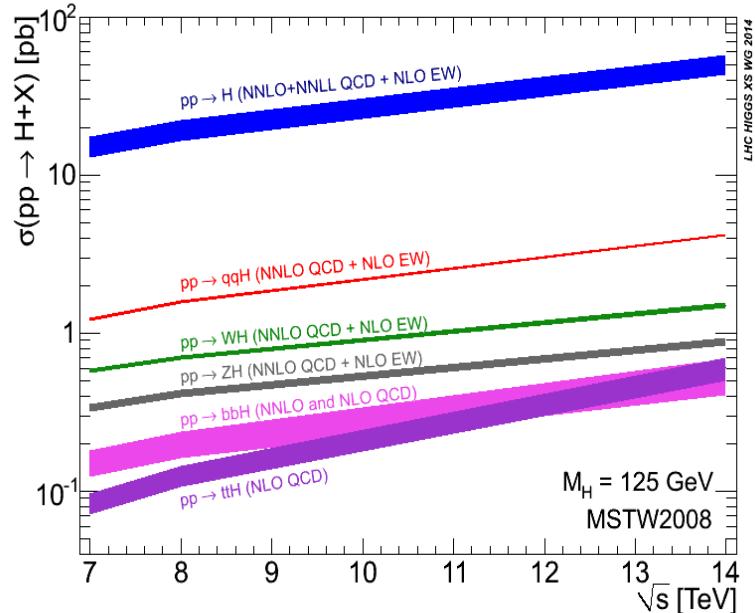
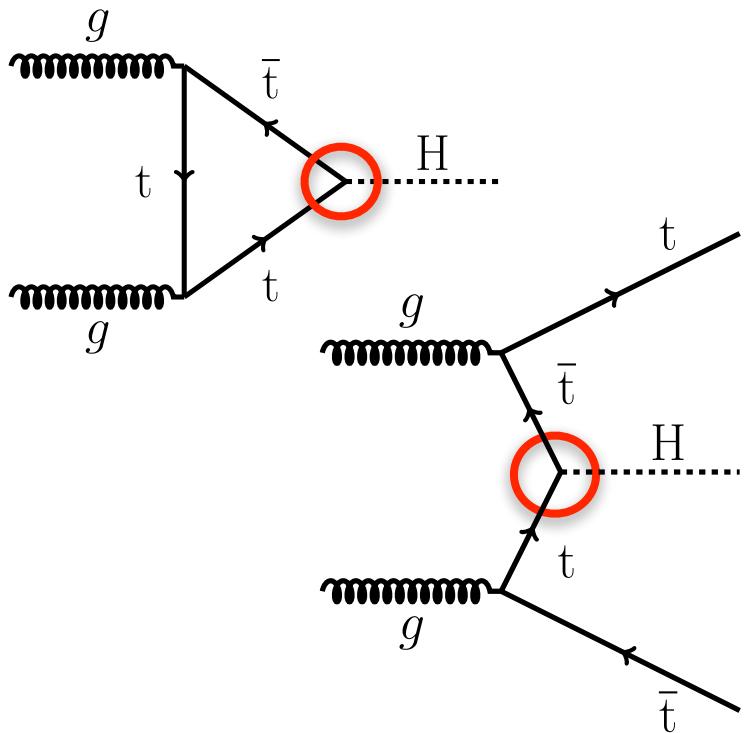


- Possible to study spin-correlations using top decay products: **access to CP structure of production vertex**



# ttH production

- Unlike other fermions, top-Higgs coupling cannot be probed with one of the Higgs boson decay mode
- Access to  $Y_t$  with **ggF +  $H \rightarrow \gamma\gamma$  decay**
- **Model-dependent constraints on  $Y_t$ , assuming no BSM particles in the loop**

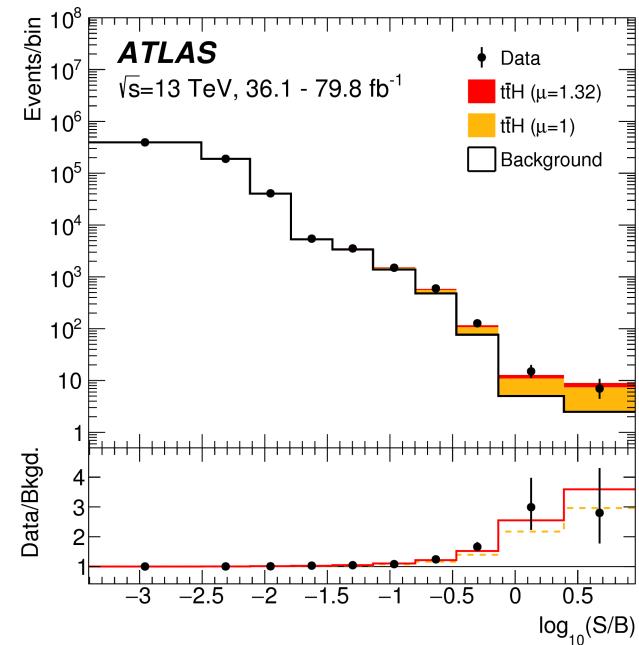
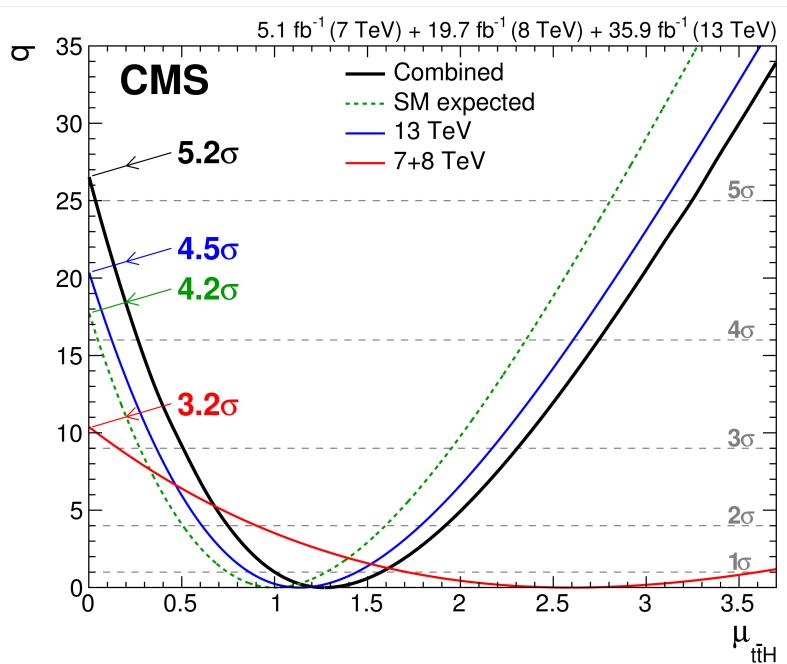


- **ttH => direct probe to study  $Y_t$**
- Large cross-section boost from 8 to 13 TeV  
=> **LHC Run 2 offers ideal conditions to study this production mode**
- **Target as much decay modes as possible** for top + Higgs to optimally exploit small cross-section (0.5 pb, ~1% ggF)

# ttH production: first observations

- First observations achieved independently by ATLAS and CMS in Spring 2018, combining Run 1 + partial Run 2 datasets

Phys. Rev. Lett. 120,  
231801 (2018)  
[arXiv:1804.02610](https://arxiv.org/abs/1804.02610)  
[hep-ex]



- One of the LHC Run 2 highlights so far

## Higgs boson comes out on top

New results from CMS and ATLAS experiments reveal how strongly the Higgs boson interacts with the heaviest known elementary particle, the top quark

4 JUNE, 2018

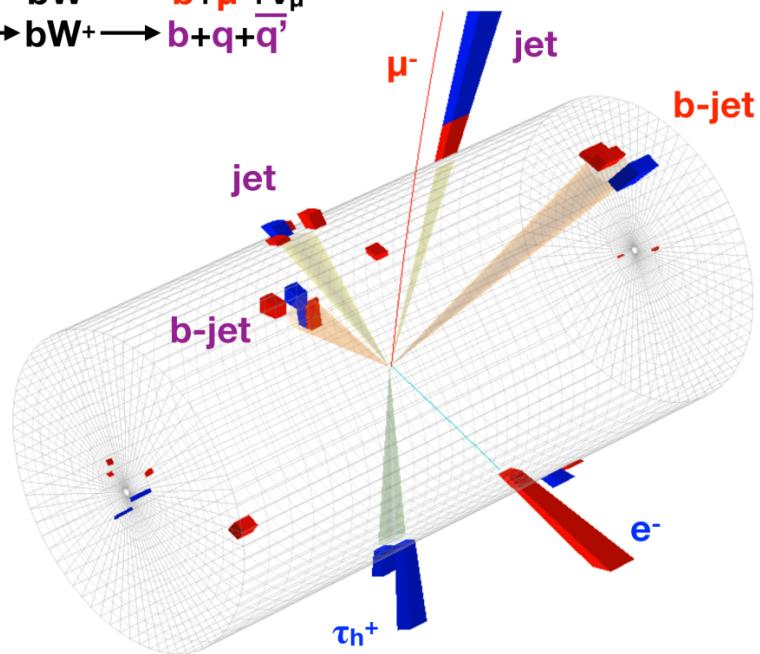
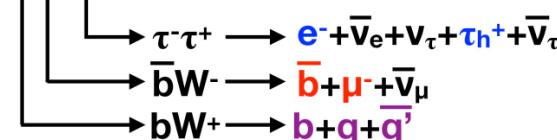
Phys. Lett. B 784 (2018) 173  
[arXiv:1804.02610](https://arxiv.org/abs/1804.02610) [hep-ex]

<https://home.cern/news/news/physics/higgs-boson-comes-out-top>

# ttH analysis challenges

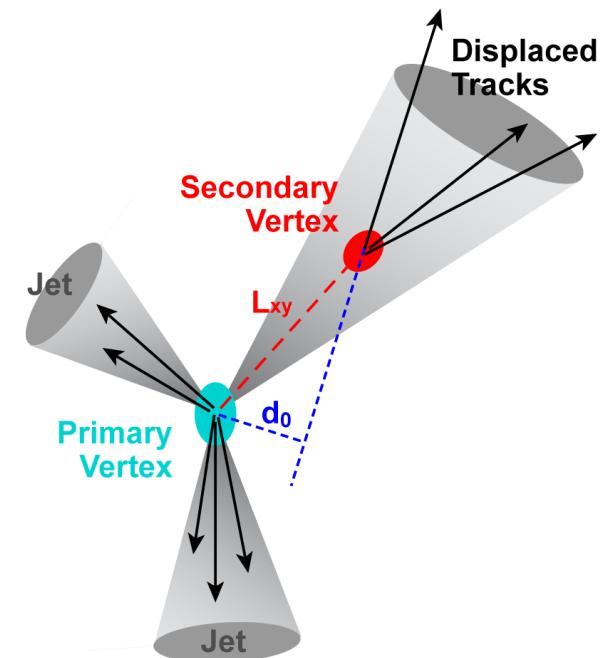
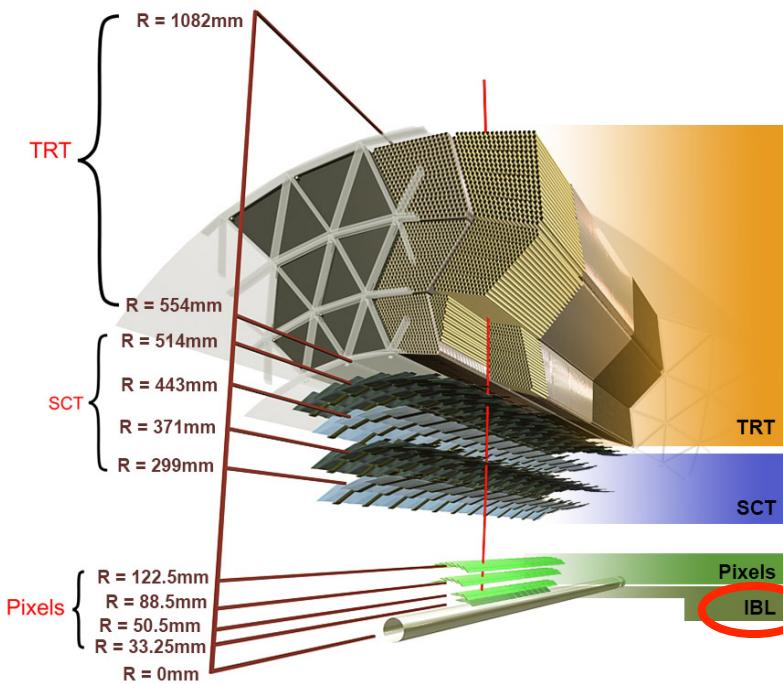
- Large multiplicity of objects in the final states: jets, leptons, photons...
- Not straightforward to reconstruct Higgs boson: MET and/or combinatorics
- Challenging backgrounds to model and to discriminate
- Extensive use of **MVA discriminants** for object identification and signal extraction
- Complex analyses with specific challenges:
  - ttH  $H \rightarrow bb$  (fully hadronic + 1-2 lepton final states)
  - ttH  $H \rightarrow WW/ZZ/\tau\tau$  (multilepton final states)
  - ttH  $H \rightarrow \gamma\gamma$
  - ttH  $H \rightarrow ZZ \rightarrow 4l$

$pp \rightarrow t\bar{t}H$



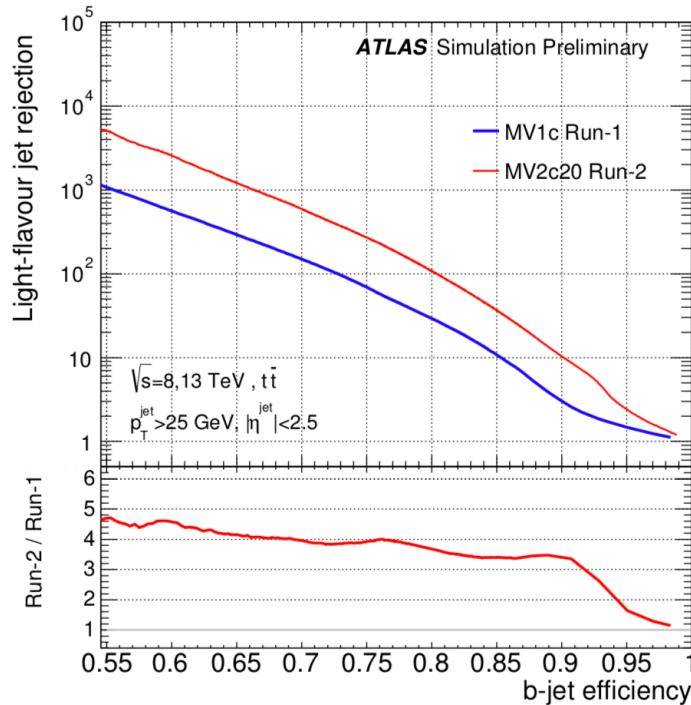
# b-tagging

- b quarks produced in top decays => b-tagging plays a major role in all ttH analyses
- b-taggers in ATLAS and CMS exploit **larger lifetime + mass of B-hadrons**
- b-tagging performance strongly correlated with ability to **reconstruct displaced secondary vertices + impact parameter resolution**



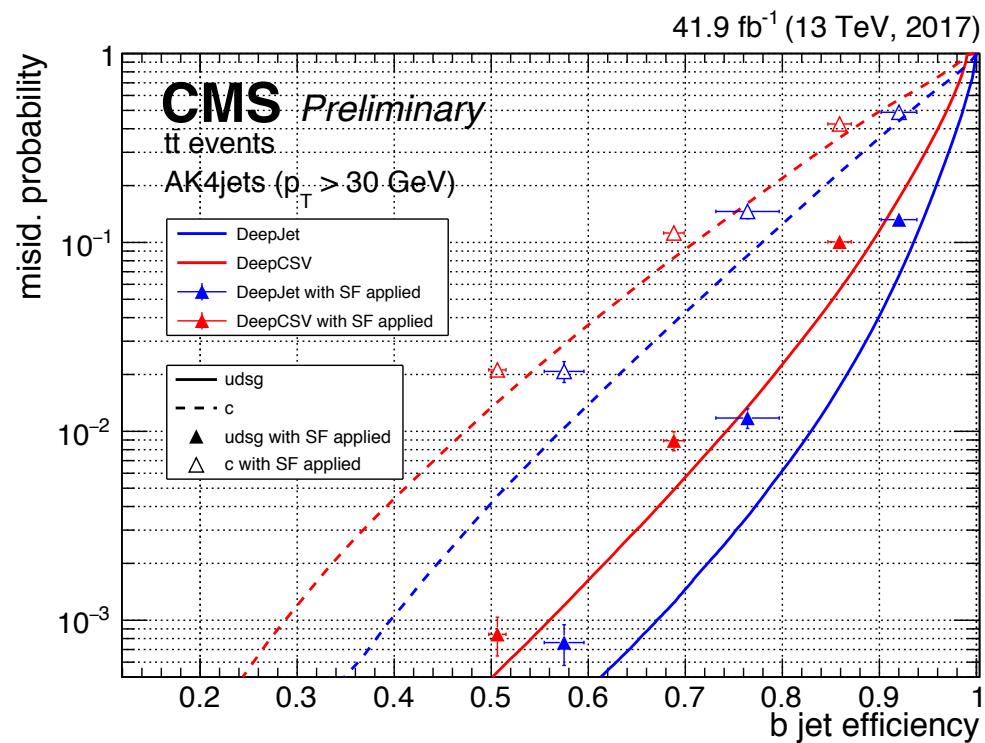
- To face **harsher Run 2 PU conditions**, both ATLAS and CMS upgraded their pixel detectors, **adding innermost pixel layer**:
  - ATLAS IBL installed during LS1
  - CMS Phase 1 pixel upgrade during EYETS 2016-2017

# b-tagging

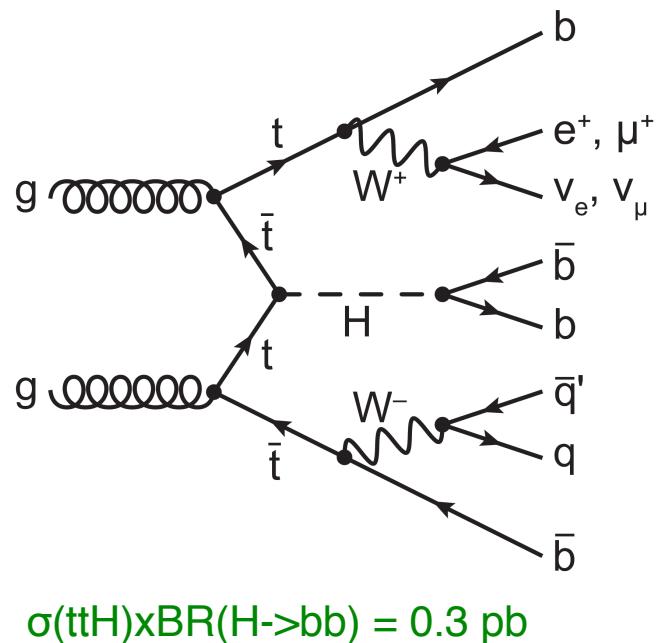
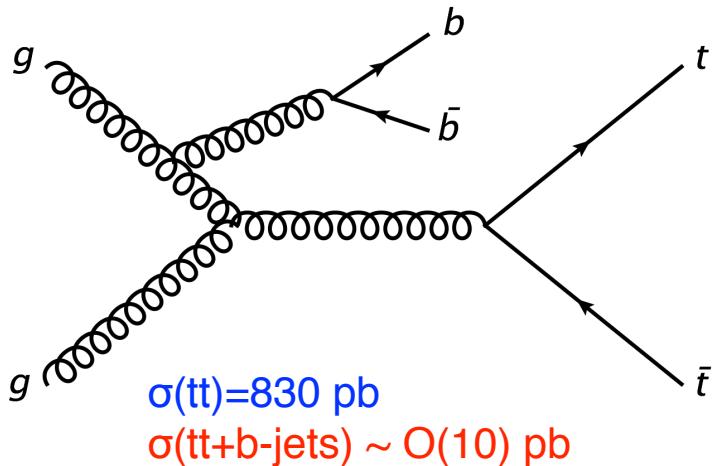


- New CMS pixel layer faced DC/DC converters failures of ~5% modules in 2017 (no issue in 2018)**
- Thanks to continuous improvements in b-tagging MVA taggers, overall CMS b-tagging performance maintained or improved over full Run 2**

- b-tagging in ATLAS based on BDT trained in ttbar simulation to discriminate b-jets from mix of light- and c-jets**
- Improved performance obtained for the whole Run 2 dataset thanks to upgraded detector + updated MVA discriminant**



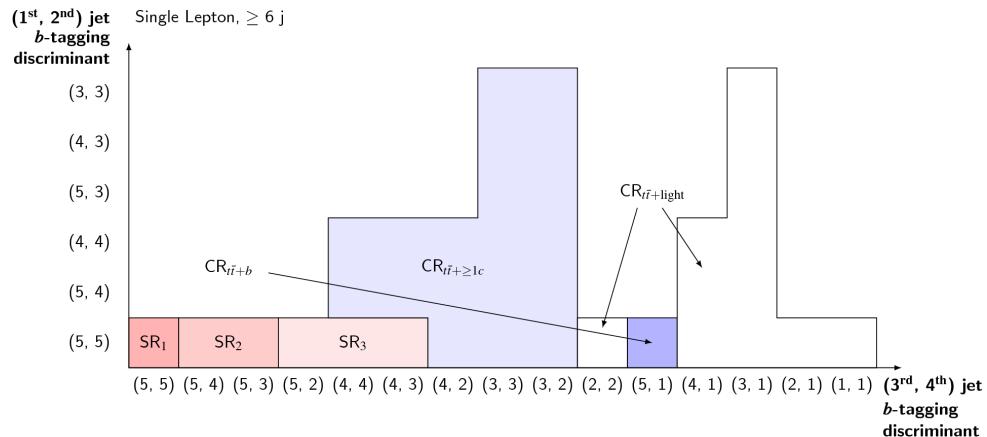
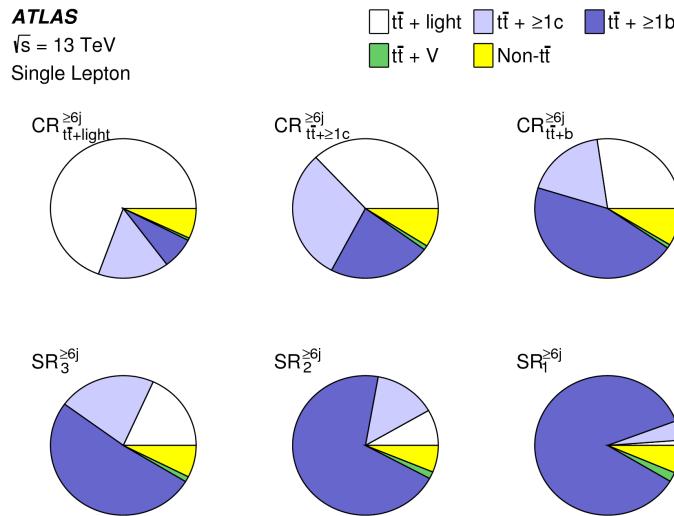
- Focus on H $\rightarrow$ bb + 1 or 2 top leptonic decays  
**=> strongly reduces QCD multijet contamination**
- **Main challenges:**  
 Very large tt+jets background  
 Large theory uncertainties on tt+bb modeling



