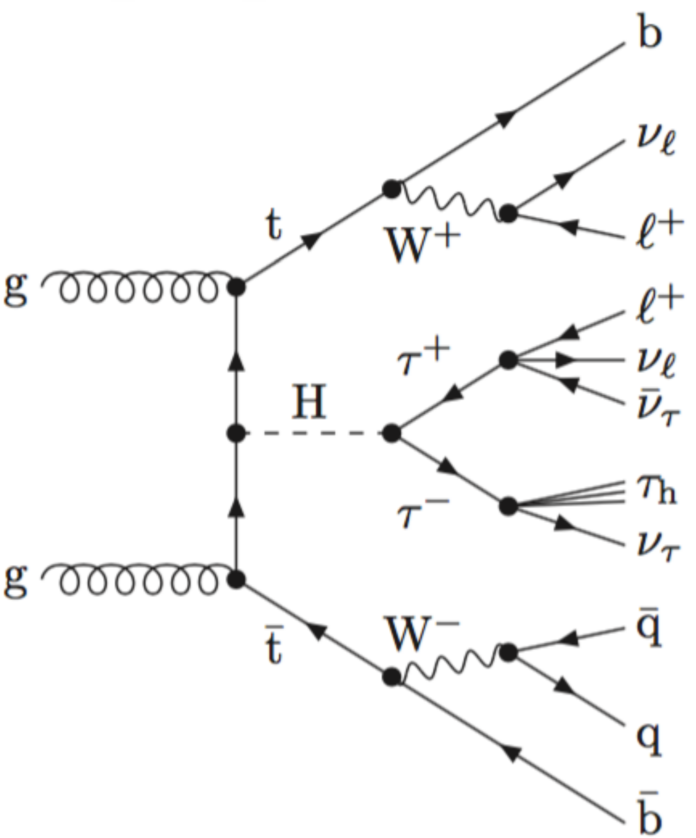
Search for ttH in final states with a τ lepton at $\sqrt{s}$=13 TeV in CMS

Thomas Strebler

Laboratoire Leprince-Ringuet
Ecole Polytechnique / CNRS-IN2P3

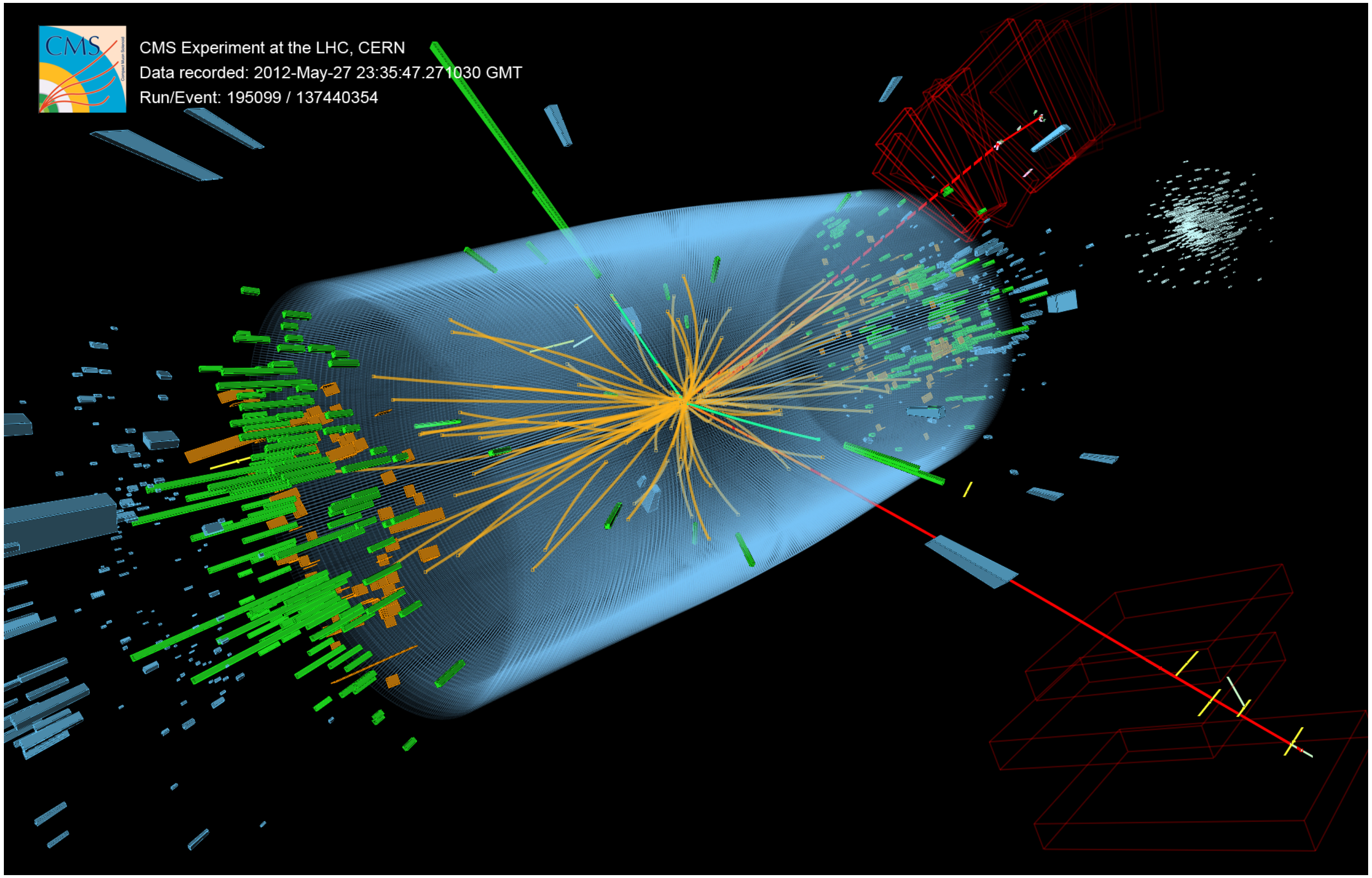
on behalf of the CMS collaboration



GDR Terascale, Montpellier
July 4th, 2017

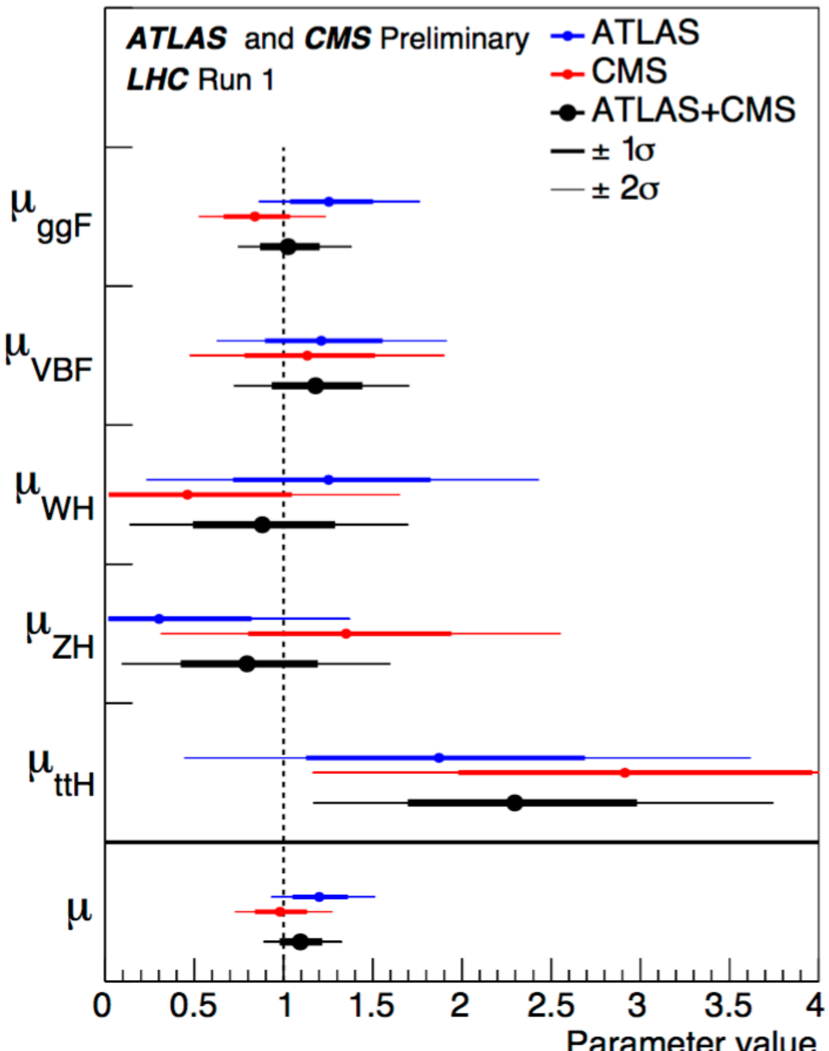
Introduction

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



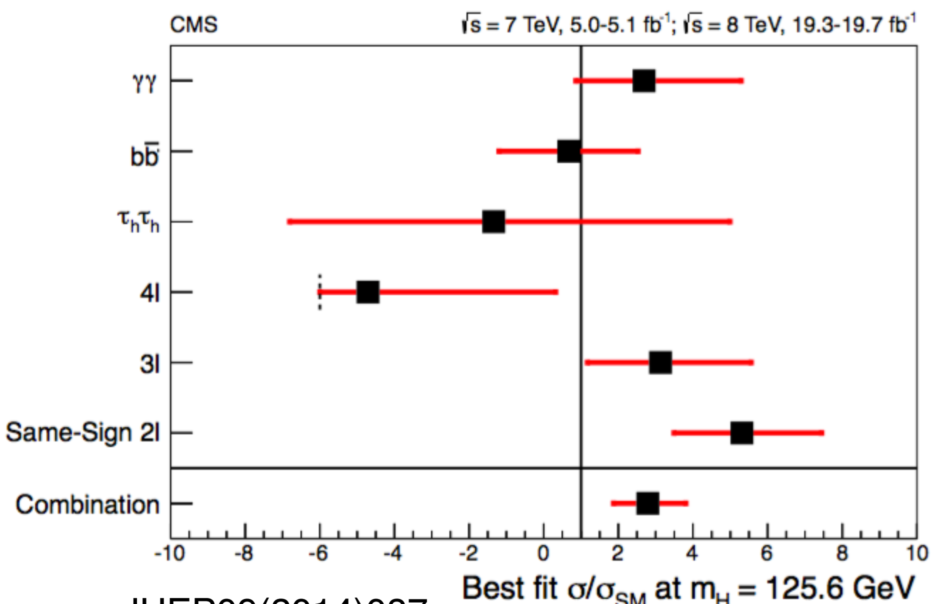
- Almost **all accessible decay modes have been observed** by both ATLAS and CMS collaborations during Run 1

ttH in Run 1



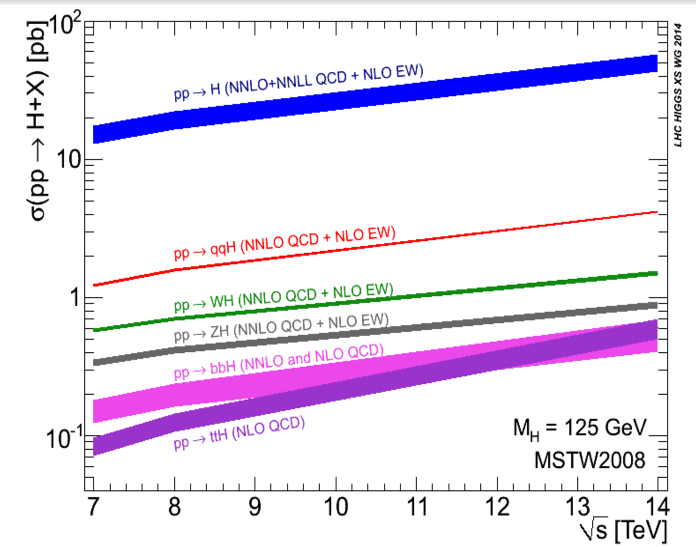
JHEP08(2016)045
arXiv:1606.02266 **[hep-ex]**

- **Associated production modes** started to be scrutinized in Run 1 already
- ATLAS+CMS combination gave **strong evidence for ttH production**: 4.4σ (2σ) observed (expected) significance
- 2σ excess over the SM prediction

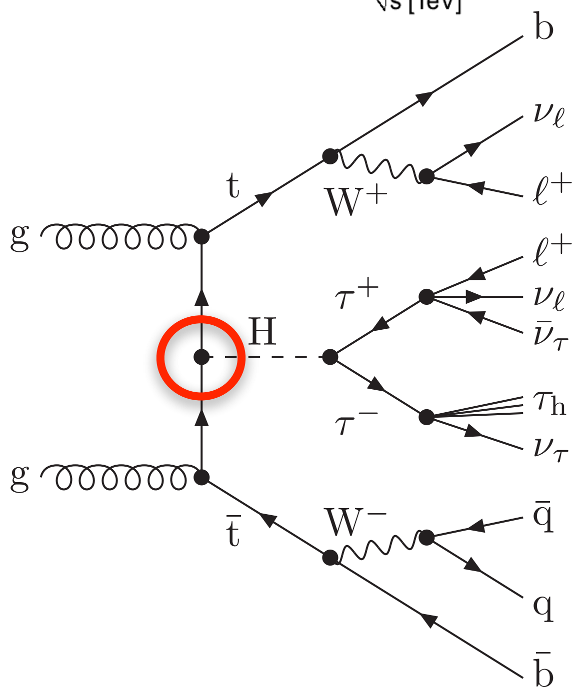


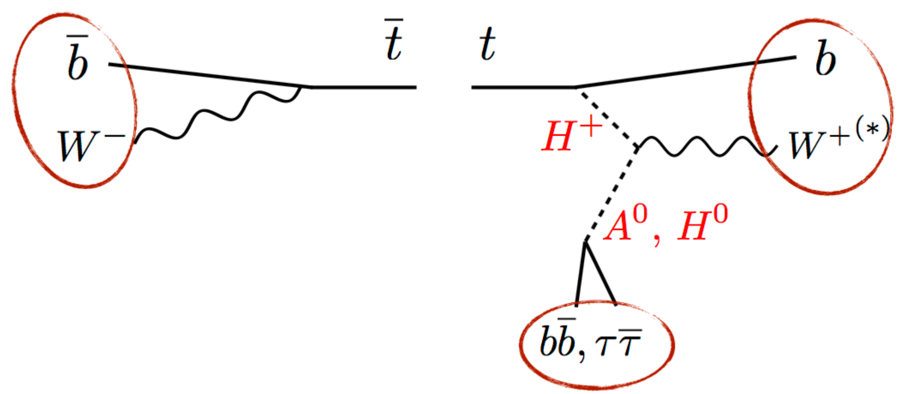
JHEP09(2014)087
arXiv:1408.1682 **[hep-ex]**

