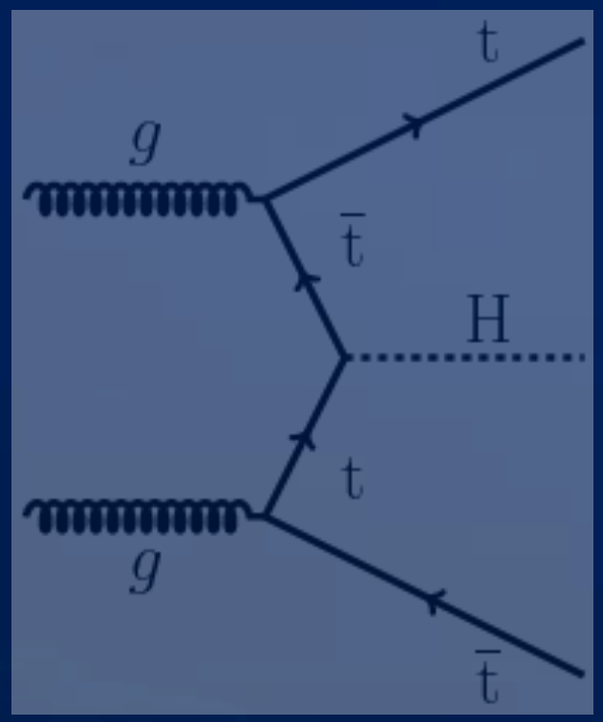
Latest CMS results on Higgs boson production in association with top quarks ($t\bar{t}H$)



Saranya Ghosh
(DPhP, IRFU, CEA Paris-Saclay),

For the CMS collaboration

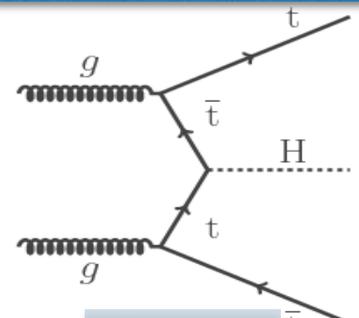
EPS-HEP 2017 : Venice, Italy : 7TH JULY 2017

INTRODUCTION

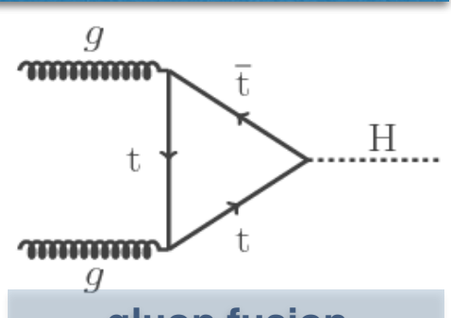
- Production of the Higgs boson in association with a pair of top quarks (ttH) probes the **top quark Yukawa coupling**
  - ➡ Direct handle on the **ttH vertex at tree level** : gluon fusion involves a loop
  - ➡ Cross-section of ttH is **~508 fb at 13 TeV**, roughly **4 times** the value at 8 TeV, increased potential for discovery
- Searches for **ttH at CMS** in Run 2 : based on decay channel of the Higgs
  - ➡ **Multi-leptonic (H⟶ZZ* , H⟶WW*, H⟶ττ) :** higher rate, multi-lepton final state with low background
  - ➡ **τ_H (H⟶ττ , also H⟶ZZ* , H⟶WW*):** One τ decaying hadronically, significant 'fake' background
  - ➡ **H⟶γγ :** small branching ratio, but clean final state (low systematic uncertainties). Excellent diphoton mass resolution
  - ➡ **H⟶bb̄ :** high branching ratio, but complex multi-jet final state

Separate Talk by Daniel Salerno

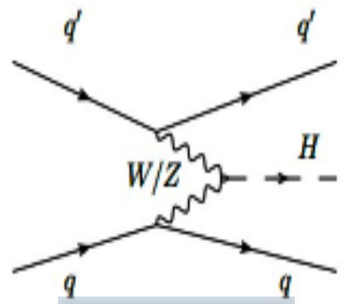
Major Higgs production modes at LHC (13 TeV)



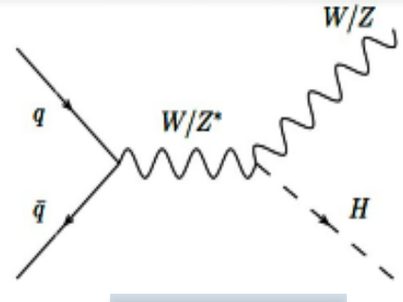
ttH (1%)



gluon fusion ggH(87%)

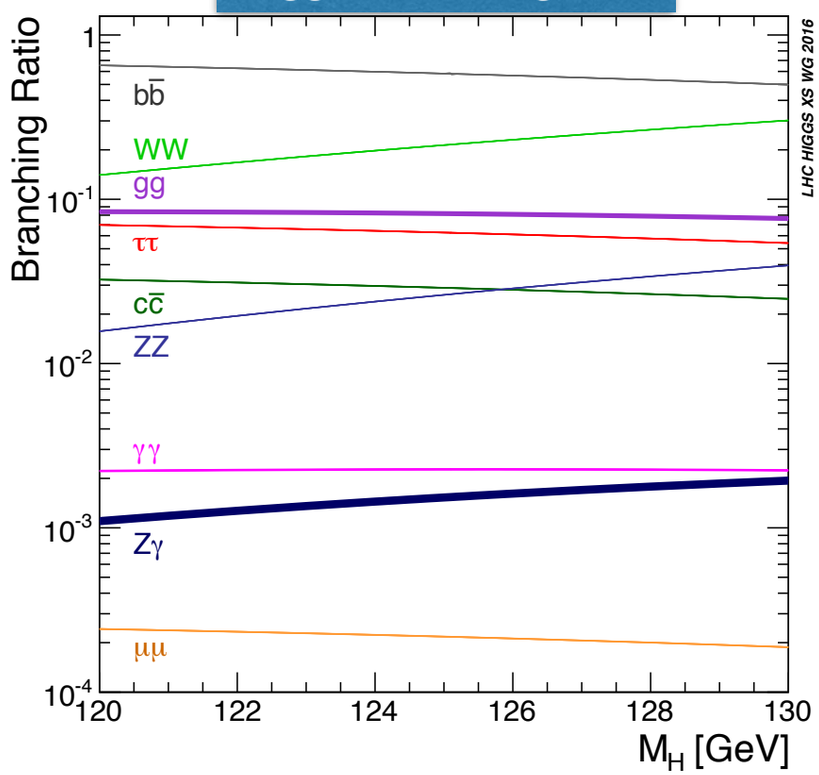


VBF (7%)



VH (4%)

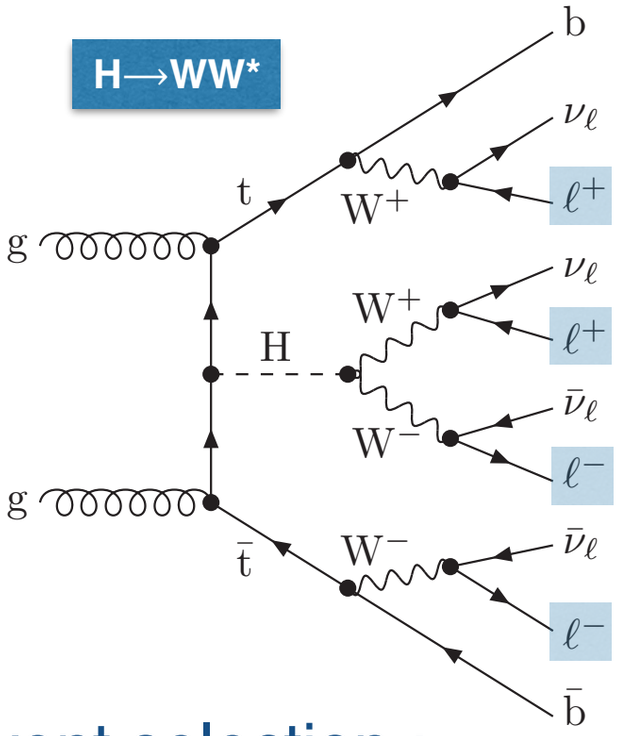
Higgs branching ratio

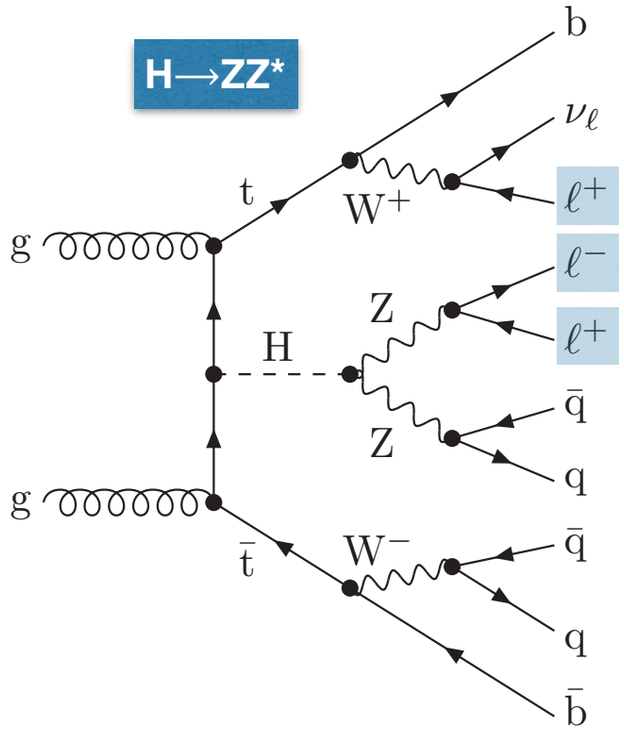


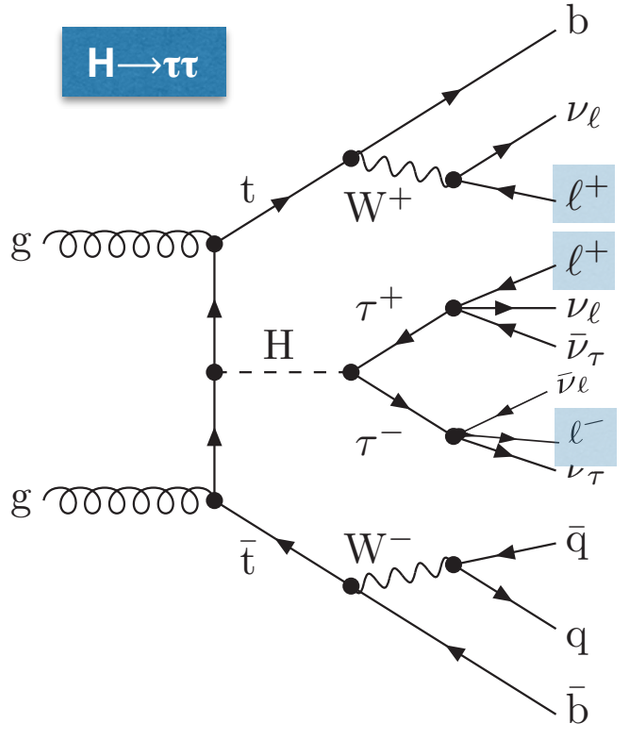
tt̄H Multi-Leptonic

CMS PAS HIG-17-004

- Multi-lepton final states from **H⟶WW*, H⟶ZZ* , H⟶ττ** (for τ leptonic decays)







