SM Higgs boson measurements with fermionic final states in ATLAS

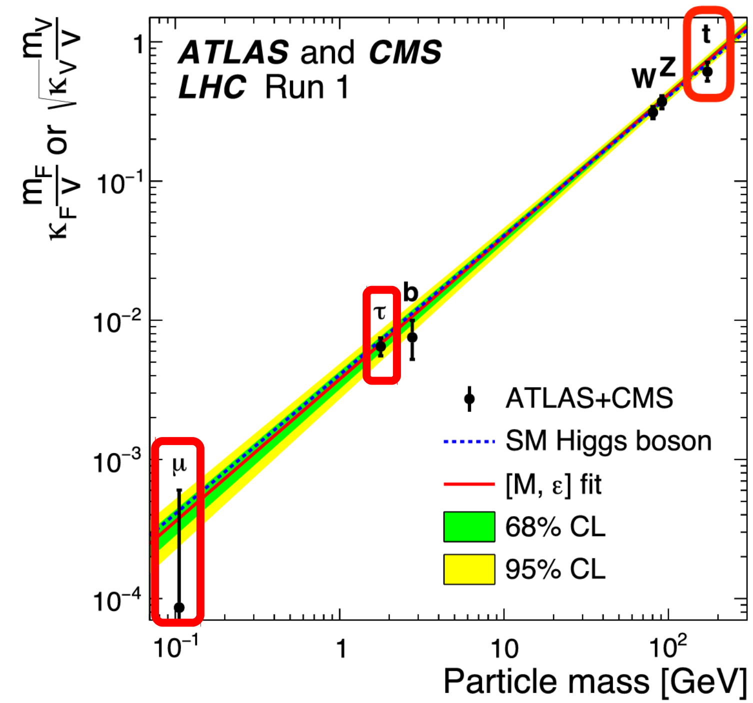
Podberezko Pavel
Budker Institute of Nuclear Physics,
Novosibirsk State University
on behalf of ATLAS collaboration

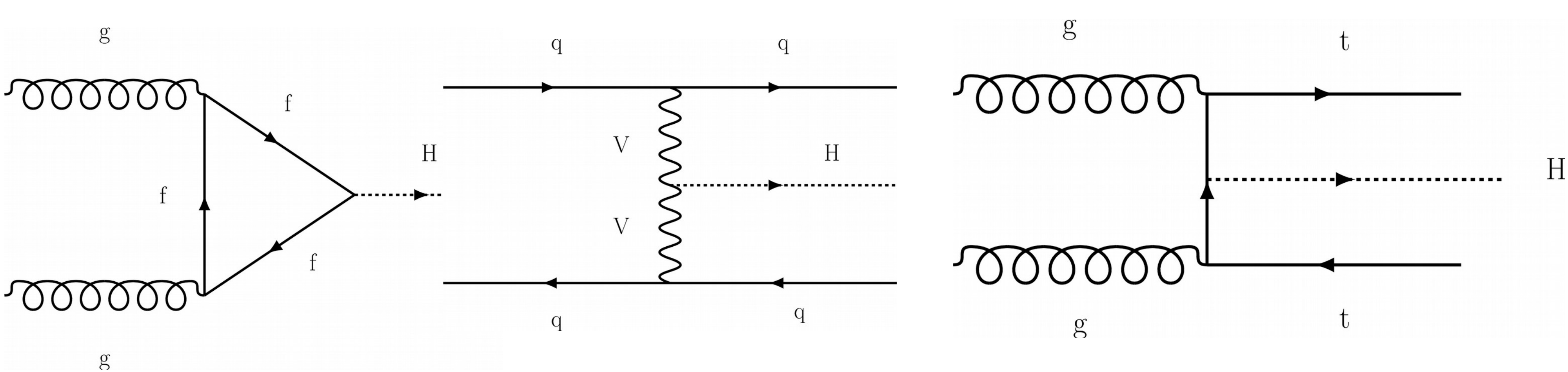
LHC Days in Split, Croatia
17th September 2018

Introduction

ATLAS measurements of the Higgs interactions with fermions are presented

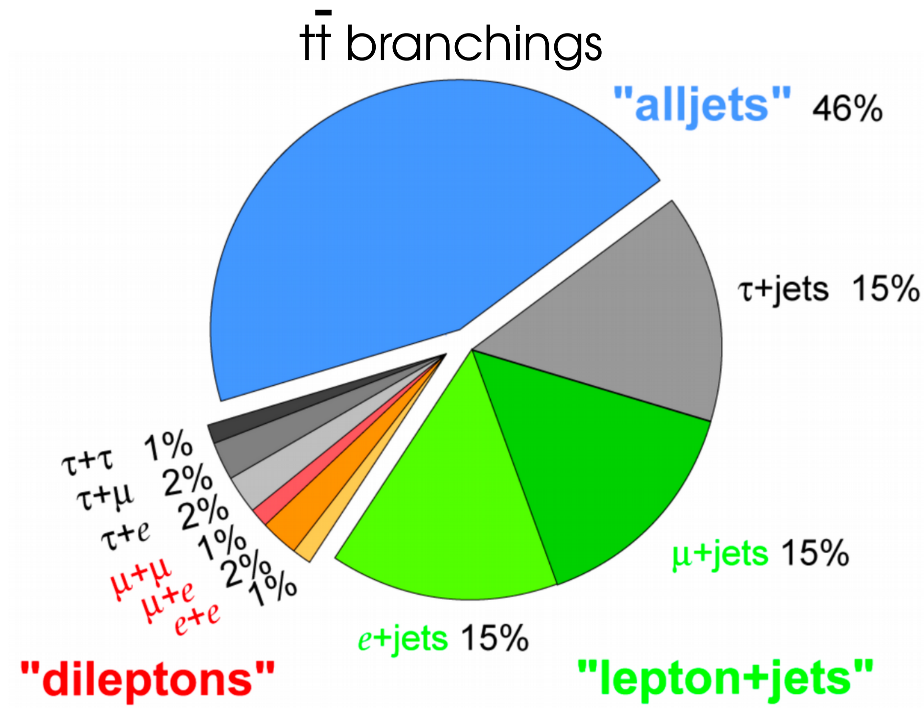
- Establish Higgs affinity to fermionic masses
- Probe Yukawa interactions of leptons and quarks with Higgs
- Final states $\mu^+\mu^-$, $\tau^+\tau^-$, $c\bar{c}$, $J/\psi\gamma$ for all accessible production modes (mainly ggF and VBF) and ttH production for all accessible final states
- Study of H interaction with b-quark is in the separate talk by Louis





ttH, General overview

- Direct measurement of top Yukawa coupling which also contributes to ggF H production and $H \to \gamma\gamma$ decay through loops
- Experimental signature depends on $t\bar{t}$ and H decays
- Four analyses combined:
  - ttH, $H \to bb$ (0, 1 or 2 e/μ)
  - ttH to multilepton ($H \to WW^*, \tau\tau, ZZ^*$)
  - ttH, $H \to \gamma\gamma$ (0 or ≥ 1 e/μ)
  - ttH, $H \to ZZ^* \to 4$ e/μ
- Challenges: $\sigma_{ttH} = 0.5$ pb, $\sigma_{tt} = 830$ pb at 13 TeV, combinatorics of leptons and jets from top decays

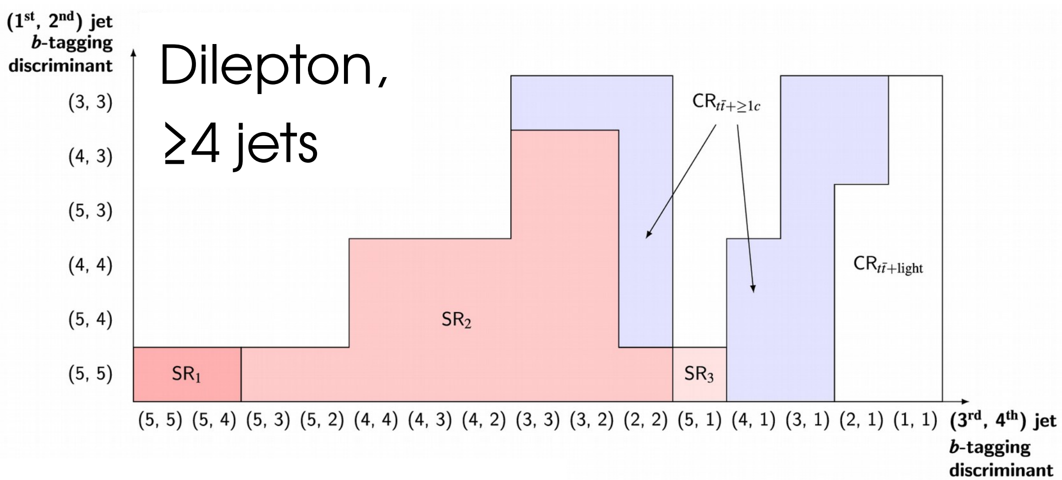


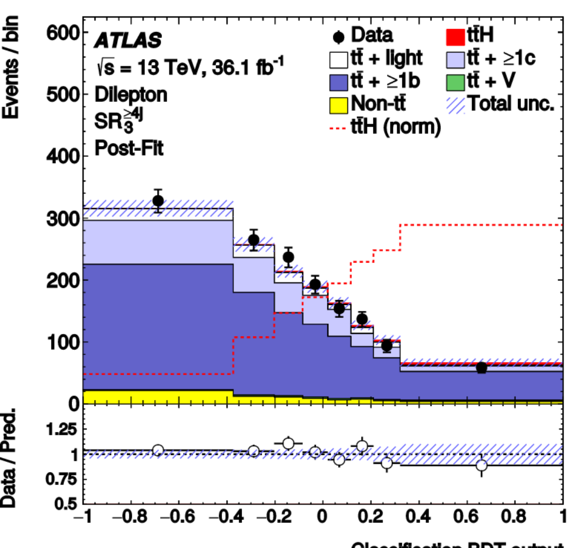
ATLAS history of the analysis:
- Slight excess in Run-1 with μ=1.7±0.8 (JHEP05(2016)160); combination with CMS gave 2.3σ deviation from SM (ATLAS-CONF-2015-044)
- ttH cross-section increased by a factor of 4 in Run-2. Results with 36.1 fb^{-1} of 13 TeV data gave μ=1.2±0.3 (consistent with SM) and brought evidence for ttH with significance: 4.2σ obs. (3.8σ exp.) (Phys. Rev. D 97, 072003)
- Results with update of some channels to 80 fb^{-1} are presented here