Motivations to study ttH



- **ttH => direct probe to study Y_t**
- Large cross-section boost from 8 to 13 TeV (x4): 0.51pb @ 13 TeV
- **Expected to be observed by the end of LHC Run 2** but important to cover as much decay modes as possible
- Final states also sensitive to BSM physics, in particular 2HDM models
  - **arXiv: 1602.06198 [hep-ph]**
  - **arXiv: 1703.06834 [hep-ph]**



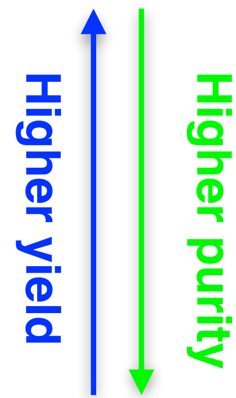


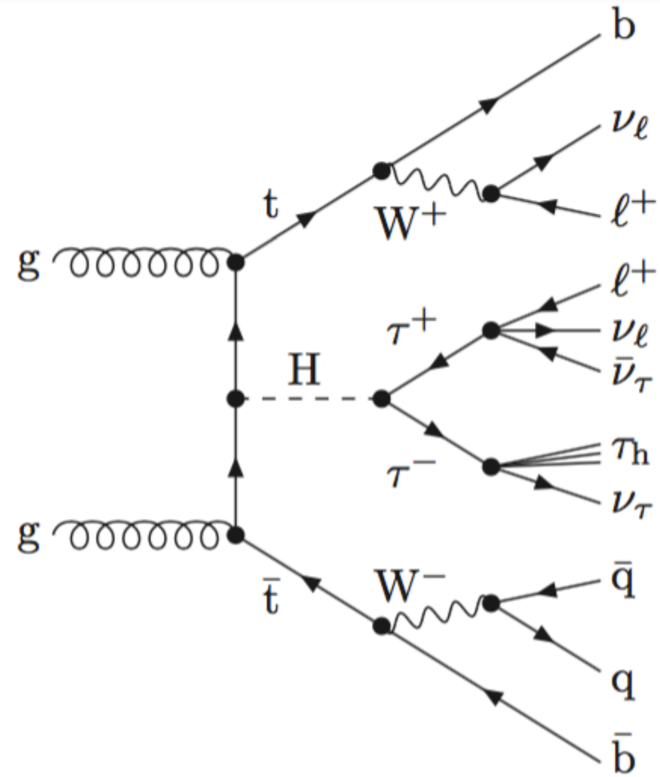
Complex final states

- Look for 2 tops + 1 Higgs

- **Top decays:**
  - t->blv : 1 b-jet + 1 lepton + MET
  - t->bqq : 1 b-jet + 2 light jets

- **Higgs decay:**
  - H->bb : 2 b-jets
  - H->$\tau\tau$: 1-2 τ_h (+ lepton + MET)
  - H->WW/ZZ : leptons (+ jets + MET)
  - H->$\gamma\gamma$: 2 photons
  - H->ZZ->4l: 4 leptons



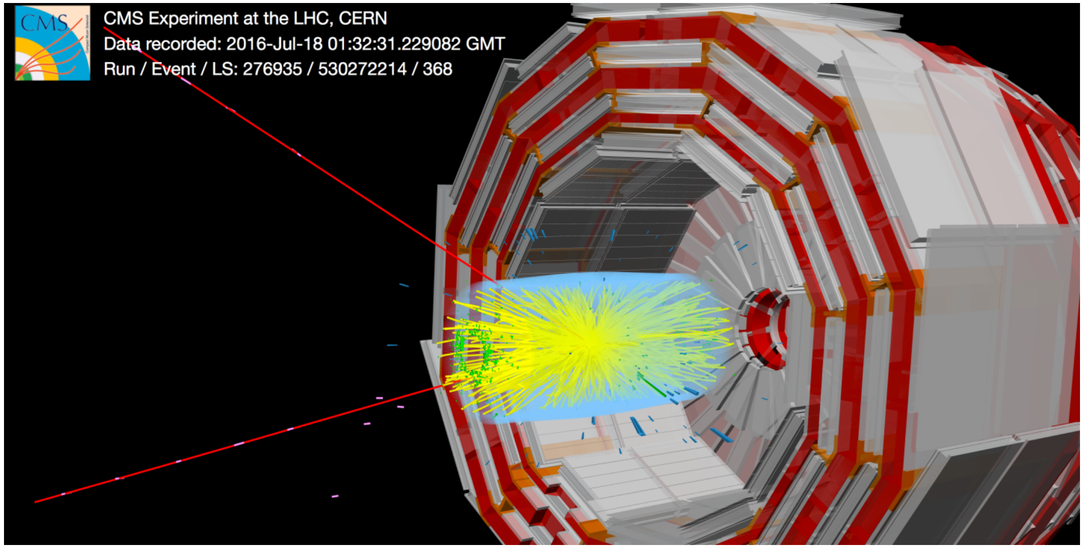


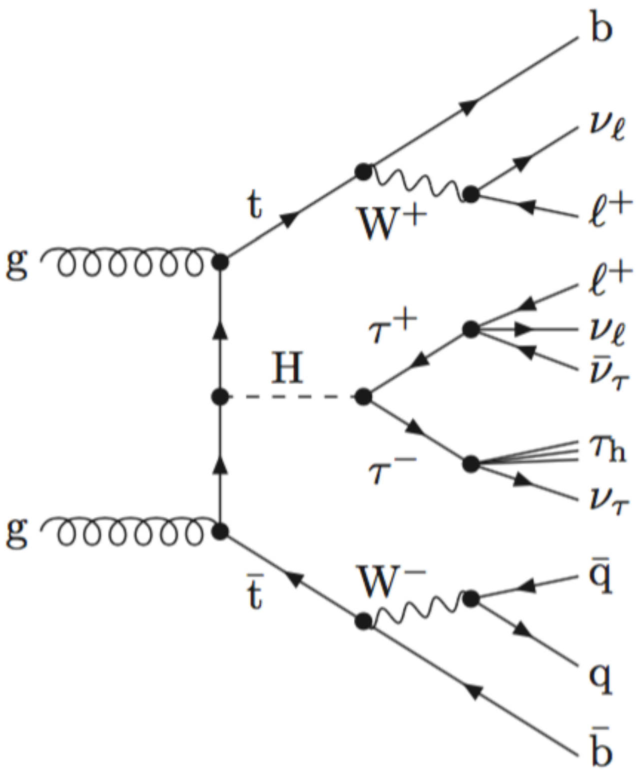
- Example: ttH, t->blv, t->bqq, H->$\tau\tau$->lτ_h
  Signature with 2 b-jets + 2 light jets + 2 leptons + 1τ_h => almost every SM particle

- Signal extraction can be challenging: **extensive use of MVA discriminants**

Matrix Element Method

- Combines observables from reconstructed objects **y** (leptons, τ_h, jets, MET)…





…with theoretical description in terms of particules, described by phase-space point **x** (τ before decay, quarks, neutrinos)

- Used to define event weights

$$w_\Omega(\mathbf{y}) \propto \sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a, Q) f(x_b, Q)}{x_a x_b s} \delta^2(x_a P_a + x_b P_b - \sum p_k) |\mathcal{M}_\Omega(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

Matrix Element Method

- **Event weight computed for hypothesis Ω** (Ω=ttH, ttV…), using observables **y** as inputs and **integrating** over unmeasured or poorly measured quantities **x**

Parton density functions

Hard-scattering matrix element: computed at LO with MadGraph, OpenLoops…

$$w_\Omega(\mathbf{y}) \propto \sum_p \int d\mathbf{x} dx_a dx_b \frac{f(x_a, Q) f(x_b, Q)}{x_a x_b s} \delta^2(x_a P_a + x_b P_b - \sum p_k) |\mathcal{M}_\Omega(\mathbf{x})|^2 W(\mathbf{y}||\mathbf{x})$$

Numerical integration: performed with VEGAS, Markov chain integrator…

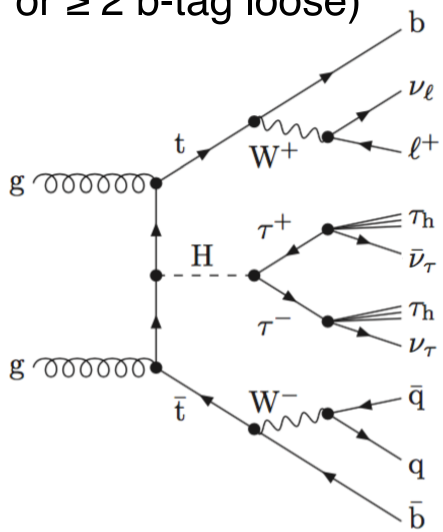
Momentum conservation (+ on-shell conditions): used to reduce dimensionality

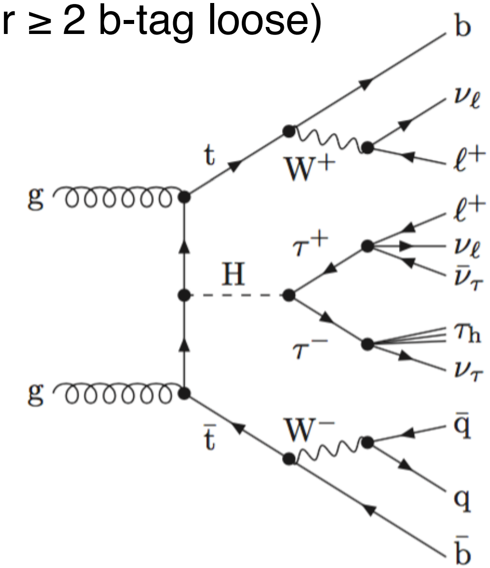
Transfer functions: probability to measure **y** given a parton configuration **x**

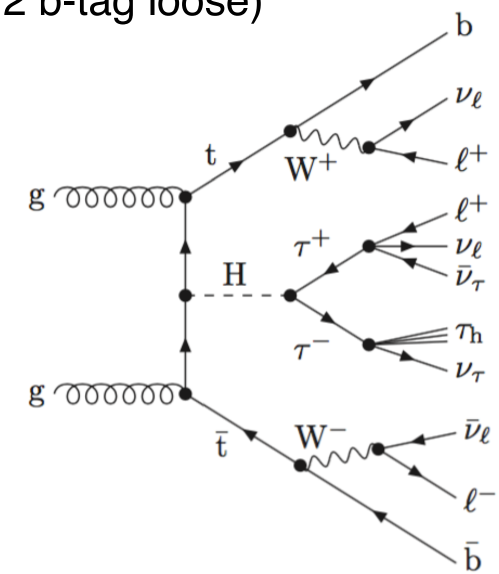
- The **final discriminant** used is in principle $\mathcal{L}(\mathbf{y}) = \dfrac{w_S(\mathbf{y})}{w_S(\mathbf{y}) + w_B(\mathbf{y})}$

ttH τ+X CMS PAS HIG-17-003

- Complements ttH multilepton analysis (CMS PAS HIG-17-004) with **final states with ≥1 τ_h**
- Benefits from **dedicated TauID with smaller isolation cone size (ΔR=0.3)** wrt default TauID (ΔR=0.5)
- **Three main channels:**
  - **1l+2τ_h** + ≥ 3 jets (≥ 1 b-tag medium or ≥ 2 b-tag loose)
  - **2lSS+1τ_h** + ≥ 3 jets (≥ 1 b-tag medium or ≥ 2 b-tag loose)
  - **3l+1τ_h** (with Z-veto) + ≥ 2 jets (≥ 1 b-tag medium or ≥ 2 b-tag loose)