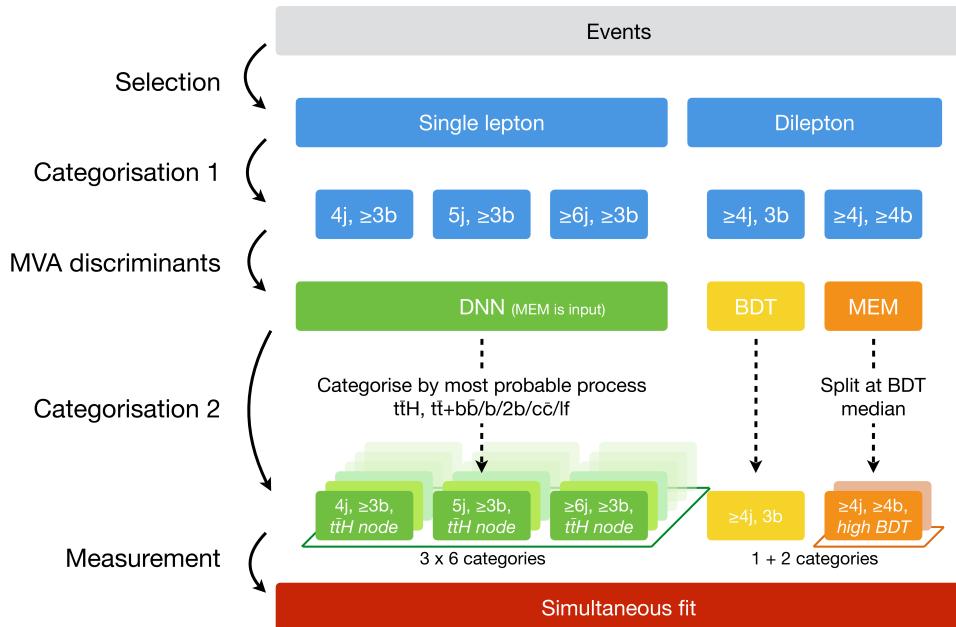
- **Similar strategy for both ATLAS and CMS analyses:**
  - target lepton+jets and dilepton ttbar decays
  - categorization based on number of jets / b-jets + boosted events for ATLAS
  - build discriminants to separate **signal** from **background** in each category

# Event classification

- Event classification in ATLAS analysis directly based on b-tagging discriminants of jets

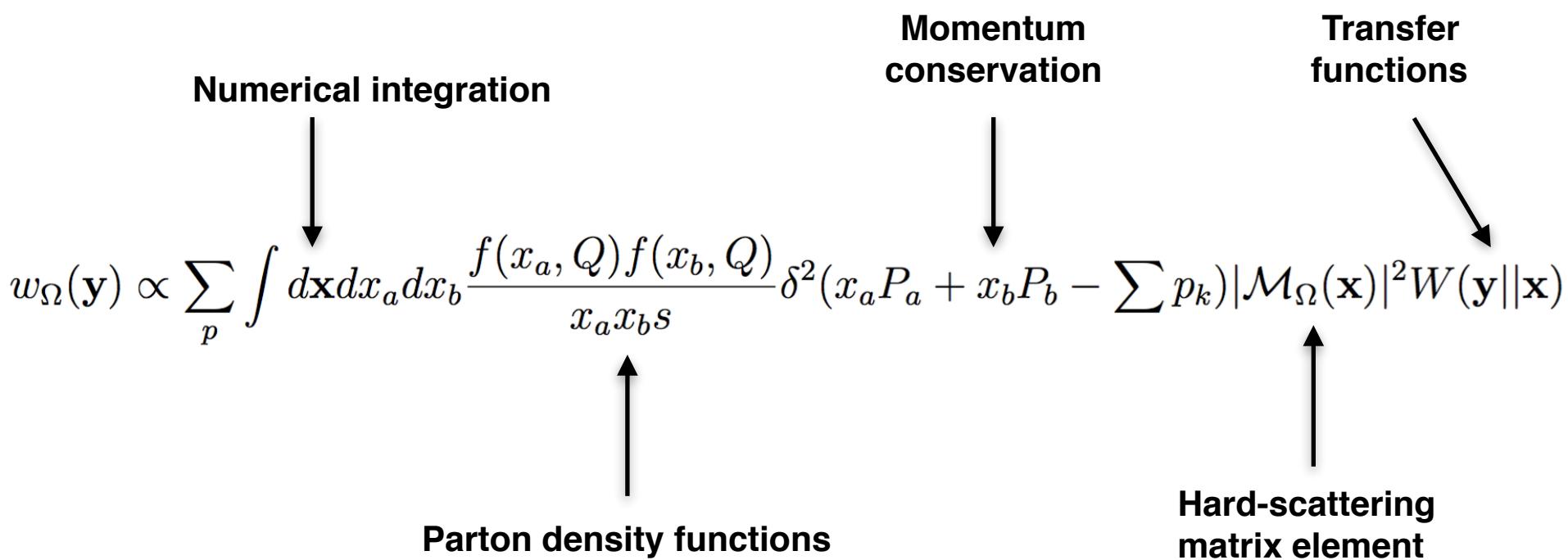


- Event classification in CMS analysis more sophisticated:
  - SL: based on **DNN multi-classifier**
  - DL: based on **BDT**
- **Background-dominated categories used to constrain various background components**



# Matrix Element Method

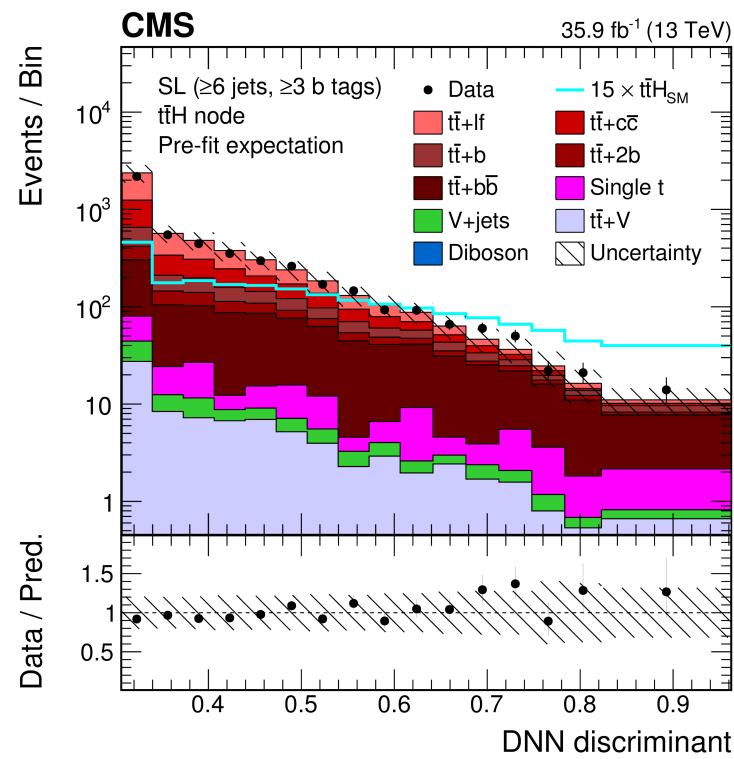
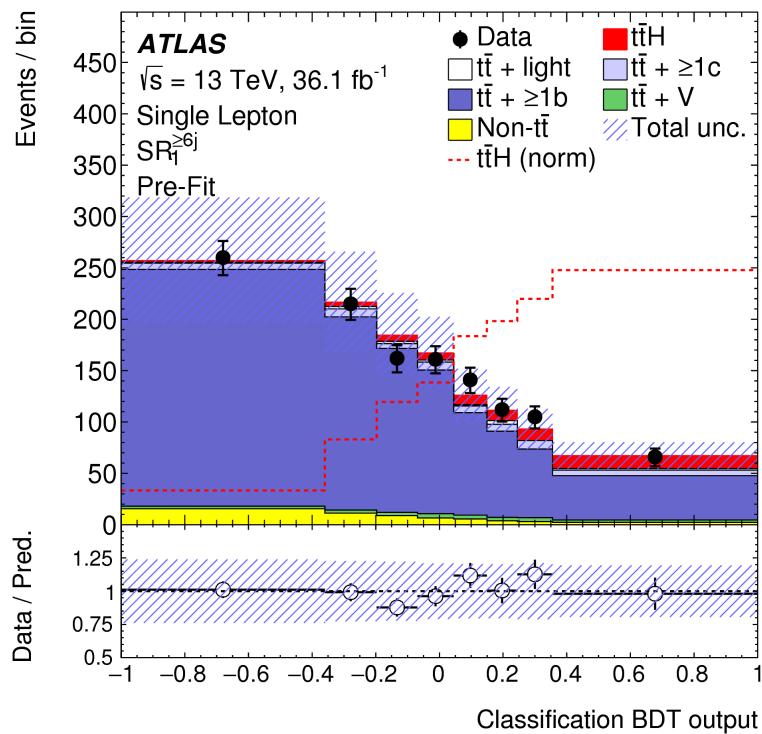
- Signal extraction discriminant optimised to separate **ttH signal** from irreducible **tt+bb background**, based on the **Matrix Element Method**
- **Event weight computed for hypothesis  $\Omega$**  ( $\Omega = \text{ttH}, \text{ttbb}, \text{ttV} \dots$ ), using observables  $\mathbf{y}$  as inputs and **integrating** over unmeasured or poorly measured quantities  $\mathbf{x}$

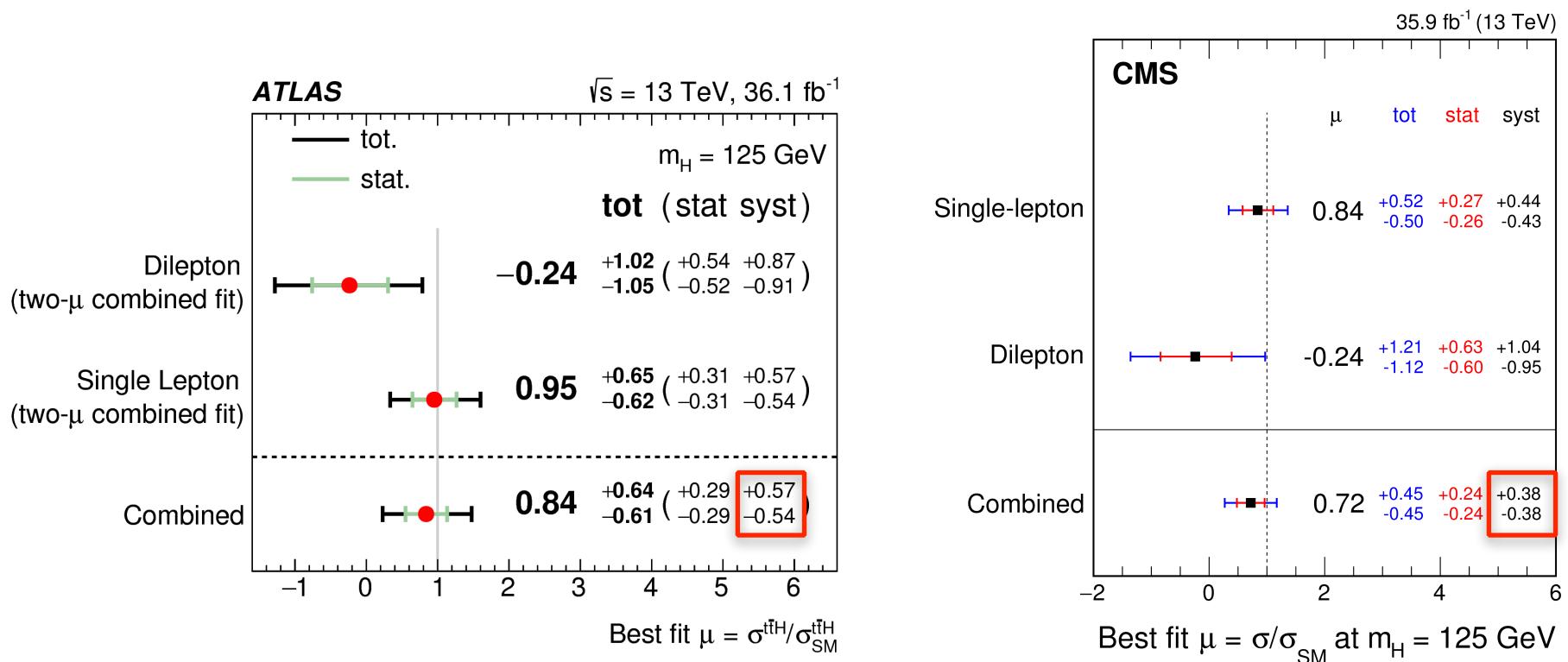


- Signal and background weights combined into **likelihood ratio**

# Signal extraction

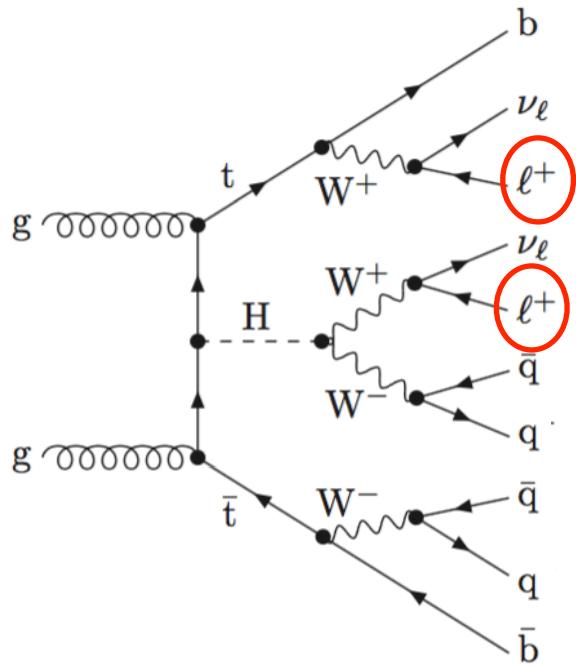
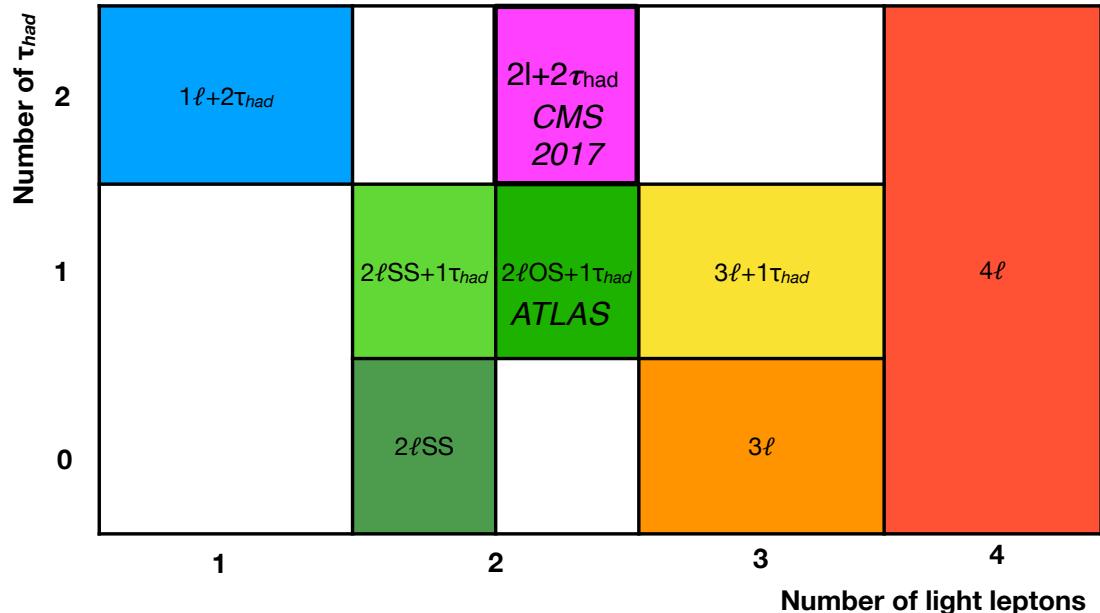
- ATLAS analysis exploits distribution of **classification BDT** using as inputs:
  - general kinematic variables
  - variables based on event reconstruction BDT
  - likelihood-discriminant
  - MEM
- CMS signal extraction in SL categories based on **DNN output of corresponding node** (including MEM as input + first layers for event reconstruction)
- In DL categories, based on **MEM discriminant** (after BDT-based categorisation)





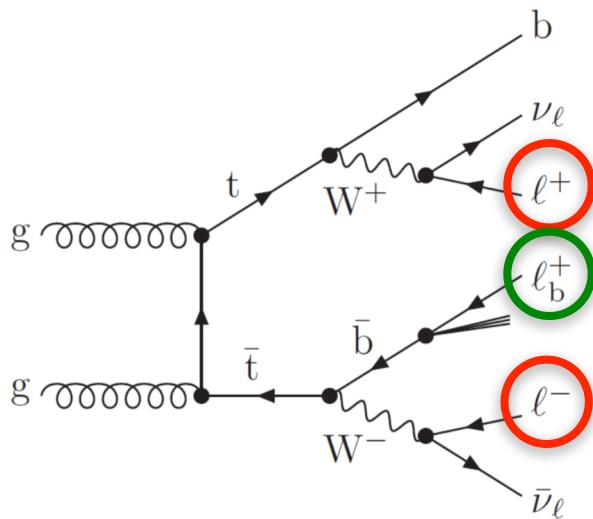
- **Significance:** ATLAS  $1.4\sigma$  obs. ( $1.6\sigma$  exp.) CMS  $1.6\sigma$  obs. ( $2.2\sigma$  exp.)
- **Similar results**, compatible with SM
- **Systematically limited** by modelling of tt+heavy-flavour background

- Focus on  $H \rightarrow WW/ZZ/\tau\tau$  decays
- Requires at least **2 same-sign leptons** (electrons, muons, hadronic tau decays)  
**=> strongly reduces most of SM backgrounds**
- Categories considered:

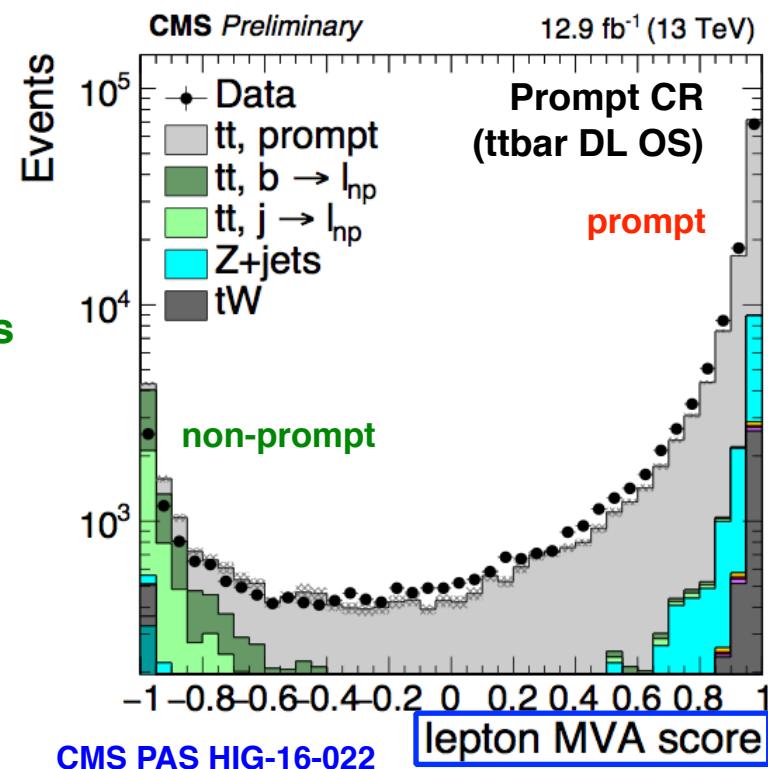


- Main sources of background:
  - **irreducible**: ttV, di-boson
  - **reducible**: fake  $\tau_{had}$ , non-prompt leptons and charge mis-ID (data-driven)

# Lepton identification



- Events with **non-prompt leptons** have a sizable contribution in all categories
- Contamination reduced using **MVA-based selections**
- BDT inputs:**
  - isolation
  - vertex
  - lepton ID
  - jet variables
- Residual background with **non-prompt leptons** evaluated using **tight-to-loose fake rate method**
- Main systematics:**
  - lepton selection efficiency
  - fake rate measurement
  - detailed flavour composition of fake leptons

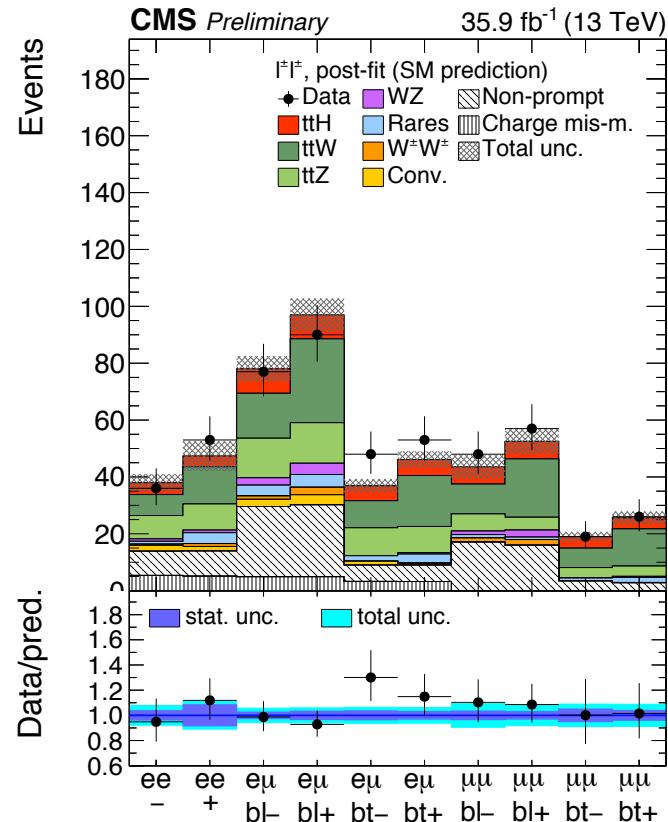


# Signal extraction

- CMS exploits subcategorisation of high stat categories based on:**
  - lepton charges
  - b-tagging requirements

- Directly exploited in signal extraction discriminant in ATLAS**

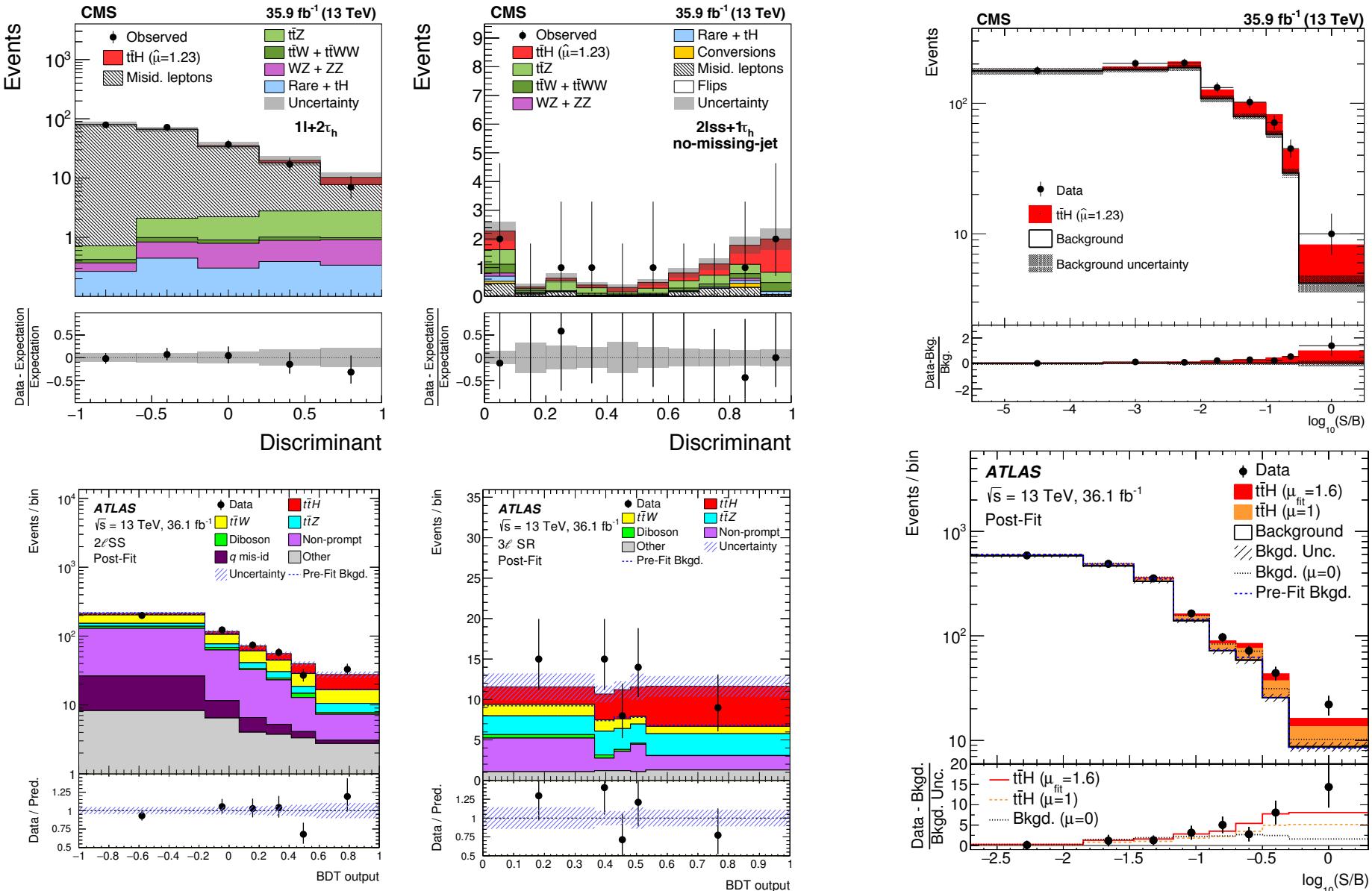
- Various MVA methods used for signal extraction in the different categories, using as inputs:**
  - lepton + jet kinematics
  - b-tagging
  - MEM or pseudo-ME