Phys. Rev. D 97, 072016

$t\bar{t}H$, H → bb, 36.1 fb^{-1}

- Largest H branching ratio, but huge background from tt+heavy flavor jets
- 1- and 2-lepton channels split into 9 signal and 10 control regions based on jet and b-jet multiplicity and quality



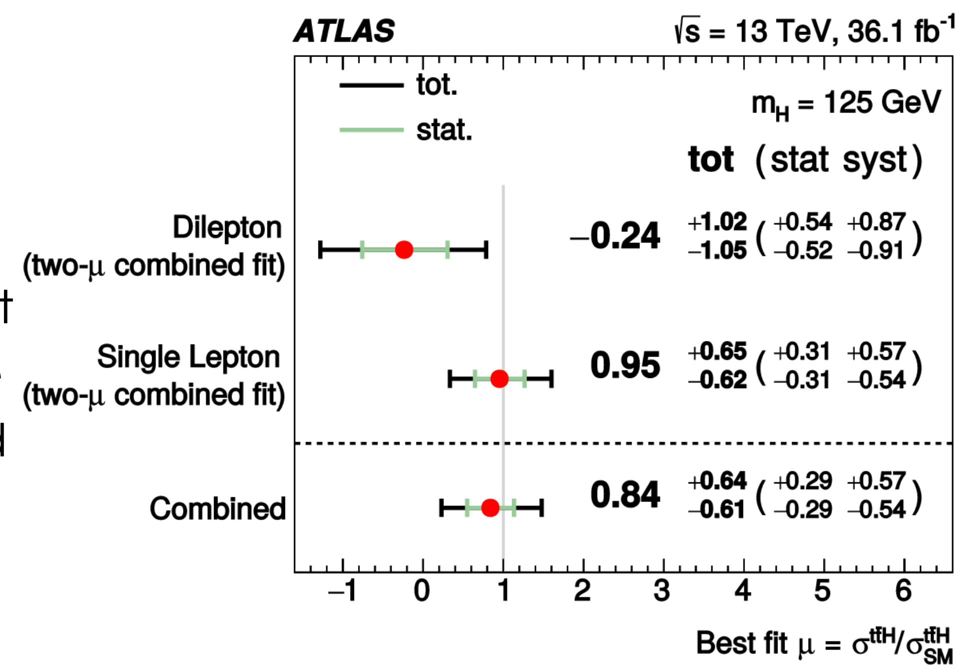


- Different ttH vs. tt discriminants and event kinematic variables are combined into Boosted Decision Tree (BDT) in each signal region

- Signal extracted with binned profile likelihood fit to all SRs and CRs.
- Main systematics:
  - tt + ≥ 1b background modelling (generator comparison): **±0.46 – Main limitation for this analysis!**
  - Background model stat. unc.: +0.29, -0.31

$$\mu = 0.84 \pm 0.29\,(\text{stat.})\ ^{+0.57}_{-0.54}\,(\text{syst.}) = 0.84^{+0.64}_{-0.61}$$

Significance: 1.4σ (1.6σ expected)

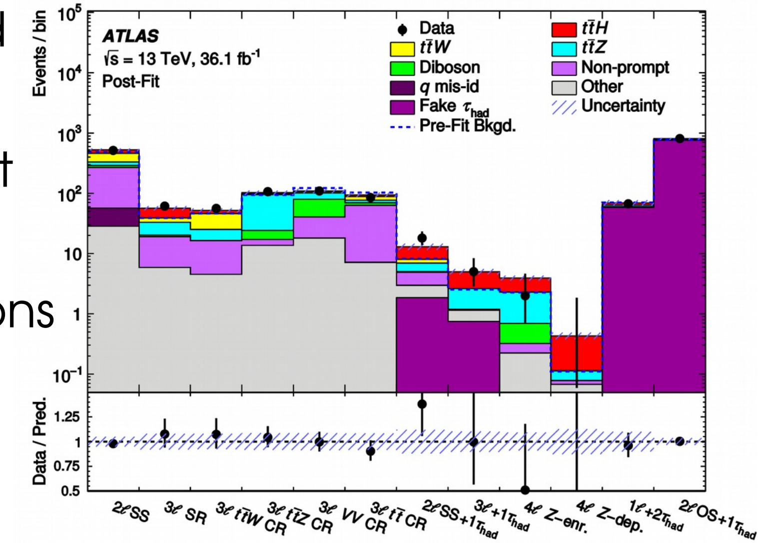


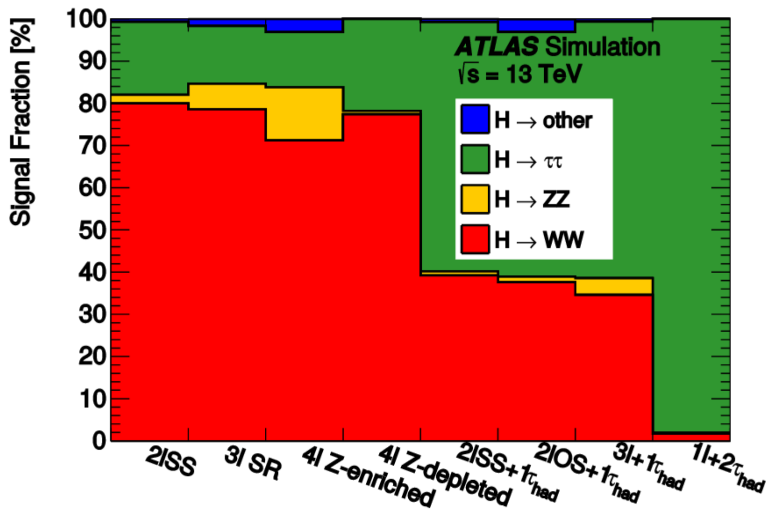
Phys. Rev. D 97, 072003

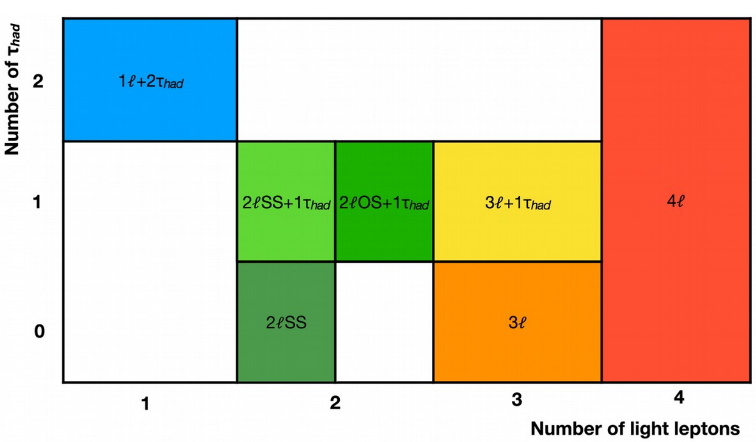
ttH, Multilepton, 36.1 fb^{-1}

- Combination of seven channels with different e/μ and τ multiplicities
- Irreducible backgrounds (ttW, ttZ) estimated from MC and validated in data
- Reducible backgrounds (non-prompt e/μ and fake hadronic τ) estimated from data

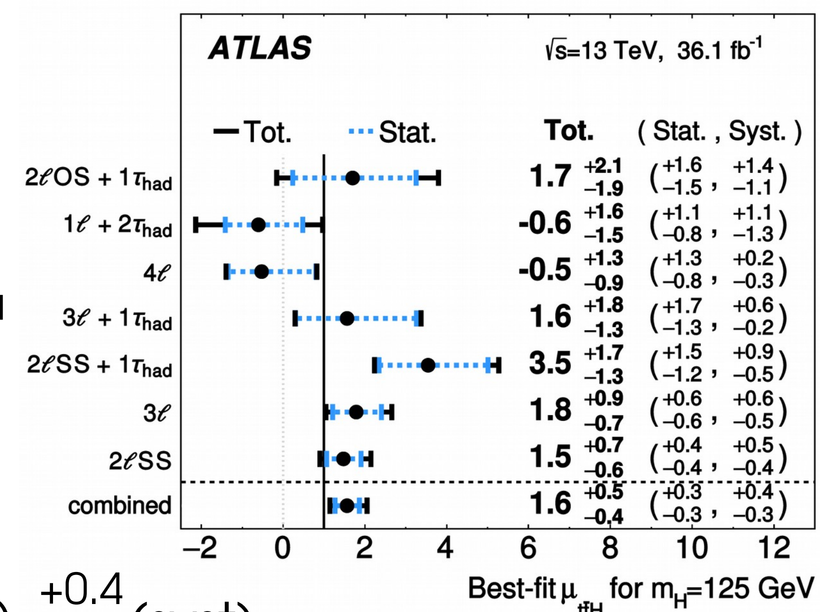
- Signal is extracted with a binned profile likelihood fit to all signal and background regions







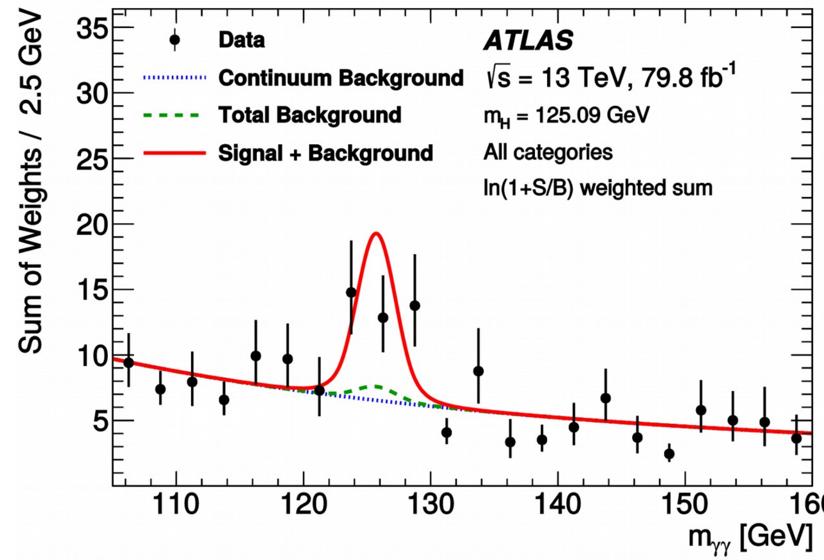
- Main systematics:
  - ⋆ Jet energy scale/resolution: +0.18/-0.15
  - ⋆ Non-prompt e/μ estimates: +0.15/-0.13