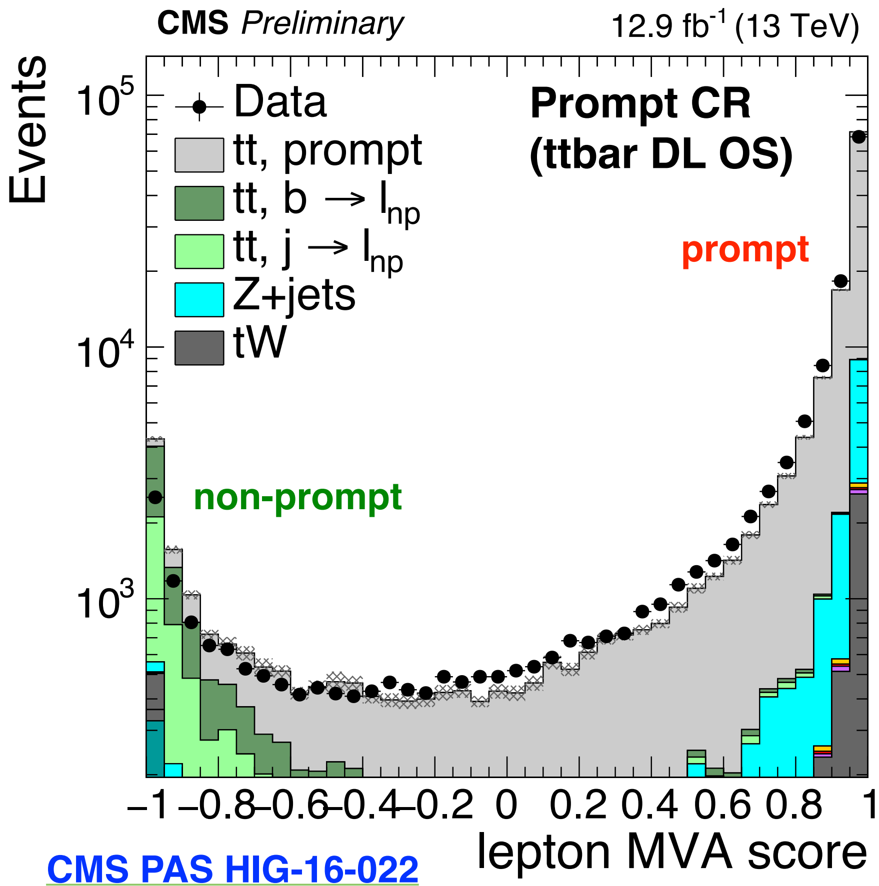






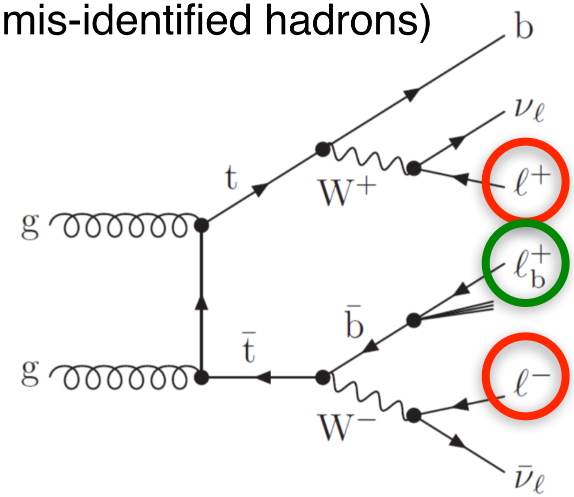
- **Large tt + fake τ_h background** (data-driven)
- **Main sources of background:**
  - **irreducible:** ttV, di-boson (from MC)
  - **reducible:** non-prompt leptons and charge mis-ID (data-driven)

Non-prompt leptons background



CMS PAS HIG-16-022

- **Leptons** selected with **MVA** trained to discriminate **prompt leptons** (from W, Z or τ decays) from **non-prompt leptons** (from b-jets and mis-identified hadrons)



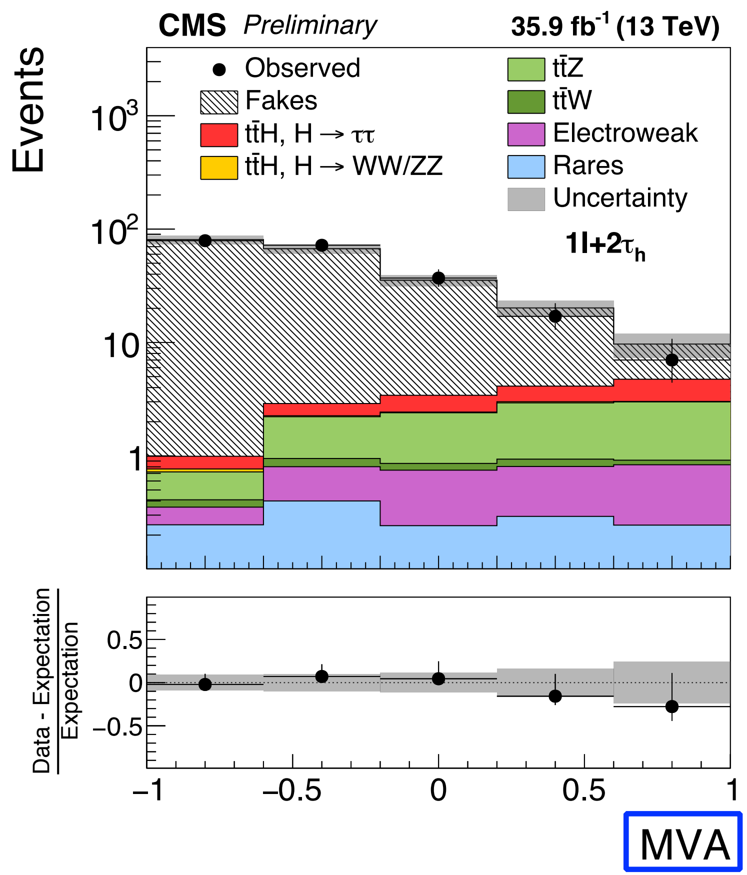
- **Inputs**:
  - isolation
  - vertex
  - lepton ID
  - jet variables
- Performance validated in data control regions (CR)

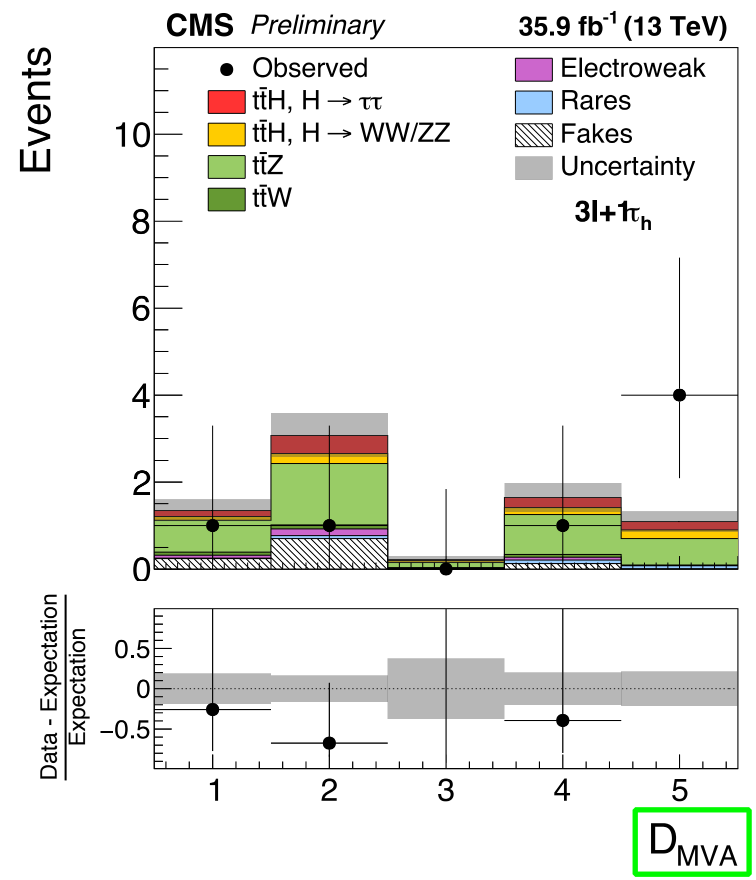
- Residual background with non-prompt leptons evaluated using **tight-to-loose fake rate method**
- In **1l+2τ_h channel**, extended to cover also **jets faking τ_h**

Signal extraction

CMS PAS HIG-17-003

- **1l+2τ_h channel:** BDT trained to discriminate ttH signal from tt background
- **3l+1τ_h channel**: 2 BDTs trained separately to discriminate ttH / ttV and ttH / tt
  2D distribution remapped into a **1D discriminant** according to S/B



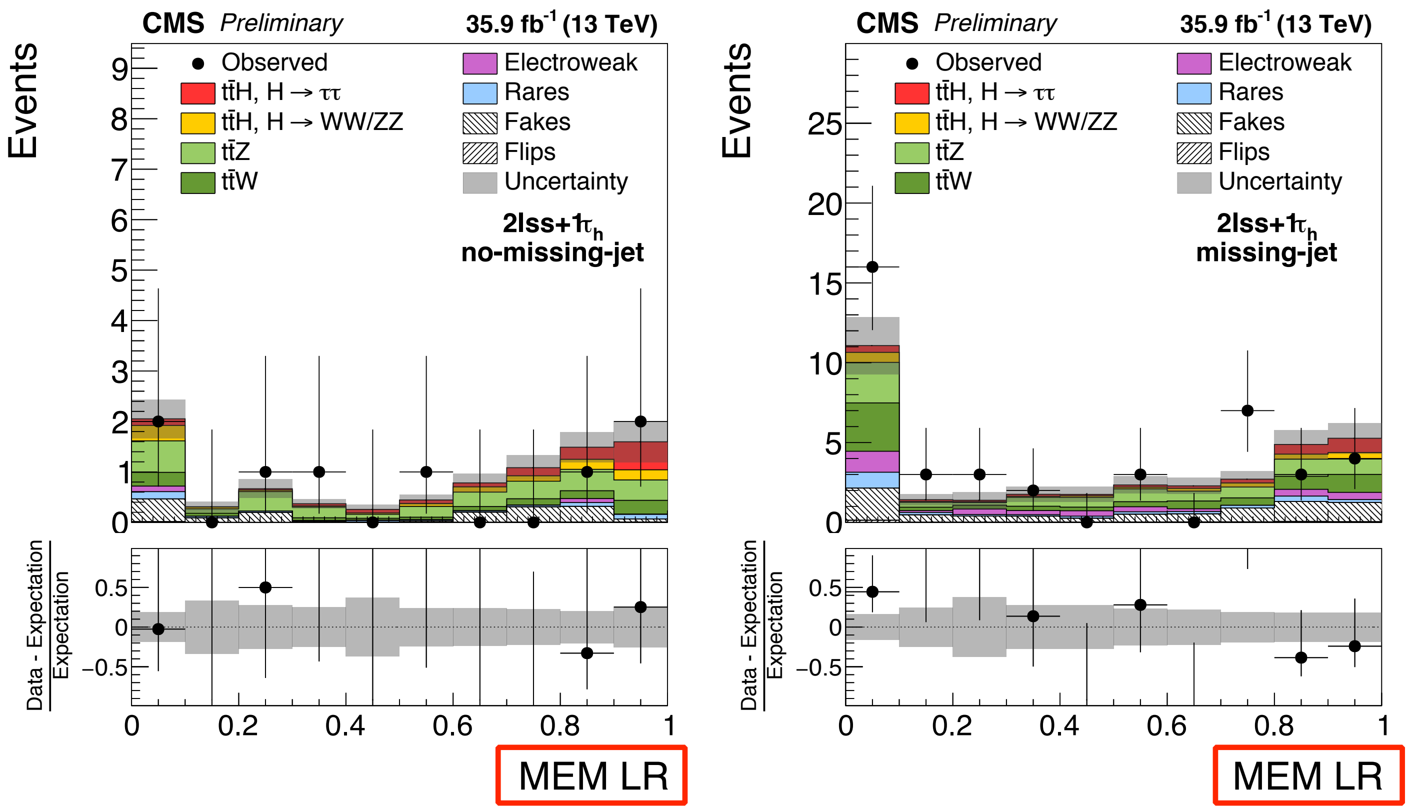


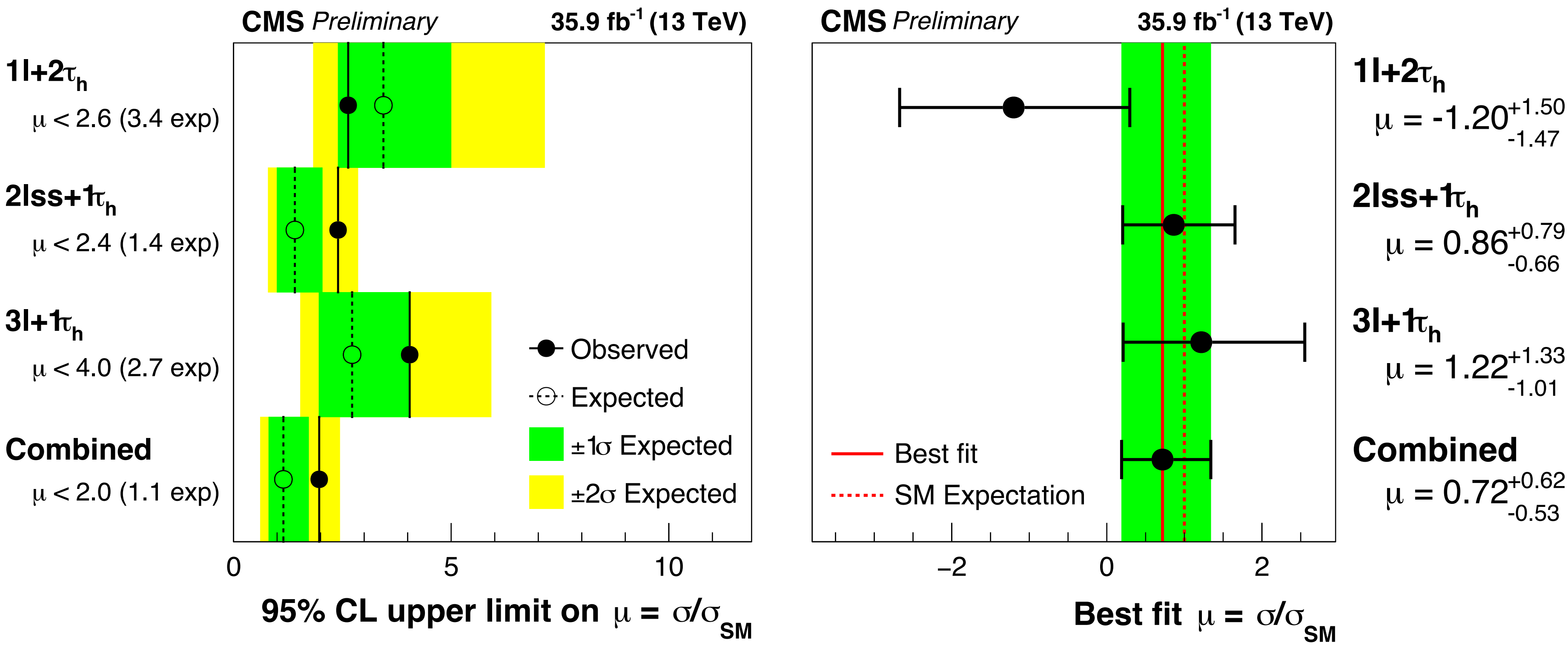
Signal extraction

CMS PAS HIG-17-003

- **2lSS+1τ_h channel:** **MEM discriminant** optimized to discriminate ttH H->$\tau\tau$ signal from ttZ Z->$\tau\tau$ + ttZ Z->ll + tt w/ non-prompt lepton