*Discriminants used for signal extraction in analysis of 2016 dataset*

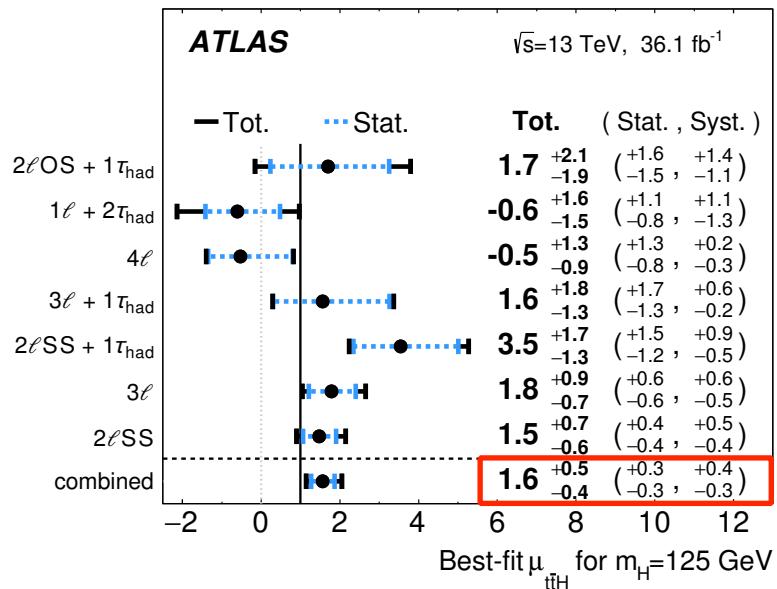
ATLAS	2 $\ell$ SS	3 $\ell$	4 $\ell$	1 $\ell$ +2 $\tau_{\text{had}}$	2 $\ell$ SS+1 $\tau_{\text{had}}$	2 $\ell$ OS+1 $\tau_{\text{had}}$	3 $\ell$ +1 $\tau_{\text{had}}$
BDT trained against Discriminant	Fakes and $t\bar{t}V$ 2×1D BDT	$t\bar{t}$ , $t\bar{t}W$ , $t\bar{t}Z$ , VV 5D BDT	$t\bar{t}Z$ / - Event count	$t\bar{t}$ BDT	all BDT	$t\bar{t}$ BDT	- Event count
CMS	2 $\ell$ SS	3 $\ell$	4 $\ell$	1 $\ell$ +2 $\tau_{\text{had}}$	2 $\ell$ SS+1 $\tau_{\text{had}}$	3 $\ell$ +1 $\tau_{\text{had}}$	-
BDT trained against Discriminant	Fakes and $t\bar{t}V$ 2×1D BDT	Fakes and $t\bar{t}V$ 2×1D BDT	- Event count	$t\bar{t}$ BDT	- MEM	Fakes and $t\bar{t}V$ 2×1D BDT	

# Visualising the signal

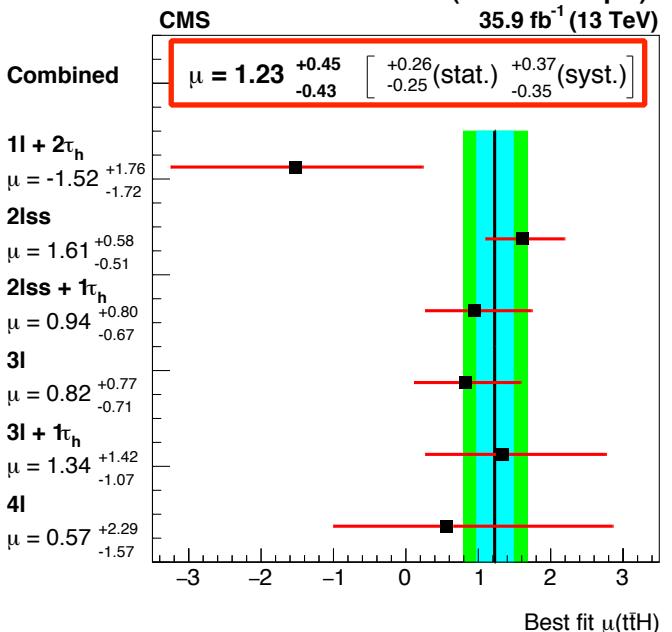


- 2016 dataset:

ATLAS 4.1 $\sigma$  obs. (2.8 $\sigma$  exp.)



CMS 3.2 $\sigma$  obs. (2.8 $\sigma$  exp.)



- Similar statistical and systematic uncertainties

- New 2017 CMS preliminary result with ttV floating normalisations

$$\begin{array}{ll} 2016 & \mu = 1.04^{+0.50}_{-0.36} \\ 2017 & \mu = 0.75^{+0.46}_{-0.43} \\ 2016+2017 & \mu = 0.96^{+0.34}_{-0.31} \end{array}$$

2.7 $\sigma$  obs. (2.7 $\sigma$  exp.)

1.7 $\sigma$  obs. (2.9 $\sigma$  exp.)

3.2 $\sigma$  obs. (4.0 $\sigma$  exp.)

$$\mu_{t\bar{t}Z} = 1.69^{+0.39}_{-0.33} \quad \mu_{t\bar{t}W} = 1.42^{+0.34}_{-0.33}$$

Trend consistent with dedicated ttV measurements in CMS + ATLAS

- Higgs decay modes with high mass resolution are "cleaner":
  - events can be selected with **high purity**
  - tt and H part of the event can be **cleanly separated**

- General strategy:** reconstruct Higgs boson decay + tag production mode with dedicated categories

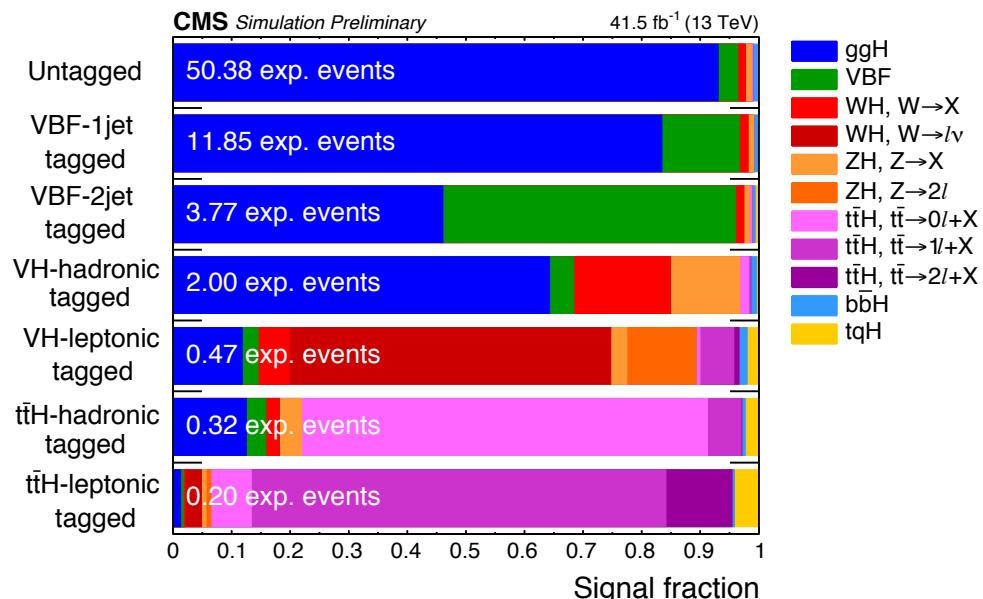
- H->4l:**

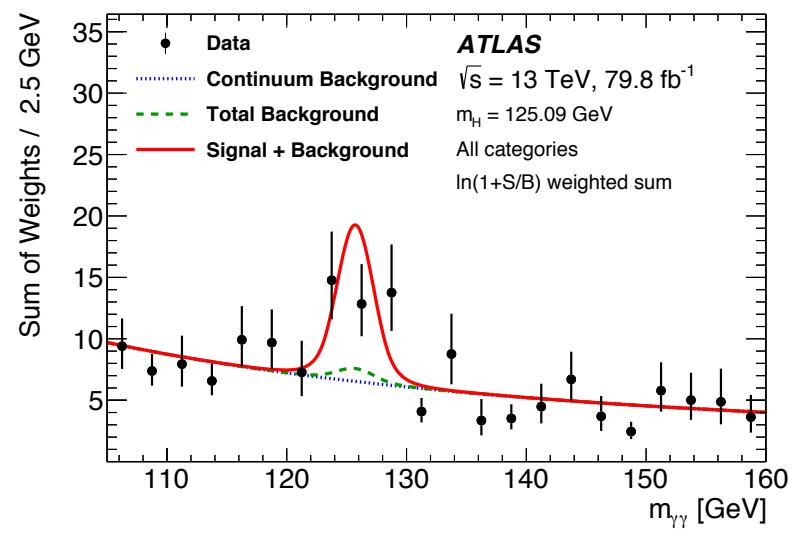
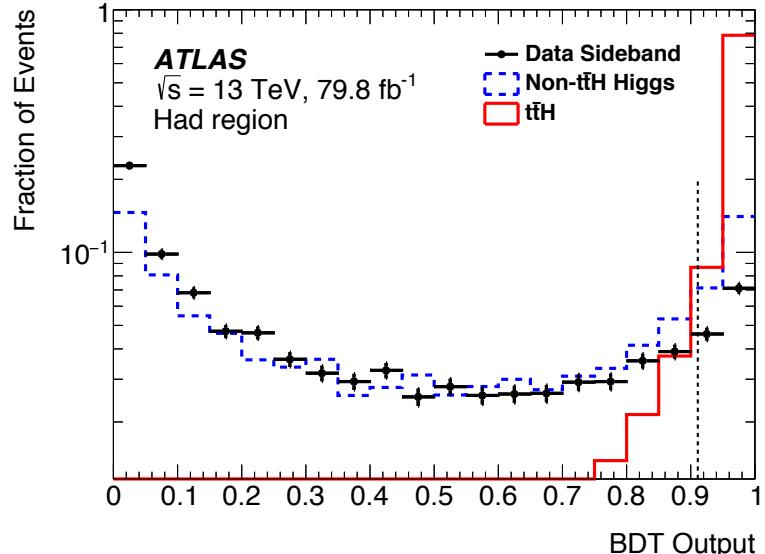
**ATLAS:** 4l +  $\geq 1$  b-jet + ( $\geq 1$  l +  $\geq 2$  jets OR  $\geq 4$  jets + **BDT-based categorization**)

**CMS:** 4l +  $\geq 1$  l (not VH-tagged) OR  $\geq 4$  jets,  $\geq 1$  b-jet

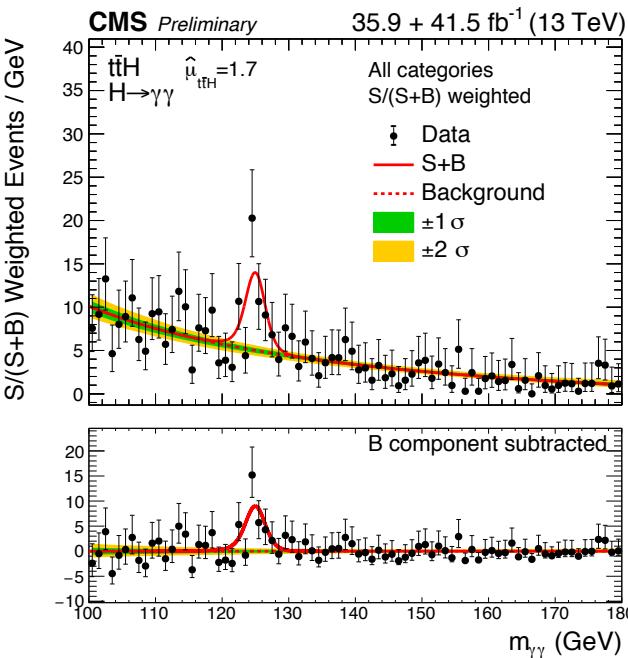
- Both in CMS and ATLAS:**

0.5 signal event expected in mass window around 125 GeV / 0 event observed

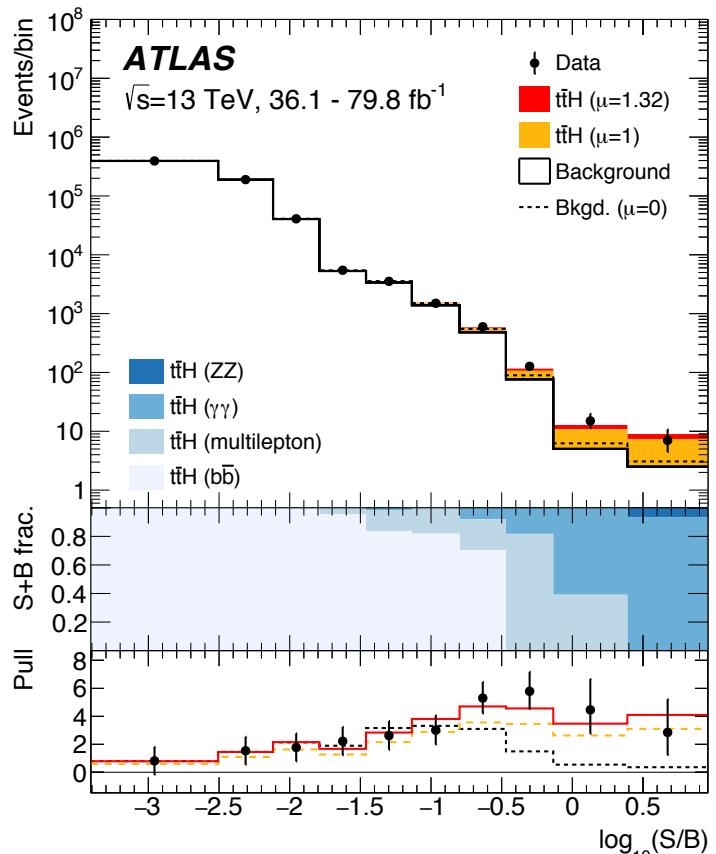
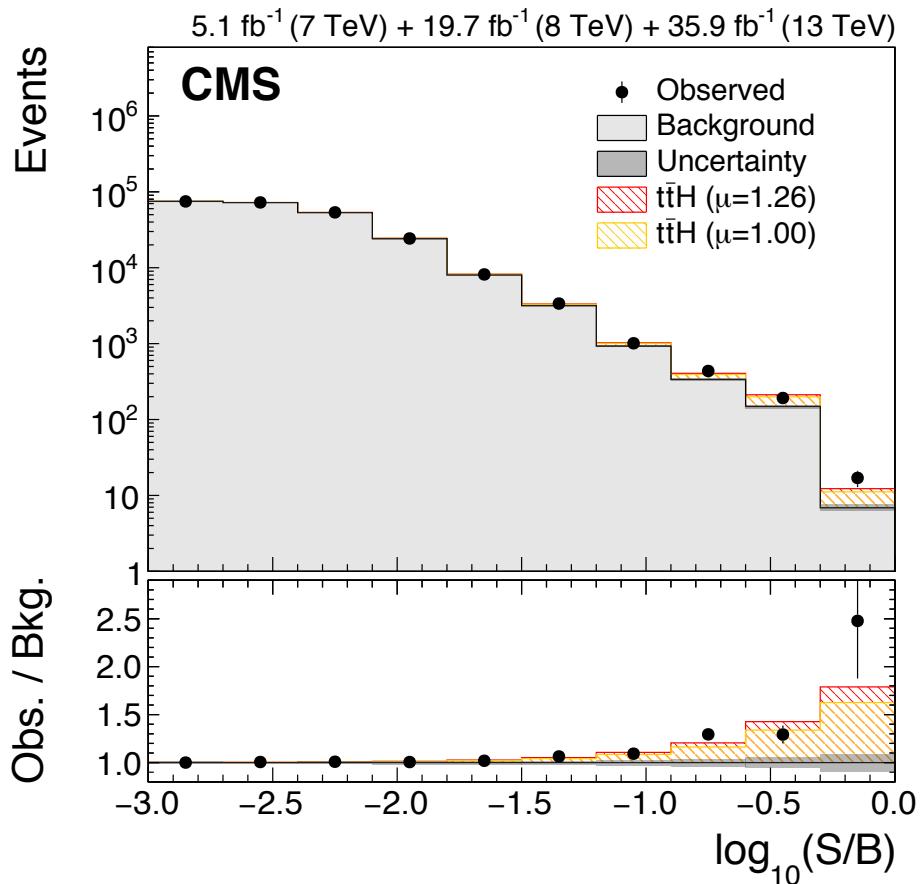




- Two channels defined: ttH hadronic and leptonic
- BDTs used to reject backgrounds + define categories with increasing purities
- Diphoton mass  $m(\gamma\gamma)$  used for signal extraction
- Results based on data collected up to 2017



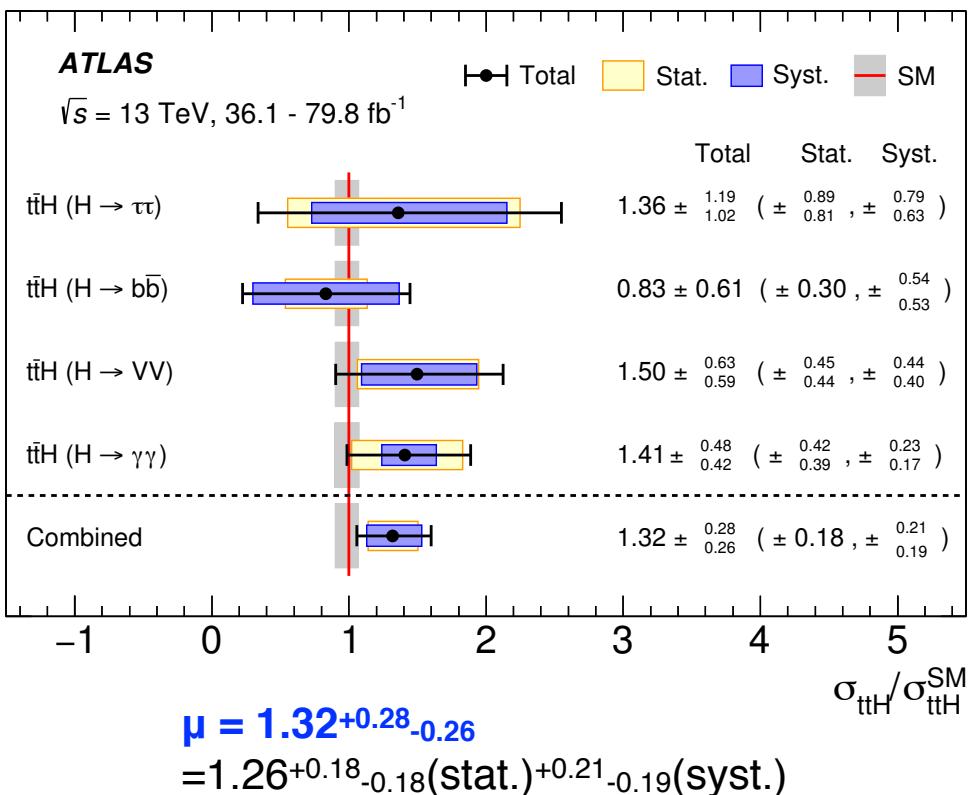
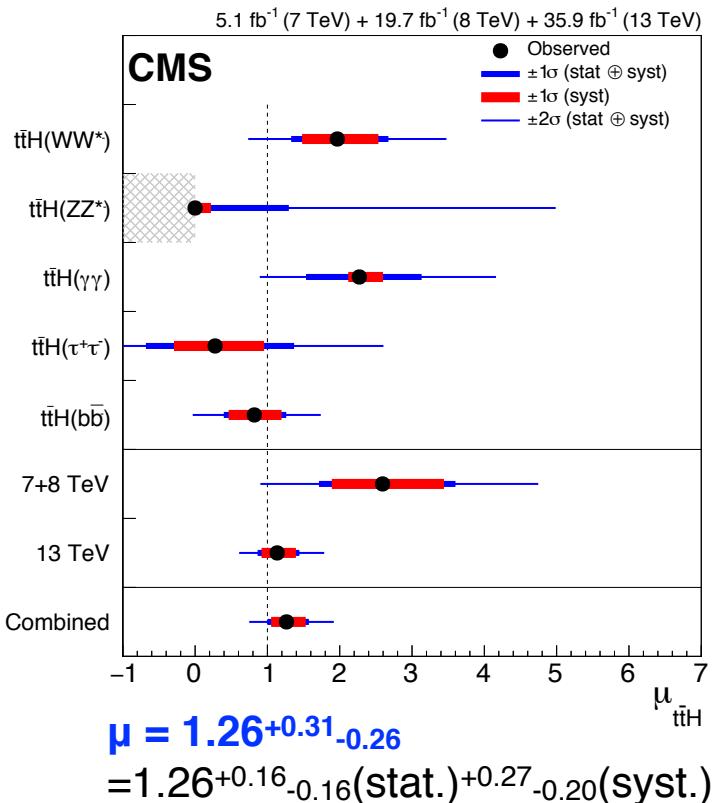
- Combinations of all available Run 1 + Run 2 ttH analyses released in Spring 2018:
  - CMS: Run 1 + Run 2 2016 datasets
  - ATLAS: Run 1 + Run 2 2015-2017 dataset



# ttH observation

Phys. Rev. Lett. 120 (2018) 231801 arXiv:1804.02610 [hep-ex]  
 Phys. Lett. B 784 (2018) 173 arXiv:1806.00425 [hep-ex]

- Measured signal strengths compatible with the SM expectation within error



- First observations of ttH process!