- Event selection :

➡ **Same-sign dilepton** channel (**2LSS**):

- **2 leptons** (**e** or **μ**) with **same sign** : further categorisation : **ee, μμ, eμ**
- **≥ 4 jets**

➡ **Three lepton** channel (**3L**):

- **3 leptons**
- **≥ 2 jets**

➡ **Four lepton** channel (4**L**):

- **≥ 4 leptons**
- **≥ 2 jets**

➡ **B Tag jets**: at least **1 jet** passing **medium** WP or **2 jets** passing **loose** WP of B tag algorithm

➡ **Z veto** : based on $m_{\ell\ell}$, E_T^{miss}, H_T^{miss}

➡ τ_H **veto** : veto events containing reconstructed τ_H ; for orthogonality with analysis for tt̄H in the τ_H channel

➡ 2LSS and 3L categories have further sub-categories based on charge of leptons, B Tag criteria

ttH MULTI-LEPTONIC : SIGNAL EXTRACTION

Expected and observed yields after the selection in 2LSS sub-categories, 3L and 4L categories

	$\mu\mu$	eμ	ee	3L	4L
$t\bar{t}W$	51.0 ± 0.6 (stat.) ± 6.9 (syst.)	72.8 ± 0.7 (stat.) ± 10.2 (syst.)	20.5 ± 0.4 (stat.) ± 3.1 (syst.)	32.8 ± 1.0 (stat.) ± 4.9 (syst.)	
$t\bar{t}Z/\gamma^*$	17.7 ± 0.8 (stat.) ± 2.9 (syst.)	47.3 ± 1.6 (stat.) ± 9.0 (syst.)	17.5 ± 1.0 (stat.) ± 3.6 (syst.)	49.8 ± 3.9 (stat.) ± 11.1 (syst.)	2.15 ± 0.24 (stat.) ± 0.44 (syst.)
WZ	4.2 ± 0.6 (stat.) ± 4.1 (syst.)	7.0 ± 0.8 (stat.) ± 6.8 (syst.)	1.8 ± 0.4 (stat.) ± 1.7 (syst.)	9.1 ± 0.9 (stat.) ± 4.0 (syst.)	
Rare SM bkg.	4.2 ± 1.5 (stat.) ± 3.0 (syst.)	13.3 ± 1.9 (stat.) ± 9.3 (syst.)	4.8 ± 1.1 (stat.) ± 3.6 (syst.)	8.8 ± 4.3 (stat.) ± 5.9 (syst.)	0.27 ± 0.16 (stat.) ± 0.19 (syst.)
WWss	3.5 ± 0.6 (stat.) ± 2.5 (syst.)	4.1 ± 0.6 (stat.) ± 3.2 (syst.)	1.4 ± 0.3 (stat.) ± 1.2 (syst.)		
Conversions		7.8 ± 2.5 (stat.) ± 2.3 (syst.)	3.6 ± 3.5 (stat.) ± 1.7 (syst.)	5.3 ± 1.2 (stat.) ± 4.0 (syst.)	
Charge mis-meas.		16.4 ± 0.2 (stat.) ± 9.1 (syst.)	10.5 ± 0.2 (stat.) ± 5.9 (syst.)		
Non-prompt leptons	38.7 ± 1.6 (stat.) ± 20.5 (syst.)	61.8 ± 2.0 (stat.) ± 13.0 (syst.)	17.7 ± 1.1 (stat.) ± 5.4 (syst.)	30.8 ± 1.5 (stat.) ± 10.9 (syst.)	
All backgrounds	120.3 ± 2.5 (stat.) ± 11.7 (syst.)	231.2 ± 4.3 (stat.) ± 13.3 (syst.)	77.9 ± 4.0 (stat.) ± 9.0 (syst.)	137.3 ± 6.2 (stat.) ± 12.4 (syst.)	2.42 ± 0.28 (stat.) ± 0.56 (syst.)
$t\bar{t}H$ signal	20.1 ± 0.5 (stat.) ± 2.1 (syst.)	27.9 ± 0.5 (stat.) ± 3.0 (syst.)	8.0 ± 0.3 (stat.) ± 1.1 (syst.)	19.5 ± 1.0 (stat.) ± 3.0 (syst.)	1.00 ± 0.09 (stat.) ± 0.11 (syst.)
Data	150	268	89	148	3

- Background estimation:
  - ➡ **Signal like final states :** $t\bar{t}V$, Di-boson : estimated from MC simulation
  - ➡ **Others :** Non-prompt leptons (largely from $t\bar{t}$), charge mis-measured leptons : Data driven estimation
- Signal extraction :
  - ➡ Multivariate **BDT** discriminants trained in simulated events to separate the **signal** from **$t\bar{t}V$** backgrounds and also **non-prompt ($t\bar{t}$)** backgrounds (not for statistically limited 4L)
  - ➡ 2 separate BDTs trained using kinematical observables.
    - η of leptons, jet multiplicity, distance between lepton & jet, m_T, E_T^{miss}
    - **Against $t\bar{t}$:** 2LSS : jet BDT discriminator score that is used to identify jets from top decay
    - **Against $t\bar{t}V$:** leading, trailing lepton p_T ;

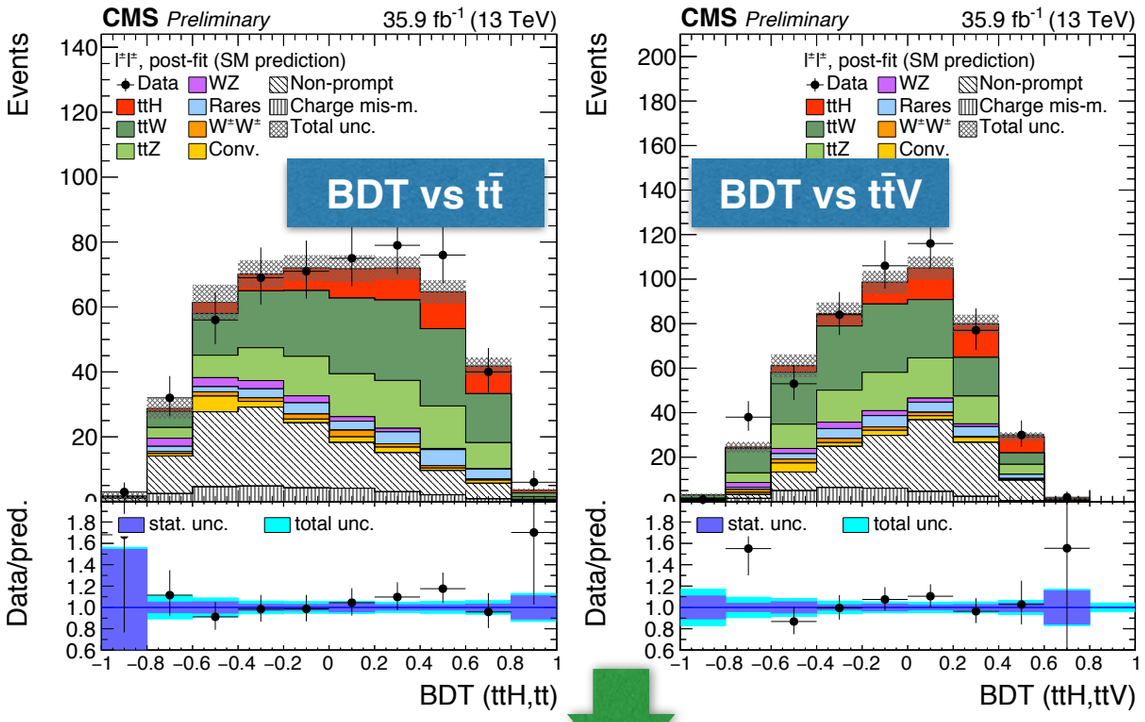
      for **2LSS** : jet BDT discriminator score that is used to identify jets from Higgs decay

      for **3L** category : MEM weight : $$w_{i,\alpha}(\Phi') = \frac{1}{\sigma_\alpha} \int d\Phi_\alpha \cdot \delta^4\left(p_1^\mu + p_2^\mu - \sum_{k>2} p_k^\mu\right) \cdot \frac{f(x_1,\mu_F) f(x_2,\mu_F)}{x_1 x_2 s} \cdot \left|\mathcal{M}_\alpha(p_k^\mu)\right|^2 \cdot W(\Phi'|\Phi_\alpha)$$