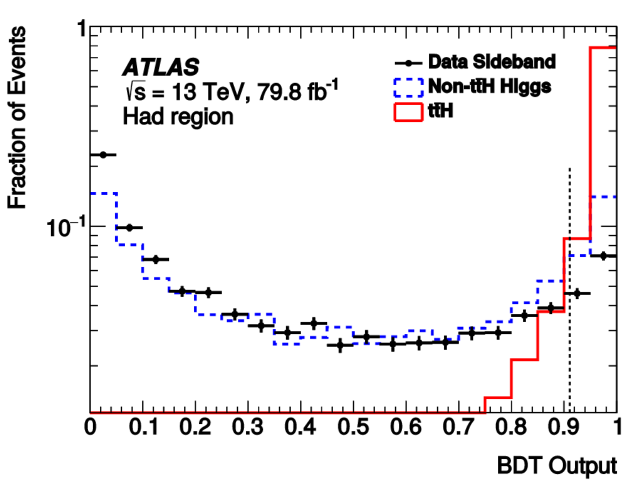


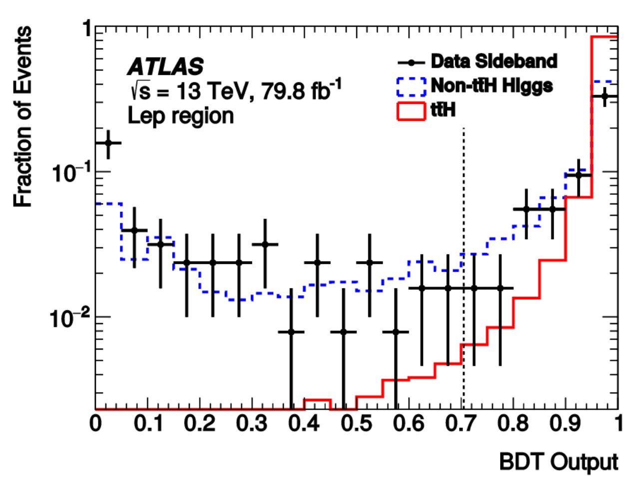
- $\mu = 1.6 \pm 0.3\ (\text{stat})\ ^{+0.4}_{-0.3}\ (\text{syst})$
- Significance: 4.1σ (2.8σ expected)

ttH, H → γγ, 79.8 fb^{-1}

- 50% improvement in the sensitivity w.r.t. the previous analysis: introduction of BDT and new reconstruction algorithms
- Categories: 0 lepton (hadronic) and ≥1 lepton (leptonic)
- Further categorisation with XGBoost BDT:
  - Inputs: 4-vector information of photons (pt, $m_{\gamma\gamma}$), jets, Etmiss, lepton and b-tag
  - Training: ttH MC vs main background (γγ and ttγγ data-driven) and other background (from MC)
  - 4 hadronic and 3 leptonic categories





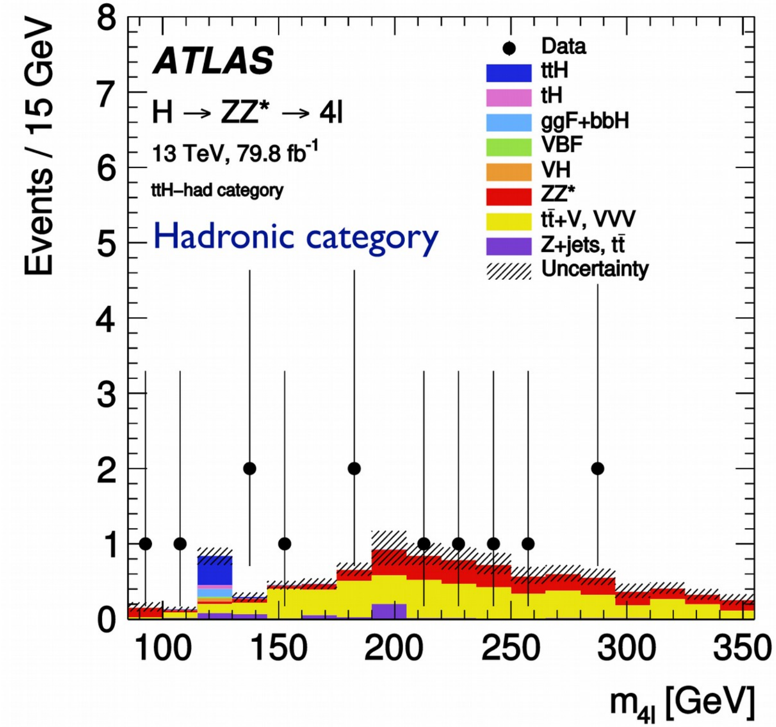


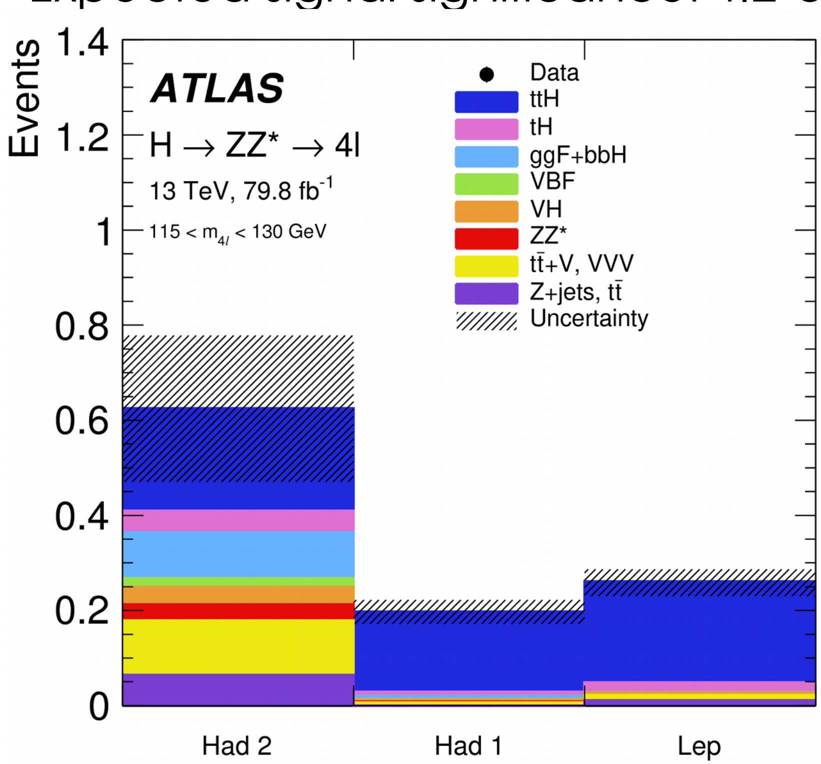
- Combined unbinned fit to 7 $m_{\gamma\gamma}$ with double-sided Crystal-Ball for signal and exponential or power law for bkg
- Significance = 4.1σ (3.7σ expected)
- $\mu = 1.6 \,^{+0.42}_{-0.38}\,\text{(stat)}\,^{+0.23}_{-0.17}\,\text{(syst)}$
- Main systematics:
  - ttH parton shower 8%
  - photon energy resolution 6%

arXiv:1806.00425

ttH, H → ZZ* → 4 e/μ, 79.8 fb^{-1}

- Very small branching ratio, but clean final state
- Categories to select ttH:
  - ≥1 b-tagged jet
  - ≥ 4 jets (hadronic tt) or 1 lepton + ≥ 2 jets (leptonic tt)
- BDT in hadronic category using as inputs: kinematic variables, matrix element discriminant (ttH vs ttV)
- Expected signal significance: 1.2 σ



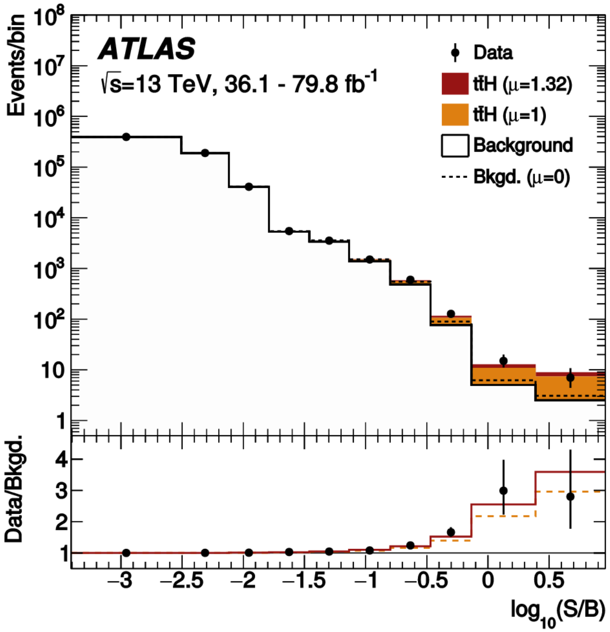


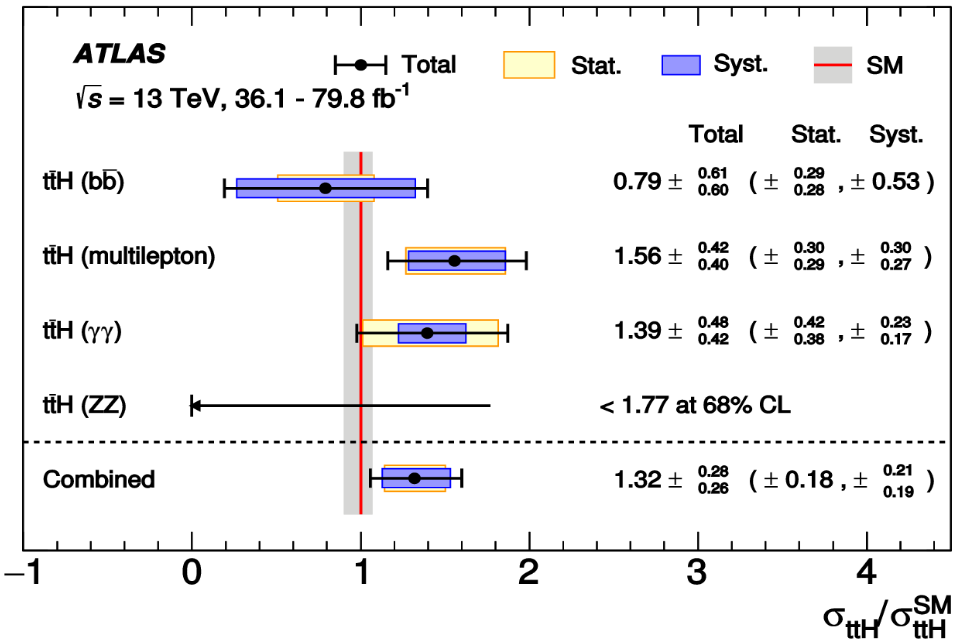
- No events observed in signal region 115 GeV < m4l < 130 GeV
- Expected 0.6 events