- Events split between two subcategories based on presence of **jets compatible with W->qq decay**: MEM integration performed on direction of missing jet if needed



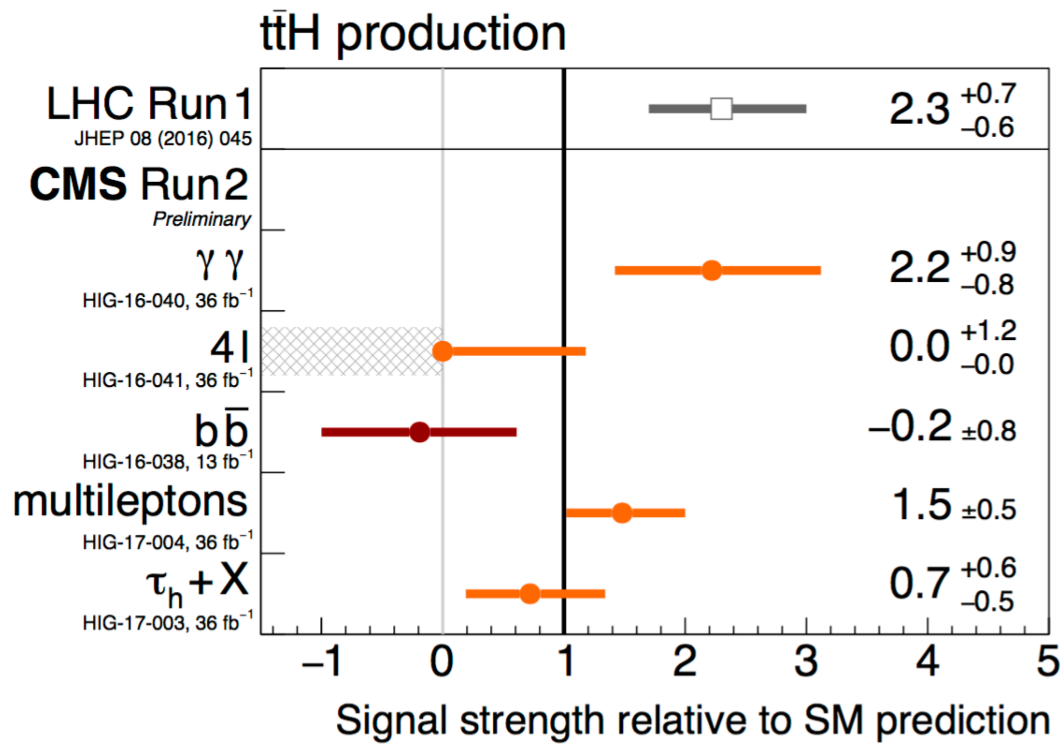


- Slight excess observed over the background-only hypothesis:
  1.4σ observed significance (1.8σ expected)

- Combined results compatible with the Standard Model expectation

Conclusion

- **ttH search in final states with τ** complements coverage of ttH phase space in CMS
- Benefits from **dedicated signal extraction methods** optimized for final states with τ
- Not unambiguous evidence by itself but **consistent with other ttH searches**
- **Combination with other ttH searches** will help to get closer to 5σ discovery
- **With 2017 dataset, statistics expected to increase by a factor 2**: improvement in signal extraction + systematics evaluation under study to get optimal result



Back-up

Matrix Element Method

- **Advantages**
  - In principle, optimal combination of **theoretical information** (matrix element) with **detector resolution** (transfer function)
  - **Can treat complex final states** with several relevant observables (jets, τ, top quarks…), including polarization + non-reconstructed objects
  - **No training required**

- **Drawback**
  - Demanding in terms of computing ressources for MC integration
    => **implementations on GPU's** currently under development: **~30 times faster** from preliminary results
    see http://www.roma1.infn.it/conference/GPU2016/pdf/talks/Grasseau.pdf

- **Complex ttH final states ideal playground for MEM:** already used in several ttH analyses (H→bb, multilepton, τ+X)

- **BDT inputs 1l+2τ**
  - The invariant mass and ΔR separation of the two reconstructed τ_h.
  - The transverse momenta of the two reconstructed τ_h.
  - The observable H_T^{miss}, computed according to Eq. (1).
  - The average ΔR separation between any pair of jets.
  - The multiplicity of jets, with and without b-tagging criteria applied.

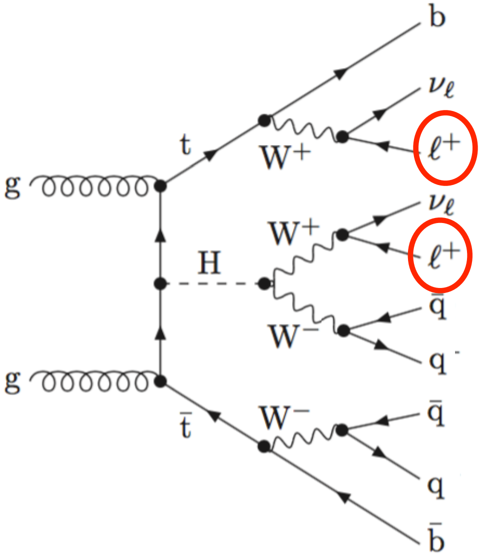
- **BDT inputs 3l+1τ**
  - The transverse momenta of the leading lepton and of the trailing lepton.
  - The maximum $|\eta|$ of the two leading leptons.
  - The multiplicity of jets.
  - The ΔR separation of the leading and of the subleading lepton with respect to the nearest jet.
  - The transverse mass of the leading lepton and the missing transverse energy vector.
  - The observable H_T^{miss}.
  - The average ΔR separation between any pair of jets.

ttH multilepton [CMS PAS HIG-17-004]

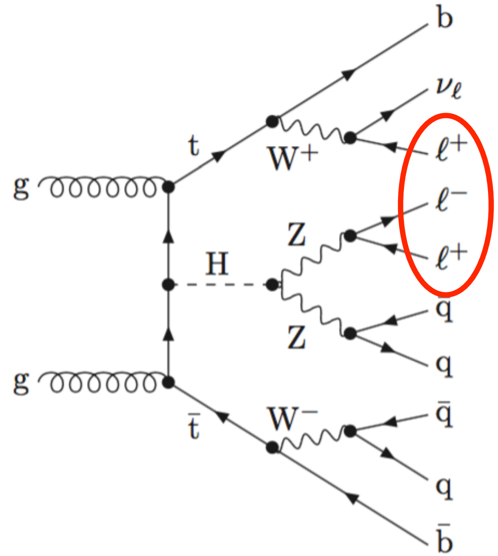
- **Multilepton final states from H→WW/ZZ + τ_h-veto**

- **Three main channels:**

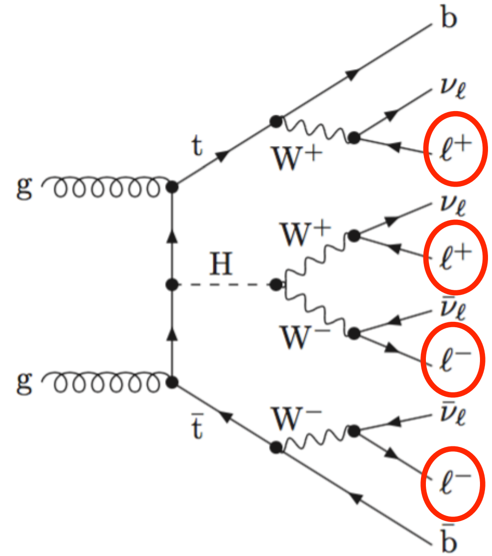
  - **2 same-sign leptons**
    + ≥ 4 jets (≥ 1 b-tag medium or ≥ 2 b-tag loose)



  - **3 leptons** (with Z-veto)
    + ≥ 2 jets (≥ 1 b-tag medium or ≥ 2 b-tag loose)



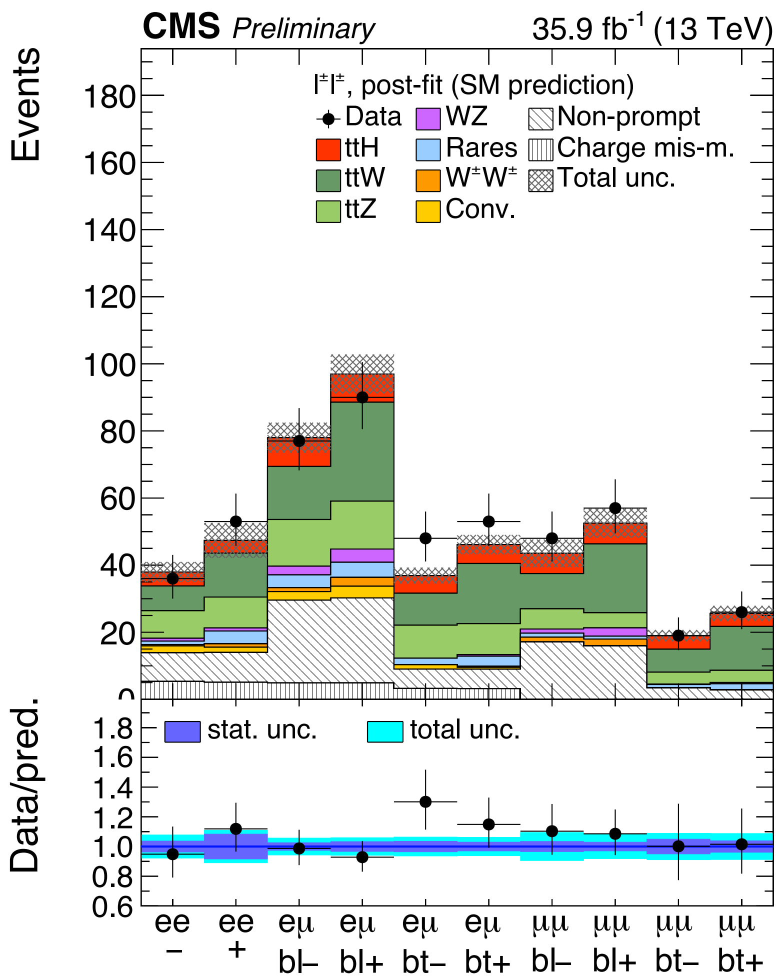
  - **4 leptons** (with Z-+ H→4l-veto)
    + ≥ 2 jets (≥ 1 b-tag medium or ≥ 2 b-tag loose)

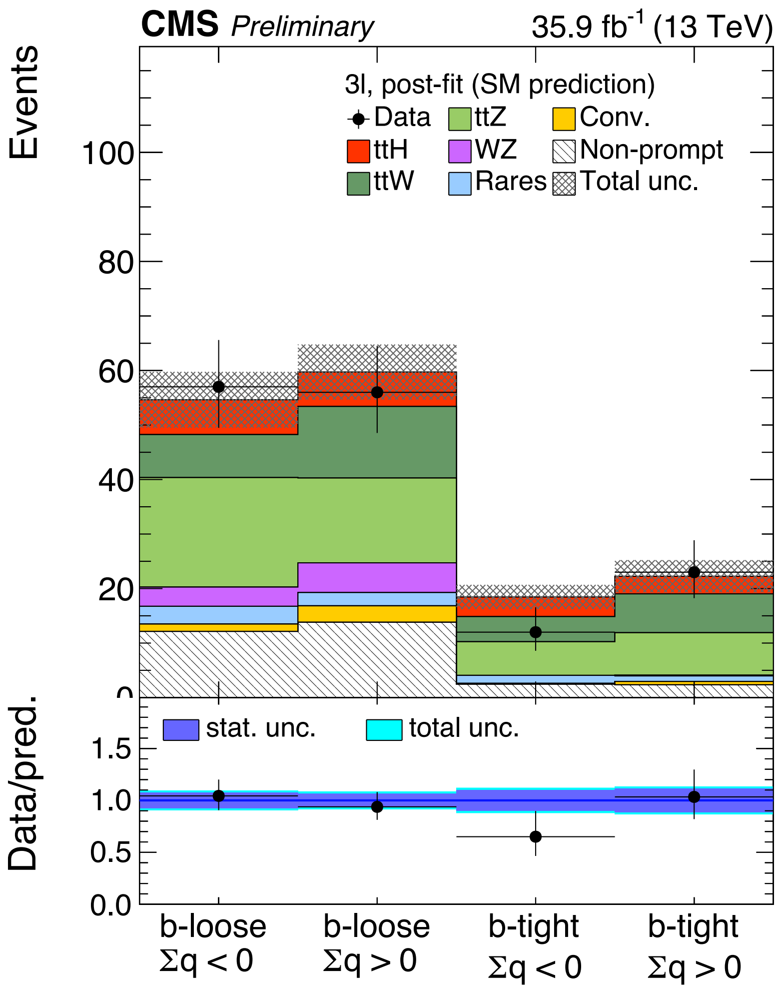


- **Main sources of background:**
  - **irreducible:** ttV (from MC), di-boson (normalization from data)
  - **reducible:** non-prompt leptons and charge mis-ID (data-driven)