Run 1	$3.2\sigma$ obs. ( $1.2\sigma$ exp.)
Run 2 (2016)	$4.5\sigma$ obs. ( $4.1\sigma$ exp.)
Run 1+ Run 2	<b><math>5.2\sigma</math> obs. (<math>4.2\sigma</math> exp.)</b>

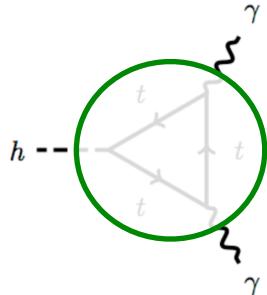
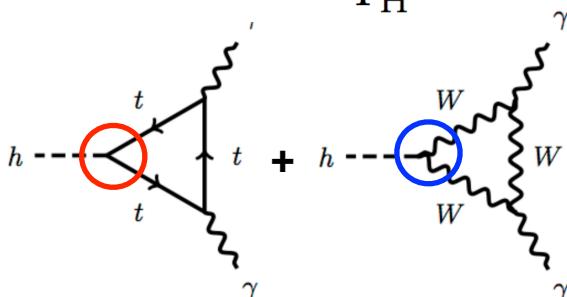
Run 1	$2.3\sigma$ obs. ( $1.5\sigma$ exp.)
Run 2 (2015-2016)	$4.2\sigma$ obs. ( $3.8\sigma$ exp.)
Run 2 (2015-2017)	$5.8\sigma$ obs. ( $4.9\sigma$ exp.)
Run 1+ Run 2	<b><math>6.3\sigma</math> obs. (<math>5.1\sigma</math> exp.)</b>

# Coupling measurements: the $\kappa$ framework

- Usually combined analysis of different final states needed to further disentangle contributions from different Higgs boson couplings

- $\kappa$  = coupling modifier

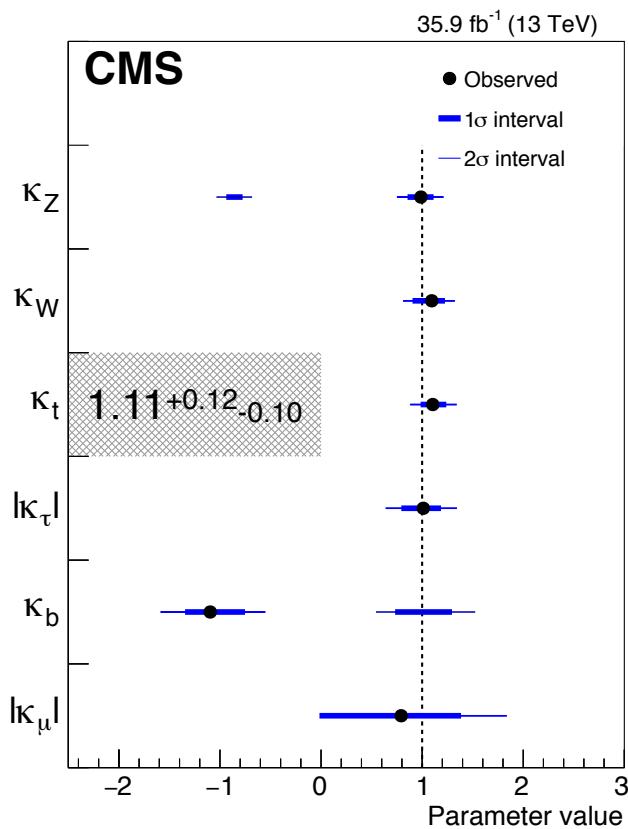
$$\sigma_i \cdot \text{BR}^f = \frac{\sigma_i(\vec{\kappa}) \cdot \Gamma^f(\vec{\kappa})}{\Gamma_H}$$



Production	Loops	Interference	Effective scaling factor	Resolved scaling factor
$\sigma(\text{ggH})$	✓	$b - t$	$\kappa_g^2$	$1.04 \cdot \kappa_t^2 + 0.002 \cdot \kappa_b^2 - 0.038 \cdot \kappa_t \kappa_b$
$\sigma(\text{VBF})$	-	-		$0.73 \cdot \kappa_W^2 + 0.27 \cdot \kappa_Z^2$
$\sigma(\text{WH})$	-	-		$\kappa_W^2$
$\sigma(\text{qq/qg} \rightarrow \text{ZH})$	-	-		$\kappa_Z^2$
$\sigma(\text{gg} \rightarrow \text{ZH})$	✓	$Z - t$		$2.46 \cdot \kappa_Z^2 + 0.47 \cdot \kappa_t^2 - 1.94 \cdot \kappa_Z \kappa_t$
$\sigma(\text{ttH})$	-	-	$\kappa_t^2$	
$\sigma(\text{gb} \rightarrow \text{WtH})$	-	$W - t$		$2.91 \cdot \kappa_t^2 + 2.40 \cdot \kappa_W^2 - 4.22 \cdot \kappa_t \kappa_W$
$\sigma(\text{qb} \rightarrow \text{tHq})$	-	$W - t$		$2.63 \cdot \kappa_t^2 + 3.58 \cdot \kappa_W^2 - 5.21 \cdot \kappa_t \kappa_W$
$\sigma(\text{bbH})$	-	-		$\kappa_b^2$
Partial decay width				
$\Gamma_{ZZ}$	-	-		$\kappa_Z^2$
$\Gamma_{WW}$	-	-		$\kappa_W^2$
$\Gamma_{\gamma\gamma}$	✓	$W - t$	$(\kappa_\gamma^2)$	$1.59 \cdot \kappa_W^2 + 0.07 \cdot \kappa_t^2 - 0.67 \cdot \kappa_W \kappa_t$
$\Gamma_{\tau\tau}$	-	-		$\kappa_\tau^2$
$\Gamma_{bb}$	-	-		$\kappa_b^2$
$\Gamma_{\mu\mu}$	-	-		$\kappa_\mu^2$
Total width for $\text{BR}_{\text{BSM}} = 0$				
$\Gamma_H$	✓	-	$\kappa_H^2$	$0.58 \cdot \kappa_b^2 + 0.22 \cdot \kappa_W^2 + 0.08 \cdot \kappa_g^2 + 0.06 \cdot \kappa_\tau^2 + 0.026 \cdot \kappa_Z^2 + 0.029 \cdot \kappa_c^2 + 0.0023 \cdot \kappa_\gamma^2 + 0.0015 \cdot \kappa_{Z\gamma}^2 + 0.00025 \cdot \kappa_s^2 + 0.00022 \cdot \kappa_\mu^2$

# top-Higgs coupling measurements

- Both CMS and ATLAS released combined measurements of Higgs boson couplings based on Run 2 dataset
- Results assuming resolved loops with no BSM contributions



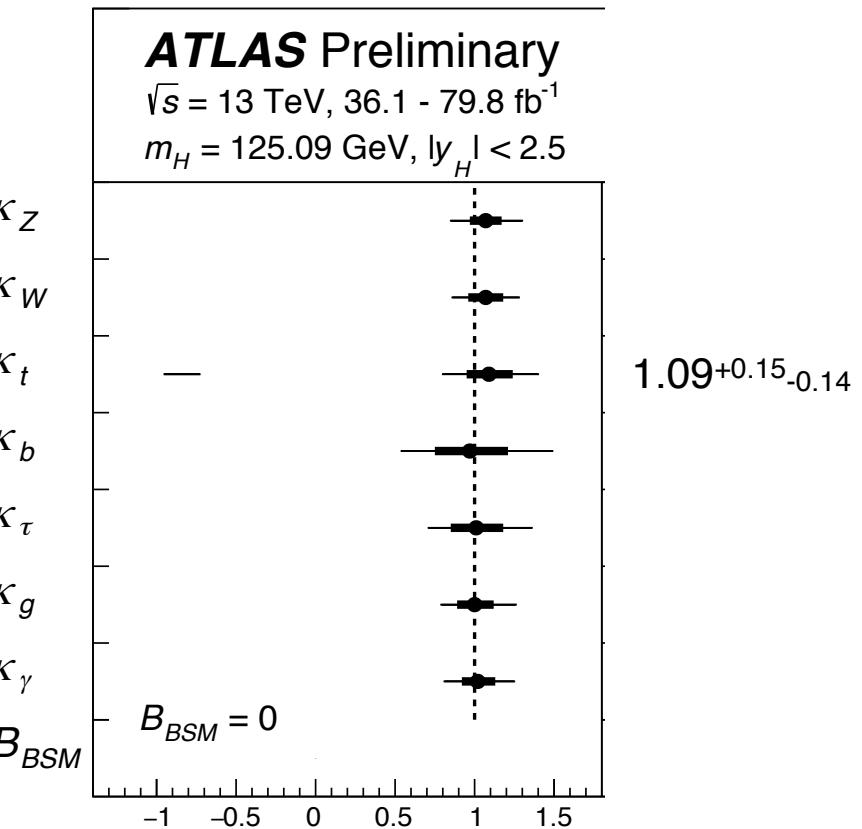
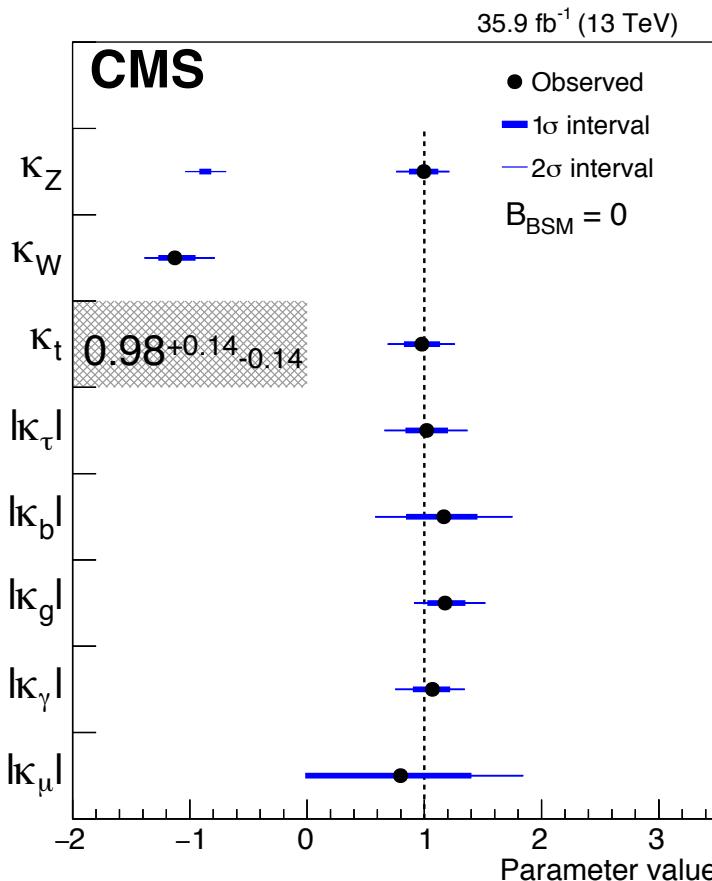
## ATLAS

Parameter	Result
$\kappa_Z$	$1.07^{+0.11}_{-0.10}$
$\kappa_W$	$1.04 \pm 0.10$
$\kappa_b$	$1.00^{+0.24}_{-0.22}$
$\kappa_t$	$1.03^{+0.12}_{-0.11}$
$\kappa_\tau$	$1.04^{+0.17}_{-0.16}$
$\kappa_\mu$	< 1.63 at 95% CL.

- Reduction of uncertainty on  $\kappa_t$  by ~40% wrt ATLAS+CMS Run 1 combination!

# top-Higgs coupling measurements

- Results with effective coupling modifiers to gluons and photons = constraint only from ttH



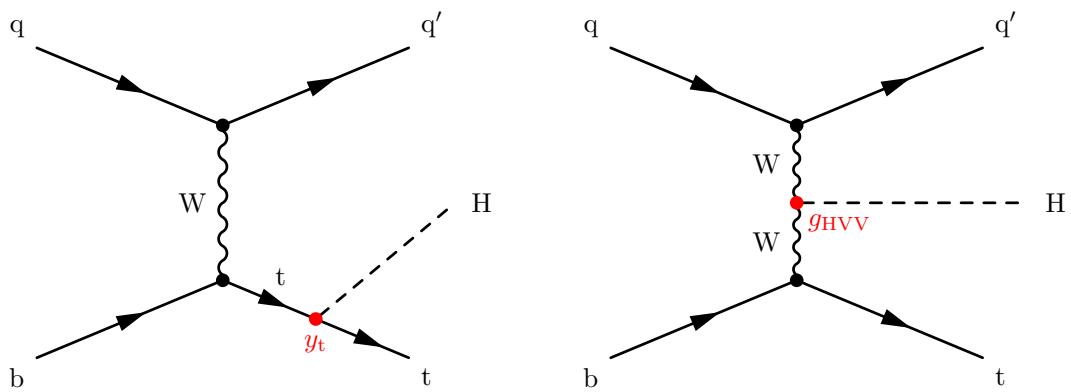
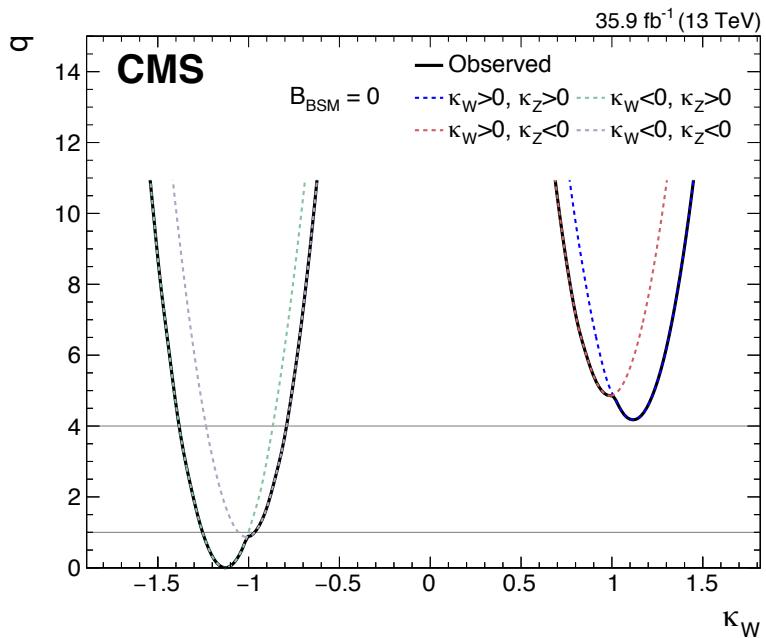
- Constraint on  $\kappa_t$  ~25% looser than with resolved assumption
- But CMS fits  $\kappa_t * \kappa_W$  negative ?!?

# Single-top + Higgs production

- Single-top + Higgs associated production very rare SM process:  
~10 times smaller than ttH

- But very sensitive to  $\kappa_t^* \kappa_W$  sign due to tree-level interference between diagrams:

SM cross-section \* 10 for  $\kappa_t^* \kappa_W = -1$



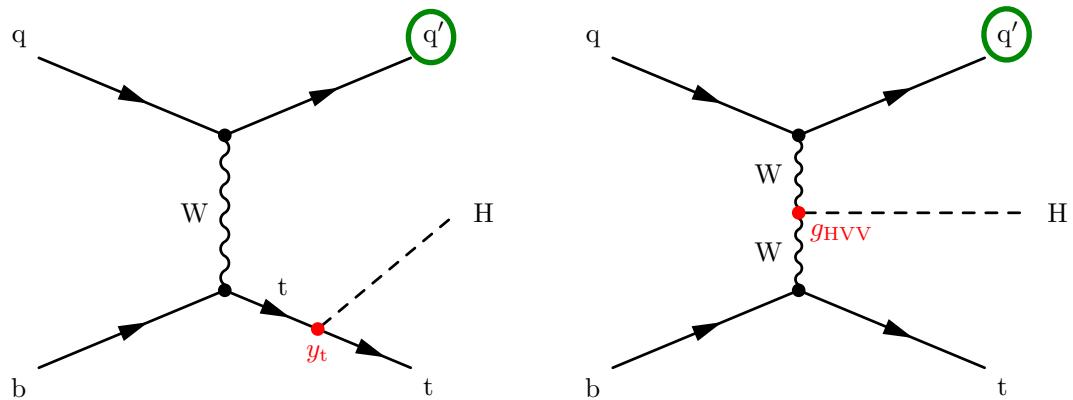
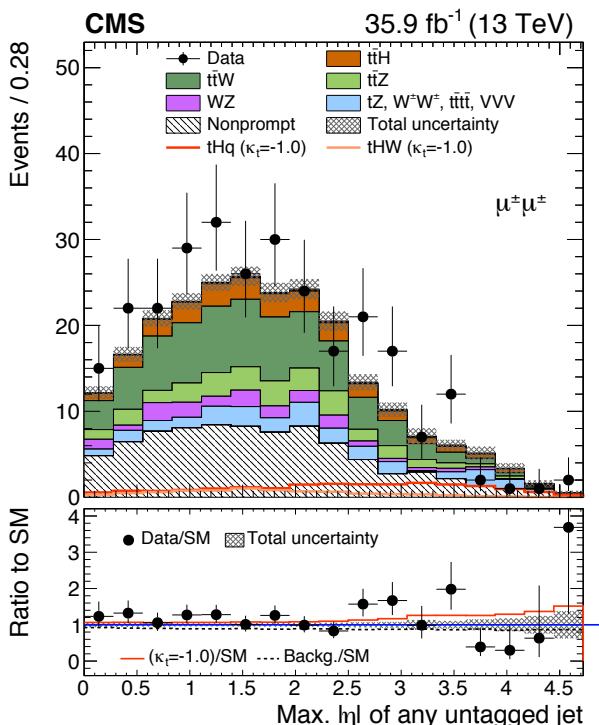
$$\sigma_{tHq} = (2.63 \kappa_t^2 + 3.58 \kappa_V^2 - 5.21 \kappa_t \kappa_V) \sigma_{tHq}^{\text{SM}},$$

$$\sigma_{tHW} = (2.91 \kappa_t^2 + 2.31 \kappa_V^2 - 4.22 \kappa_t \kappa_V) \sigma_{tHW}^{\text{SM}}.$$

- CMS negative measurement of  $\kappa_W$  related to ttH  $H \rightarrow \gamma\gamma$  observed excess:
  - no constraint from  $H \rightarrow \gamma\gamma$  loop decay
  - $|\kappa_t|$  constrained by other ttH categories
  - excess fitted with increase of single-top + Higgs cross-section in ttH categories

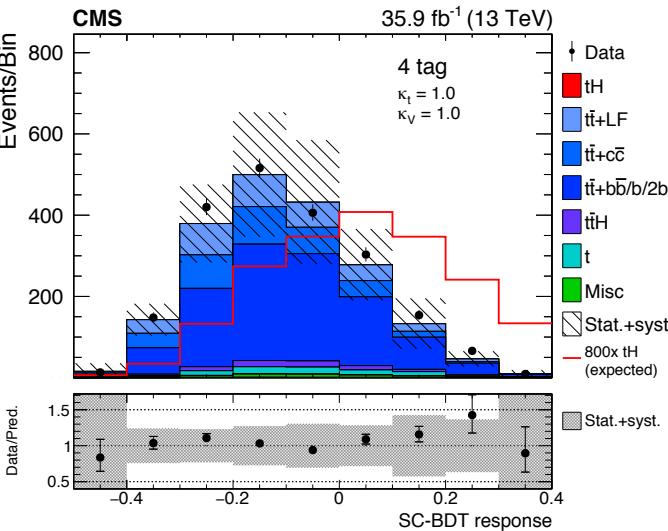
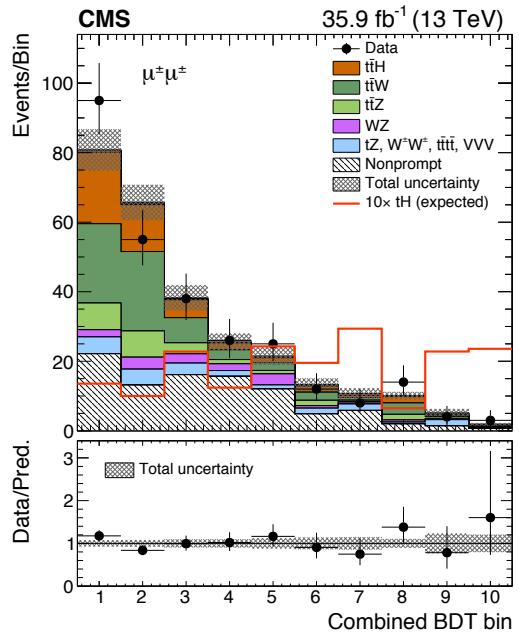
# Single-top + Higgs production

- Dedicated search performed by CMS: combination of dedicated  $tH$   $bb$  + multilepton analyses +  $\gamma\gamma$  reinterpretation
- Distinctive feature of  $tHq$  wrt  $ttH$ : additional forward jet

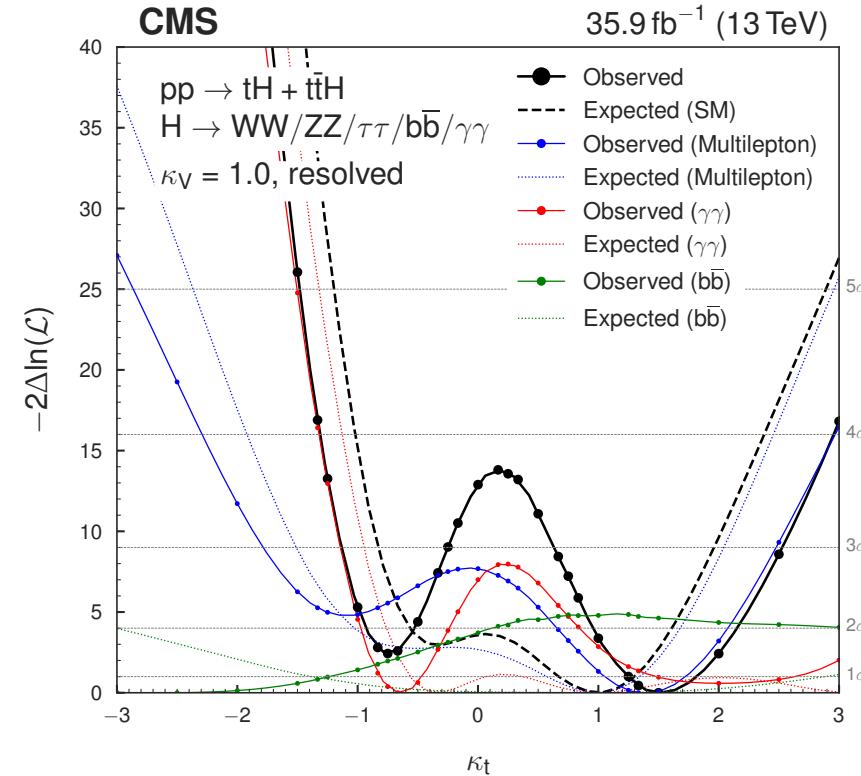


- Multilepton analysis similar to  $ttH$  w/ 2lss + 3l categories but:
  - considers forward jet for event selection
  - makes use of BDTs including inputs sensitive to forward jets + trained to distinguish  $tHq$  vs  $ttV$  /  $tHq$  vs  $tt\bar{b}$
- $bb$  analysis focuses on single lepton + 3/4 b-jets events:
  - makes use of jet assignment BDTs for event reconstruction under  $tHq$  /  $tHW$  /  $tt\bar{b}$  hypotheses
  - signal classification BDT used to distinguish  $tHq+tHW$  from  $tt\bar{b}$
  - flavour classification BT used to distinguish  $tt+$ light from  $tt+$ heavy flavour in dilepton control region

# Single-top + Higgs production

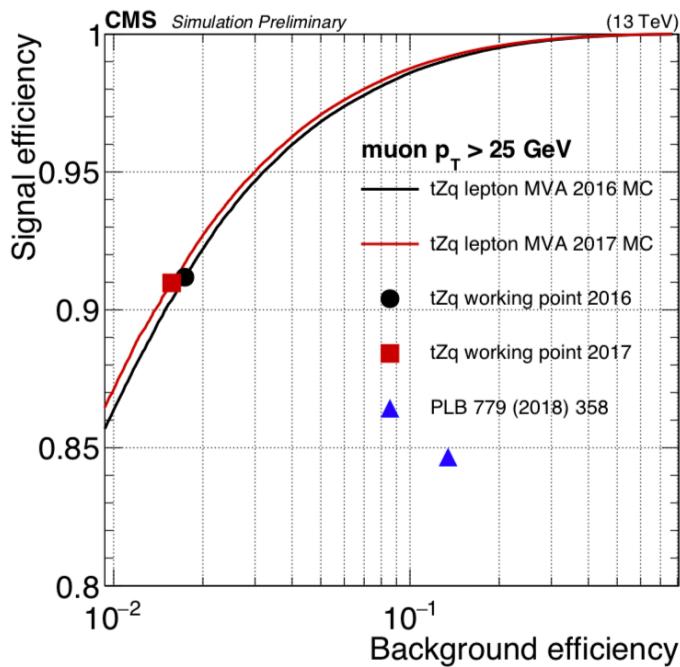


- $\kappa_t$  dependance on normalisation + shape of ttH / tHq / tHW fully taken into account for signal extraction including impact on branching ratios
- Sensitivity driven by multilepton channels for  $\kappa_t > 0$  and  $\gamma\gamma$  channels for  $\kappa_t < 0$  with positive value favoured