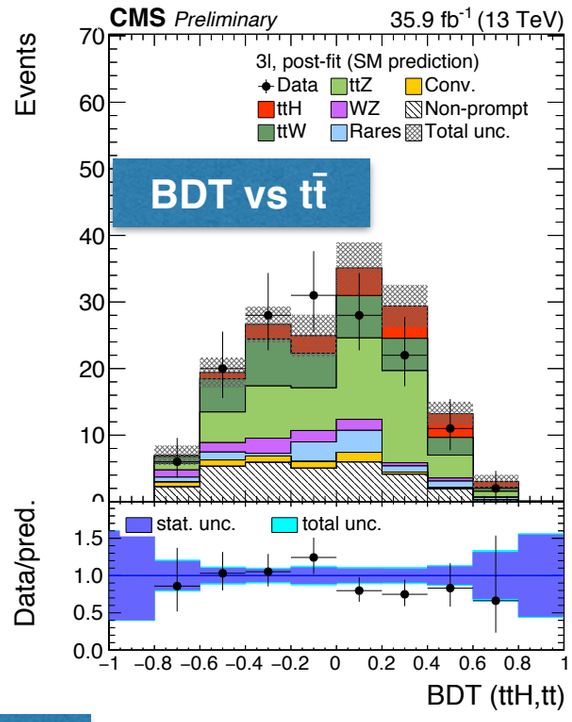
ttH Multi-Leptonic : Signal Extraction

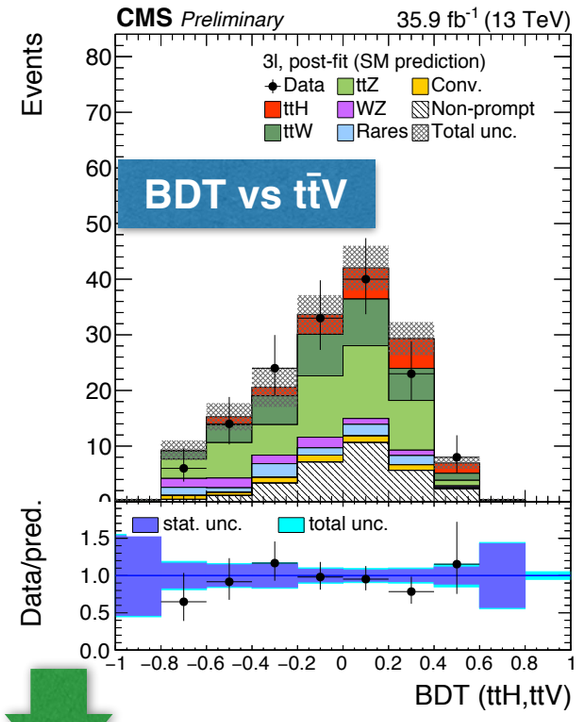
- Distribution of discriminating BDT outputs after fit to data, combined from different sub-categories, for 2LSS and 3L categories

2LSS

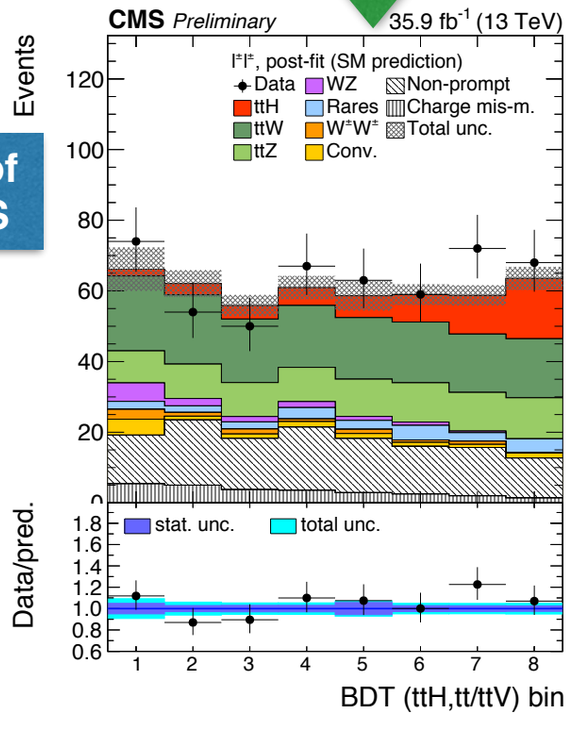


3L

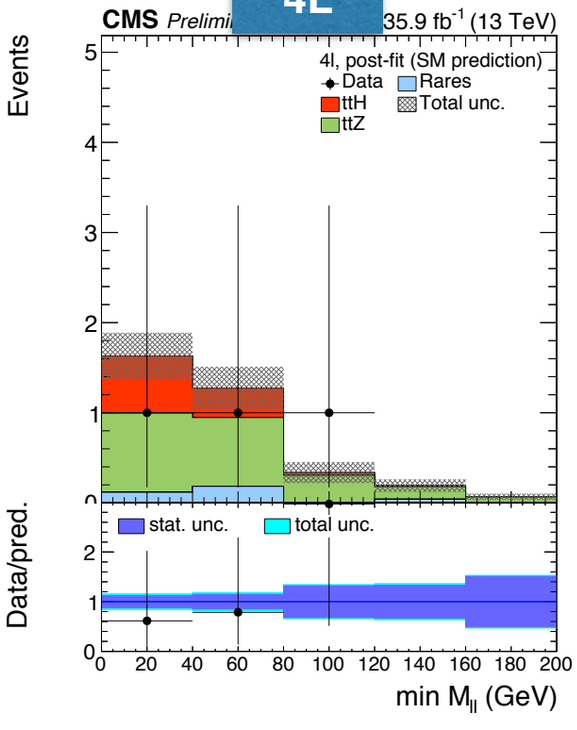


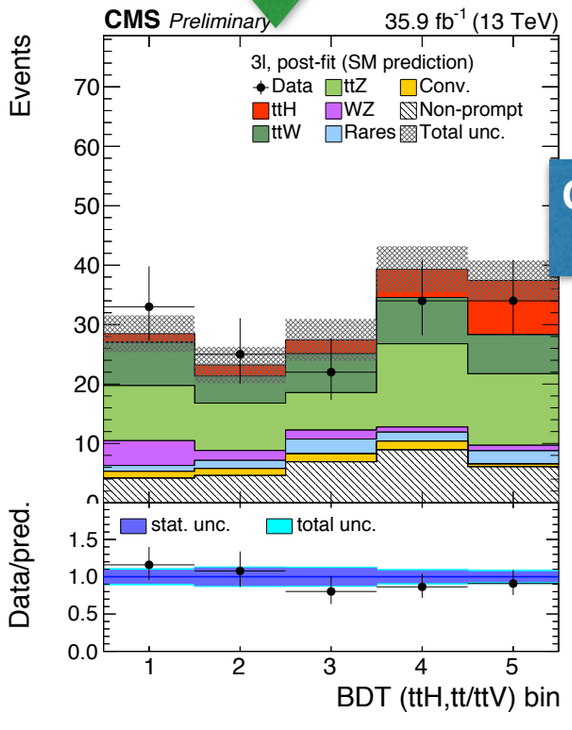


Combination of BDTs for 2LSS



4L

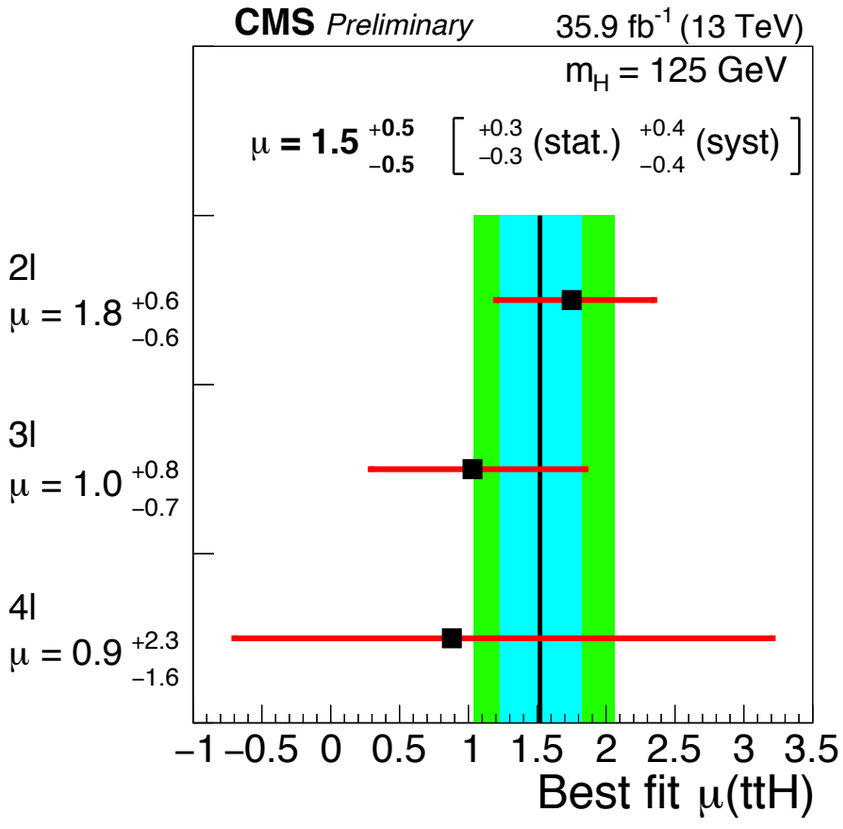




Combination of BDTs for 3L

ttH MULTI-LEPTONIC : RESULTS

- Results based on **35.9 fb^{-1}** of data collected during **2016** and **combination** with **2.3 fb^{-1}** collected in **2015**



Signal strength measurement (2016 + 2015 dataset)

Category	Observed μ fit $\pm1\sigma$	Expected μ fit $\pm1\sigma$
Same-sign di-lepton	$1.7\,(-0.5)\,(+0.6)$	$1.0\,(-0.5)\,(+0.5)$
Three lepton	$1.0\,(-0.7)\,(+0.8)$	$1.0\,(-0.7)\,(+0.8)$
Four lepton	$0.9\,(-1.6)\,(+2.3)$	$1.0\,(-1.6)\,(+2.4)$
Combined (2016 data)	$1.5\,(-0.5)\,(+0.5)$	$1.0\,(-0.4)\,(+0.5)$
Combined (2015 data) [42]	$0.6\,(-1.1)\,(+1.4)$	$1.0\,(-1.1)\,(+1.3)$
Combined (2015+2016 data)	$1.5\,(-0.5)\,(+0.5)$	$1.0\,(-0.4)\,(+0.5)$

Limits on signal strength (2016 dataset)

Category	Observed limit	Expected limit $\pm1\sigma$
Same-sign di-lepton	2.8	$0.9\,(-0.3)\,(+0.4)$
Three lepton	2.5	$1.4\,(-0.4)\,(+0.7)$
Four lepton	5.9	$4.9\,(-1.7)\,(+3.1)$
Combined	2.5	$0.8\,(-0.2)\,(+0.3)$