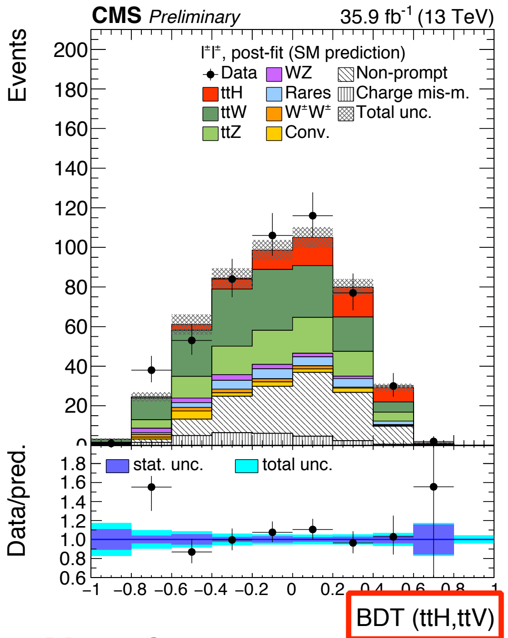
ttH multilepton [CMS PAS HIG-17-004]

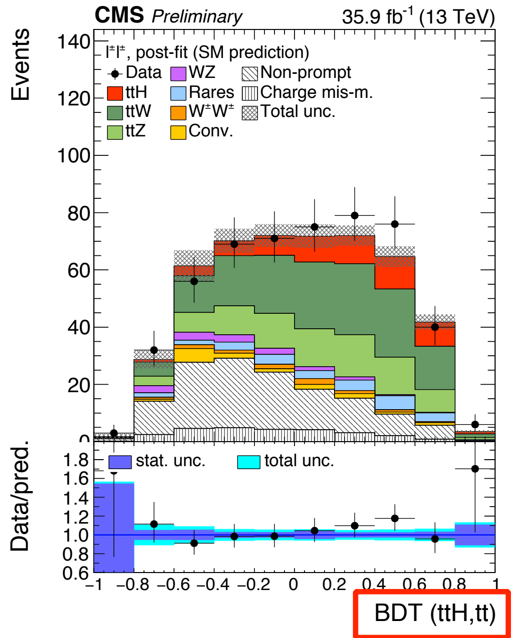
- Subcategorization based on **lepton flavor, lepton charge, b-tagging**



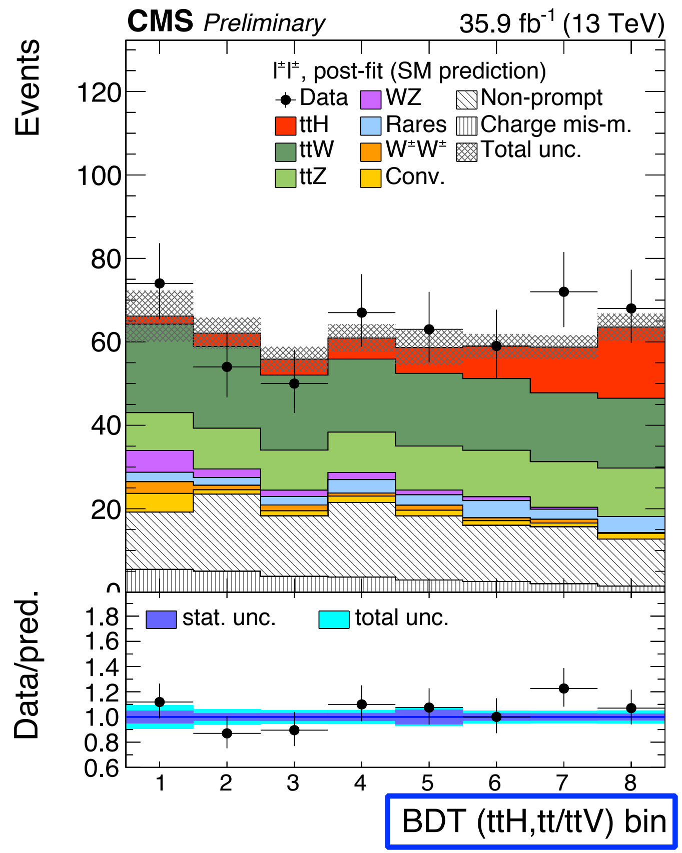


ttH multilepton [CMS PAS HIG-17-004]



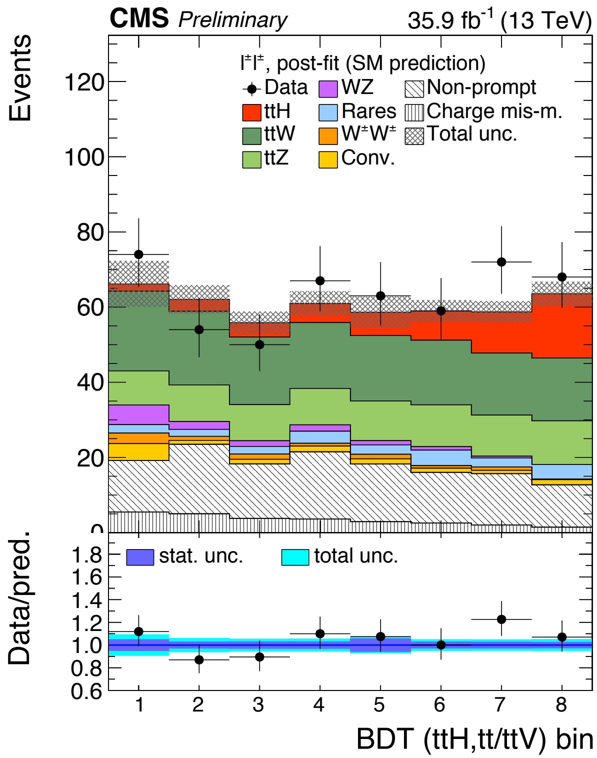


- **Signal extraction in 2lSS + 3l categories based on 2 BDTs** trained to discriminate **ttH / ttV** and **ttH / tt**
  Inputs: jet multiplicity, lepton/jet angular separation, MET, lepton p_T

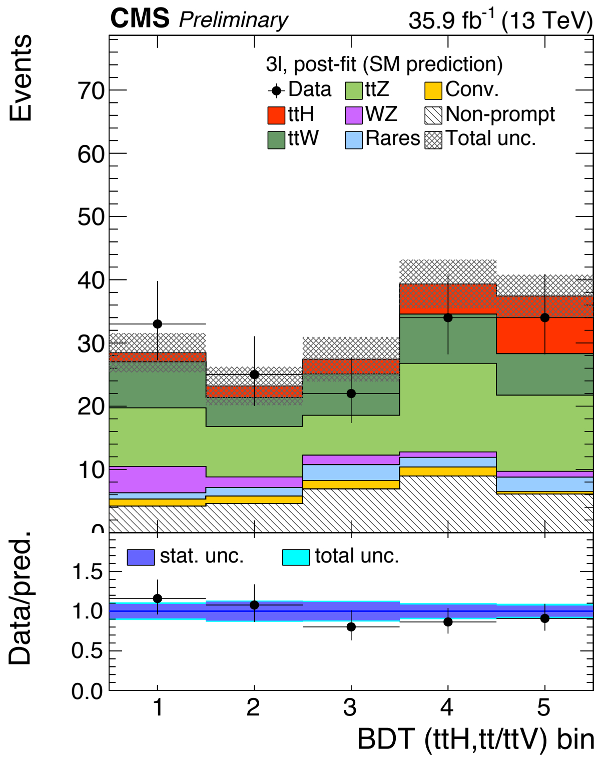


- **New for analysis on full 2016 dataset:**
  - hadronic top MVA-tagger used as input for ttH / tt 2lSS BDT
  - Hj MVA-tagger (jets from H decay) used as input for ttH / ttV 2lSS BDT
  - MEM LR ttH / ttV used as input for ttH / ttV 3l BDT
- **2D BDT** distributions remapped into a **1D discriminant** with increasing S/B
- Counting experiment in 4l category

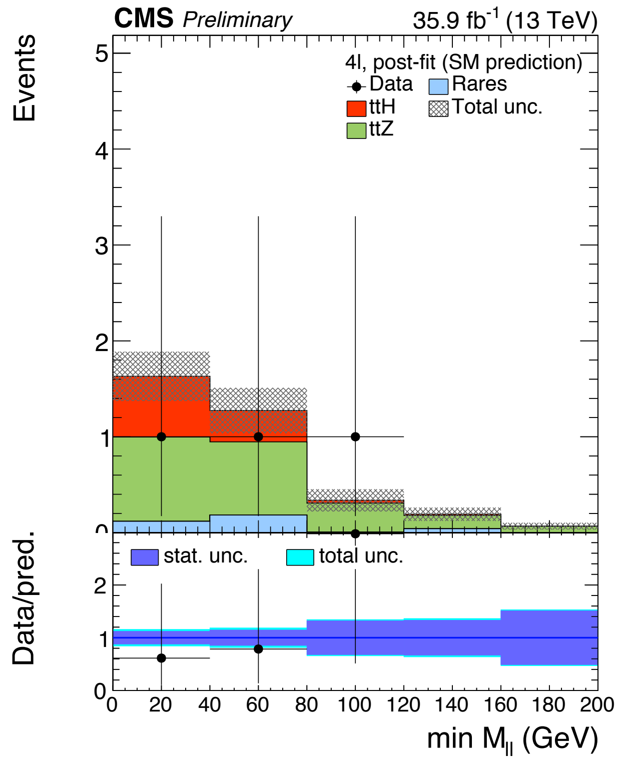
- **Final discriminants**



2lSS



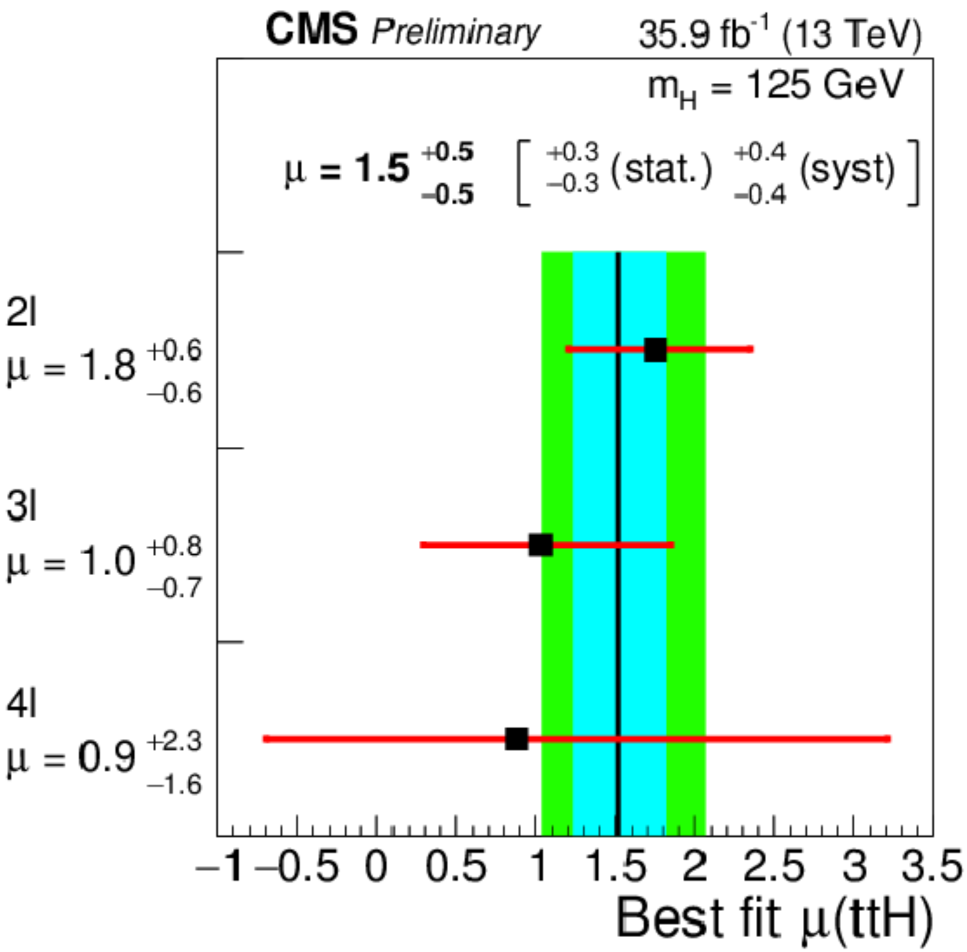
3l



4l (counting experiment)

ttH multilepton

CMS PAS HIG-17-004



- **3.3σ observed significance**
  (2.5σ expected)

μ(ttW)>1?