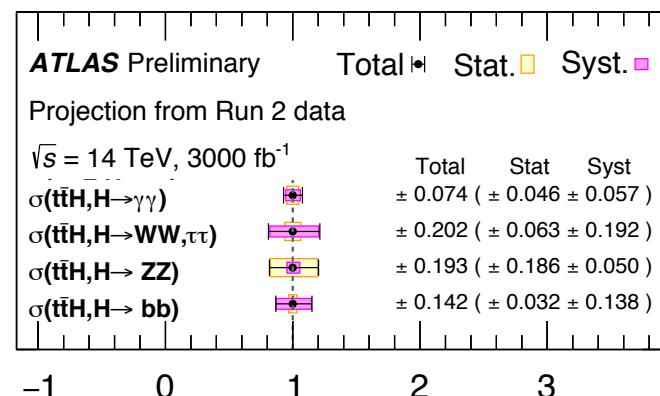


# Prospects

- $H \rightarrow bb$  and multilepton channels systematics limited by now  
 $\Rightarrow H \rightarrow \gamma\gamma + 4l$  (+  $\mu\mu$ ?) channels will ultimately become golden ones to study ttH production
- Study of differential cross-sections in ttH final states may cast some light on exact nature of Higgs couplings:  $pT(H)$  spectrum to constrain CP structure of top Yukawa coupling or Higgs self-coupling



## CMS tZq observation seminar



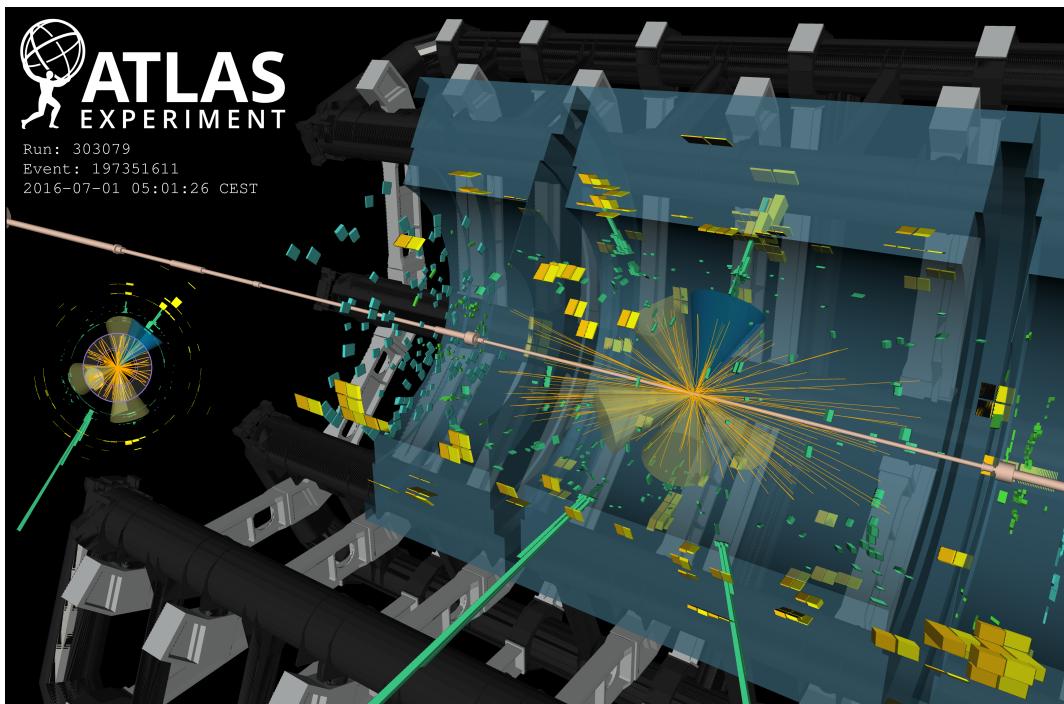
Cross section norm. to SM value

[ATLAS-PHYS-PUB-2018-054](#)  
[CMS-PAS-FTR-18-011](#)  
[CMS-PAS-FTR-18-019](#)

- But still, statistics will not increase fast before some time now
- In the meantime, makes sense to keep trying improving other channels as well: recent example of CMS tZq observation where switch to MVA-based lepton selection reduced by a factor 2 the systematics uncertainty!

# Conclusion

- Studying the top-Higgs coupling is one of the keys to better understand the mechanism of mass generation for fermions
- ttH process offers a direct probe to study it
- Observation of this process at LHC already before the end of Run 2!  
=> still possible to produce outstanding physics results beyond the Higgs discovery
- Possible thanks to significant analysis efforts relying heavily on MVA techniques, some of them exploiting cutting edge machine-learning developments
- Beyond ttH observation, detailed properties of top-Higgs coupling still to be studied in more details at LHC and some significant physics results are yet to come in that area



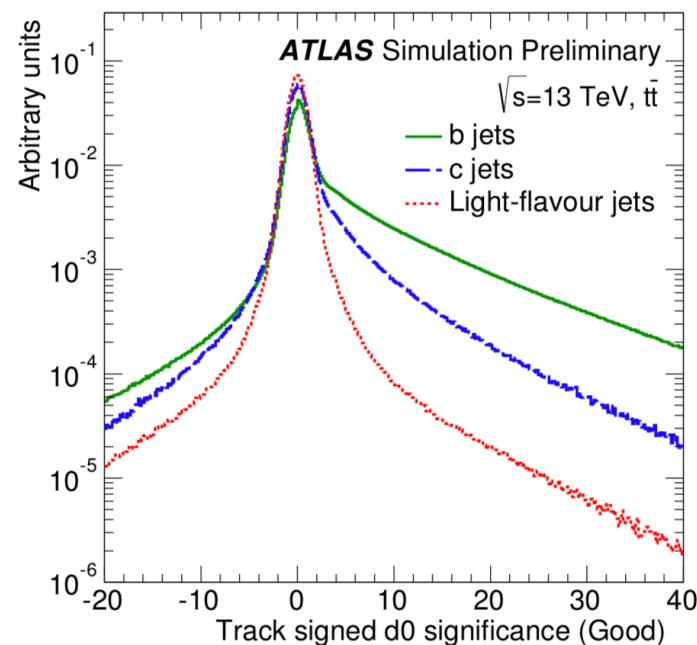
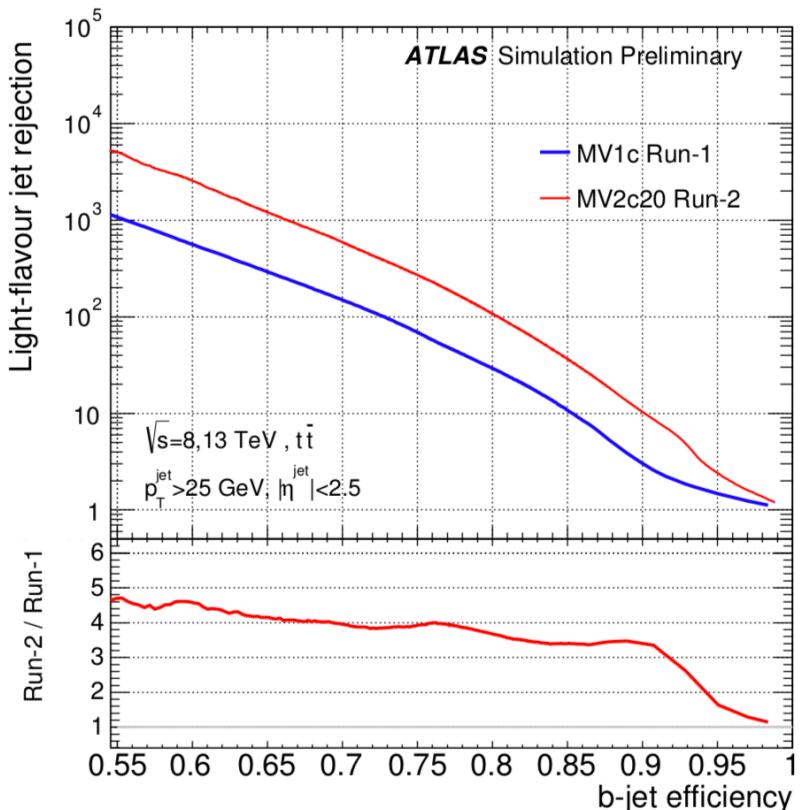
---

# **Back-up**

# b-tagging in ATLAS Run 2

- Relies on three kinds of variables:

- transverse and longitudinal impact parameter significance of individual tracks
- presence of a secondary vertex and related properties
- reconstruction of B-hadron decay chain, including tertiary charm decay vertices

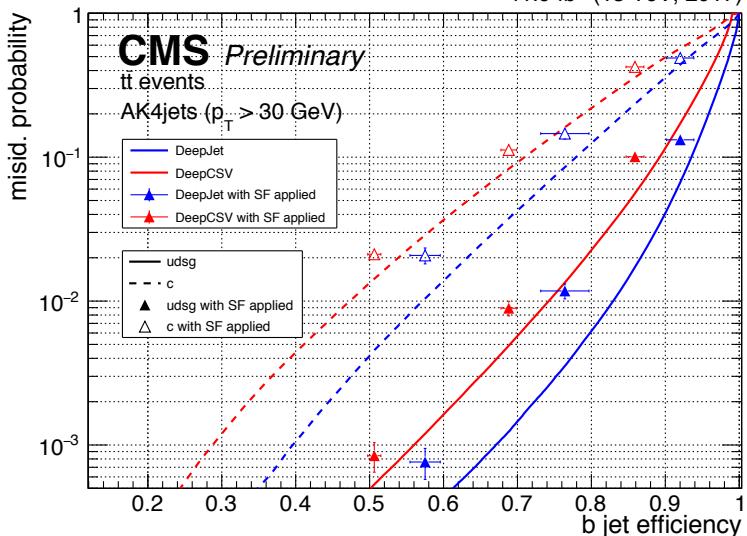
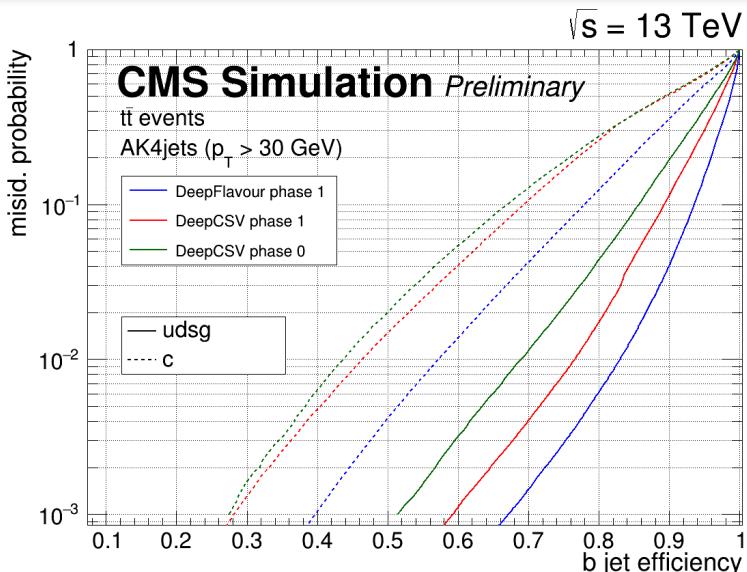


- Boosted Decision Tree trained in ttbar simulation to discriminate b-jets from mix of light- and c-jets
- Improved performance obtained for the whole Run 2 dataset thanks to upgraded detector + updated MVA discriminant

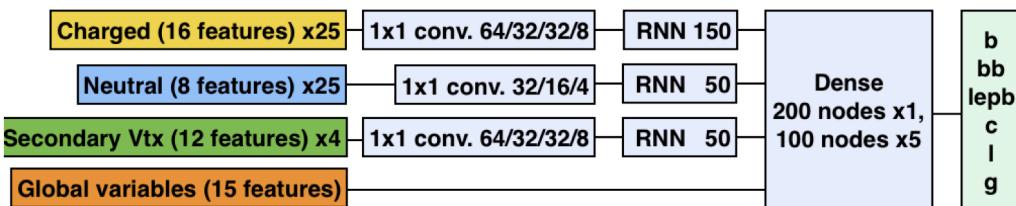
# b-tagging in CMS Run 2

CMS-DP-2018-033 , CMS-DP-2018-058

JINST 13 (2018) P05011 arXiv:1712.07158 [hep-ex]

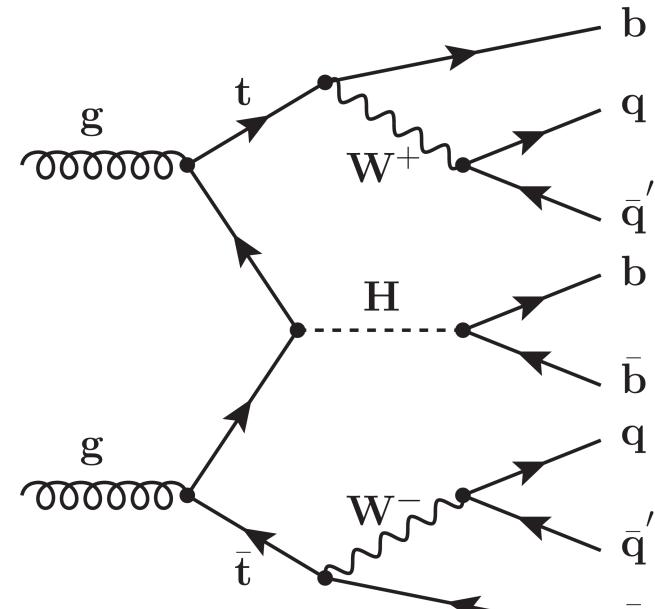
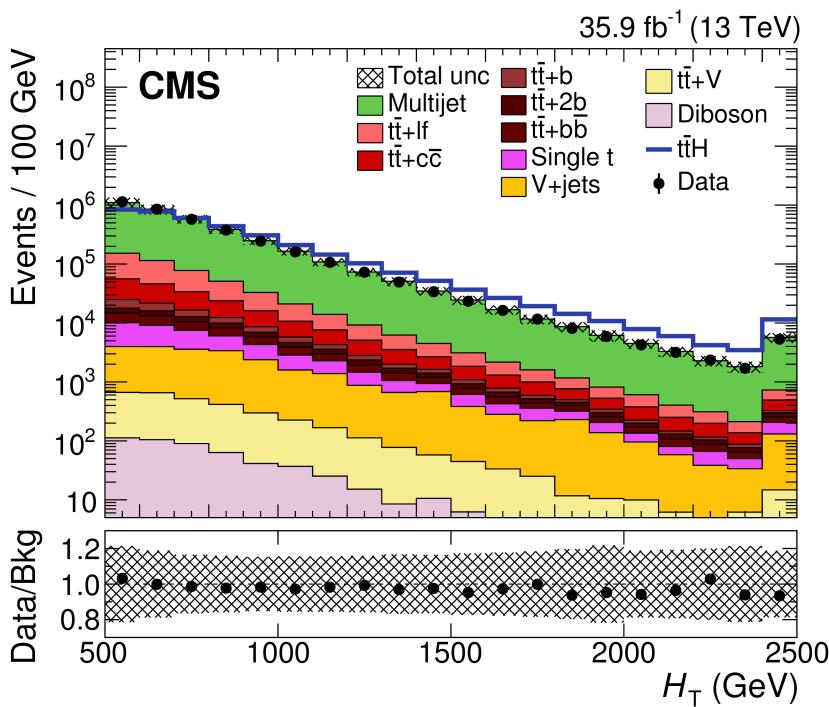


- Relies on similar input variables as in ATLAS (up to 2018)
- Combined using different MVA methods:
  - Run 1 CSV likelihood-based
  - 2015-2016 CSVv2 neural network
  - 2017 DeepCSV deep neural network
  - 2018 DeepJet deep neural network multiclassifier with low-level PF candidate inputs



- New pixel layer faced DC/DC converters failures of ~5% modules in 2017 (no issue in 2018)
- Thanks to continuous improvements in b-tagging MVA taggers, overall CMS b-tagging performance maintained or improved over full Run 2

- Targets highest branching ratio:  
 $H \rightarrow bb + 2$  top hadronic decays
- Most recent result by CMS based on 2016 dataset  
Run 1 result in this channel by ATLAS only
- Trigger:  $\geq 6$  jets,  $\geq 1$  or  $2$  b-tagged jets,  $H_T > 450/400$  GeV



- Event selections:
  - no lepton
  - $\geq 6$  jets
  - $\geq 2$  b-tagged jets
  - $H_T > 500$  GeV
  - at least a pair of light jets compatible with  $W$  mass
- Final state overwhelmed by QCD multijet production

# Quark/gluon discriminant

CMS-DP-2016-070

- All LO light jets expected to be quarks from W decays  
**=> MVA quark/gluon discriminant used to reduce QCD background**

- Likelihood-based q/g discriminant computed for each jet using:
  - #jet constituents
  - minor axis of jet ellipsis

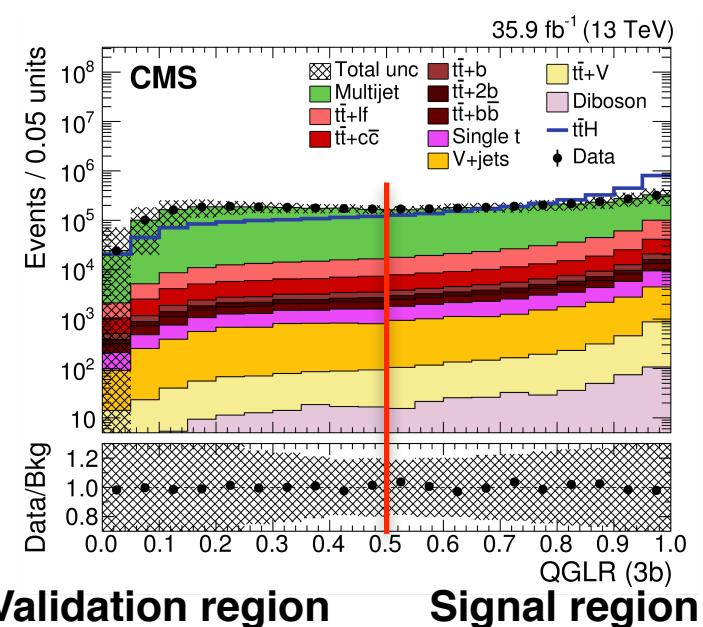
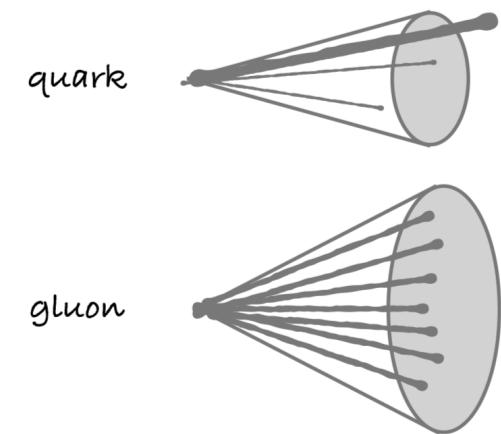
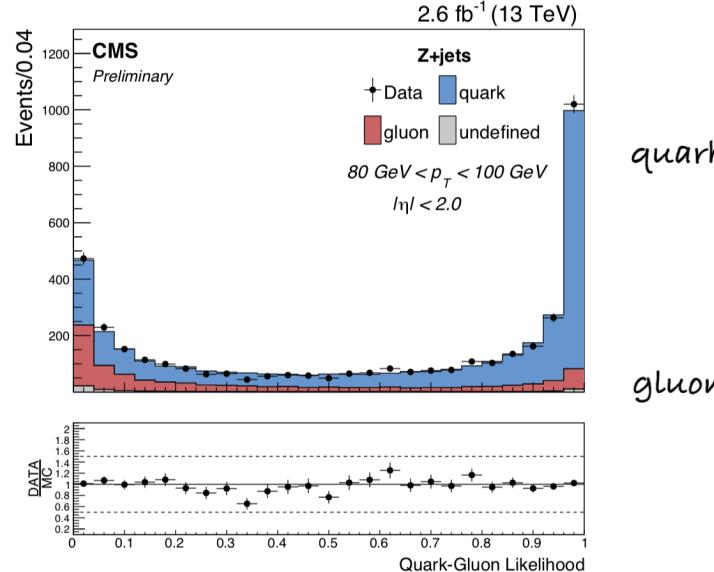
$$p_T D = \frac{\sqrt{\sum_i p_{T,i}^2}}{\sum_i p_{T,i}}$$

- Combined into an **event-level likelihood ratio**

$$q_{\text{LR}(N_1, N_2)} = \frac{L(N_1, 0)}{L(N_1, 0) + L(N_2, N_1 - N_2)}$$

$$L(N_q, N_g) = \sum_{\text{perm}} \left( \prod_{k=i_1}^{i_{N_q}} f_q(\zeta_k) \prod_{m=i_{N_q+1}}^{i_{N_q+N_g}} f_g(\zeta_m) \right)$$

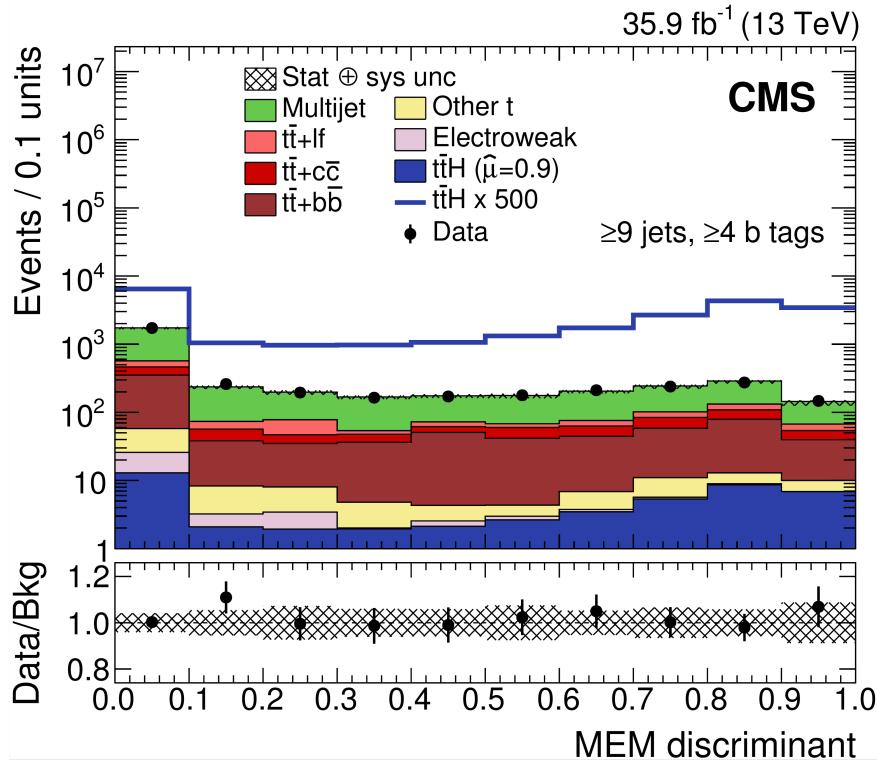
- QCD shape from control region / normalisation unconstrained for signal extraction**