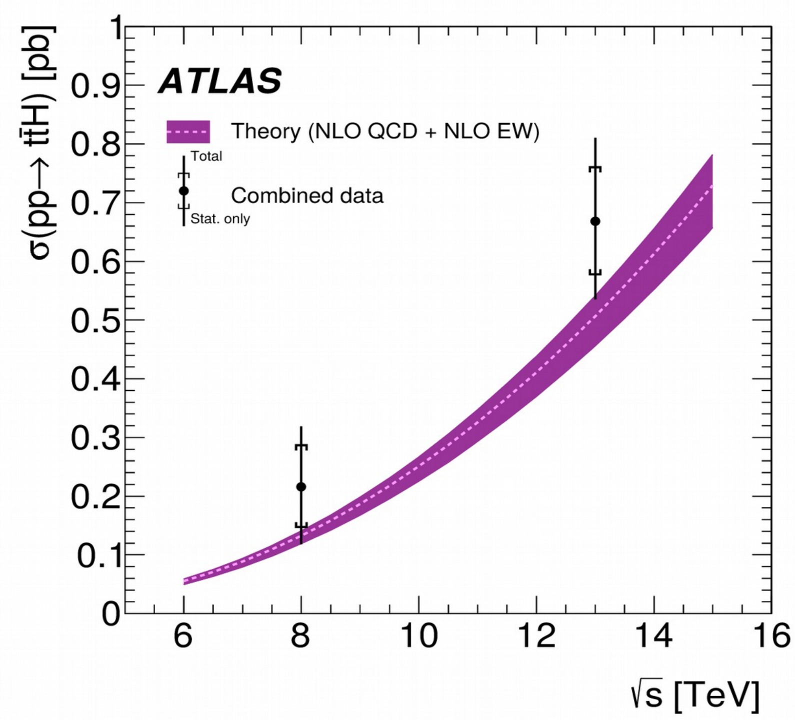
arXiv:1806.00425

ttH, Combination, 36.1 – 79.8 fb^{-1}

- Simultaneous fit to signal and control regions of individual analyses
- Contribution from non-ttH Higgs fixed to SM
- Cross-section for pp → ttH extracted assuming SM branching ratios
- Significance is 5.8σ (4.9σ) observed (expected)
  with Run-1: 6.3σ (5.1σ) observed (expected)
- Measured: σ(pp→ttH) = 670 ± 90 (stat) +110/-100 (syst) fb
  Standard Model: 507 +35/-50 fb





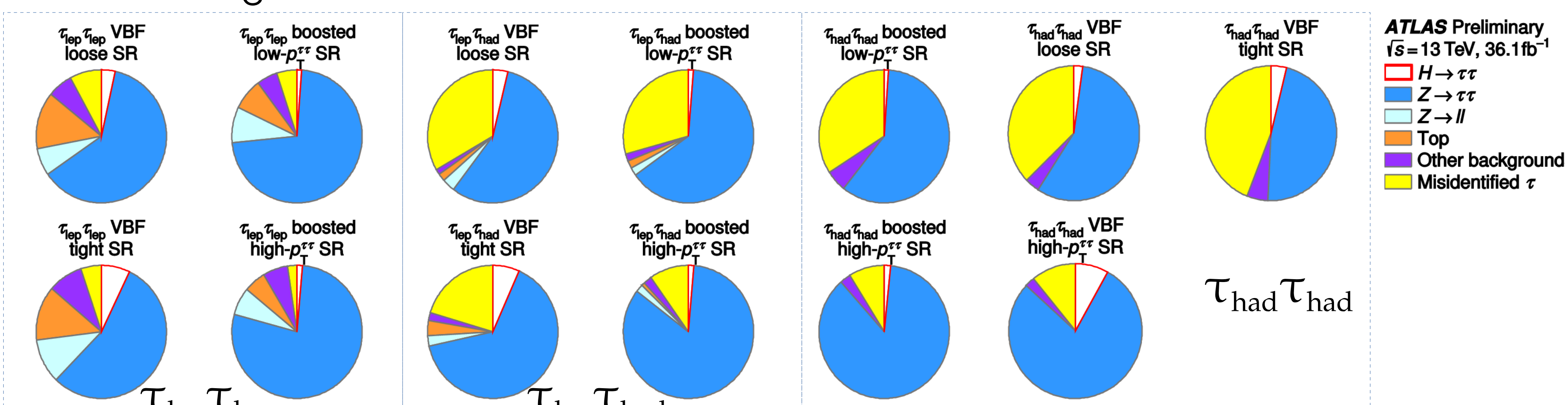


$H \rightarrow \tau^+\tau^-$, Signal and Background

To enhance signal sensitivity several regions are defined

- 5 categories based on production mode: VBF (splits into loose, tight, high-pt), boosted (‘low’ pt, high pt)
- 3 categories based on final state: $\tau_{lep}\tau_{lep}$, $\tau_{lep}\tau_{had}$, $\tau_{had}\tau_{had}$.
- Overall (3 x 5) – 2(high pt VBF only for $\tau_{had}\tau_{had}$) = 13 Signal regions
- Main backgrounds: $Z\tau\tau$, jet-faking-τ, Zll, top
- 6 Control regions to constrain normalization of Zll and top backgrounds
- Dominant irreducible background $Z\tau\tau$ accounts for 50 – 90% of total background and is estimated from MC

ATLAS history of the analysis:
- Evidence in Run-1 with 4.5 (3.4) obs (exp) σ (JHEP04(2015)117)
- Observation on Run-1 dataset in combination with CMS: 5.5 (5.0) σ (JHEP08(2016)045)
- Now observation with combination of Run-1 and 36.1 fb^{-1} of Run-2 ATLAS only dataset: 6.4 (5.4) σ

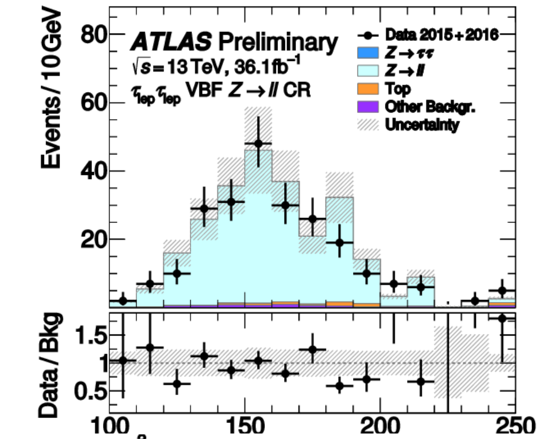


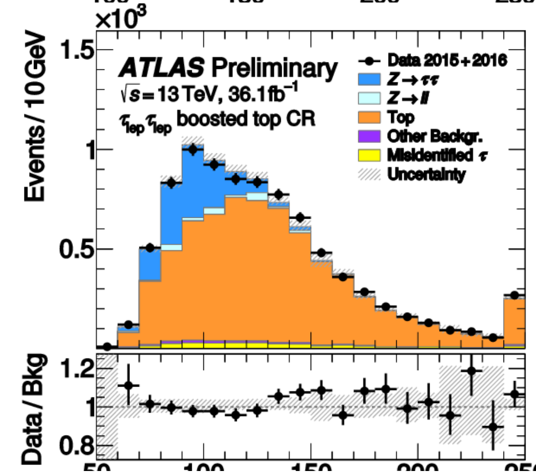
36.1 fb^{-1}

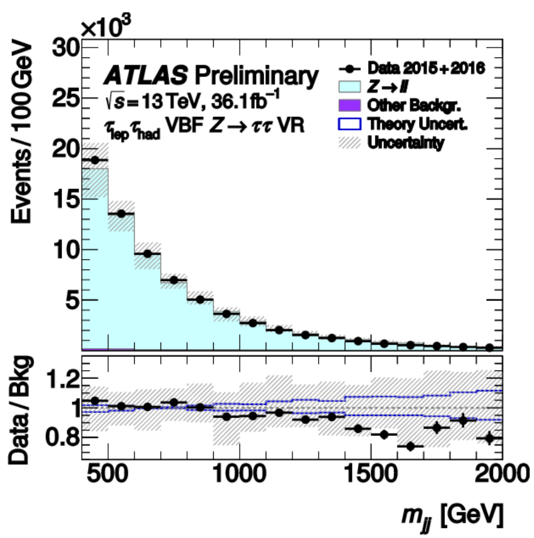
ATLAS-CONF-2018-021

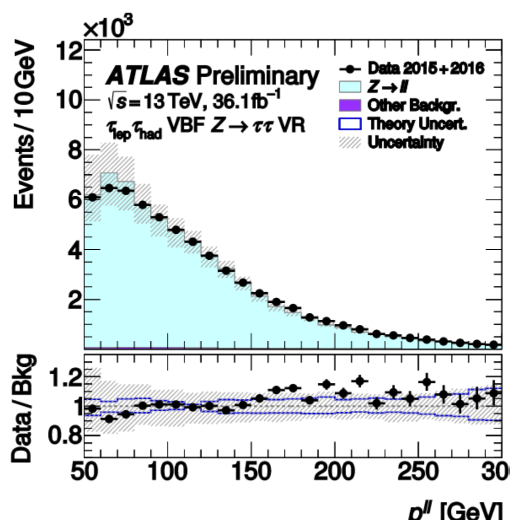
$H \rightarrow \tau^+\tau^-$, Background estimation