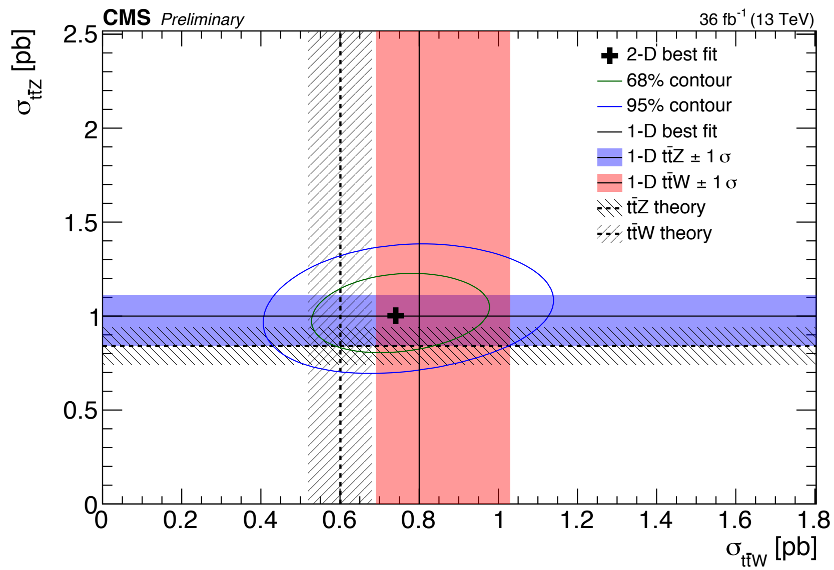
- **Consistent pattern of μ(ttW)>1 across ATLAS & CMS, Run 1 & Run 2:**
  - for several analyses (ttH search vs ttV measurement, cut-based vs MVA…)
  - not the case for μ(ttZ) nor other searches in 2lSS (e.g. SUSY or WW same-sign)

- **ttW measurement (assuming μ(ttH)=1)**

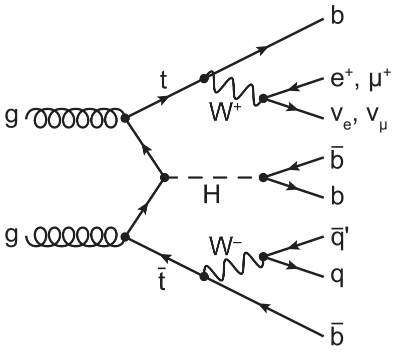
μ(ttV) = **1.3 ± 0.6**	CMS	5 fb^{-1}	7 TeV	PRL 110 (2013) 172002
μ(ttW) = **1.7 ± 0.5**	ATLAS	20 fb^{-1}	8 TeV	JHEP 11 (2015) 172
μ(ttW) = **1.9 ± 0.6**	CMS	20 fb^{-1}	8 TeV	JHEP 01 (2016) 096
μ(ttW) = **2.5 ± 1.4**	ATLAS	3 fb^{-1}	13 TeV	EPJC77 (2017) 40
μ(ttW) = **1.3 ± 0.3**	CMS	36 fb^{-1}	13 TeV	CMS PAS TOP-17-005

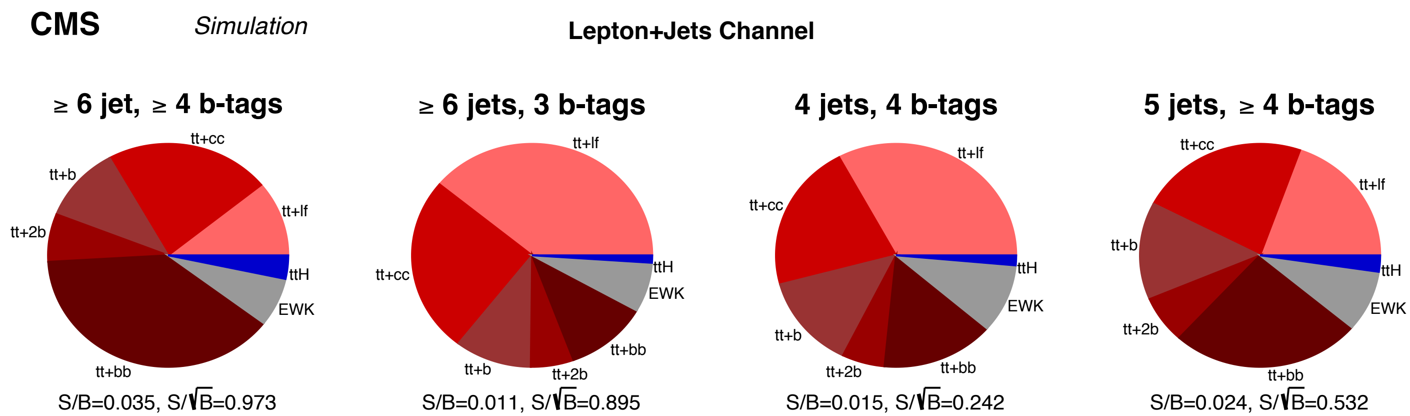
CMS PAS TOP-17-005

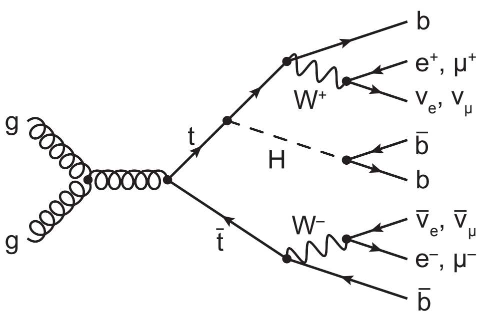
- **ttH multilepton cross-check with free floating ttW & ttZ**
  - including in fit CRs with 2l+3jets and 3l on-Z
  - fitted μ(ttW) and μ(ttZ) compatible at 1σ with SM
  - μ(ttH)=1.3 ± 0.5

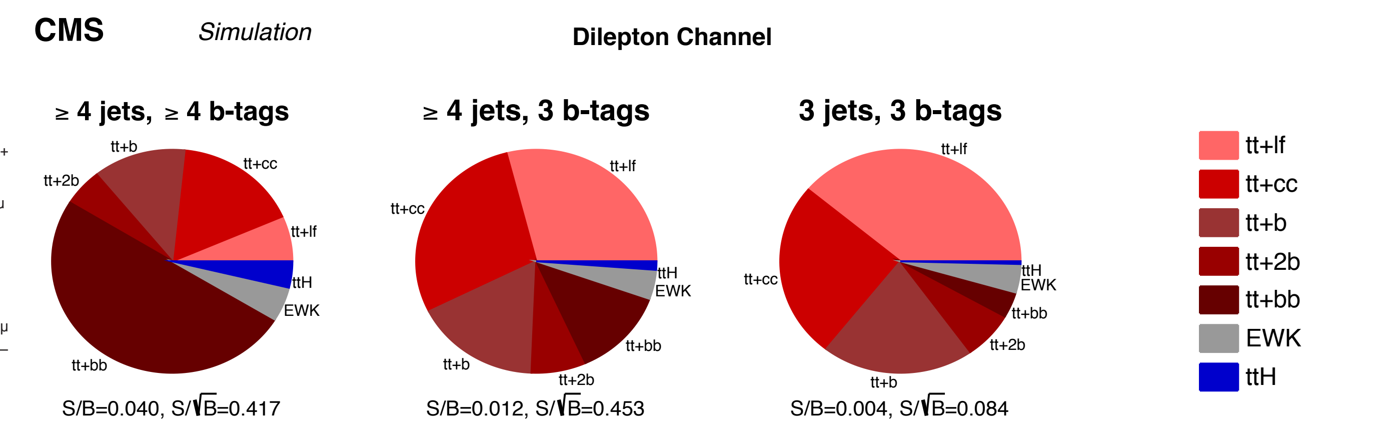


- **Two channels considered:**
  - lepton+jets: 1 lepton + ≥ 4 jets
  - dilepton: 2 OS leptons + ≥ 2 jets
- Categories based on **jet multiplicity + b-tagging**

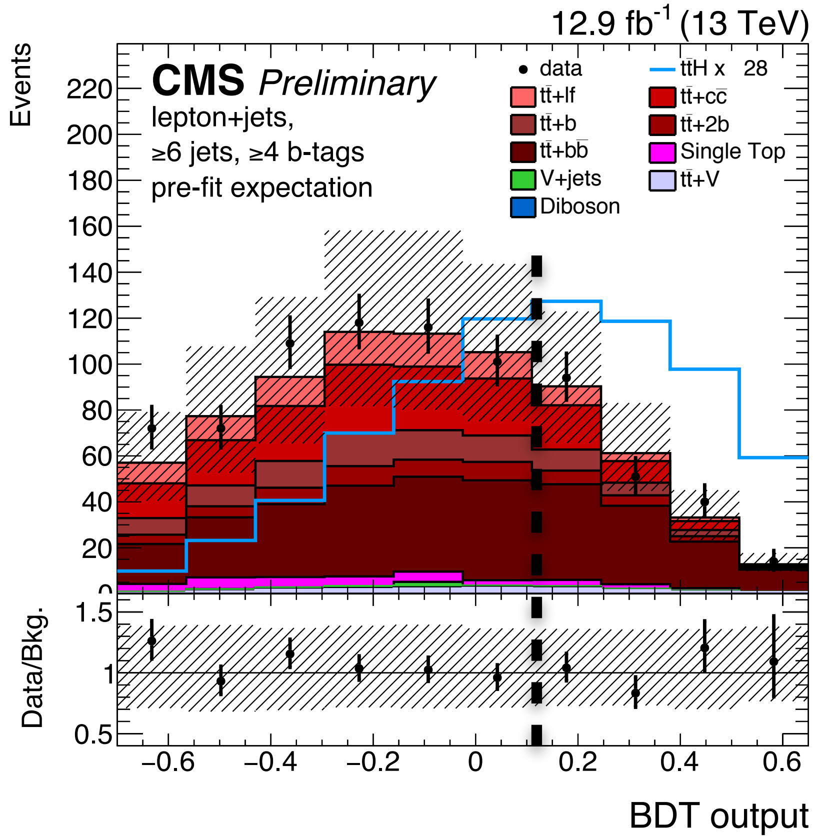


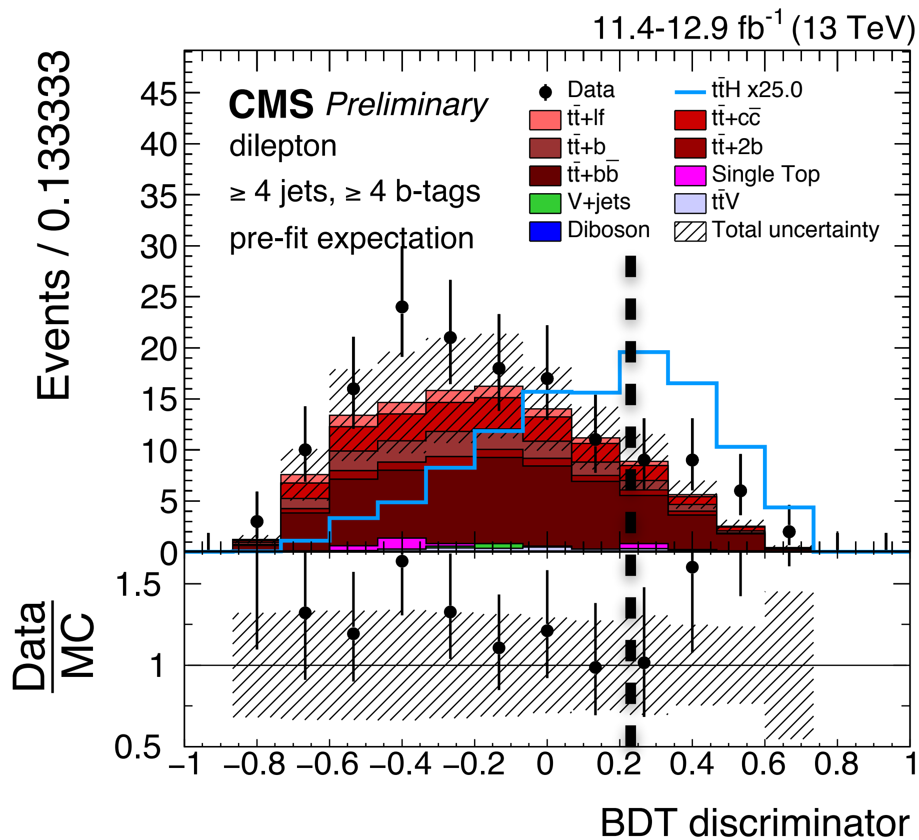






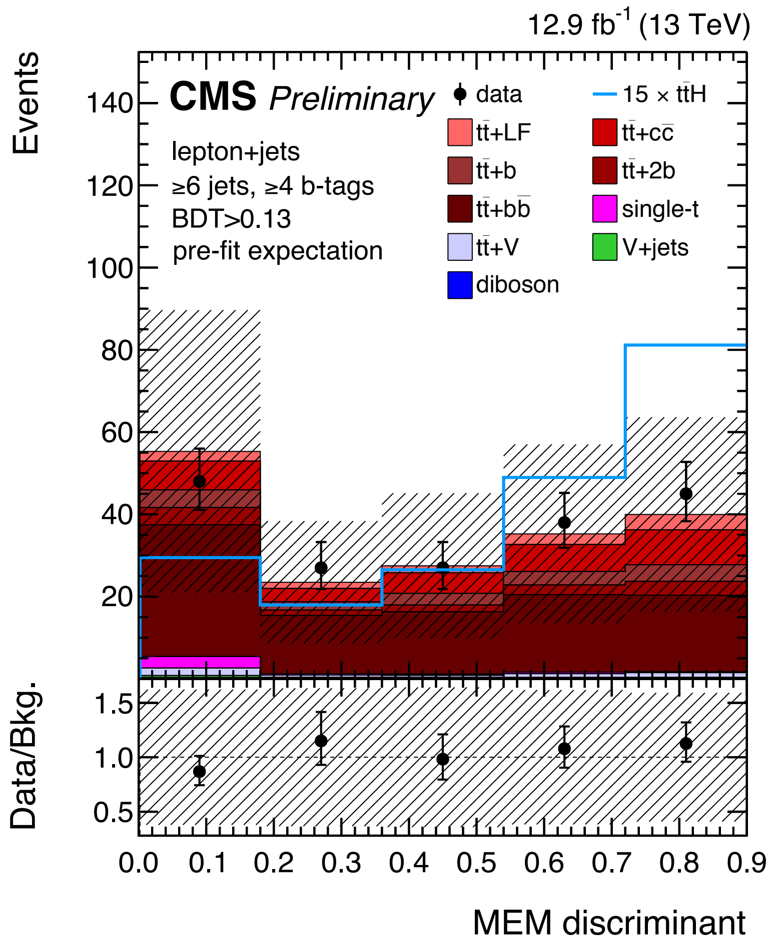
- **Further categorization based on BDT outputs** (except in dilepton 3 jet, 3 b-tag): low/high BDT regions