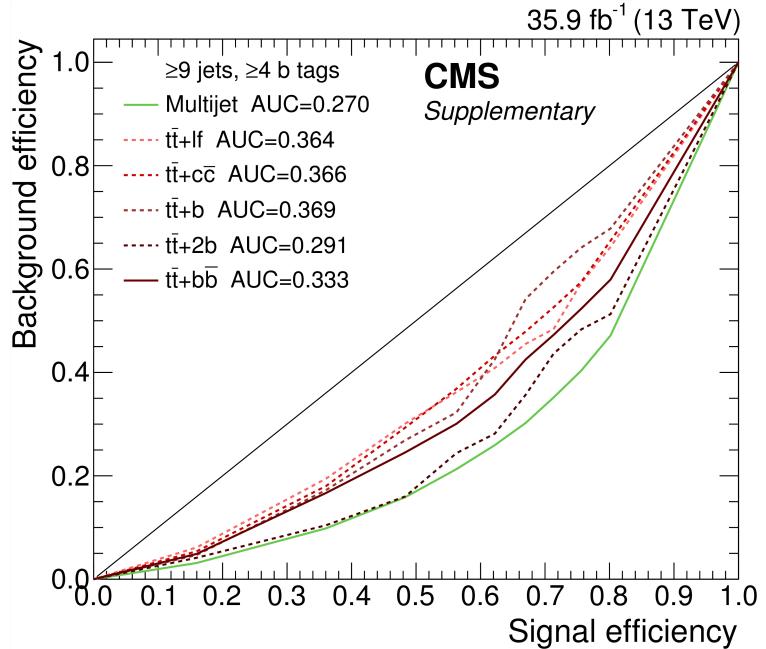
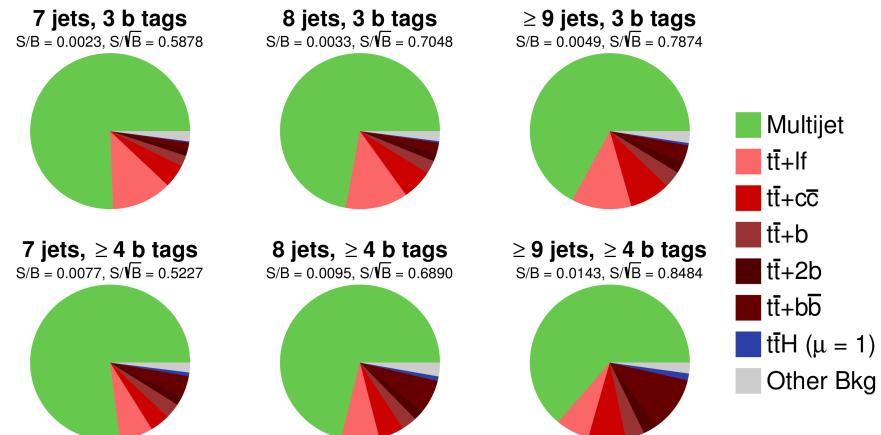
# ttH H $\rightarrow$ bb fully hadronic

JHEP 06 (2018) 101 arXiv:1803.06986 [hep-ex]

- Categories based on jet + b-jet multiplicity
- MEM computation optimised for each category: performant both against QCD and tt+bb

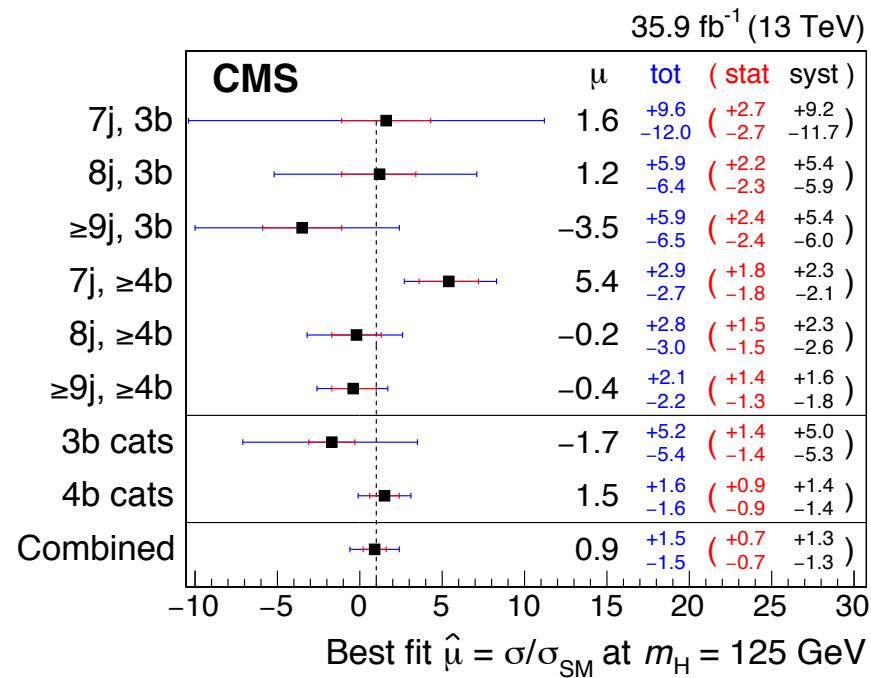


CMS Supplementary



# ttH $H \rightarrow bb$ fully hadronic

JHEP 06 (2018) 101 arXiv:1803.06986 [hep-ex]

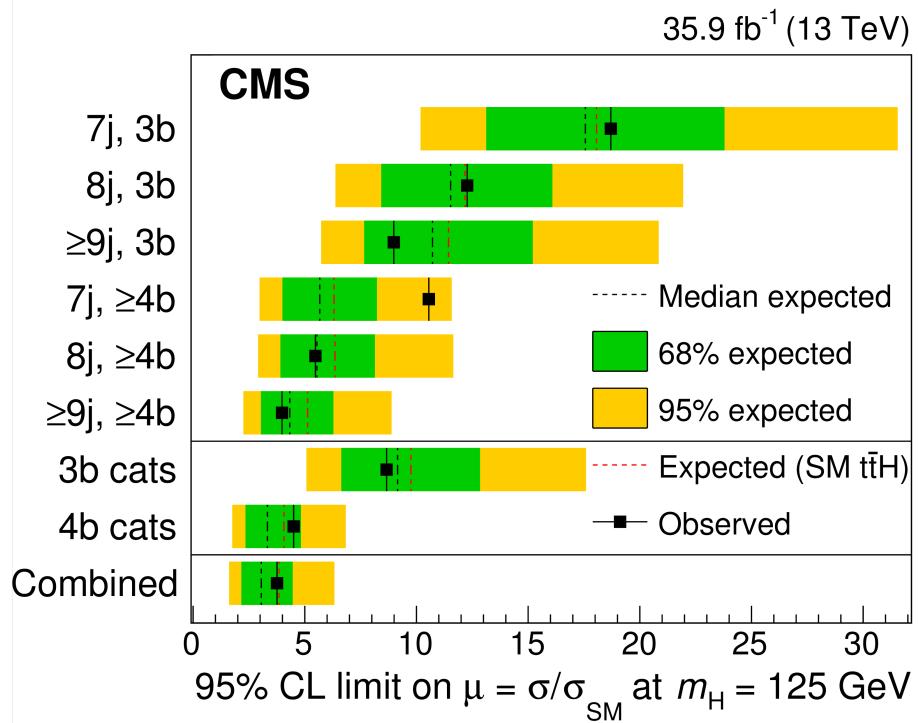


- Best-fit  $\mu$**

$$\hat{\mu} = 0.9 \pm 0.7 \text{ (stat)} \pm 1.3 \text{ (syst)} = 0.9 \pm 1.5 \text{ (tot)}$$

- Limited impact on CMS ttH observation but **impressive efforts which can be exploited in other ttH / fully hadronic analyses**

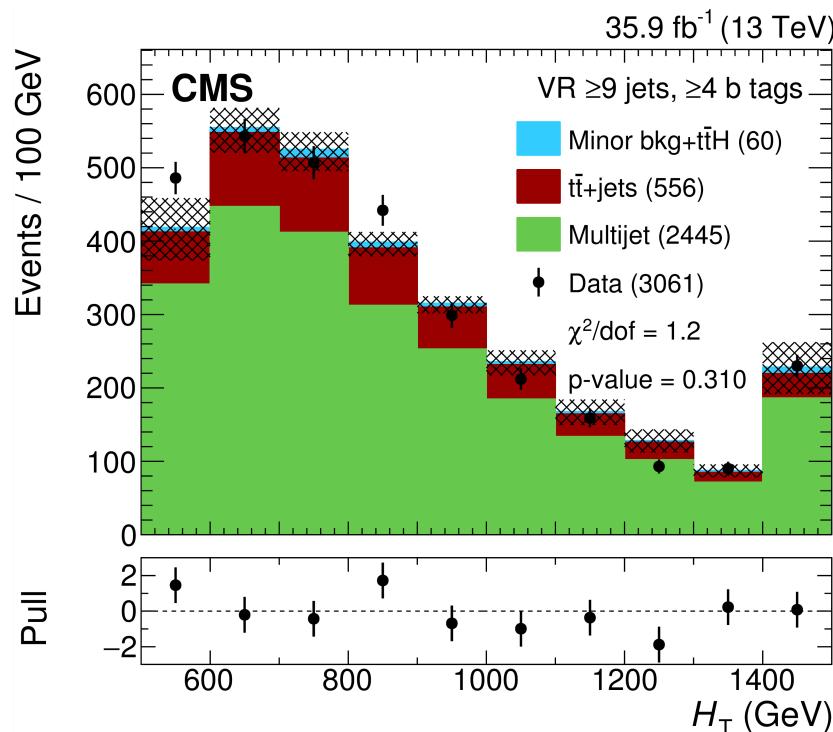
- Sensitivity of analysis limited by large QCD contribution and associated uncertainty:** 20-30% improvement with fixed QCD normalisation
- 95% CL upper limit  $\mu < 3.8$  obs. (3.1 exp.)**  
ATLAS Run 1  $\mu < 6.4$  obs. (5.4 exp.)



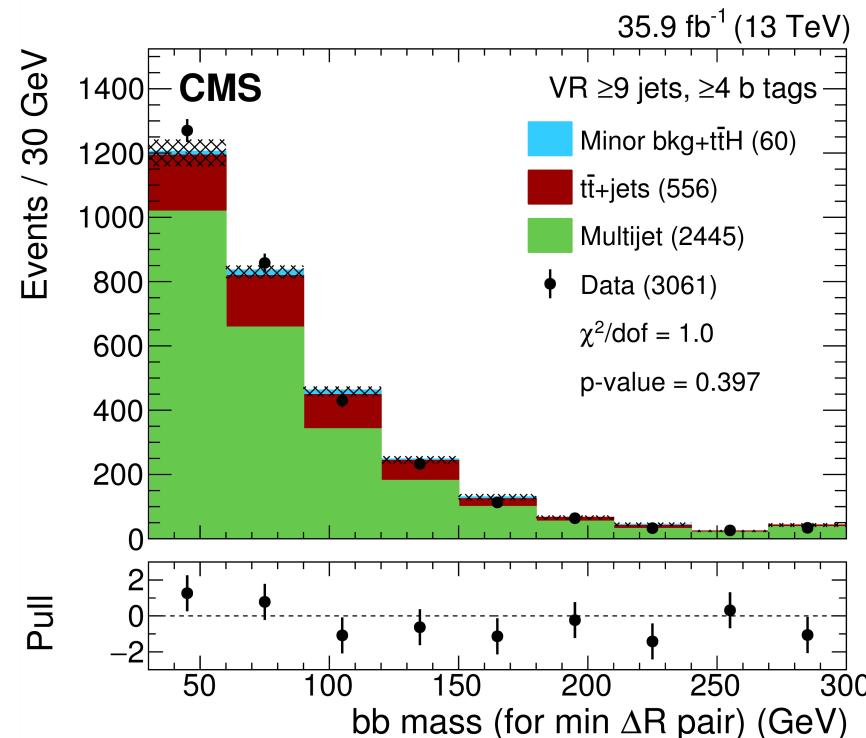
# ttH H $\rightarrow$ bb fully hadronic

JHEP 06 (2018) 101 arXiv:1803.06986 [hep-ex]

- QCD background very difficult to model in simulation => **data-driven estimate**
- **QCD shape from control region / normalisation unconstrained** for signal extraction
- **Validation in QCD-enriched region**

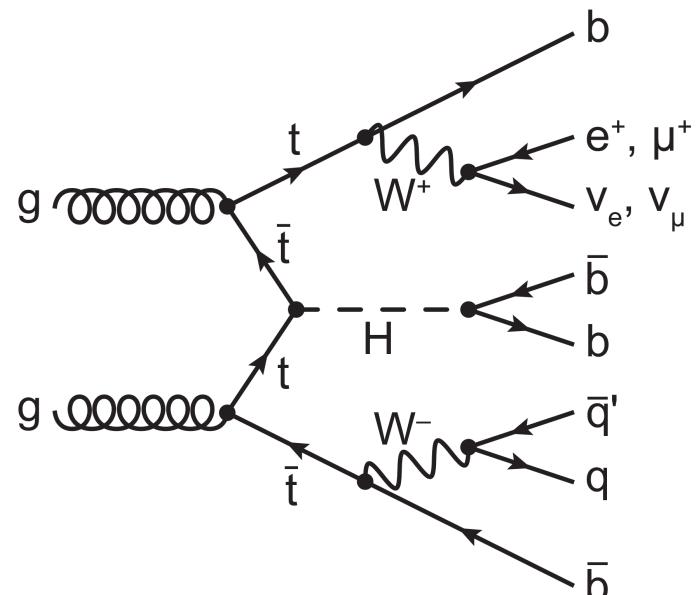
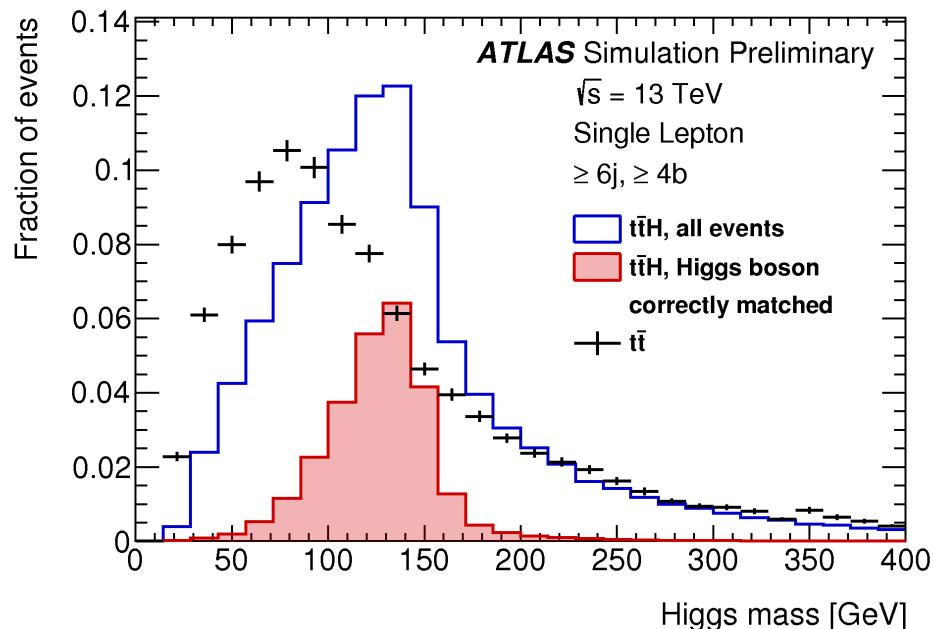


	$N_{CSVM} = 2$ $N_{CSVL} \geq 3$	$N_{CSVM} \geq 3$
QGLR > 0.5	CR (to extract distribution)	SR (final analysis)
QGLR < 0.5	Validation CR (to validate distribution)	VR (comparison with data)



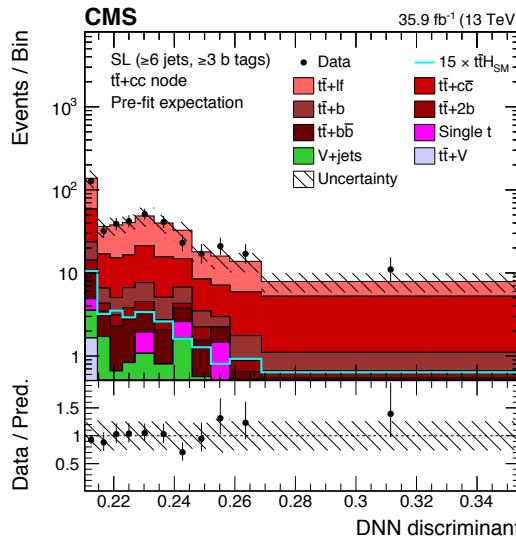
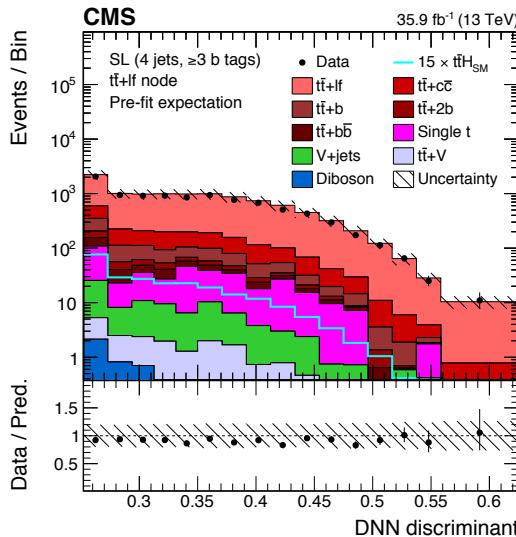
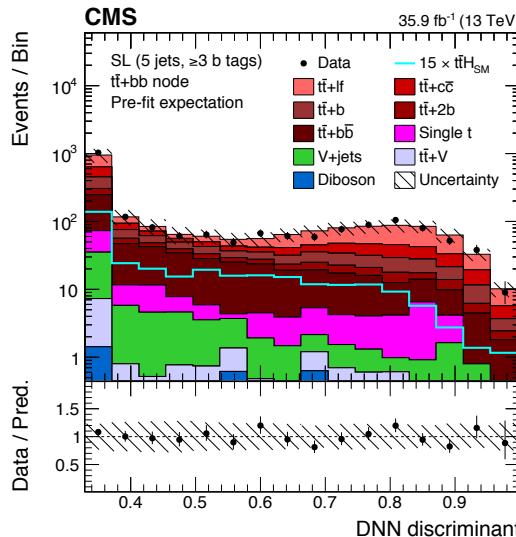
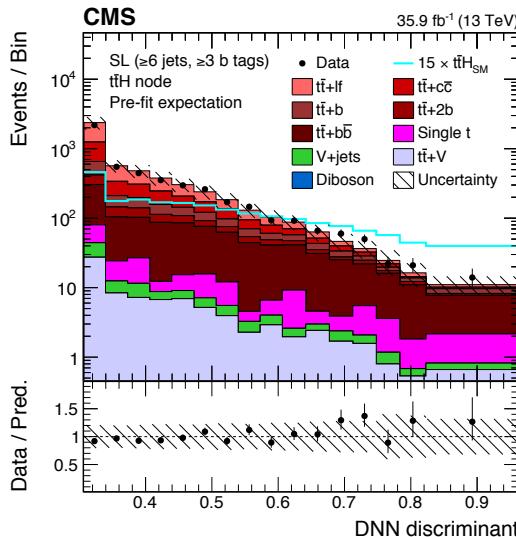
# Event reconstruction with MVA

- Association between **reconstructed objects** and **expected LO objects** attempted to mitigate combinatorics
- ATLAS/CMS analysis makes use of **reconstruction BDT/DNN** trained to **discriminate correct/wrong associations**



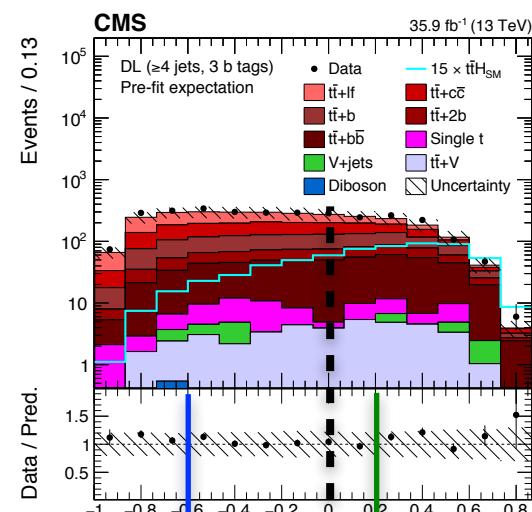
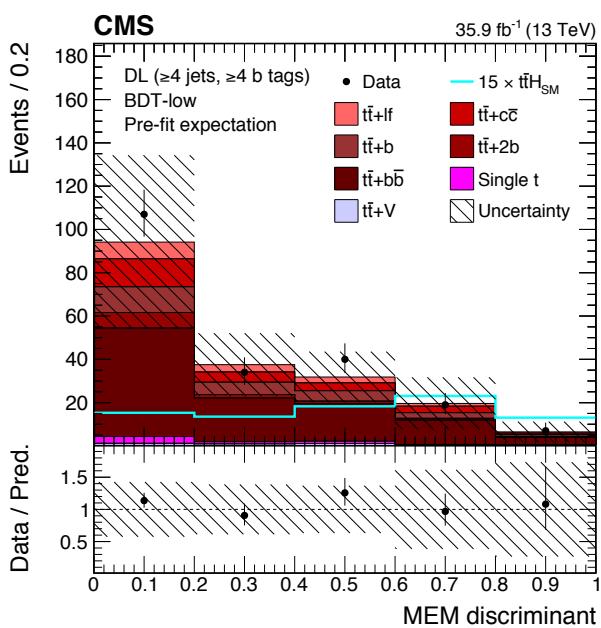
- Used to define kinematic variables with additional discrimination powers:  
e.g. Higgs mass = mass of two b-jets associated to the Higgs boson in the association with highest reconstruction BDT score

- Signal extraction in CMS analysis in SL channels



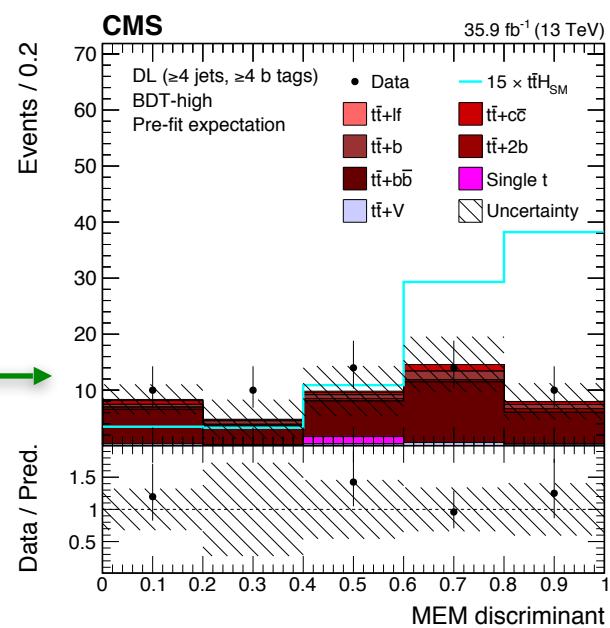
- Events categorised based on most likely process, using DNN probabilities: assigned to ttH/tt+bb/tt+lf/tt+cc nodes
- Distribution of DNN probability used for final fit

- Signal extraction in CMS analysis in DL channels



$$P_{s/b} = \frac{w(\vec{y}|t\bar{t}H)}{w(\vec{y}|t\bar{t}H) + k_{s/b} w(\vec{y}|t\bar{t}+bb)}$$

- **BDT** trained in each category to discriminate ttH signal from inclusive tt+jets background



- **MEM discriminant** used in each SR BDT subcategory, optimized to discriminate ttH signal from irreducible tt+bb background