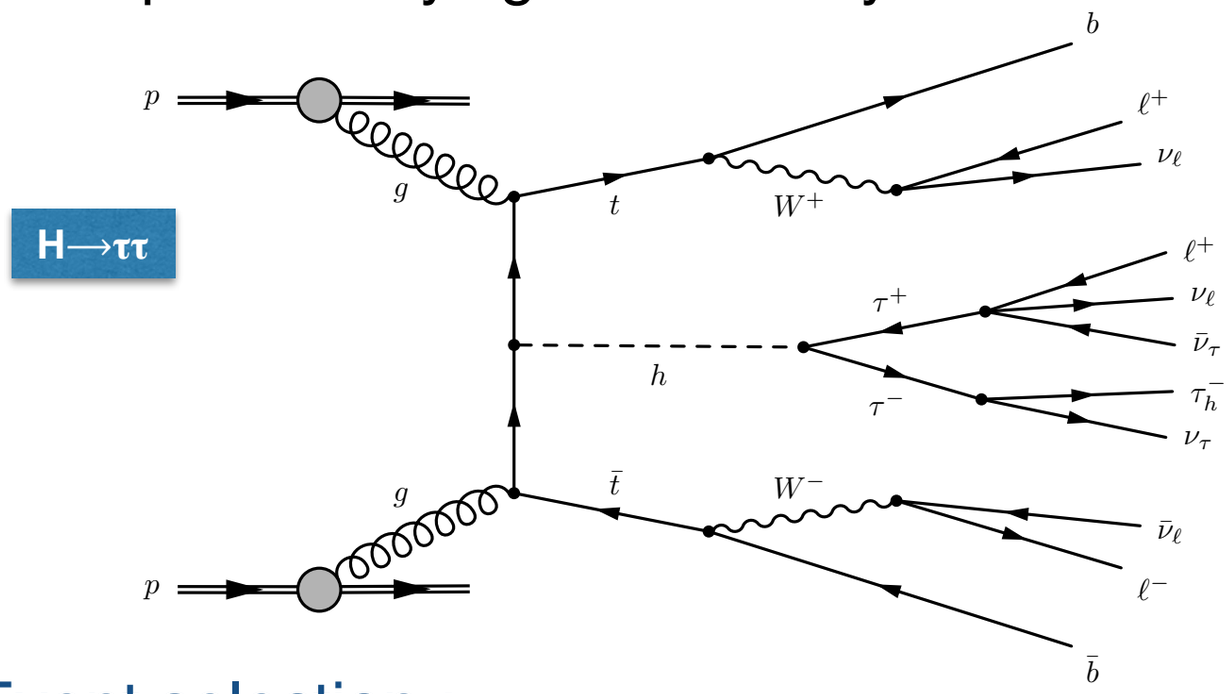
- Main sources of systematic uncertainties:
  - ➡ Lepton selection efficiency, modelling of non-prompt backgrounds, charge mis-identification, diboson background modelling, theoretical uncertainties

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

ttH TAU-HADRONIC (τ_H)

CMS PAS HIG-17-003

- Tau-lepton decaying hadronically in final state : **H⟶ττ** (also H⟶WW*, H⟶ZZ*)



τ_H identification: Dedicated MVA based discriminant (isolation, impact parameter, τ lifetime) with "hadrons plus strips" (HPS) algorithm to identify

$$\tau^\pm \to h^\pm \nu_\tau$$
$$\tau^\pm \to h^\pm \pi^0 \nu_\tau$$
$$\tau^\pm \to h^\pm \pi^0 \pi^0 \nu_\tau$$
$$\tau^\pm \to h^\pm h^\mp h^\pm \nu_\tau$$

- Event selection :

➡ **1ℓ + 2 τ_H** :

- **1 lepton (e or μ)**
- **2 τ_H** opposite charge, 'tight' WP of τ_H ID, p_T>30,20 GeV
- **≥3 hadronic jets**

➡ **3ℓ + 1 τ_H** :

- **≥3 leptons** of
- **1 τ_H** sum of charge with ℓ = 0, 'medium' WP of τ_H ID, p_T>20 GeV
- **≥2 hadronic jets**

➡ **2ℓss + 1 τ_H** :

- **2 leptons** of **same sign**
- **1 τ_H** opposite charge wrt ℓ, 'medium' WP of τ_H ID, p_T>20 GeV
- **≥3 hadronic jets**
- **sub-categories :**
  - **no-missing-jet** : full reconstruction of $t\bar{t}H \to bW\bar{b}W\tau\tau \to bjj\bar{b}\ell\nu\ell\bar{\nu}_\ell\nu_\tau\tau_h\bar{\nu}_\tau$
  - **missing-jet** : one jet missing : outside acceptance overlapping

➡ **B Tag jets**: at least **1 jet** passing **medium** WP or **2 jets** passing **loose** WP

➡ **Z veto** : based on $m_{\ell\ell}$, E_T^{miss}, H_T^{miss}

$t\bar{t}H$ TAU-HADRONIC (τ_H) : SIGNAL EXTRACTION

Expected and observed yields after the selection in the different categories

Process	$1\ell + 2\tau_h$	$3\ell + 1\tau_h$
$t\bar{t}H$, $H \to \tau\tau$	2.84 ± 1.35	1.01 ± 0.65
$t\bar{t}H$, $H \to WW$	0.07 ± 0.04	0.63 ± 0.29
$t\bar{t}H$, $H \to ZZ$	0.02 ± 0.01	0.09 ± 0.04
$t\bar{t}Z$	4.07 ± 0.56	3.78 ± 0.62
$t\bar{t}W$	0.21 ± 0.05	0.24 ± 0.05
Electroweak	1.10 ± 1.05	0.32 ± 0.05
Fake	20.98 ± 3.87	1.07 ± 0.34
Other	0.54 ± 0.23	0.24 ± 0.08
Total expected background	26.91 ± 3.84	5.65 ± 0.85
SM expectation	29.85 ± 4.07	7.38 ± 1.10
Observed data	24	7

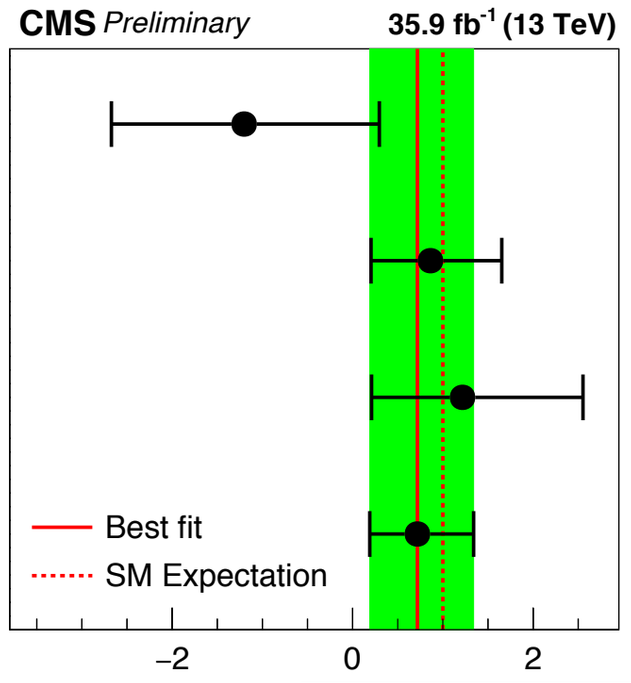
Process	$2\ell ss + 1\tau_h$ "no-missing-jet"	$2\ell ss + 1\tau_h$ "missing-jet"
$t\bar{t}H$, $H \to \tau\tau$	1.38 ± 0.89	2.86 ± 1.68
$t\bar{t}H$, $H \to WW$	1.03 ± 0.47	2.09 ± 1.01
$t\bar{t}H$, $H \to ZZ$	0.06 ± 0.03	0.06 ± 0.04
$t\bar{t}Z$	3.07 ± 0.46	8.33 ± 1.08
$t\bar{t}W$	1.10 ± 0.15	7.18 ± 0.80
Electroweak	0.21 ± 0.19	3.73 ± 3.39
Fake	1.66 ± 0.52	7.80 ± 2.51
Charge flip	0.05 ± 0.01	0.39 ± 0.10
Other	0.50 ± 0.20	2.44 ± 1.01
Total expected background	6.59 ± 0.88	29.87 ± 4.75
SM expectation	9.06 ± 1.33	34.88 ± 5.05
Observed data	8	41

- Background estimation:
  - ➡ **Irreducible :** $t\bar{t}V$, Di-boson+jets (estimated from MC, validated in data)
  - ➡ **Reducible :** Non-prompt leptons (largely from $t\bar{t}$), charge mis-measured leptons : Data driven estimation. Fake τ_H : data driven estimate in $1\ell+2\tau_H$ category, simulation based estimate in other categories.
- Signal extraction :
  - ➡ Signal extraction through maximum likelihood fit to the distribution of discriminating observable
  - ➡ Multivariate **BDT** discriminants to separate the **signal** from **$t\bar{t}$, $t\bar{t}V$** backgrounds, separately in each category
  - ➡ Additionally **MEM** based discriminant used in **$2\ell ss+1\tau_H$** category to discriminate signal from **$t\bar{t}V$** backgrounds

$t\bar{t}H$ Tau-Hadronic (τ_H) : Results

- Results based on **35.9 fb^{-1}** of data collected during **2016**
- Main sources of systematic uncertainties:
  - ➡ Affects yield : Charge mis-measurement, lepton & τ_H efficiency, theory uncertainty
  - ➡ Affects shape : Jet & τ_H energy scale, b-tag efficiency, fake background estimation, modelling of signal and background in simulation
- Signal significance **observed** (expected) : **1.4σ** (1.8σ)