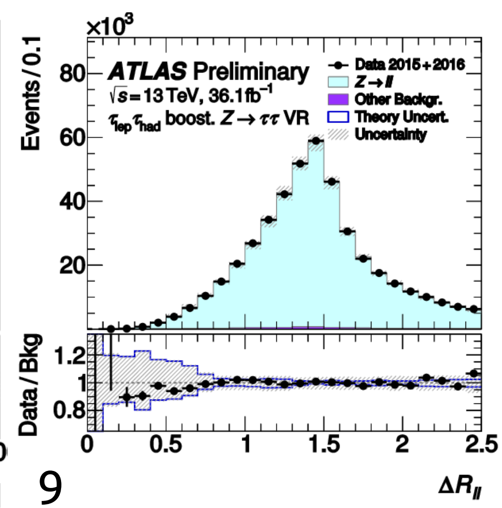
- <u>$Z\tau\tau$</u> is estimated from MC, validated in special $Z\tau\tau$ validation regions
- Separation from signal is limited by $m_{\tau\tau}$ reconstruction
- $Z\tau\tau$ normalization factor is correlated across channels and constrained by $m_{\tau\tau}$ at Z peak.
- <u>Zll</u> and <u>top</u> shapes are taken from MC, normalization is constrained from 6 control regions
- <u>Fake τ</u> are estimated using data-driven techniques.
  ⋆ W+jets and QCD multijet contributing to this bkg in $\tau_{lep}\tau_{had}$ are estimated using fake factors F = N(pass ID τ_{had})/N(fail ID τ_{had})
  ⋆ QCD multijet contributes to fakes in $\tau_{had}\tau_{had}$ and is modeled with template extracted from data by inverting opposite charges requirement and allowing relaxed N(tracks).

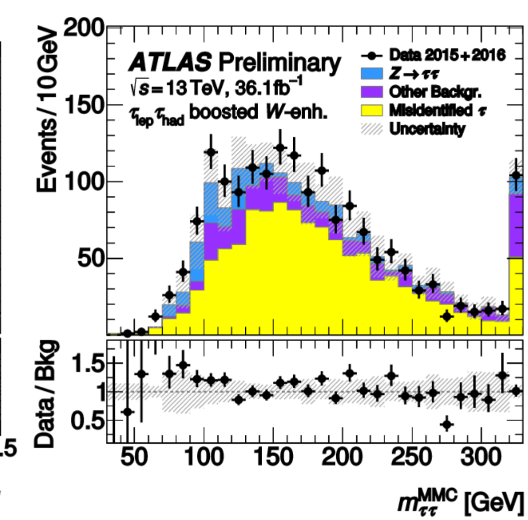






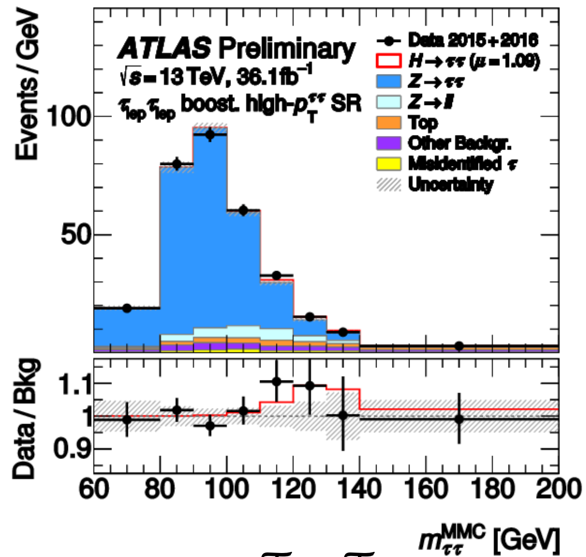




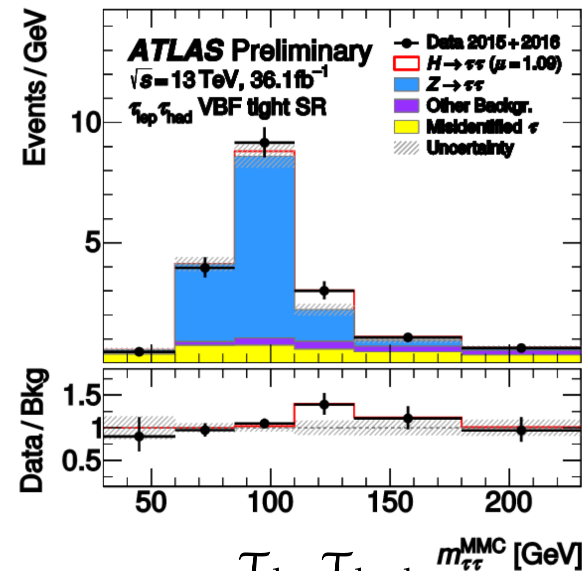


$H \rightarrow \tau^+\tau^-$, Statistical Procedure

- Simultaneous maximum likelihood fit of $m_{\tau\tau}$ to 13 signal and 6 control regions
- Parameter of interest is $\sigma_H * B(H \rightarrow \tau^+\tau^-)$, systematics are taken into account with nuisance parameters
- Significance of signal excess around 125 GeV is tested against SM background only hypothesis



$\tau_{lep}\tau_{lep}$



$\tau_{lep}\tau_{had}$

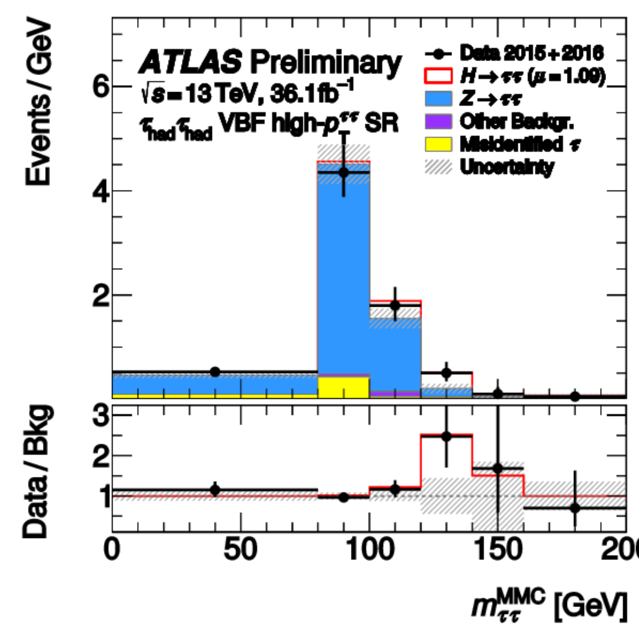
$$\mu = 1.09 \,^{+0.18}_{-0.17}\,(\text{stat.})\,^{+0.27}_{-0.22}\,(\text{syst.})\,^{+0.16}_{-0.11}\,(\text{theory syst.})$$

$$\sigma_{H\rightarrow\tau\tau} \text{ is } 3.71 \pm 0.59\,(\text{stat.})\,^{+0.87}_{-0.74}\,(\text{syst.})\,\text{pb}$$

$$\sigma^{SM}_{H\rightarrow\tau\tau} = 3.43 \pm 0.13\,\text{pb}$$

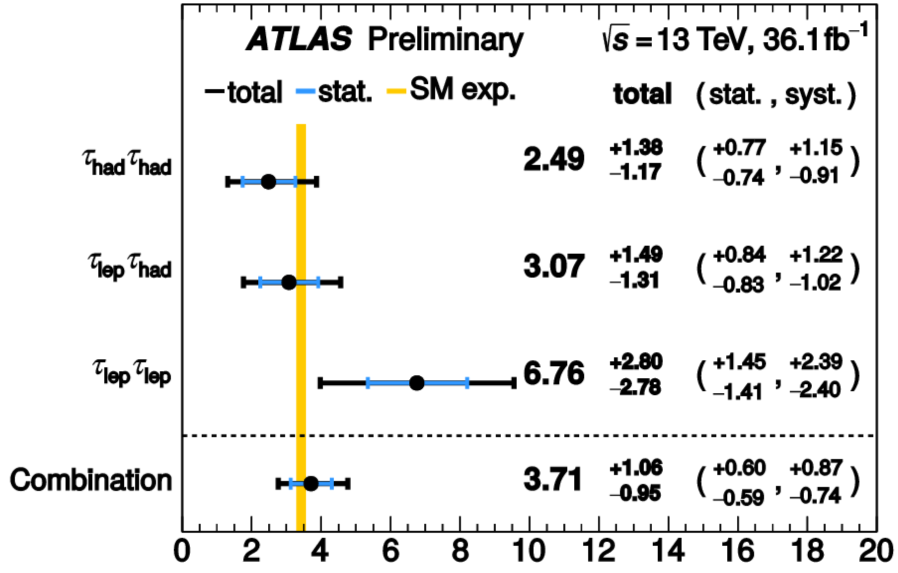
√s (TeV)	7, 8	13	Combined
Observed (σ)	4.5	4.4	6.4
Expected (σ)	3.4	4.1	5.4