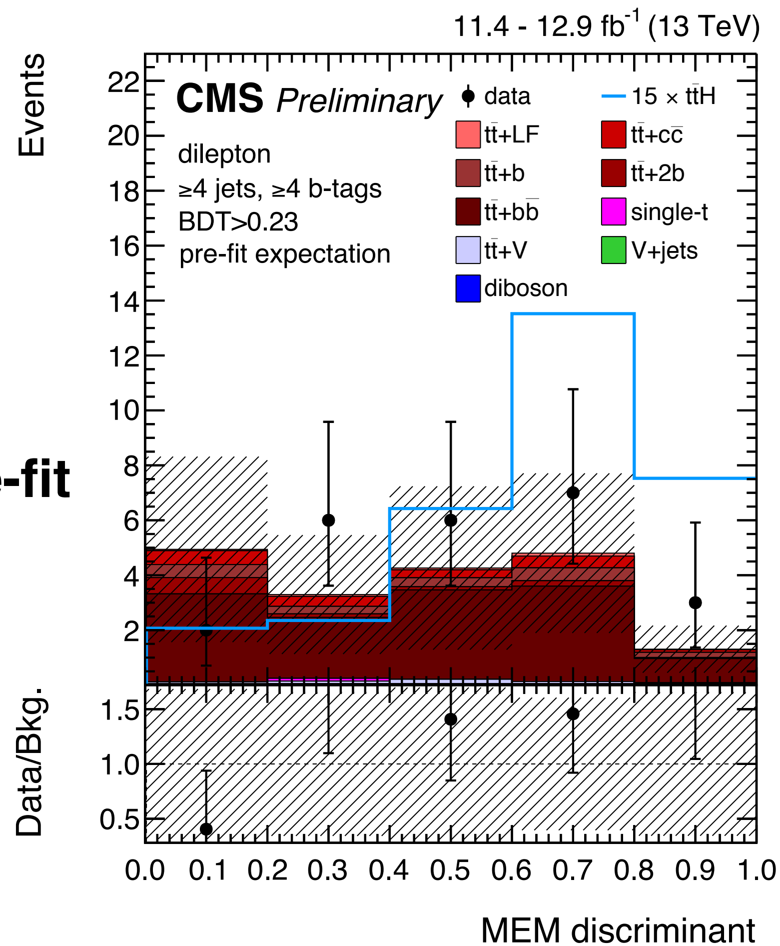


- **Fit MEM discriminant for signal extraction** in each BDT subcategories (except in dilepton 3 jet, 3 b-tag)

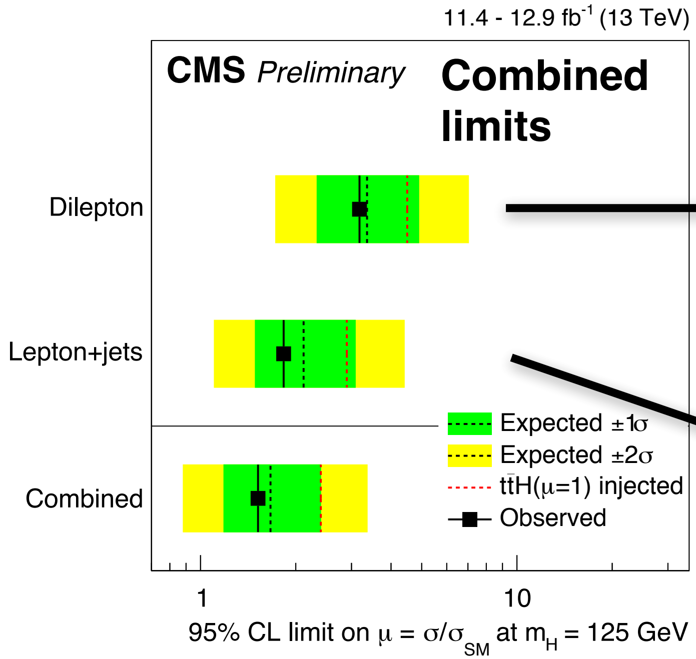
$$P_{s/b} = \frac{w(\vec{y}|t\bar{t}H)}{w(\vec{y}|t\bar{t}H) + k_{s/b} w(\vec{y}|t\bar{t}+b\bar{b})}$$

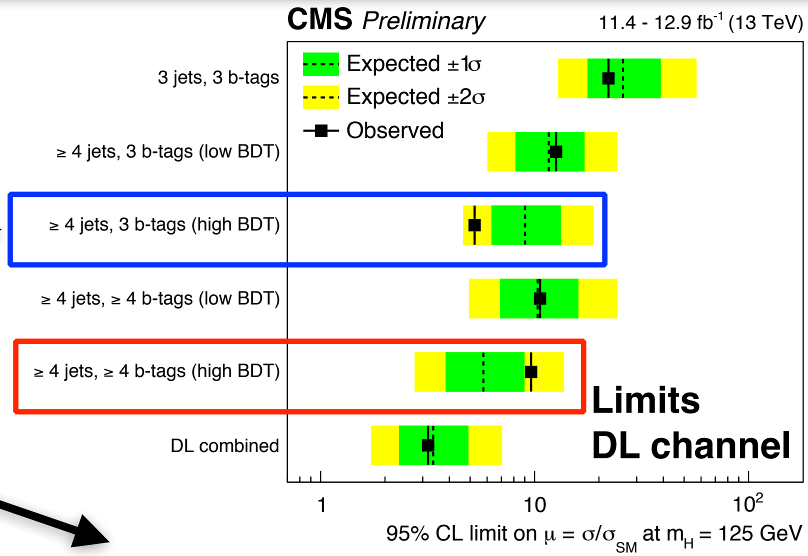


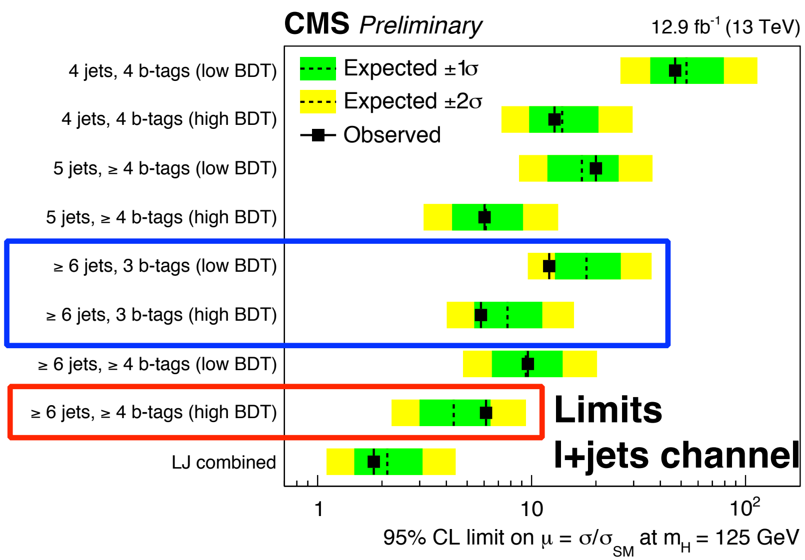
Pre-fit



ttH H→bb [CMS PAS HIG-16-038]



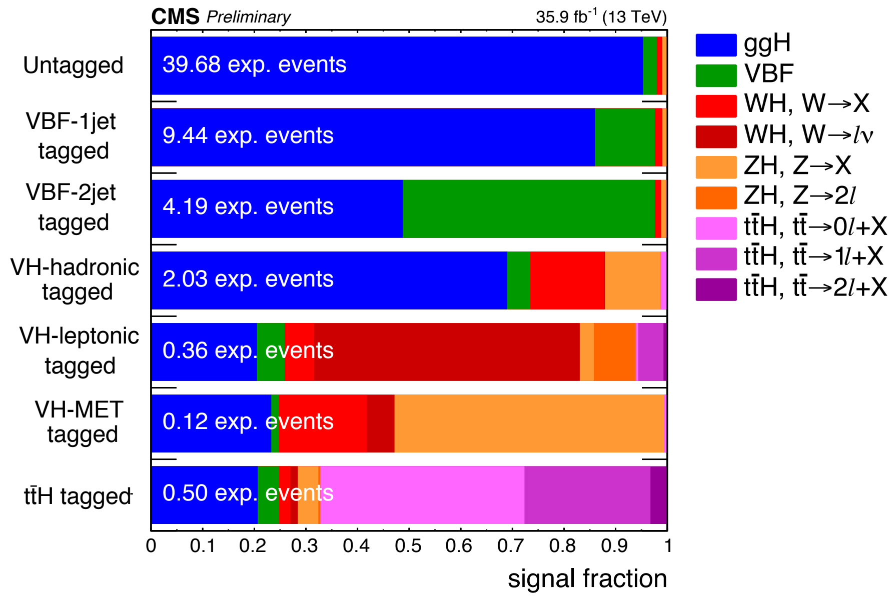


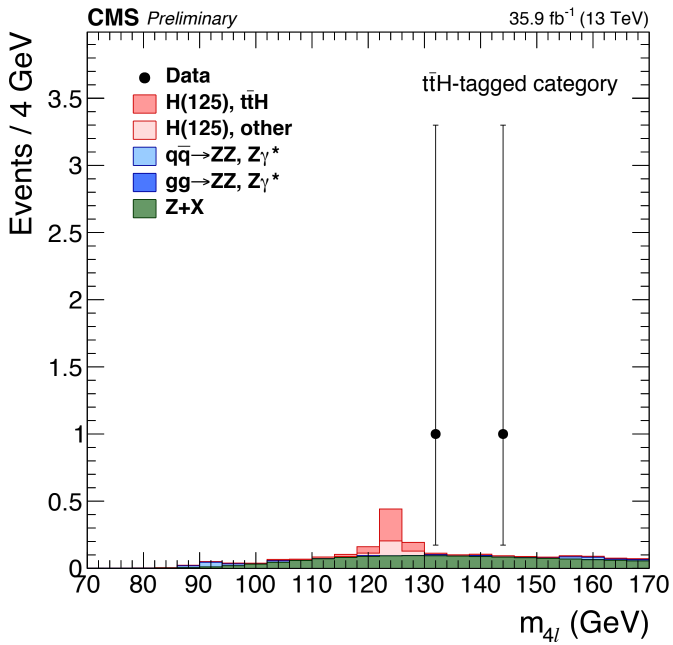


- **Upward fluctuations in most sensitive subcategories**
- **Downwards fluctuations in higher stat subcategories**
- **Best fit value** $\hat{\mu}_{obs}(ttH) = -0.19^{+0.80}_{-0.81}$

ttH H->4l CMS PAS HIG-16-041

- 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**
- **Main challenge: signal yield** (σ(ttH)xBR ~ 1fb for γγ, ~ 0.1fb for 4l)



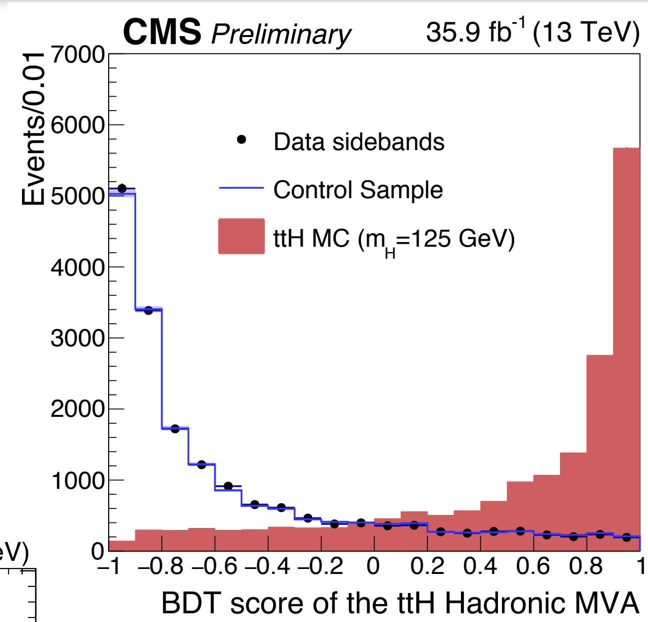


- Acceptance maximized by considering all tt decay modes:
  ttH category = 4l + ≥ 1l (not VH tagged) OR ≥ 4 jets, ≥ 1 b-jet
- **No event observed in 118<m(4l)<130 GeV range**

$$\hat{\mu}_{obs}(ttH) = 0.0^{+1.2}_{-0.0}$$

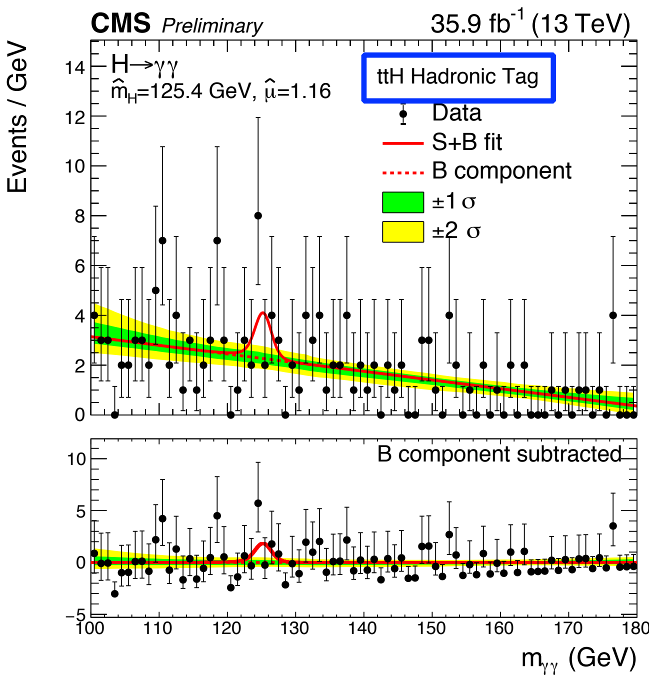
ttH H->γγ CMS PAS HIG-16-040

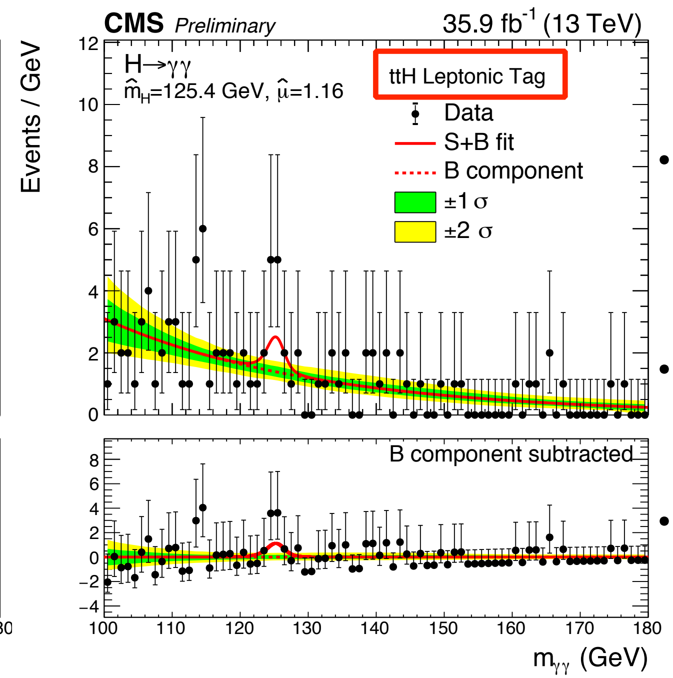
- **New for analysis on full 2016 dataset:**
  BDT trained to discriminate ttH hadr. / diphoton bkg