Signal strength and limits



1l+2τ_h
$\mu = -1.20^{+1.50}_{-1.47}$

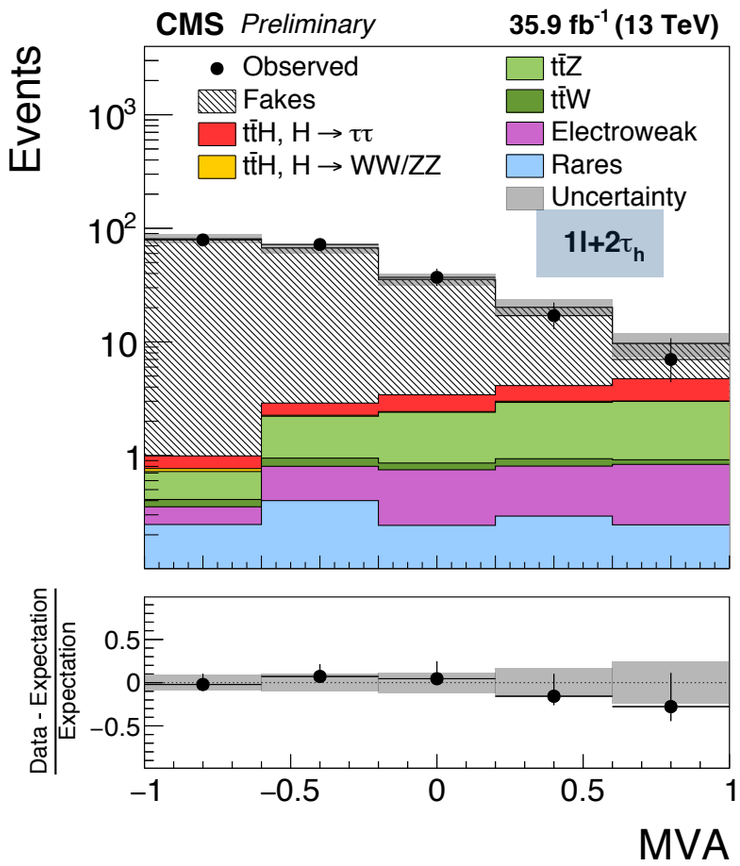
2lss+1τ_h
$\mu = 0.86^{+0.79}_{-0.66}$

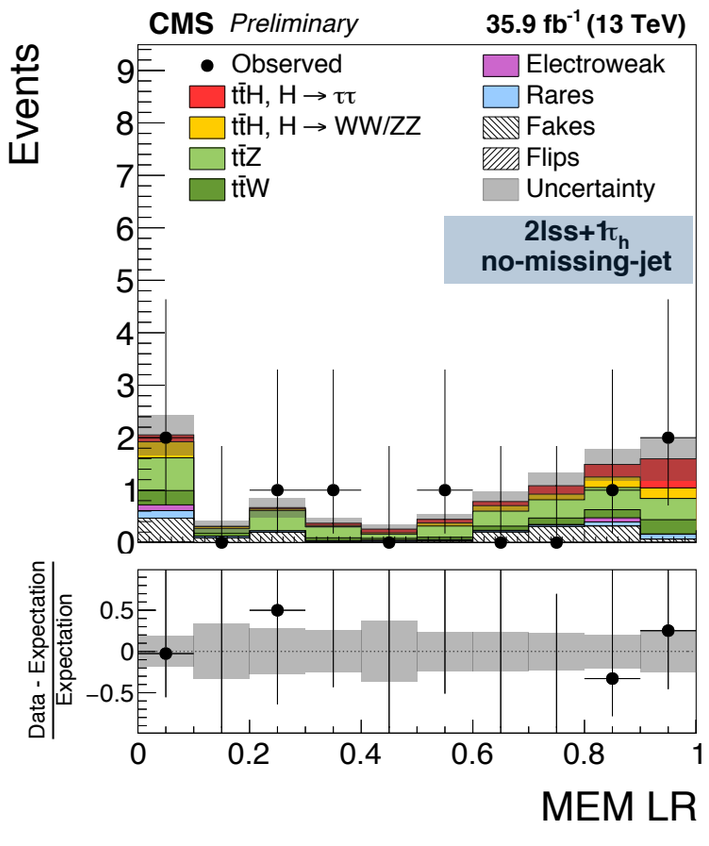
3l+1τ_h
$\mu = 1.22^{+1.33}_{-1.01}$

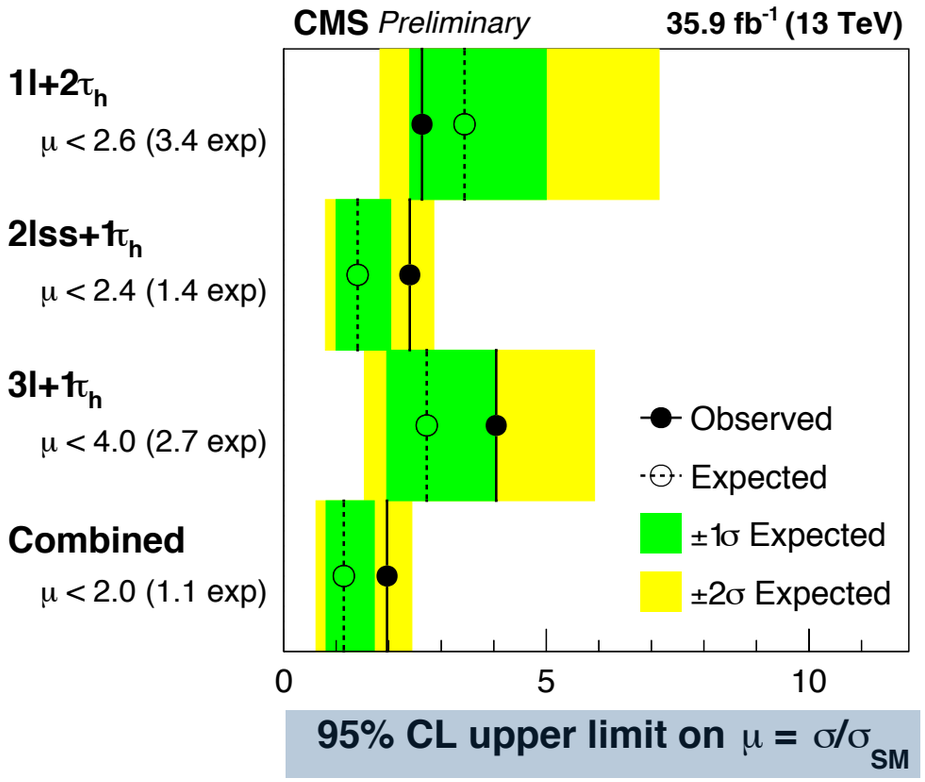
Combined
$\mu = 0.72^{+0.62}_{-0.53}$

Best fit $\mu = \sigma/\sigma_{SM}$

Signal extraction fits







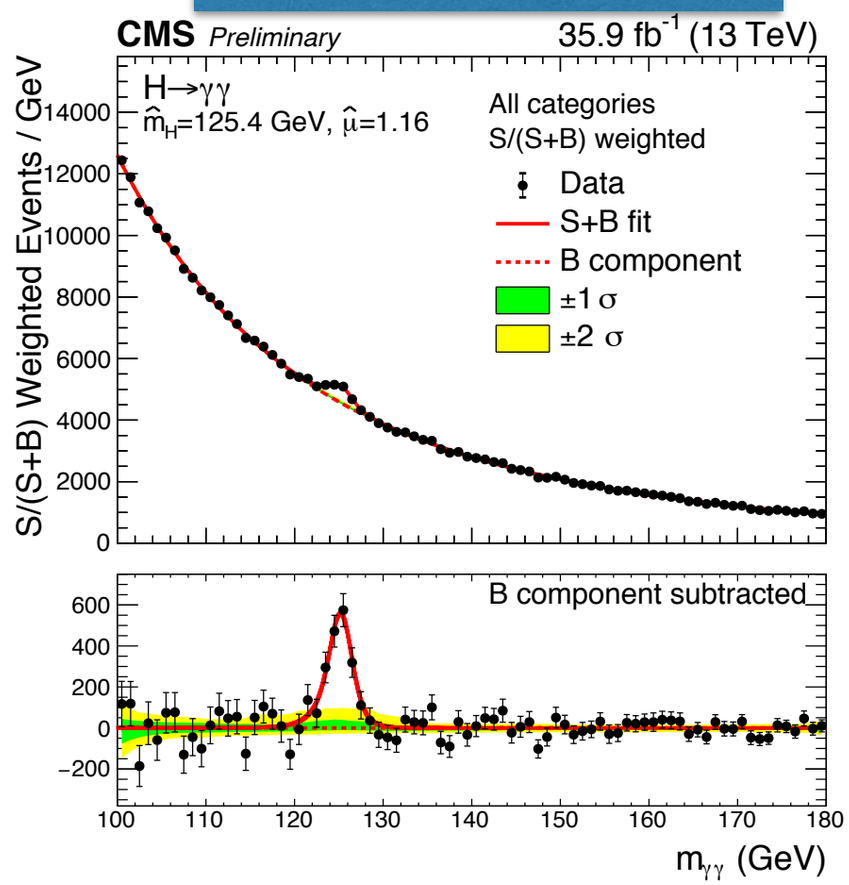
95% CL upper limit on $\mu = \sigma/\sigma_{SM}$

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

CMS PAS HIG-16-040

- A part of the general $H\rightarrow\gamma\gamma$ analysis **Talk by Michael Planner**
  - ➡ Events with **two** high p_T isolated **photons** selected
  - ➡ Narrow **peak** around m_H on top of the falling **$m_{\gamma\gamma}$ distribution**
  - ➡ Different production modes (**$t\bar{t}H$**, VBH, VH) identified based on **additional final state objects**
  - ➡ Signal, background extraction from **fit to $m_{\gamma\gamma}$ distribution**

$m_{\gamma\gamma}$ distribution (weighted)



- Overview of the $H\rightarrow\gamma\gamma$ analysis :

Photon Reconstruction → Vertex Identification → Photon, Diphoton Selection → **Event Categorisation** → Signal, Bkg extraction; Statistical analysis

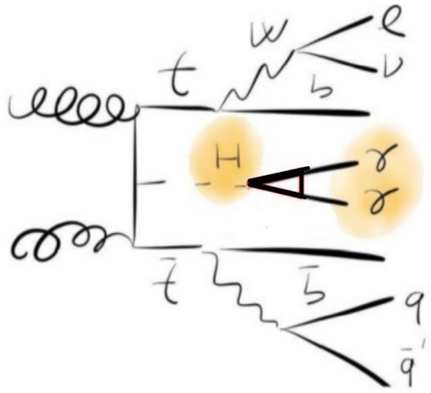
- ➡ **Event categorisation for different production modes of Higgs** based on
  - **additional final state objects** to tag **production modes**
  - mass resolution and kinematics for the ‘Untagged’ categories