$\tau_{had}\tau_{had}$

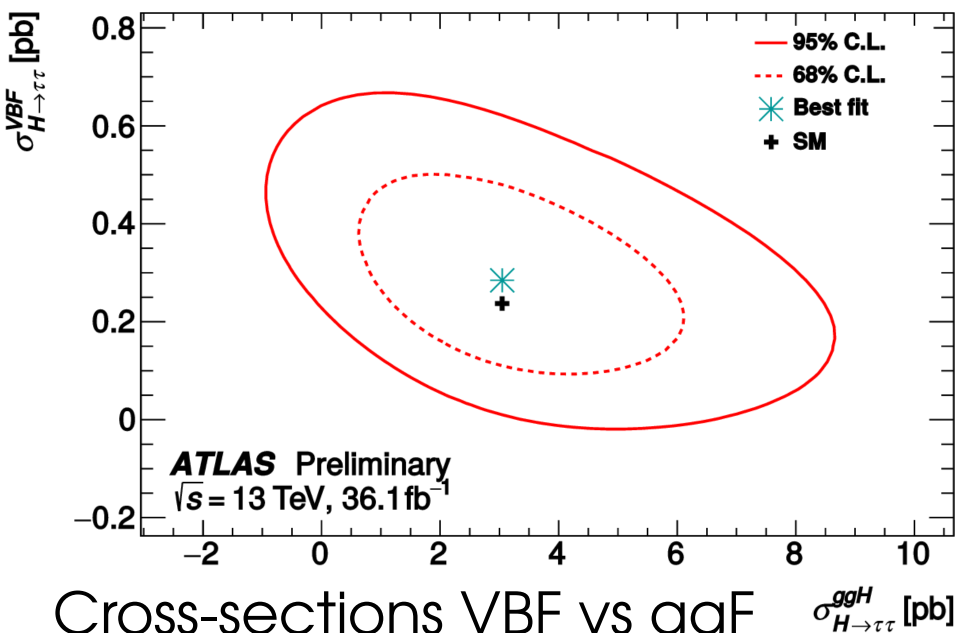


$H \rightarrow \tau^+\tau^-$, Cross-section results

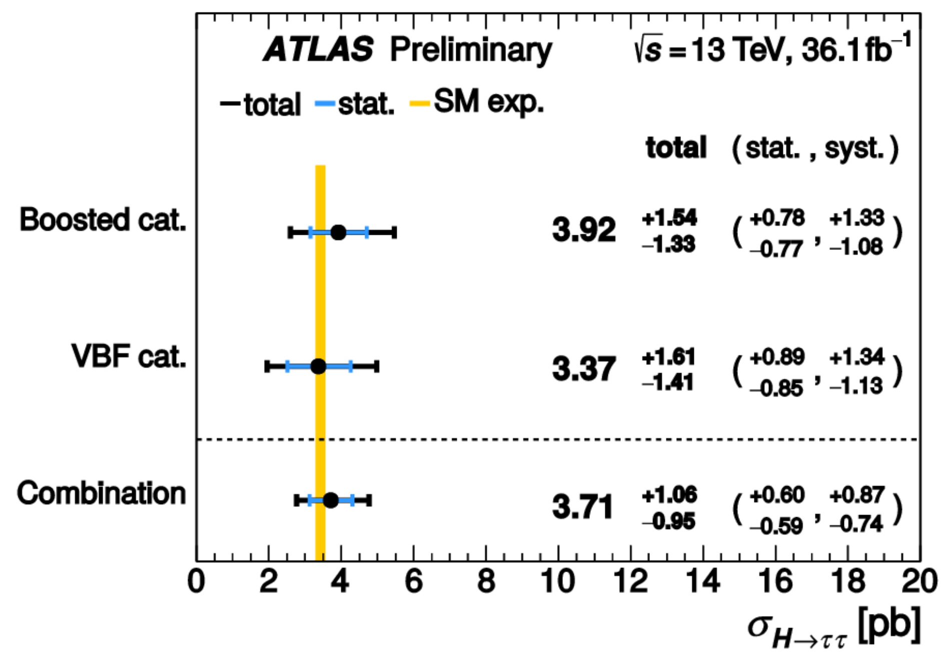
- VBF category is sensitive to VBF production and boosted is to ggF
- Results are consistent with SM
- 2d likelihood contour for the combination of all channels as a result of a two-cross-section-parameter fit
- Strongly anti-correlated results for ggF and VBF cross-sections (ρ = -52%)



Cross-sections by final states

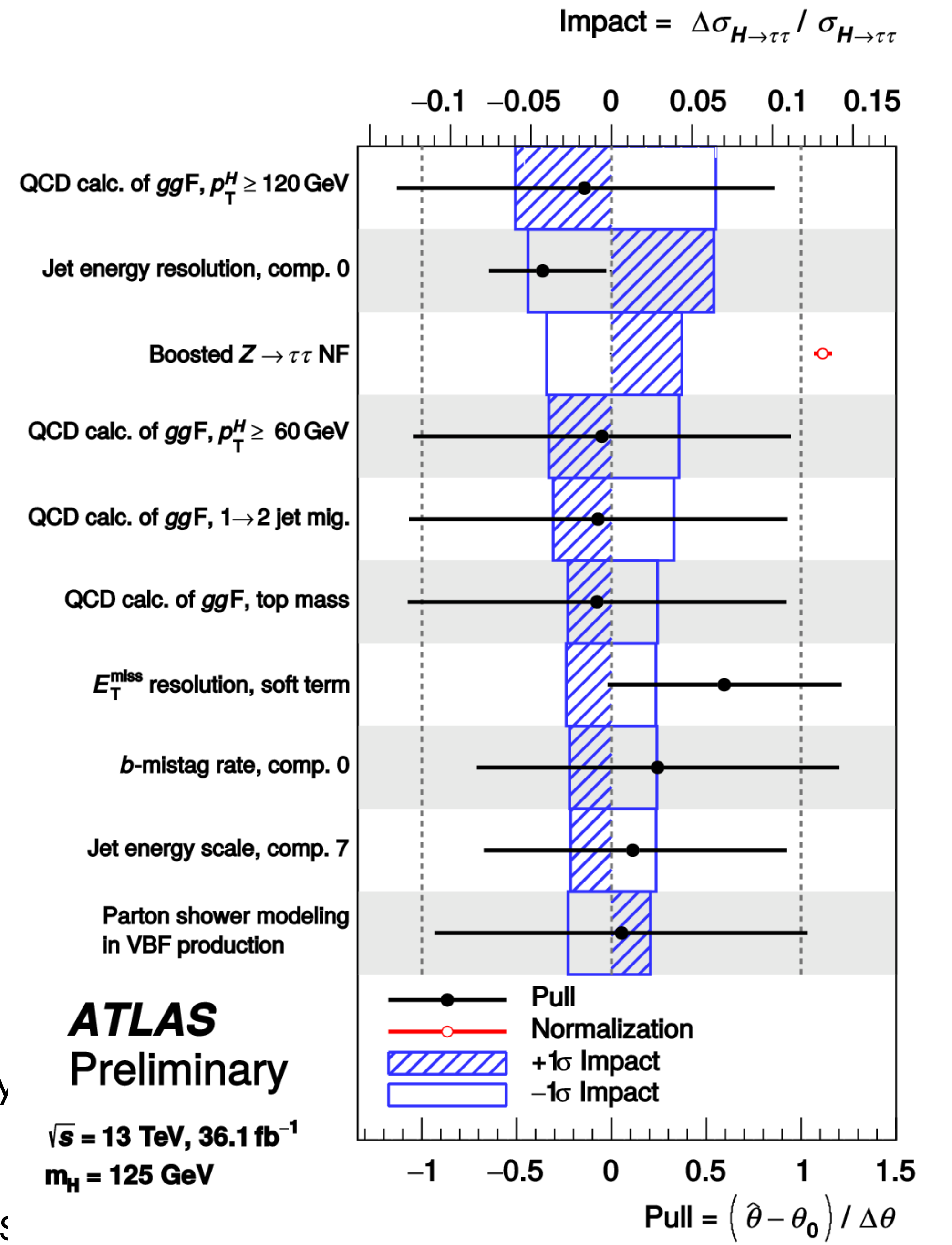


Cross-sections VBF vs ggF



$H \to \tau^+\tau^-$, Systematic uncertainties

Source of uncertainty	Impact $\Delta\sigma/\sigma_{H\to\tau\tau}$ (%) Observed	Expected
Theoretical uncert. on signal	+13.5 / −8.7	+11.9 / −7.7
Background statistics	+11 / −10	+10.2 / −9.8
Jets and E_T^{miss}	+11.5 / −9.3	+10.5 / −8.6
Background normalization	+6.8 / −4.8	+6.6 / −4.6
Misidentified τ	+4.5 / −4.2	+3.7 / −3.4
Theoretical uncert. on background	+4.6 / −3.6	+5.1 / −4.2
Hadronic taus	+4.7 / −3.0	+5.8 / −4.2
Flavour tagging	+3.3 / −2.4	+2.9 / −2.2
Luminosity	+3.3 / −2.3	+3.1 / −2.2
Electrons and muons	+1.2 / −1.0	+1.1 / −0.9
Total systematic uncert.	+24 / −20	+22 / −19
Data statistics	±16	±15
Total	+28 / −26	+27 / −25

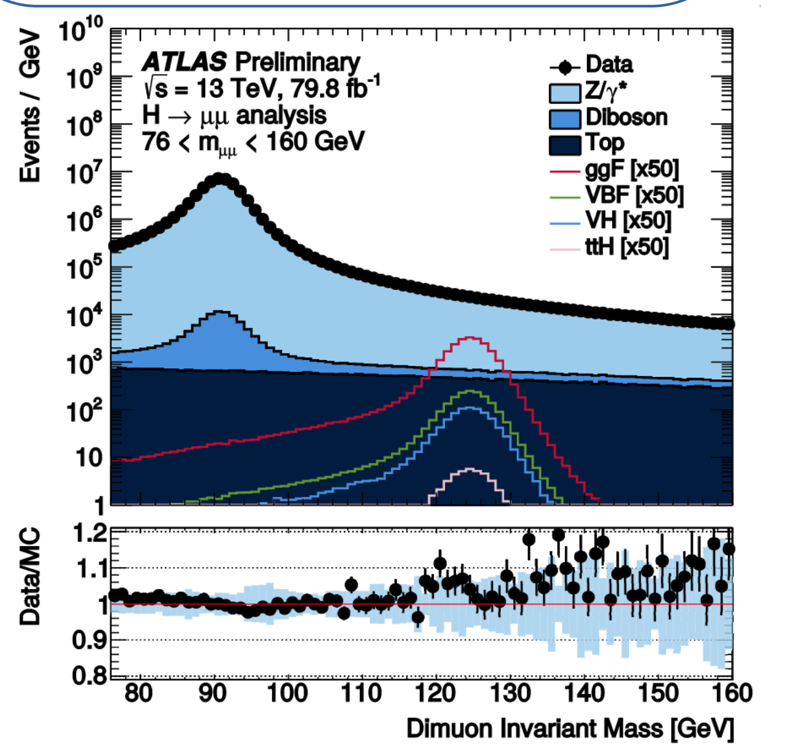


- 9 sources of theoretical unc. on signal:
  - 2 for yield due to factorization (μF) and renormalization (μR) scale variations;
  - 2 for events migration between Njets
  - 2 for migration between „low" & high pt
  - 1 for loop corrections due top mass unc.
  - 2 for acceptance unc. of ggF into VBF category
- Background statistics unc. is mainly due to the low MC statistics available for Z+jets sample in SRs phase space.