- Signal extraction method providing the best-sensitivity used in each channel

---

Channel	Method	Best-fit $\mu$ $\pm \text{tot} (\pm \text{stat} \pm \text{syst})$
Single-lepton	BDT+MEM	$1.0^{+0.69}_{-0.66}$ $\begin{pmatrix} +0.31 & +0.62 \\ -0.30 & -0.59 \end{pmatrix}$
Single-lepton	DNN	$1.0^{+0.58}_{-0.55}$ $\begin{pmatrix} +0.30 & +0.50 \\ -0.29 & -0.47 \end{pmatrix}$
Dilepton	BDT+MEM	$1.0^{+1.22}_{-1.12}$ $\begin{pmatrix} +0.65 & +1.04 \\ -0.62 & -0.93 \end{pmatrix}$
Dilepton	DNN	$1.0^{+1.38}_{-1.36}$ $\begin{pmatrix} +0.71 & +1.18 \\ -0.69 & -1.18 \end{pmatrix}$
Combined	BDT+MEM	$1.0^{+0.60}_{-0.57}$ $\begin{pmatrix} +0.28 & +0.53 \\ -0.27 & -0.51 \end{pmatrix}$
Combined	DNN	$1.0^{+0.55}_{-0.51}$ $\begin{pmatrix} +0.27 & +0.47 \\ -0.27 & -0.44 \end{pmatrix}$

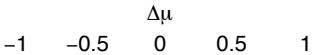
---

# ttH H $\rightarrow$ bb leptonic

Phys. Rev. D 97 (2018) 072016 arXiv:1712.08895 [hep-ex]  
 CMS-HIG-17-026 arXiv:1804.03682 [hep-ex]

Pre-fit impact on  $\mu$ :

$\square = \hat{\theta} + \Delta\theta$     $\square = \hat{\theta} - \Delta\theta$



Post-fit impact on  $\mu$ :

$\square = \hat{\theta} + \Delta\hat{\theta}$     $\square = \hat{\theta} - \Delta\hat{\theta}$

● Nuis. Param. Pull

t $\bar{t}$ +1b: SHERPA5F vs. nominal

t $\bar{t}$ +1b: SHERPA4F vs. nominal

t $\bar{t}$ +1b: PS & hadronization

t $\bar{t}$ +1b: ISR / FSR

t $\bar{t}$ H: PS & hadronization

b-tagging: mis-tag (light) NP I

$k(t\bar{t}+1b) = 1.24 \pm 0.10$

Jet energy resolution: NP I

t $\bar{t}$ H: cross section (QCD scale)

t $\bar{t}$ +1b: tt+3b normalization

t $\bar{t}$ +1c: SHERPA5F vs. nominal

t $\bar{t}$ +1b: shower recoil scheme

t $\bar{t}$ +1c: ISR / FSR

Jet energy resolution: NP II

t $\bar{t}$ +light: PS & hadronization

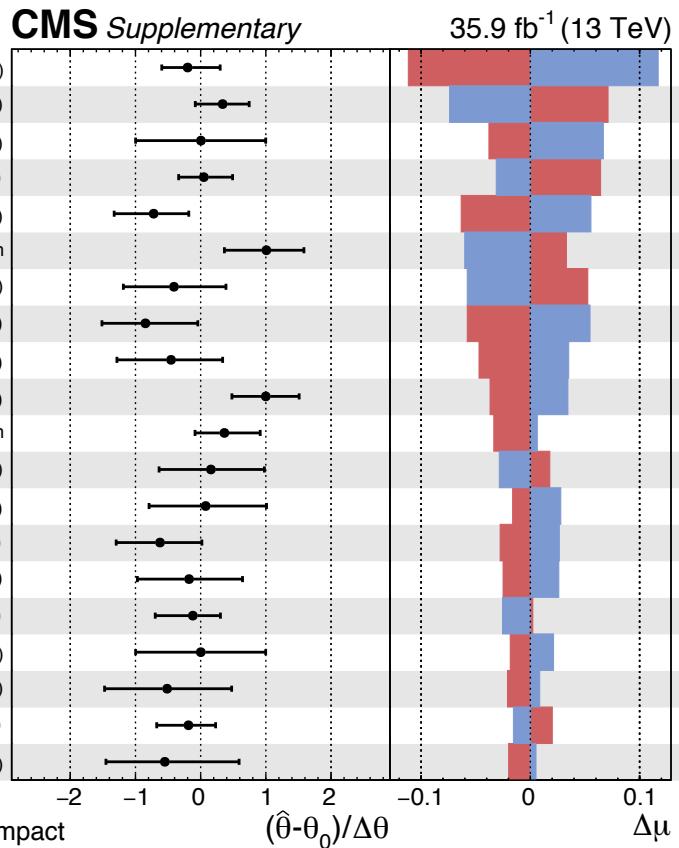
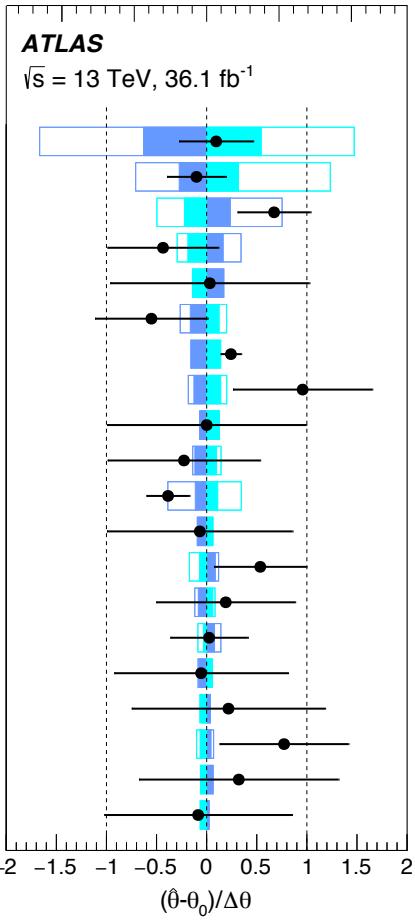
Wt: diagram subtr. vs. nominal

b-tagging: efficiency NP I

b-tagging: mis-tag (c) NP I

E $_{\text{T}}^{\text{miss}}$ : soft-term resolution

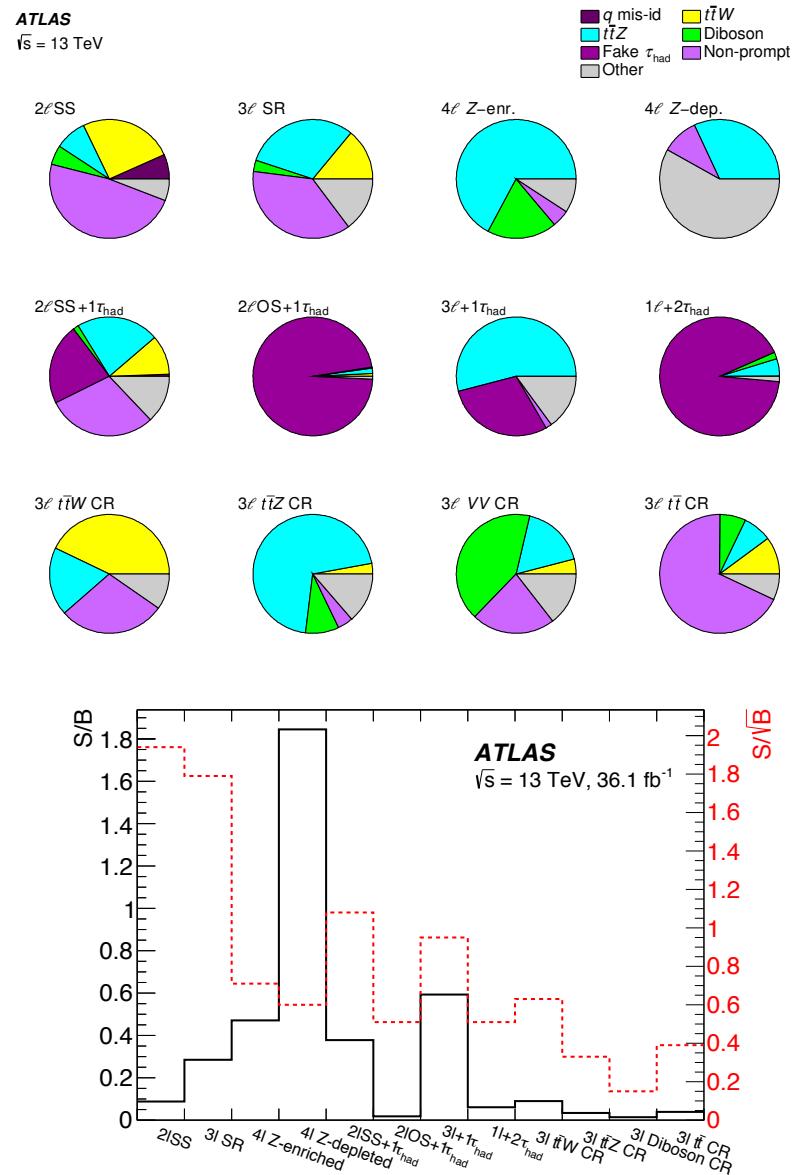
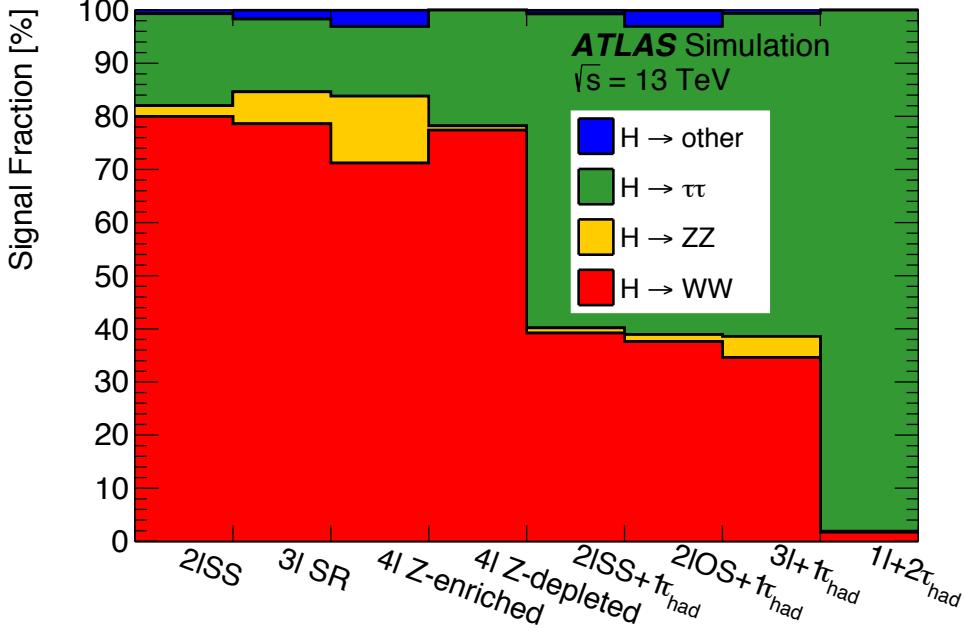
b-tagging: efficiency NP II



● Pull   ■ +1 $\sigma$  Impact   ■ -1 $\sigma$  Impact

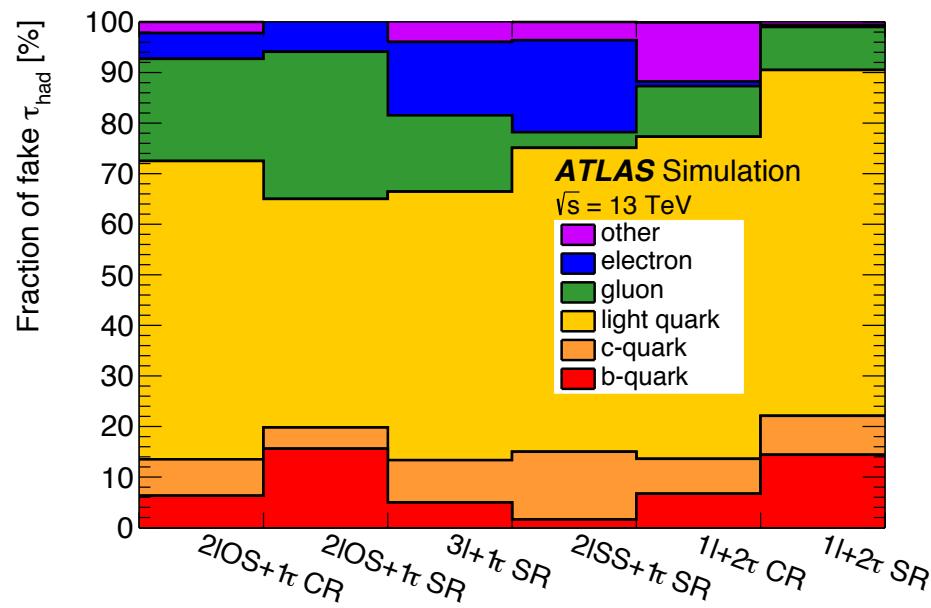
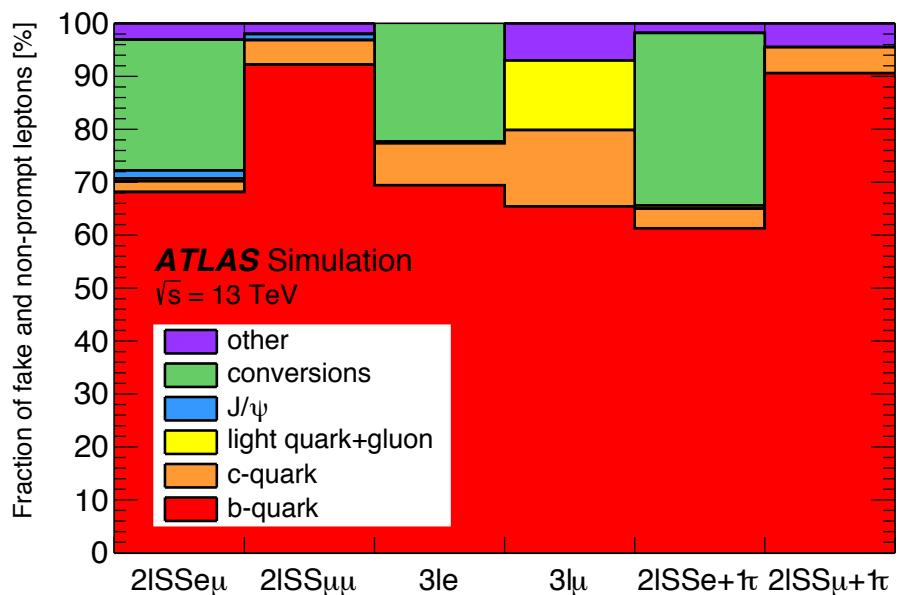
# ttH multilepton

- Categories used in ATLAS ttH multilepton analysis



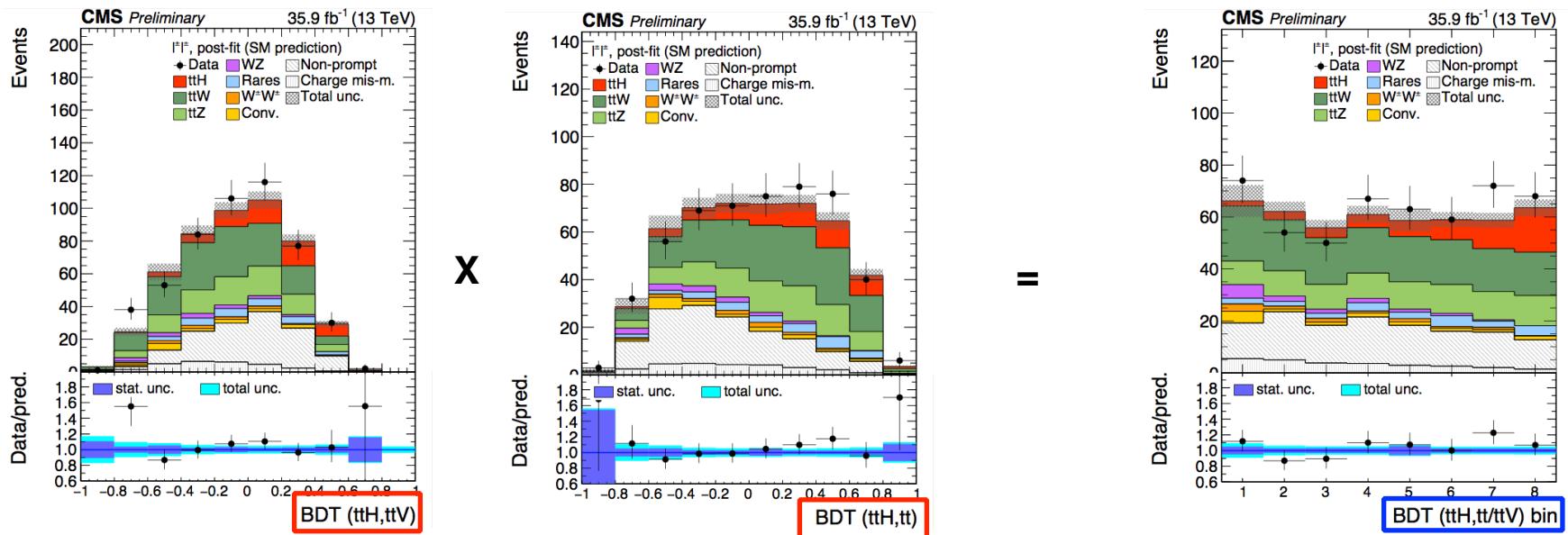
# ttH multilepton

- Sources of non-prompt leptons and fake  $\tau_{\text{had}}$  in the multilepton categories



# ttH multilepton

- Signal extraction in **2ISS + 3l + 3l1 $\tau_h$**  categories based on **2 BDTs** trained to discriminate ttH / ttV and ttH / tt remapped into a **1D discriminant** with increasing S/B



ATLAS	$2\ell\text{SS}$	$3\ell$	$4\ell$	$1\ell+2\tau_{\text{had}}$	$2\ell\text{SS}+1\tau_{\text{had}}$	$2\ell\text{OS}+1\tau_{\text{had}}$	$3\ell+1\tau_{\text{had}}$
BDT trained against Discriminant	Fakes and $t\bar{t}V$ $2\times 1\text{D BDT}$	$t\bar{t}$ , $t\bar{t}W$ , $t\bar{t}Z$ , VV $5\text{D BDT}$	$t\bar{t}Z$ / - Event count	$t\bar{t}$ BDT	all BDT	$t\bar{t}$ BDT	- Event count
CMS	$2\ell\text{SS}$	$3\ell$	$4\ell$	$1\ell+2\tau_{\text{had}}$	$2\ell\text{SS}+1\tau_{\text{had}}$	$3\ell+1\tau_{\text{had}}$	
BDT trained against Discriminant	Fakes and $t\bar{t}V$ $2\times 1\text{D BDT}$	Fakes and $t\bar{t}V$ $2\times 1\text{D BDT}$	- Event count	$t\bar{t}$ BDT	- MEM	Fakes and $t\bar{t}V$ $2\times 1\text{D BDT}$	

# ttH multilepton

- Main systematics

Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modeling (cross section)	+0.20	-0.09
Jet energy scale and resolution	+0.18	-0.15
Non-prompt light-lepton estimates	+0.15	-0.13
Jet flavor tagging and $\tau_{\text{had}}$ identification	+0.11	-0.09
$t\bar{t}W$ modeling	+0.10	-0.09
$t\bar{t}Z$ modeling	+0.08	-0.07
Other background modeling	+0.08	-0.07
Luminosity	+0.08	-0.06
$t\bar{t}H$ modeling (acceptance)	+0.08	-0.04
Fake $\tau_{\text{had}}$ estimates	+0.07	-0.07
Other experimental uncertainties	+0.05	-0.04
Simulation sample size	+0.04	-0.04
Charge misassignment	+0.01	-0.01
Total systematic uncertainty	+0.39	-0.30

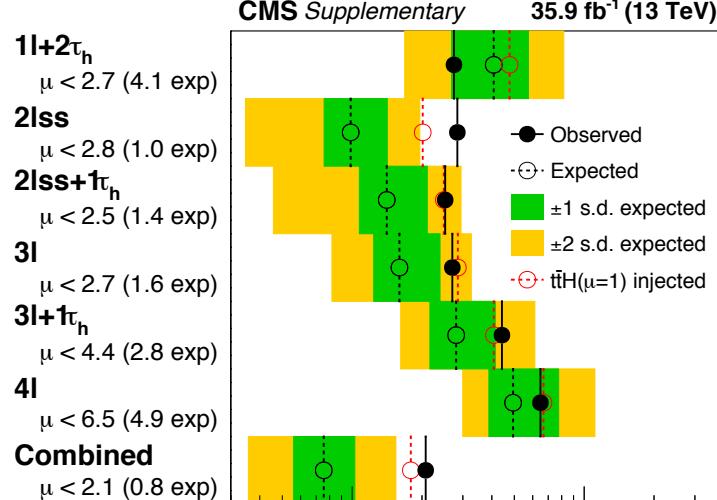
ATLAS

Source	Uncertainty [%]	$\Delta\mu/\mu$ [%]
$e, \mu$ selection efficiency	2–4	11
$\tau_h$ selection efficiency	5	4.5
b tagging efficiency	2–15 [? ]	6
Reducible background estimate	10–40	11
Jet energy calibration	2–15 [? ]	5
$\tau_h$ energy calibration	3	1
Theoretical sources	$\approx 10$	12
Integrated luminosity	2.5	5