$t\bar{t}H$ WITH $H\to\gamma\gamma$: $t\bar{t}H$ CATEGORIES

- 2 categories corresponding to $t\bar{t}H$ based on the decay of the top quarks

- **TTH Leptonic Tag:**

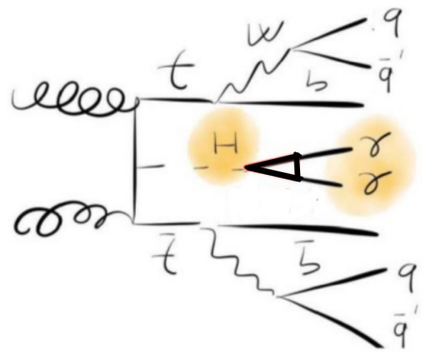
$t\bar{t} \to bl\nu_l\bar{b}q\bar{q}'$ or $t\bar{t} \to bl\nu_l\bar{b}l'\nu_{l'}$



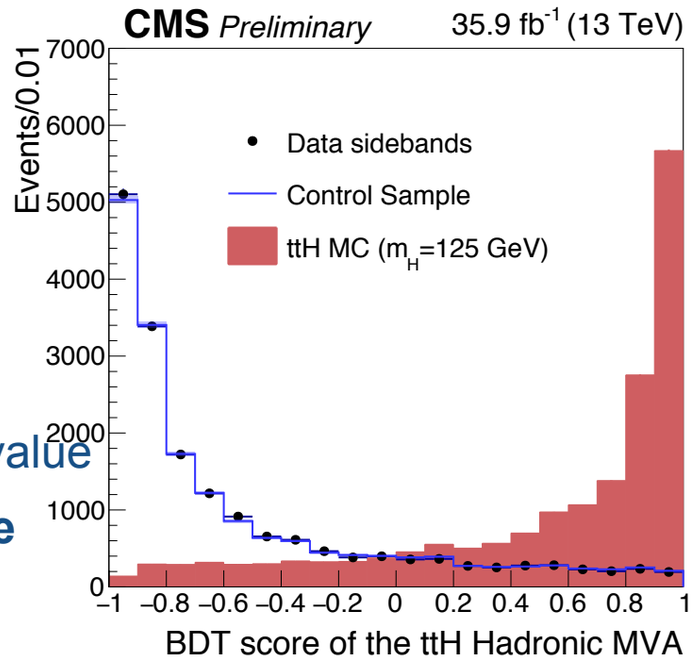
➡ Cut based approach:

- **≥ 1 lepton** (muon or electron)
- **≥ 2 jets**
- **≥ 1 B-tagged jet**
- Diphoton BDT cut

- **TTH Hadronic Tag**: $t\bar{t} \to bq\bar{q}'\bar{b}q\bar{q}'$



- Loose selection :
  - ➡ ≥ 3 jets
  - ➡ ≥ 1 B-jet (loose WP)
  - ➡ 0 leptons
- **MVA** with inputs:
  - ➡ number of jets
  - ➡ lead jet p_T
  - ➡ Max B-tag discriminant value
  - ➡ Second max B-tag value
- Diphoton BDT cut



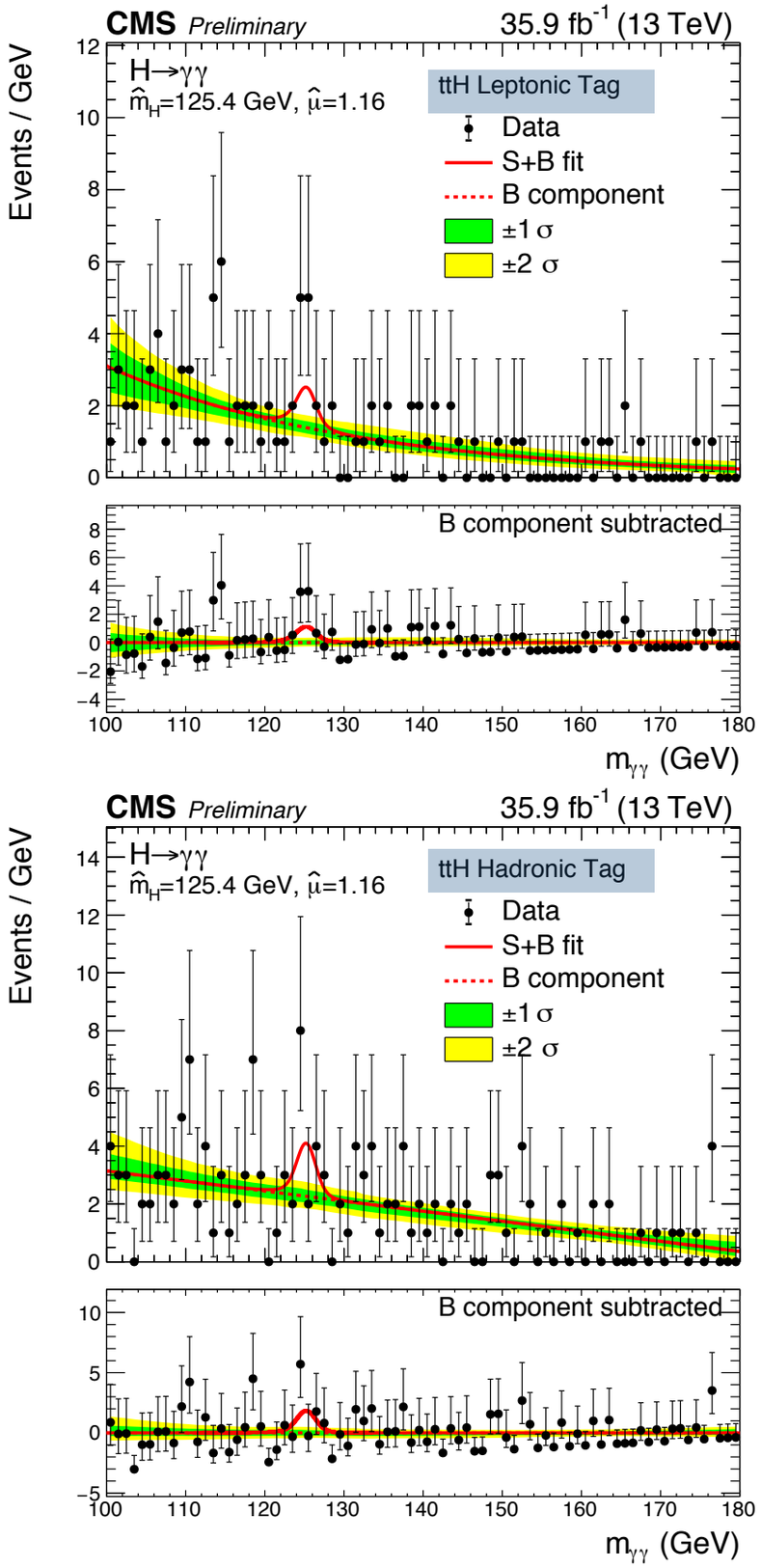
Expected signal for $t\bar{t}H$ categories : very pure in $t\bar{t}H$ contribution

Event Categories	SM 125 GeV Higgs boson expected signal													Bkg (GeV^{-1})
	Total	ggH	VBF	ttH	bbH	tHq	tHW	WH lep	ZH lep	WH had	ZH had	σ_{eff}	σ_{HM}	
ttH Hadronic	5.85	10.99 %	0.70 %	77.54 %	2.02 %	4.13 %	2.02 %	0.09 %	0.05 %	0.63 %	1.82 %	1.48	1.30	2.40
ttH Leptonic	3.81	1.90 %	0.05 %	87.48 %	0.08 %	4.73 %	3.04 %	1.53 %	1.15 %	0.02 %	0.02 %	1.60	1.35	1.50

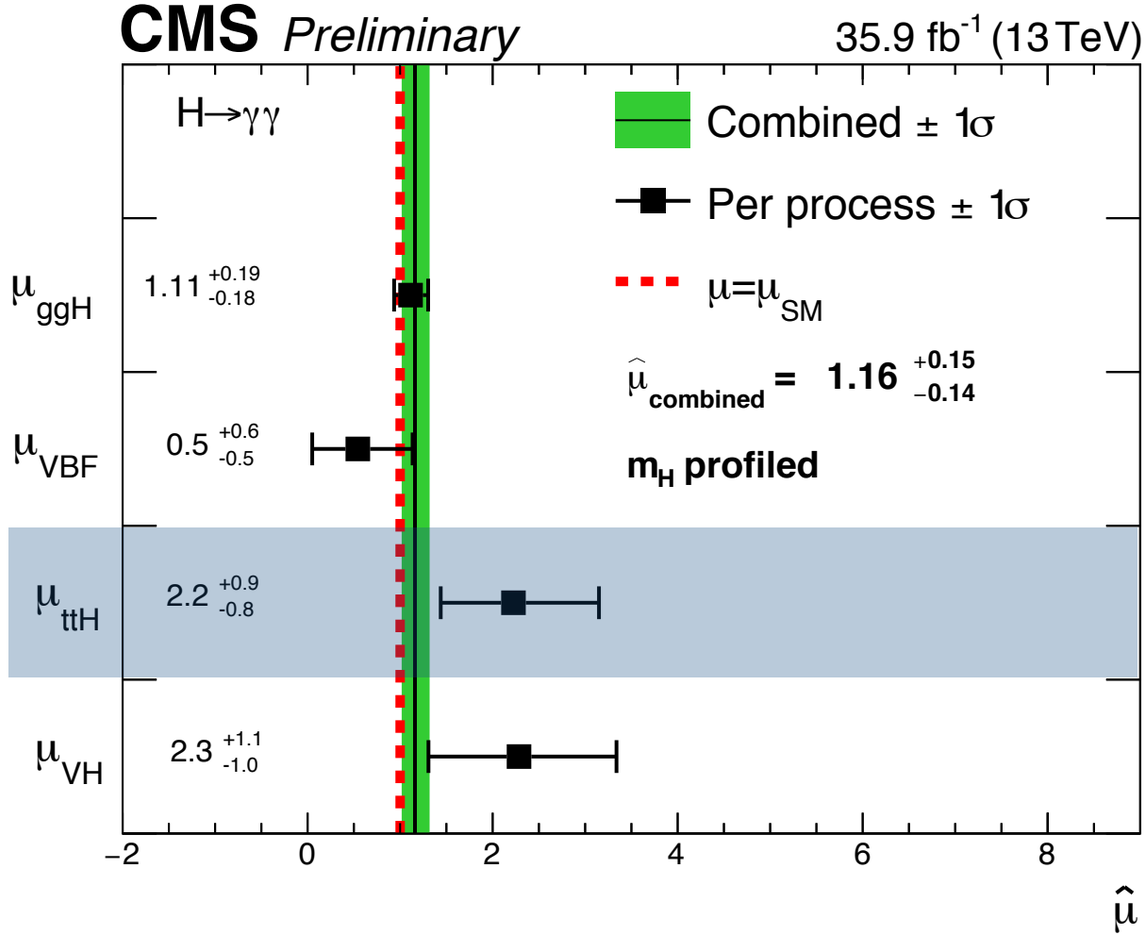
ttH WITH H→γγ : RESULTS

- Results based on **35.9 fb^{-1}** of data at **13 TeV** collected during **2016**