$H \rightarrow \mu^+\mu^-$, Analysis strategy

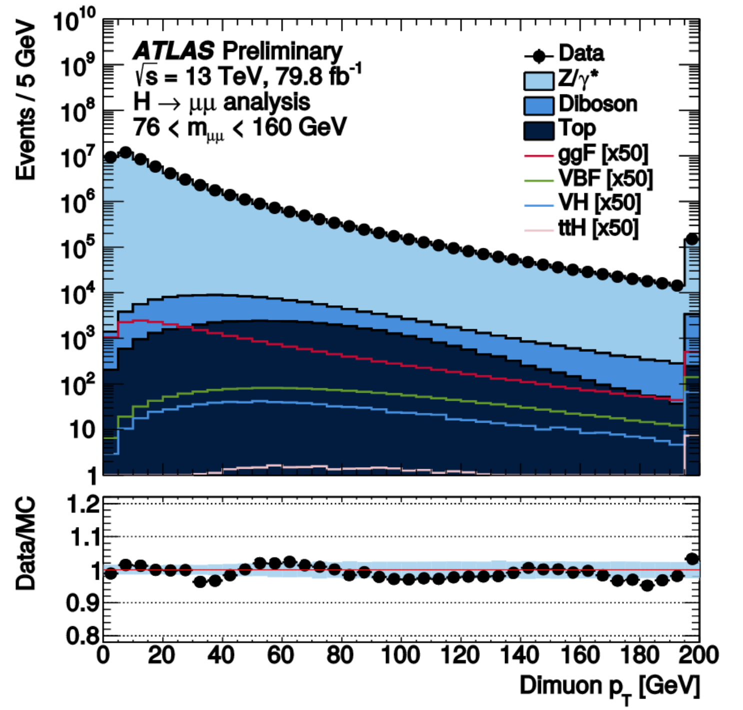
- Clean experimental signature, but very small rate compared to background
- Main background is Drell-Yan $Z/\gamma^* \rightarrow \mu\mu$
- General event selection: exactly two muons with high pt, $E_T^{miss} < 80$ GeV to suppress events with any b-jets (WZ, WW background)
- Dedicated categories for ggF and VBF to exploit kinematic differences between signal and background and increase S/B

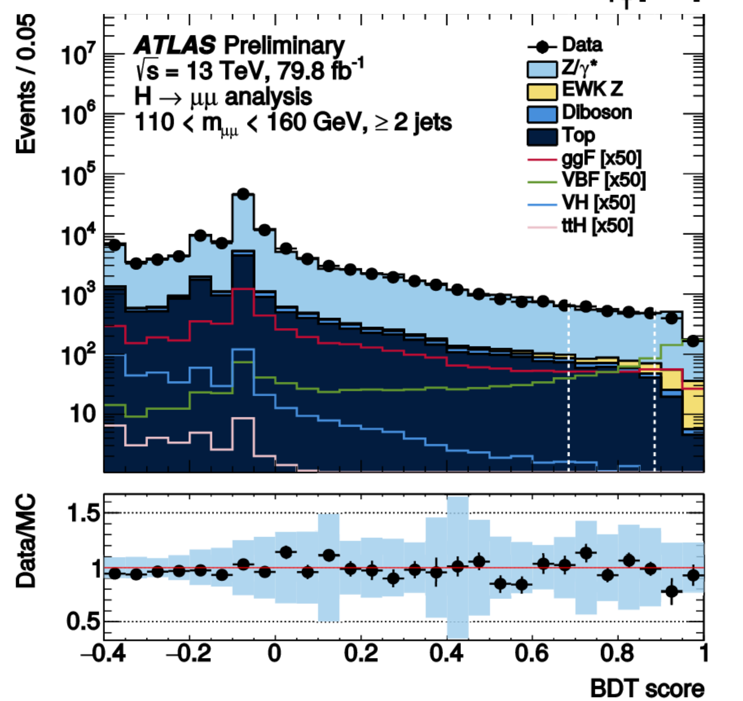
ATLAS history of the analysis:
- Run-1 analysis delivered upper limit on the signal strength: 7.1 observed (7.2 expected) for 125 GeV H. (Phys. Lett. B 738 (2014) 68)
- Corresponding result for 36.1 fb^{-1} of Run-2: $\mu < 2.8$ observed (2.9 expected) (Phys. Rev. Lett. 119 (2017) 051802)
- Results with 79.8 fb^{-1} are presented here



$H \rightarrow \mu^+\mu^-$, Categorization

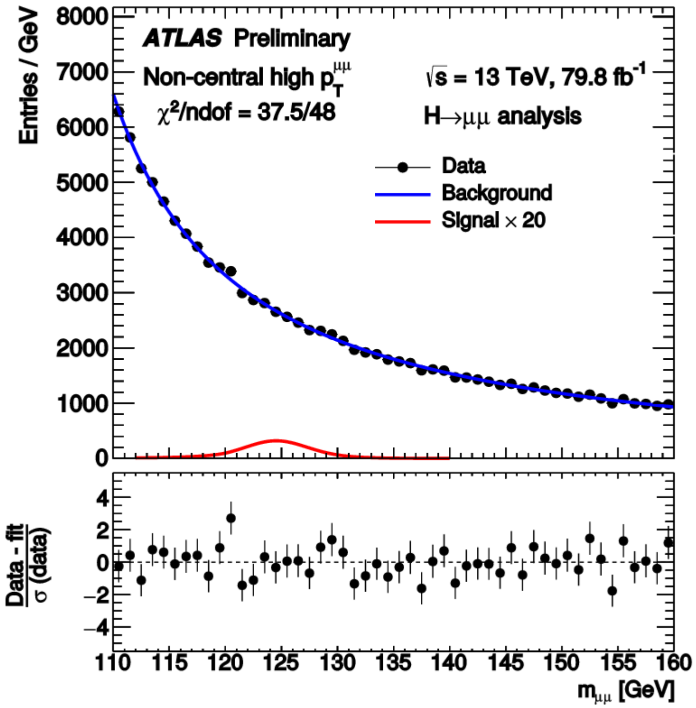
- <u>ggF</u>: signal tends to have harder pt spectrum than Drell-Yan due to QCD radiation from initial partons — events are separated into regions:
  - low pt: $pt_{\mu\mu} < 15$ GeV
  - medium pt: 15 GeV $< pt_{\mu\mu} <$ 50 GeV
  - high pt: $pt_{\mu\mu} > 50$ GeV
- <u>VBF</u>: BDT to separate signal, 14 variables (most sensitive mjj, Δηjj, ptμμ, ΔRjj).
  - BDT > 0.885 — VBF tight
  - 0.685 < BDT < 0.885 — VBF loose
  - BDT < 0.685 — ggF-like

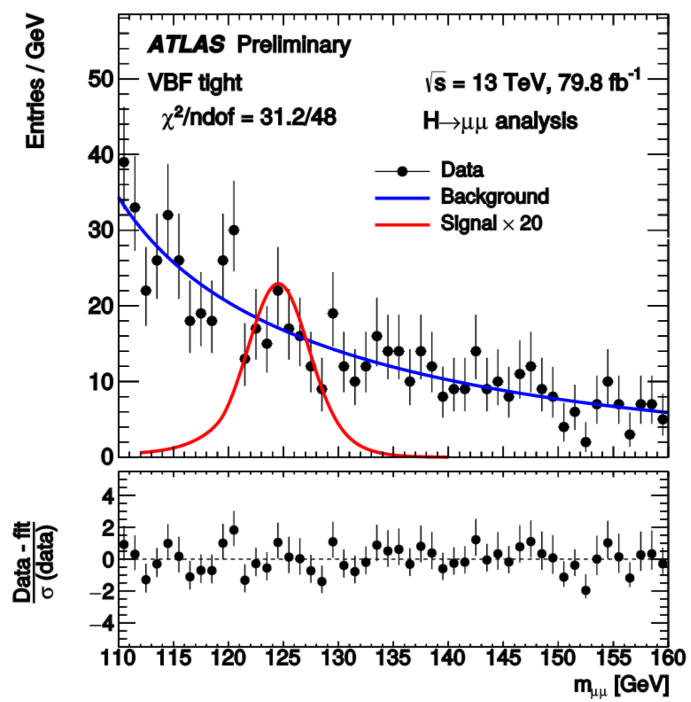




H → μ⁺μ⁻, Results

- In fit signal modelled with Crystal-Ball + Gaussian
- Background function: Breit-Wigner convolved with Gauss and exponential divided by cubic function
- No excess observed
  Expected significance: 0.9σ
  μ < 2.1 (2.0) observed (expected)