- **Two ttH sensitive channels**
  - **ttH hadronic**: $2\gamma + \geq 3$ jets (≥ 1 b-tag) + BDT > 0.75
  - **ttH leptonic**: $2\gamma + \geq 1$ lepton + ≥ 2 jets (≥ 1 b-tag)

- **Main background tt+genuine/fakes γ:** estimated from fit of $m(\gamma\gamma)$ distribution





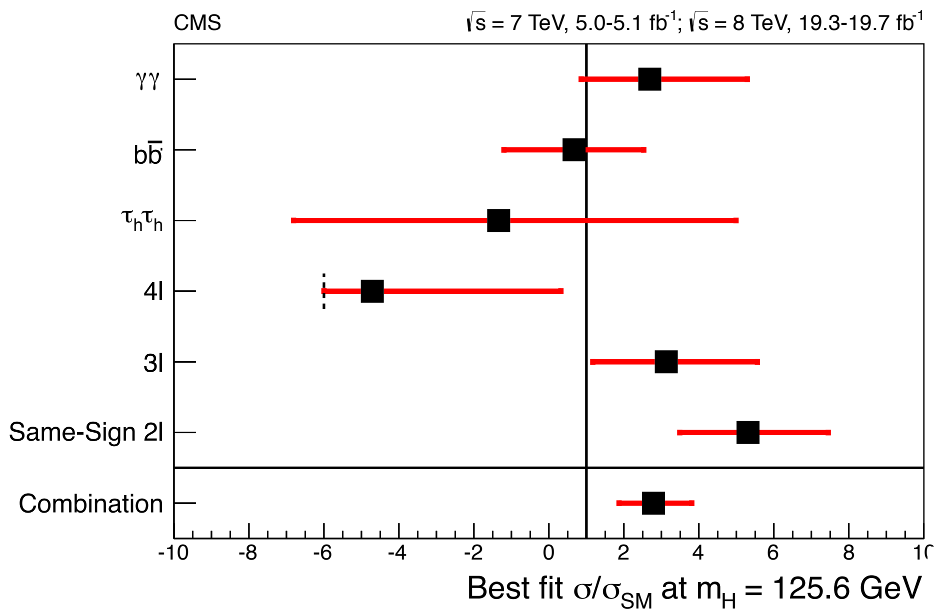
- **Diphoton mass $m(\gamma\gamma)$ used for signal extraction**

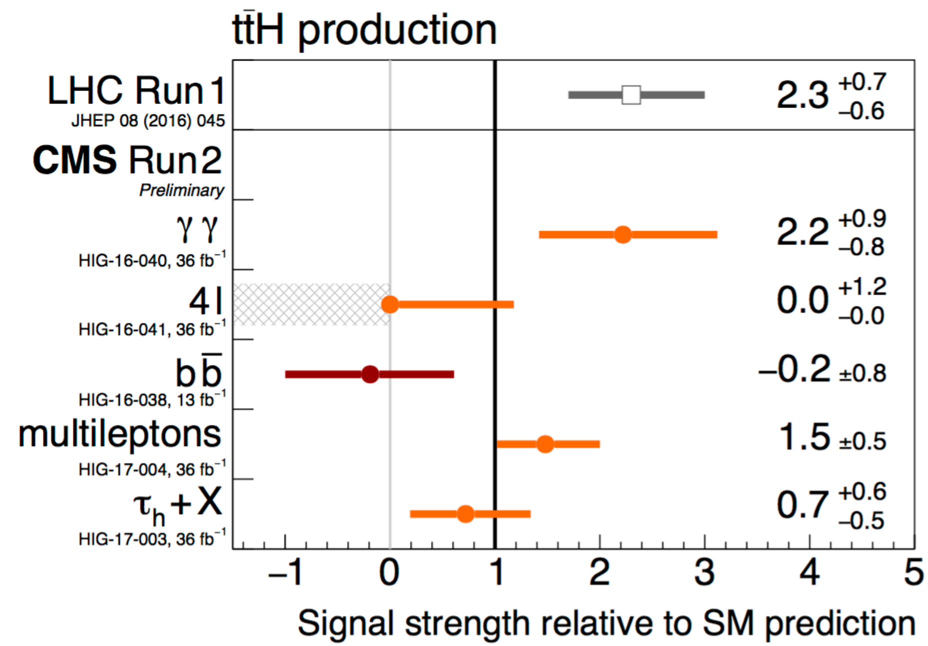
- $\hat{\mu}_{obs}(ttH) = 2.2^{+0.9}_{-0.8}$

- **3.3σ observed significance** (1.8σ expected)

Run 1 results on ttH

JHEP09(2014)087
arXiv:1408.1682 [hep-ex]





Run2

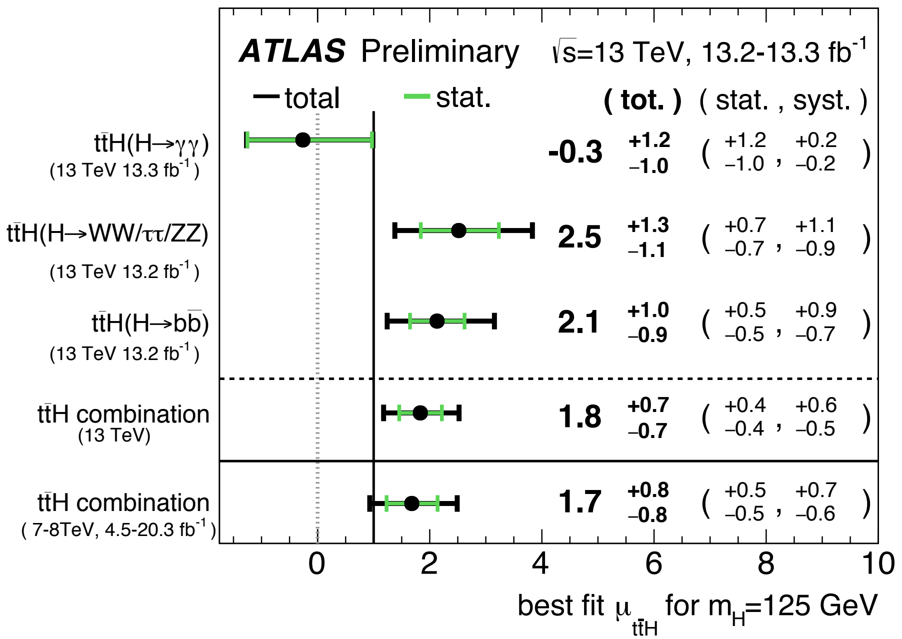
1l+2τ: μ=-1.2 +/- 1.5

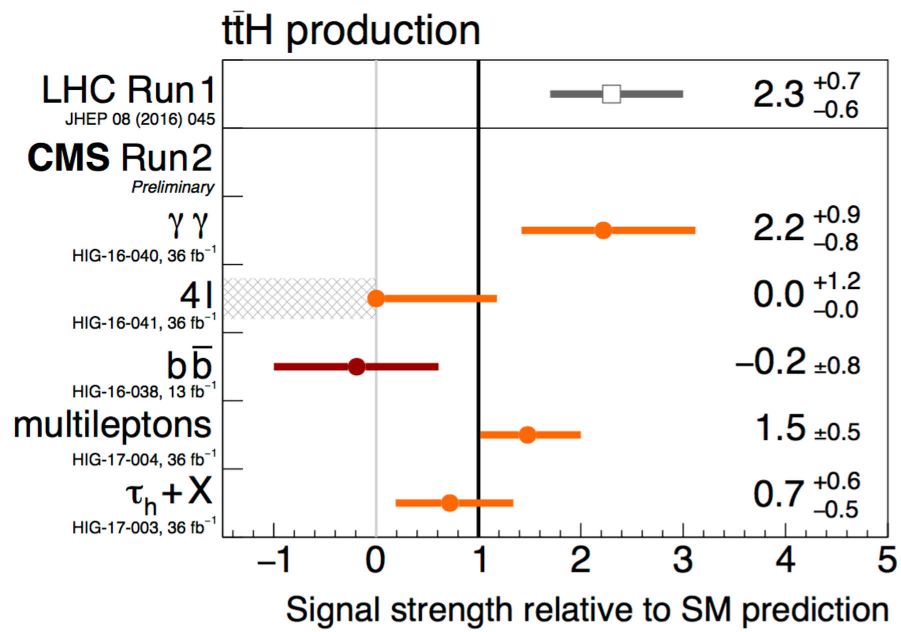
2lSS: μ=1.8 +/- 0.6

3l: μ=1.0 +0.8 / -0.7

4l: μ=0.9 +2.3 / -1.6

ATLAS results on ttH ATLAS-CONF-2016-068





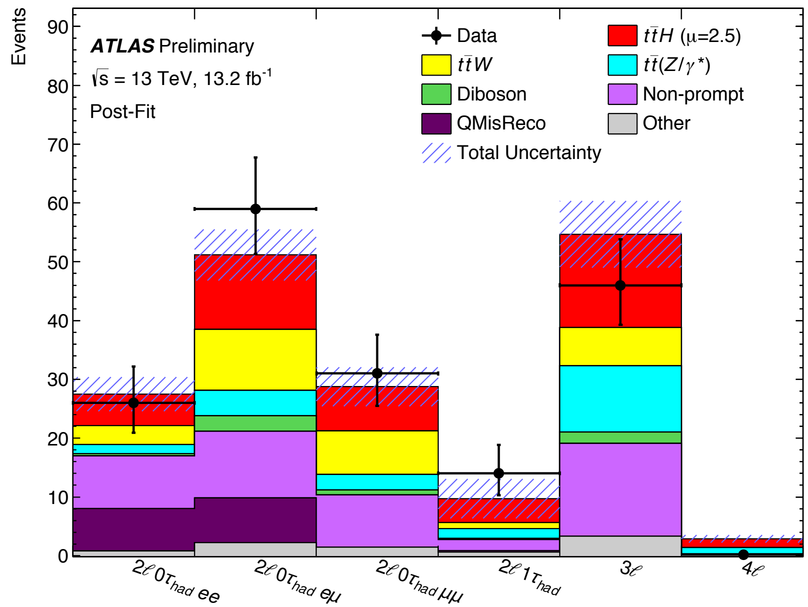
- **2.8σ observed significance**
  (1.8σ expected)

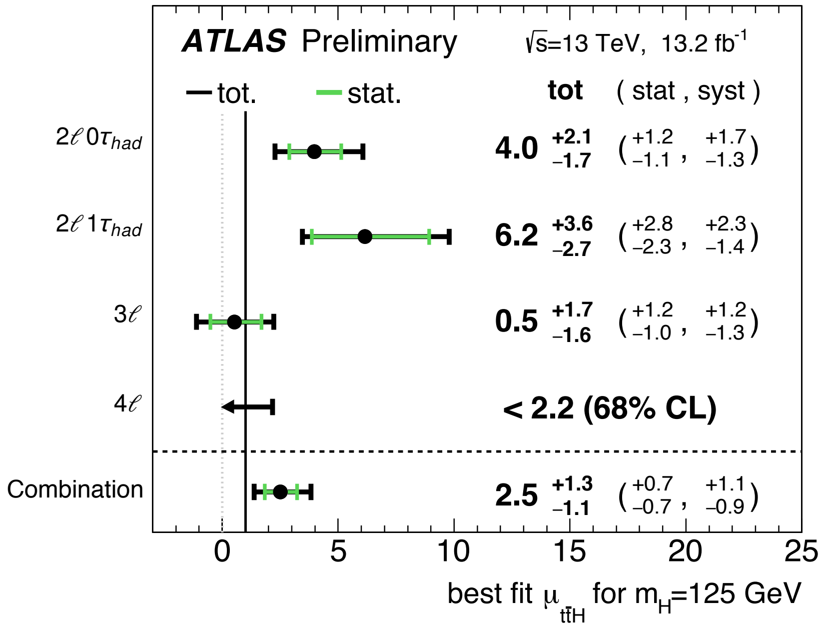
ATLAS results on ttH τ+X

ATLAS-CONF-2016-058

- **2lSS+1τ_h category included in ATLAS ttH multilepton result based on first 13.2 fb^{-1} of 2016 data:**
  Tighter jet selection n(jet)≥4 (n(jet)≥3 in CMS analysis)

- **Similar background estimation techniques:**
  Irreducible backgrounds from MC
  Reducible w/ non-prompt leptons and charge misreconstruction data-driven





- **Cut-and-count analysis**

- **Post-fit yields in 2lSS+1τ_h:**
  SM backgrounds 5.7 exp. events
  ttH 1.6 exp. events
  14 observed events

- **Consistent with 2lSS+0τ_h categories**