Signal plus background model fits to $m_{\gamma\gamma}$ distribution for ttH categories



Signal strength for each process



- Uncertainties are statistics dominated
- Signal significance **observed** (expected) : **3.3σ** (1.5σ)

SUMMARY & OUTLOOK

- Search for **$t\bar{t}H$** at **CMS** in **Run 2** (13 TeV) are being performed in different decay modes of the Higgs, with complex analysis techniques. Latest results are compatible with standard model :

  - **Multi-leptonic ($H\to ZZ^*$, $H\to WW^*$, $H\to\tau\tau$) :** 35.9 fb^{-1} (2016) + 2.3 fb^{-1} (2015)

    CMS PAS HIG-17-004

    - $\mu_{t\bar{t}H} = 1.5^{+0.5}_{-0.5}$ ($\pm$ 0.29 stat. $\pm$ 0.24 theo. $\pm$ 0.32 exp.)
    - **observed** (expected) **significance** of **3.3σ** (2.5σ)

  - τ_H **($H\to\tau\tau$, also $H\to ZZ^*$, $H\to WW^*$):** 35.9 fb^{-1} (2016)

    CMS PAS HIG-17-003

    - $\mu_{t\bar{t}H} = 0.72^{+0.62}_{-0.53}$
    - **observed** (expected) **significance** of **1.4σ** (1.8σ)

  - $H\to\gamma\gamma$: 35.9 fb^{-1} (2016)

    CMS PAS HIG-16-040

    - $\mu_{t\bar{t}H} = 2.2^{+0.9}_{-0.8}$ (stat. dominated)
    - **observed** (expected) **significance** of **3.3σ** (1.5σ)

  - $H\to b\bar{b}$: 12.9 fb^{-1} (2016)

    CMS PAS HIG-16-038

    - $\mu_{t\bar{t}H} = -0.19^{+0.8}_{-0.81}$ ($^{+0.45}_{-0.44}$ stat. $^{+0.66}_{-0.68}$ syst.)

- **Data taking** for **2017** underway, more updates to come from Run 2…

ADDITIONAL MATERIAL

SUMMARY OF RUN-1 RESULTS

- Studies of the **$t\bar{t}H$ production** in LHC **Run-1** at **CMS** were based on the different Higgs decay channels : $\gamma\gamma$, $b\bar{b}$, $\tau\tau$ (hadronic) and multi-leptonic ($H\rightarrow ZZ^*$, $H\rightarrow WW^*$, $H\rightarrow\tau\tau$ with multi-lepton final states).

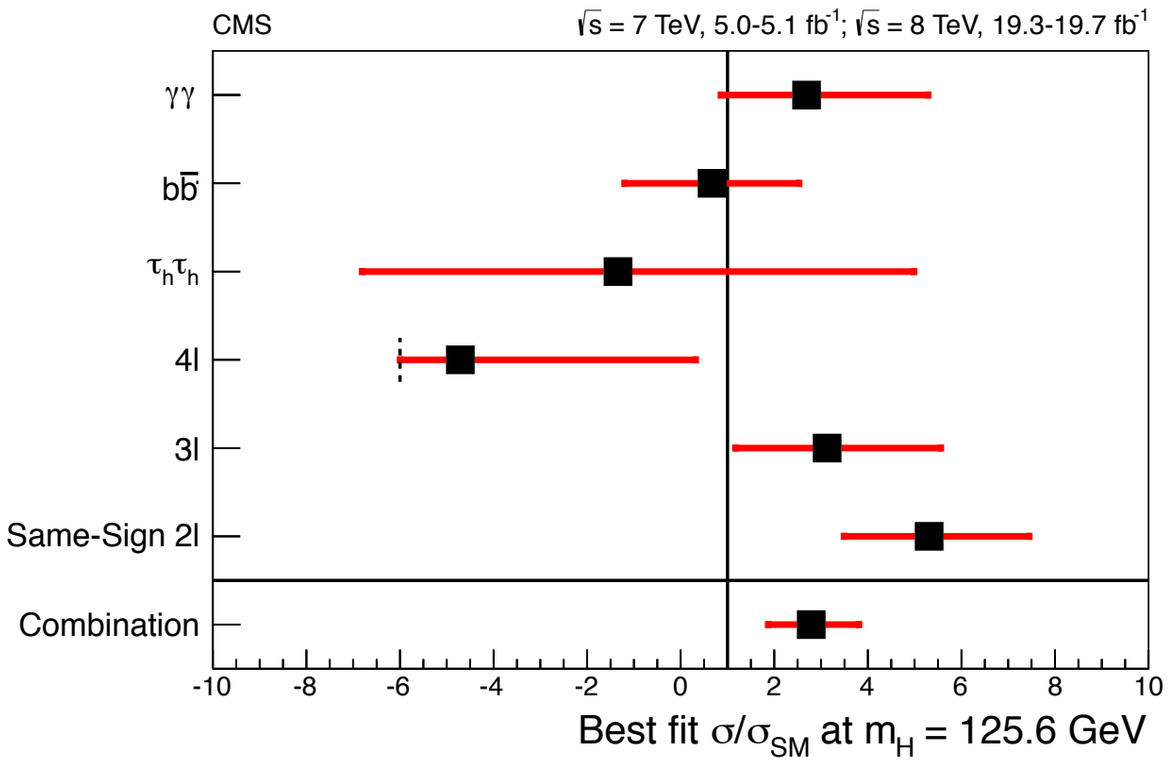
- **Combination** of different Higgs decay channels :

  ➡ **$t\bar{t}H$** combination from **Run1 CMS** measurements **:**

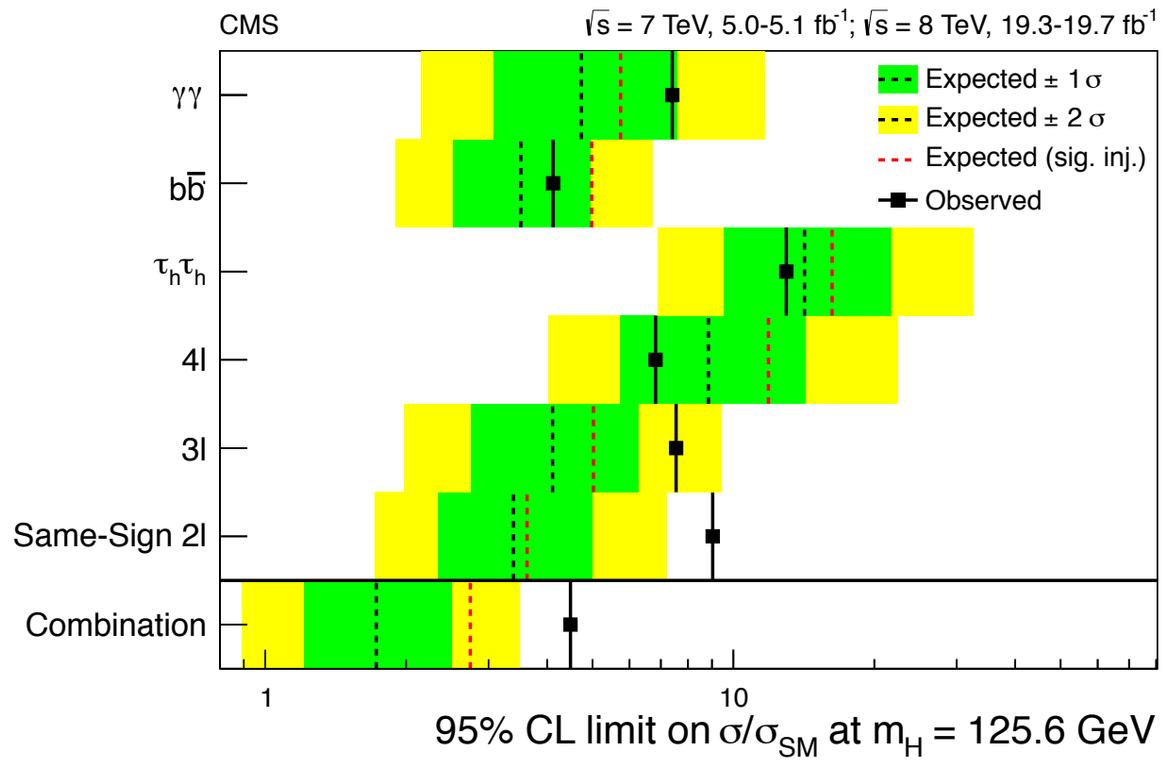
  - **$\mu_{t\bar{t}H}$ = 2.8 ± 1.0**
  - **95% CL limit = 4.5**