36.1 fb^{-1}

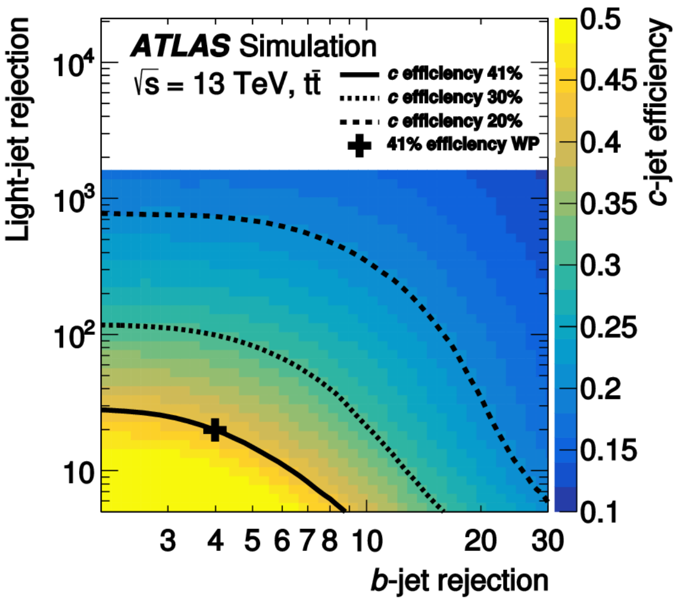
H → $c\bar{c}$

Phys. Rev. Lett. 120 (2018) 211802

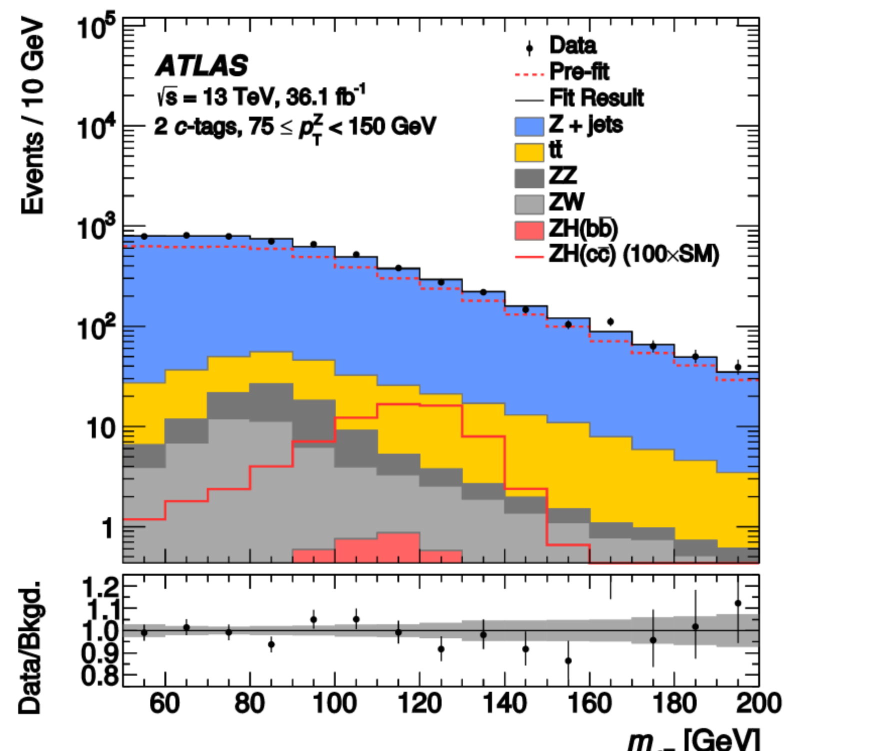
- The production mode ZH is studied. Z is tagged by decays into e^+e^- or $\mu^+\mu^-$
- BDT is trained to separate c-jets from light jets and b-jets. 41% efficiency WP selected with rejection power 4 and 20 for b- and l-jets respectively
- 4 Categories: 2 by Z pt (75 — 150 GeV, > 150 GeV) and 2 by the number of c-tags (1 or 2)
- Signal has harder Z pt distribution than Z+jets
- Joint binned maximum-profile-likelihood fit in categories to m_{cc} is performed to determine signal strength
- The observed (expected) upper limit on σ(pp→ZH)B(H→cc): 2.7 ($3.9^{+2.1}_{-1.1}$) pb, while SM value is 26 fb.

Very similar analysis H→bb will be presented today by Louis

Source	σ/σ_{tot}
Statistical	**49%**
Floating Z + jets normalization	31%
Systematic	**87%**
Flavor tagging	73%
Background modeling	47%
Lepton, jet and luminosity	28%
Signal modeling	28%
MC statistical	6%



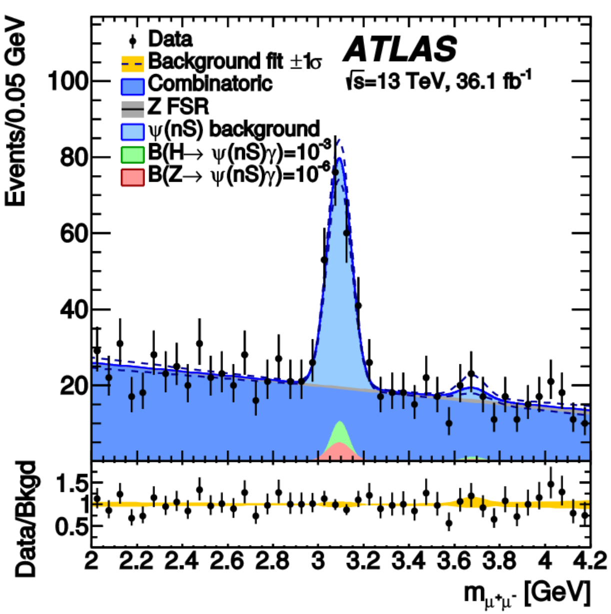
No significat excess observed so far in 36.1 fb^{-1} Run-2 dataset



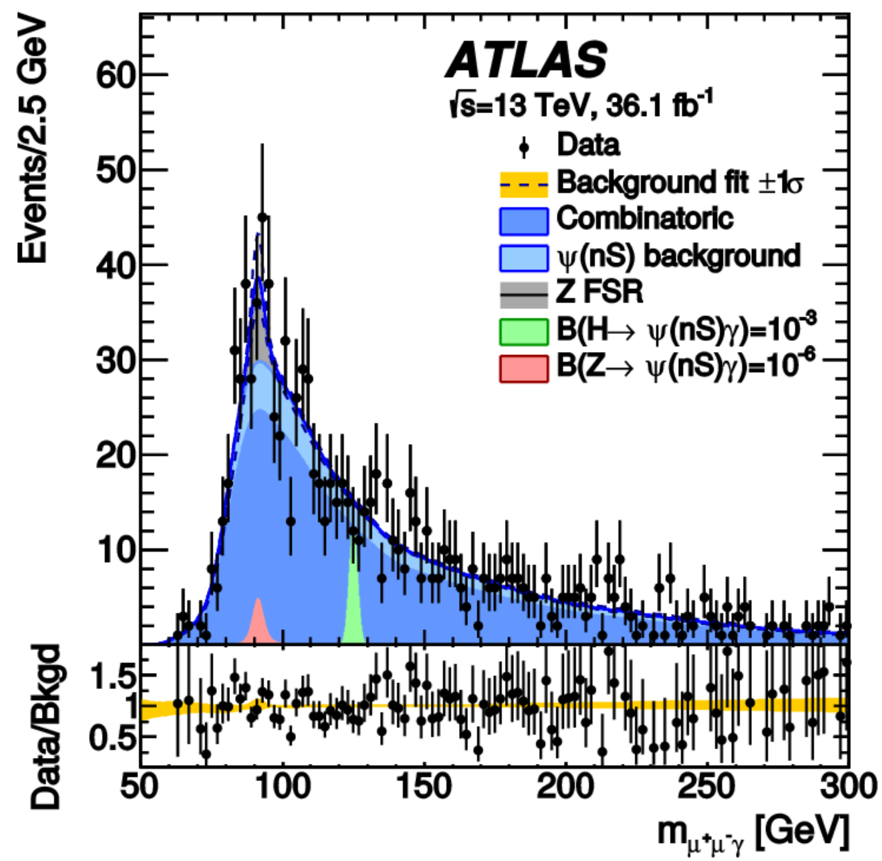
$H \rightarrow J/\psi\gamma$

- Also sensitive Yukawa coupling of H to c, but has very small branching ratio
- Only clean final state with $J/\psi \rightarrow \mu^+\mu^-$ decay is considered
- Two main sources of background:
  - non-resonant production of $\psi\gamma$. Estimated by extracting template from data and normalization from fit
  - resonant $Z \rightarrow \mu^+\mu^-\gamma$. Shape is from MC, normalization from fit to data
- Unbinned maximum likelihood fit to $m_{\mu+\mu-\gamma}$
- Upper limit on $H \rightarrow J/\psi\gamma$ branching: 3.5 * 10^{-4}, factor of 4 improvement w.r.t. Run-1 result

Source of systematic uncertainty	Yield uncertainty $H(Z) \rightarrow Q\gamma$
Total $H(Z)$ cross section	7.0% (2.9%)
Integrated luminosity	2.1%
$H(Z)$ QCD modelling	1.8% (6%)
Trigger efficiency	2.0%
Photon identification	1.4%
Muon reconstruction	2.8%
Photon energy scale	0.3%
Muon momentum scale	0.2%



No significat excess observed so far in 36.1 fb^{-1} Run-2 dataset



Conclusions

- ATLAS has an extensive program of Higgs boson properties measurements
- Higgs interactions with fermions present an important sector to test the nature of Higgs boson
- **ttH** production: provides direct sensitivity to y_t Yukawa coupling. ATLAS reached observation sensitivity with 79.8 fb^{-1}. Significance in combination with Run-1: 6.3σ (5.1σ) observed (expected).
  Measured: σ(pp→ttH) = 670 ± 90 (stat) +110/-100 (syst) fb
  Standard Model: 507 +35/-50 fb
- **H→ττ**: provides Higgs coupling measurement to 3rd generation leptons. Observation in combination of 36.1 fb-1 Run-2 data with Run-1: 6.4σ (5.4σ) observed (expected) significance.
  Measured: σ(pp→H→ττ) = 3.71 ± 0.59 (stat) +0.87/-0.74 (syst) pb
  Standard Model: 3.43 ± 0.13 pb
- **H→μμ**: allows to study Higgs couplings to 2nd generation leptons, but we are still far from reaching enough sensitivity. Signal strength limits with 36.1 fb-1 data: μ < 2.1 (2.0) observed (expected)
- **H→$c\bar{c}$ and H→J/ψγ**: both allow to study Higgs couplings to 2nd generation quarks, but current experimental sensitivity is still very far from observation.
- All results are in agreement with the Standard Model