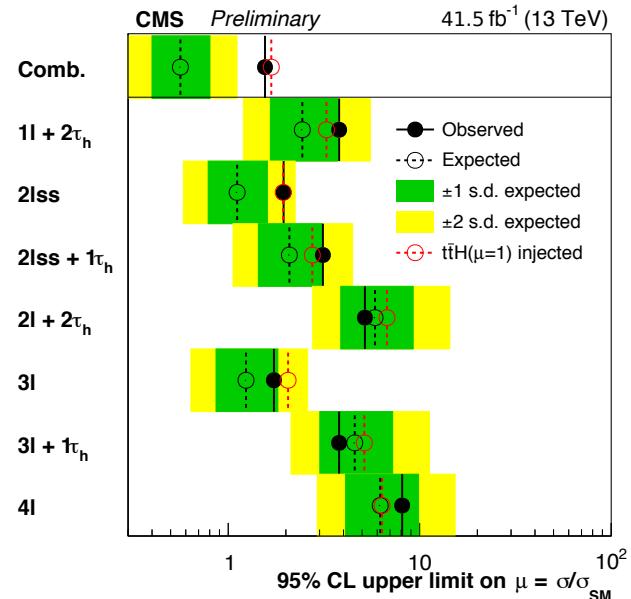
CMS

# ttH multilepton

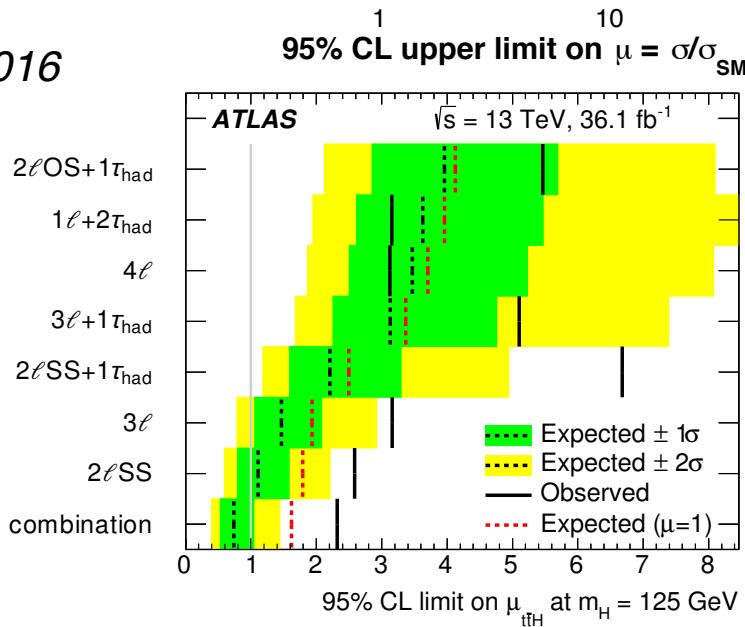
CMS 2016



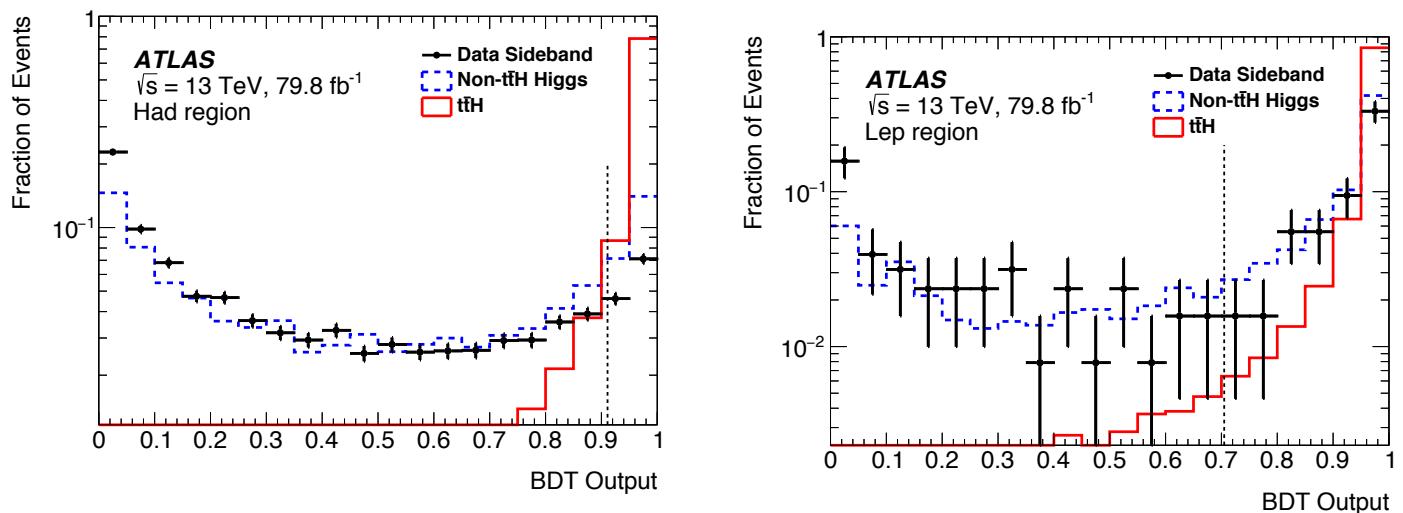
CMS 2017



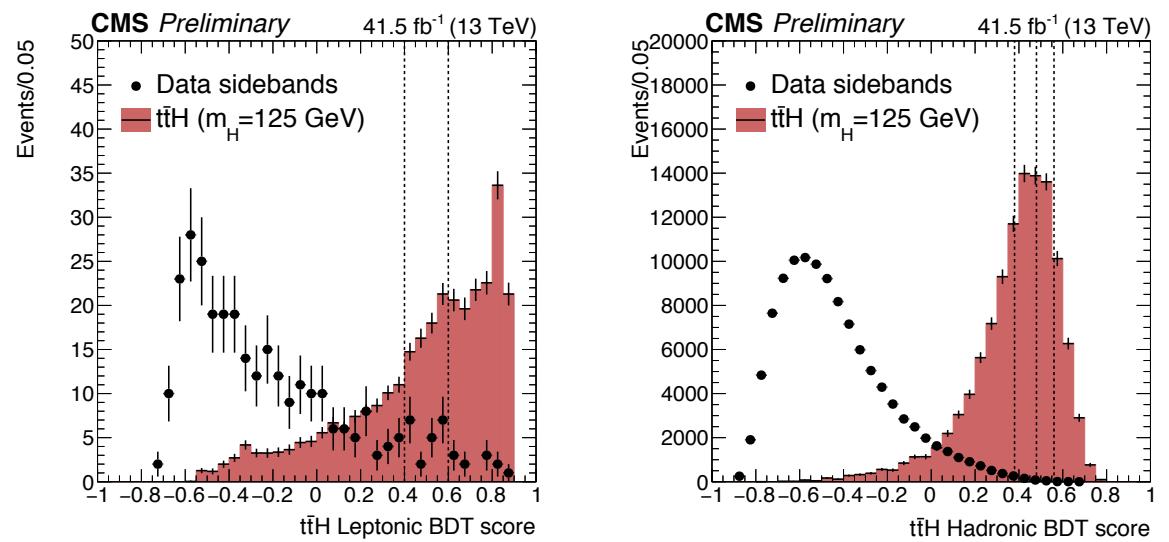
ATLAS 2016



• Sensitivity per category



- **BDTs used for event categorisation**



- Similar uncertainties, inherited from input analyses:

- ttH  $H \rightarrow bb$
- ttH multilepton
- ttH  $H \rightarrow \gamma\gamma$
- ttH  $H \rightarrow ZZ \rightarrow 4l$

Uncertainty source	$\Delta\mu$	
Signal theory	+0.15	-0.07
Inclusive ttH normalisation (cross section and BR)	+0.15	-0.07
ttH acceptance (scale, pdf, PS and UE)	+0.004	-0.004
Other Higgs boson production modes	+0.002	-0.003
Background theory	+0.14	-0.13
tt + bb/cc prediction	+0.13	-0.11
tt + V(V) prediction	+0.06	-0.06
Other background uncertainties	+0.03	-0.03
Experimental	+0.17	-0.15
Lepton (inc. $\tau_h$ ) trigger, ID and iso. efficiency	+0.08	-0.06
Misidentified lepton prediction	+0.06	-0.06
b-Tagging efficiency	+0.05	-0.04
Jet and $\tau_h$ energy scale and resolution	+0.04	-0.04
Luminosity	+0.04	-0.03
Photon ID, scale and resolution	+0.01	-0.01
Other experimental uncertainties	+0.01	-0.01
Finite number of simulated events	+0.08	-0.07
Statistical	+0.16	-0.16
Total	+0.31	-0.26

Uncertainty source	$\Delta\sigma_{ttH}/\sigma_{ttH} [\%]$
Theory uncertainties (modelling)	11.9
$t\bar{t} +$ heavy flavour	9.9
$t\bar{t}H$	6.0
Non- $t\bar{t}H$ Higgs boson production modes	1.5
Other background processes	2.2
Experimental uncertainties	9.3
Fake leptons	5.2
Jets, $E_T^{\text{miss}}$	4.9
Electrons, photons	3.2
Luminosity	3.0
$\tau$ -lepton	2.5
Flavour tagging	1.8
MC statistical uncertainties	4.4

ATLAS

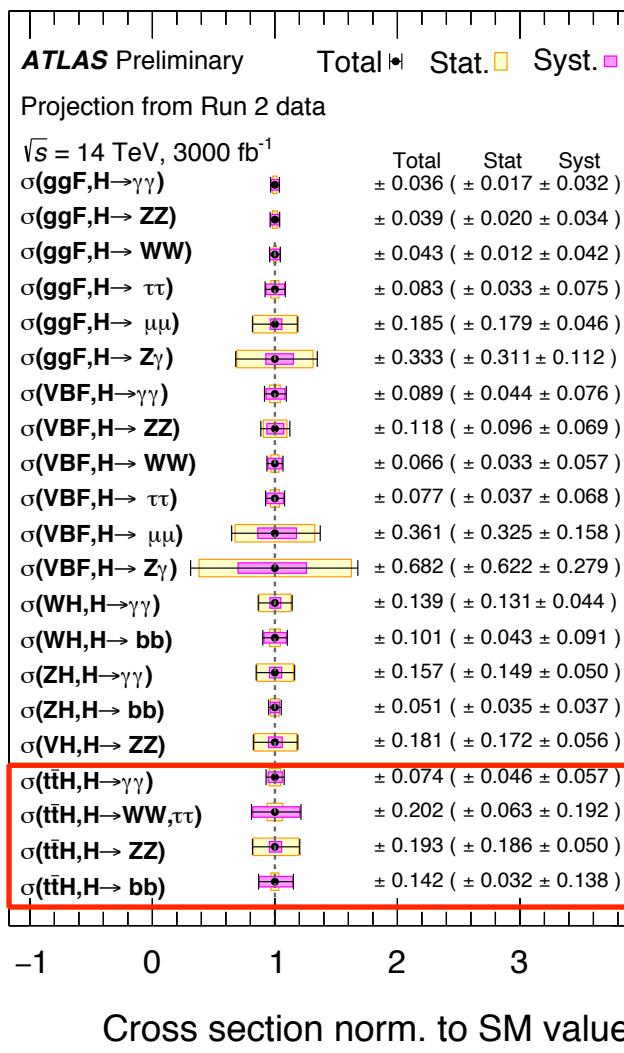
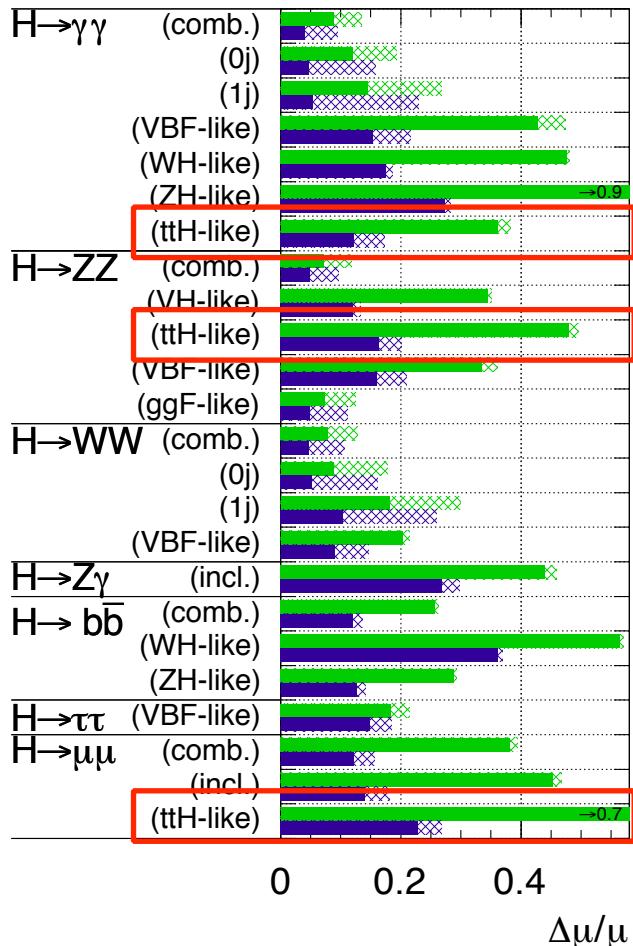
CMS

# Prospects ttH

ATLAS-PHYS-PUB-2014-016  
ATLAS-PHYS-PUB-2018-054

**ATLAS** Simulation Preliminary

$\sqrt{s} = 14 \text{ TeV}$ :  $\int L dt = 300 \text{ fb}^{-1}$ ;  $\int L dt = 3000 \text{ fb}^{-1}$



- **ttH  $H \rightarrow \gamma\gamma$**

Prod. mode	Scenario	$\Delta_{\text{tot}}/\sigma_{\text{SM}}$	$\Delta_{\text{stat}}/\sigma_{\text{SM}}$	$\Delta_{\text{exp}}/\sigma_{\text{SM}}$	$\Delta_{\text{sig}}/\sigma_{\text{SM}}$	$\Delta\mu_{\text{sig}}$
<i>top</i>	Run 2, 80 fb $^{-1}$	+0.37 -0.32	+0.34 -0.30	+0.10 -0.07	+0.10 -0.07	+0.18 -0.11
	HL-LHC S1	+0.11 -0.10	+0.05 -0.05	+0.07 -0.06	+0.07 -0.07	+0.13 -0.11
	HL-LHC S2	+0.08 -0.08	+0.05 -0.05	+0.05 -0.04	+0.04 -0.04	+0.07 -0.06

- **ttH  $H \rightarrow ZZ \rightarrow 4l$**

Prod. mode	Analysis	$\Delta_{\text{tot}}/\sigma_{\text{SM}}$	$\Delta_{\text{stat}}/\sigma_{\text{SM}}$	$\Delta_{\text{exp}}/\sigma_{\text{SM}}$	$\Delta_{\text{sig}}/\sigma_{\text{SM}}$	$\Delta_{\text{bkg}}/\sigma_{\text{SM}}$	$\Delta\mu_{\text{sig}}$
<i>t̄H</i>	Run 2, 80 fb $^{-1}$	< 5.75		-			
	HL-LHC S1	+0.246 -0.213	+0.217 -0.195	+0.056 -0.042	+0.100 -0.074	+0.020 -0.026	+0.156 -0.095
	HL-LHC S2	+0.226 -0.202	+0.217 -0.195	+0.042 -0.032	+0.047 -0.037	+0.010 -0.015	+0.074 -0.051

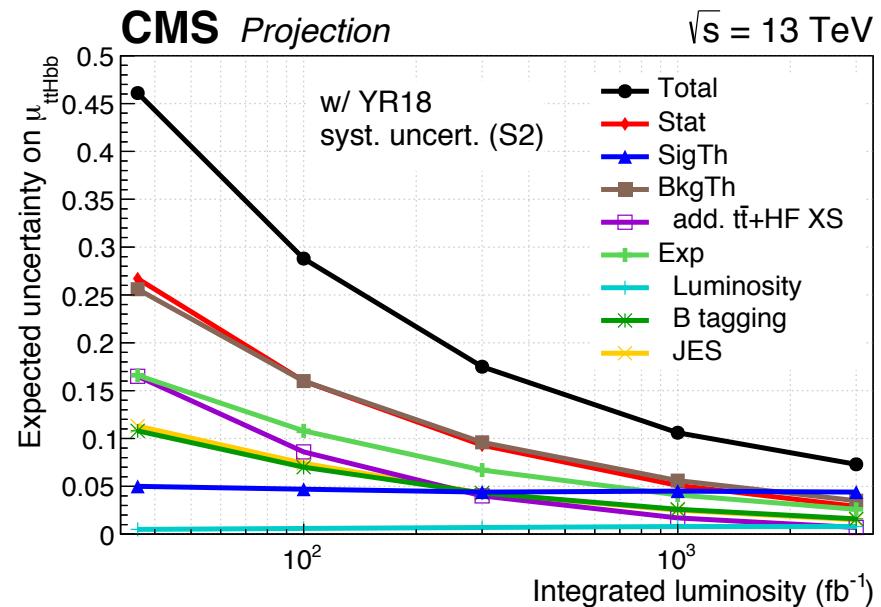
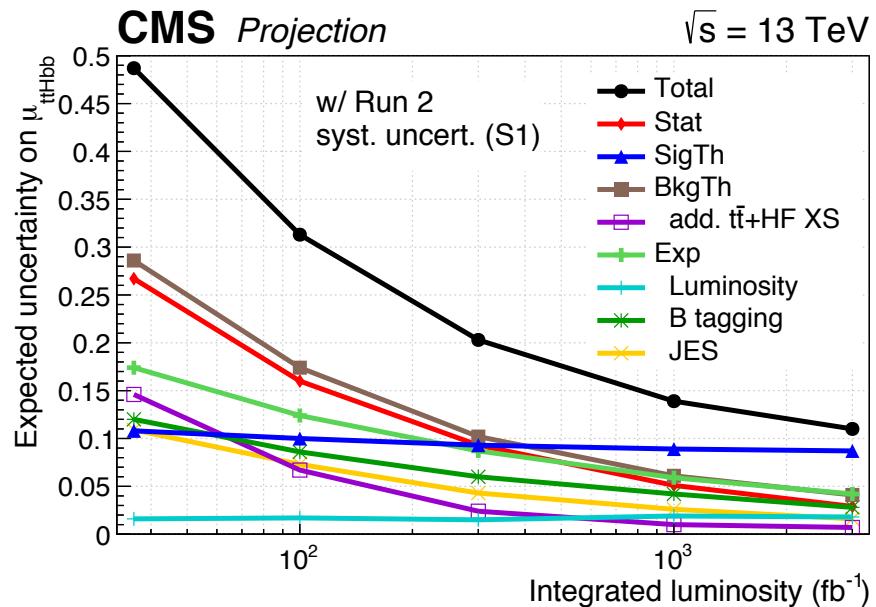
- **ttH ML +  $H \rightarrow bb$**

Final state	Scenario	$\Delta_{\text{tot}}/\sigma_{\text{SM}}$	$\Delta_{\text{stat}}/\sigma_{\text{SM}}$	$\Delta_{\text{exp}}/\sigma_{\text{SM}}$	$\Delta_{\text{sig}}/\sigma_{\text{SM}}$	$\Delta_{\text{bkg}}/\sigma_{\text{SM}}$	$\Delta\mu_{\text{sig}}$
<i>t̄H, H → ML</i> (no $\tau$ )	Run 2, 36 fb $^{-1}$	+0.40 -0.40	+0.33 -0.34	+0.15 -0.15	+0.10 -0.10	+0.13 -0.13	+0.13 -0.13
	HL-LHC S1	+0.18 -0.18	+0.04 -0.04	+0.13 -0.14	+0.08 -0.08	+0.12 -0.12	+0.11 -0.11
	HL-LHC S2	+0.17 -0.17	+0.04 -0.04	+0.12 -0.13	+0.05 -0.05	+0.09 -0.09	+0.07 -0.07
<i>t̄H, H → ML</i> (with $\tau$ )	Run 2, 36 fb $^{-1}$	+0.64 -0.64	+0.54 -0.54	+0.29 -0.29	+0.10 -0.09	+0.14 -0.13	+0.13 -0.13
	HL-LHC S1	+0.27 -0.28	+0.07 -0.07	+0.23 -0.23	+0.09 -0.08	+0.12 -0.12	+0.11 -0.11
	HL-LHC S2	+0.25 -0.25	+0.07 -0.07	+0.22 -0.22	+0.05 -0.05	+0.07 -0.07	+0.07 -0.07
<i>t̄H, H → b̄b</i> (single lepton)	Run 2, 36 fb $^{-1}$	+0.61 -0.61	+0.22 -0.22	+0.27 -0.28	+0.10 -0.09	+0.47 -0.47	+0.15 -0.15
	HL-LHC S1	+0.25 -0.20	+0.02 -0.02	+0.10 -0.10	+0.08 -0.06	+0.22 -0.17	+0.10 -0.11
	HL-LHC S2	+0.18 -0.15	+0.02 -0.02	+0.09 -0.09	+0.06 -0.05	+0.14 -0.11	+0.08 -0.07
<i>t̄H, H → b̄b</i> (dilepton)	Run 2, 36 fb $^{-1}$	+1.06 -1.08	+0.51 -0.51	+0.32 -0.31	+0.11 -0.12	+0.90 -0.92	+0.14 -0.14
	HL-LHC S1	+0.32 -0.26	+0.06 -0.06	+0.13 -0.13	+0.08 -0.07	+0.27 -0.21	+0.11 -0.09
	HL-LHC S2	+0.23 -0.20	+0.06 -0.06	+0.11 -0.11	+0.06 -0.06	+0.17 -0.15	+0.08 -0.08

# Prospects ttH H $\rightarrow$ bb

CMS-PAS-FTR-18-011

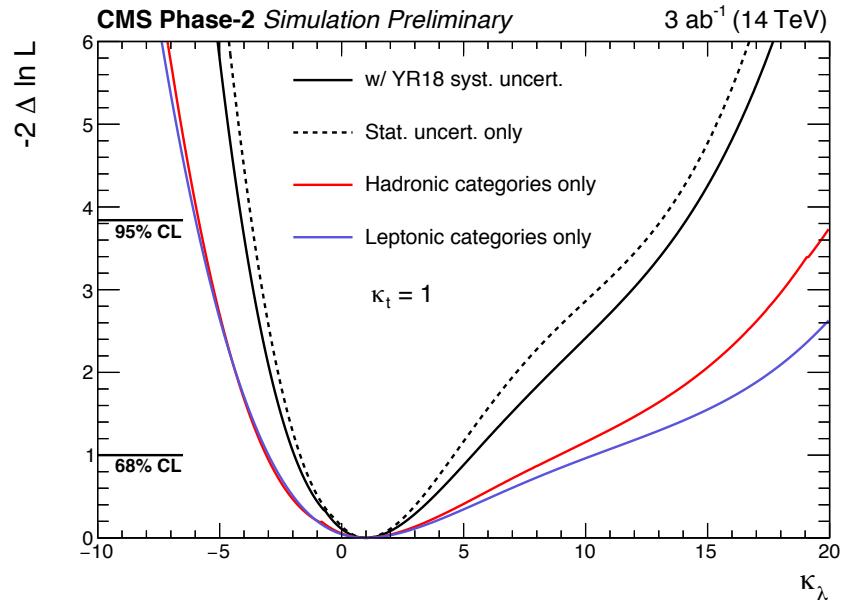
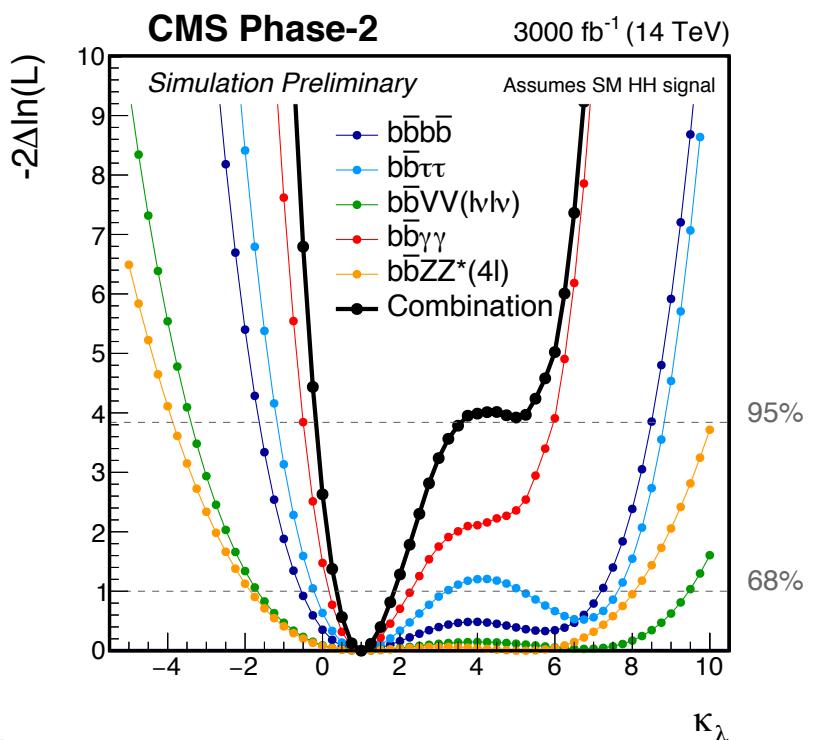
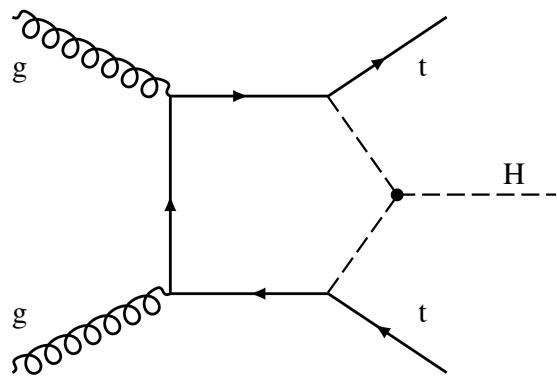
- ttH H $\rightarrow$ bb



# Prospects Higgs self-coupling

CMS-PAS-FTR-18-019  
CMS-PAS-FTR-18-020

- $t\bar{t}H \rightarrow \gamma\gamma$



$\mathcal{L}_{\text{int}} (\text{ab}^{-1})$	68% interval	95% interval
1	[-3.1, 10.9]	[-6.2, 20+]
2	[-2.2, 6.5]	[-4.6, 17.0]
3	[-1.9, 5.3]	[-4.1, 14.1]

- Direct di-Higgs searches