CMS PAS HIG-13-029
— — —
JHEP09(2014)087

Best fit signal strength : $\mu_{t\bar{t}H}$



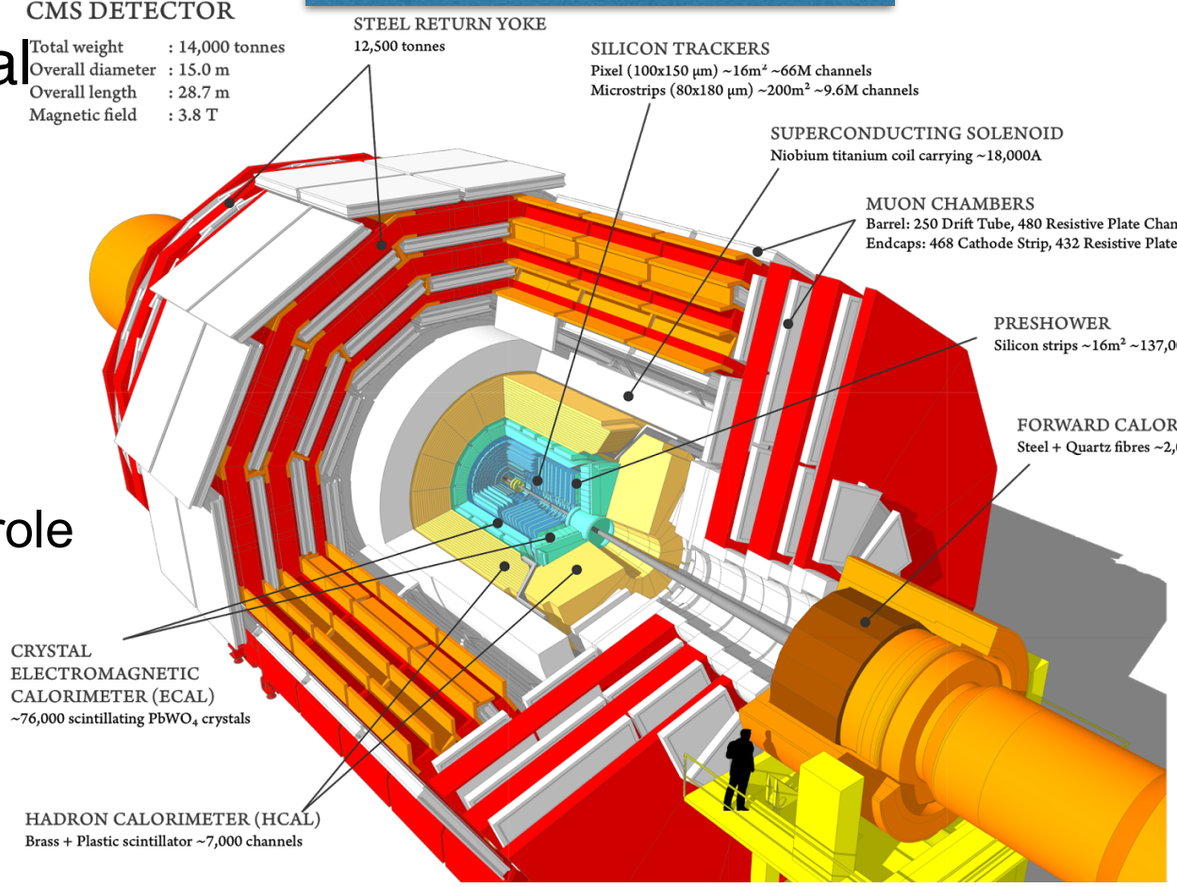
95% CL upper limit $\mu_{t\bar{t}H}$



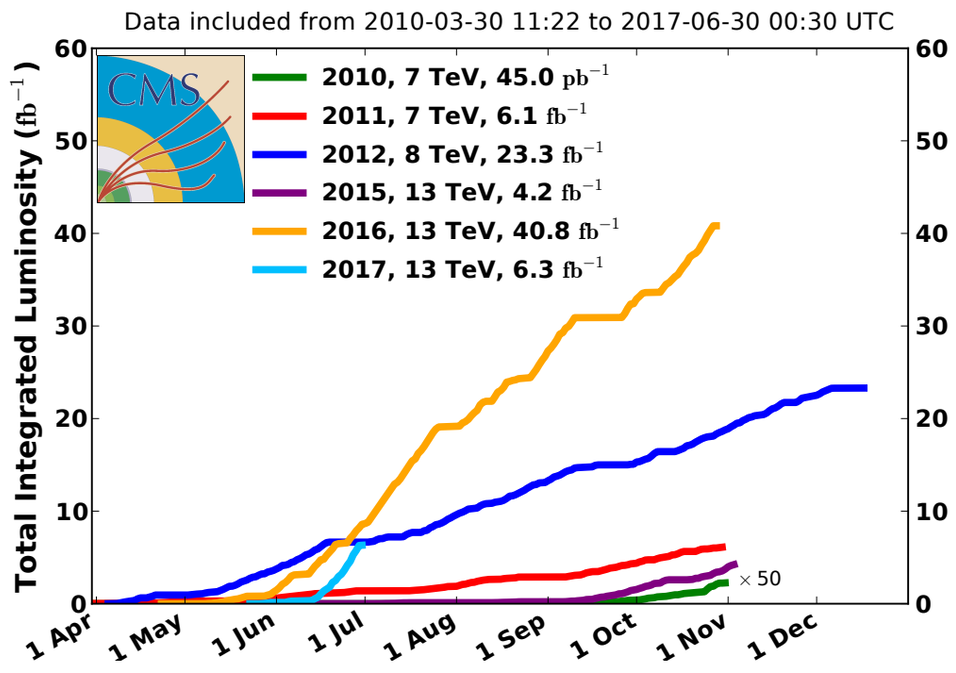
CMS AT THE LHC

- Compact Muon Solenoid (CMS) is a general purpose detector at the LHC
  - ➡ Multiple layers based around the solenoid magnet to detect different kinds of particles
  - ➡ **ttH** process produces **jets**, **leptons** along with photons, other sub detectors also play crucial role
- LHC operations for proton-proton collisions
  - ➡ At center of mass energies of 7 TeV (2010-11) and 8 TeV (2012) during Run 1
  - ➡ At center of mass energy of **13 TeV** during **Run 2** (2015-18)
  - ➡ CMS is expected to collect 100 fb^{-1} by the end of Run 2 (2018) and 300 fb^{-1} by the end of Phase 1 (2022)
- Analyses presented in **this talk** : **35.9 fb^{-1}** of data at **13 TeV** collected in **2016**

CMS detector



CMS integrated luminosity (pp)



RESULTS WITH LIMITS

- Search for **$t\bar{t}H$** at **CMS** in **Run 2** (13 TeV) are being performed in different decay modes of the Higgs, with complex analysis techniques. Latest results are compatible with standard model :

  - **Multi-leptonic ($H\to ZZ^*$, $H\to WW^*$, $H\to\tau\tau$) :** 35.9 fb^{-1} (2016) + 2.3 fb^{-1} (2015)

    CMS PAS HIG-17-004

    - **$\mu_{t\bar{t}H} = 1.5^{+0.5}_{-0.5}$** (± 0.29 stat. ± 0.24 theo. ± 0.32 exp.)
    - **observed** (expected) **significance** of **3.3σ** (2.5σ)
    - **observed** (expected) **limits** on **$\mu_{t\bar{t}H}$ = 2.5**(0.8 without SM $t\bar{t}H$)

  - **τ_H ($H\to\tau\tau$, also $H\to ZZ^*$, $H\to WW^*$):** 35.9 fb^{-1} (2016)

    CMS PAS HIG-17-003

    - **$\mu_{t\bar{t}H} = 0.72^{+0.62}_{-0.53}$**
    - **observed** (expected) **significance** of **1.4σ** (1.8σ)
    - **observed** (expected) **limits** on **$\mu_{t\bar{t}H}$ = 2.0** (2.2 with SM $t\bar{t}H$; 1.1 without SM $t\bar{t}H$)

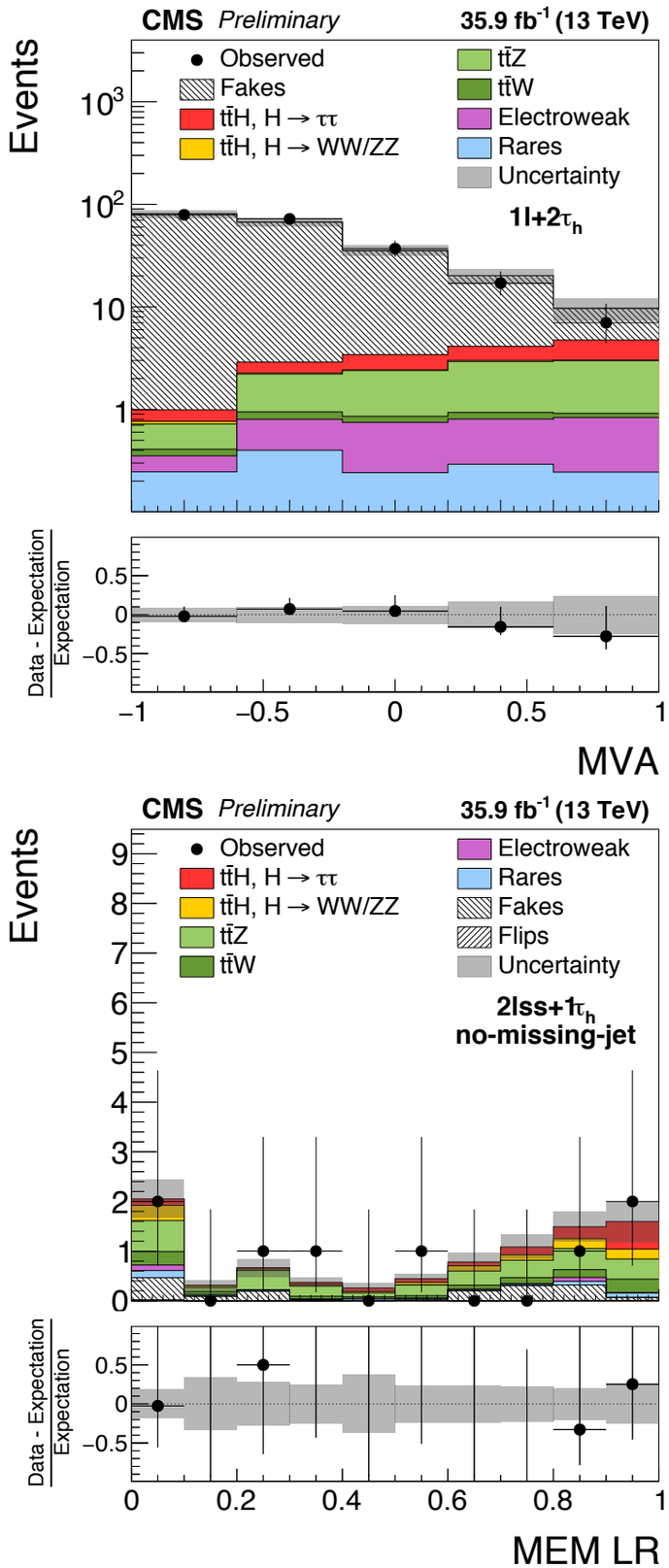
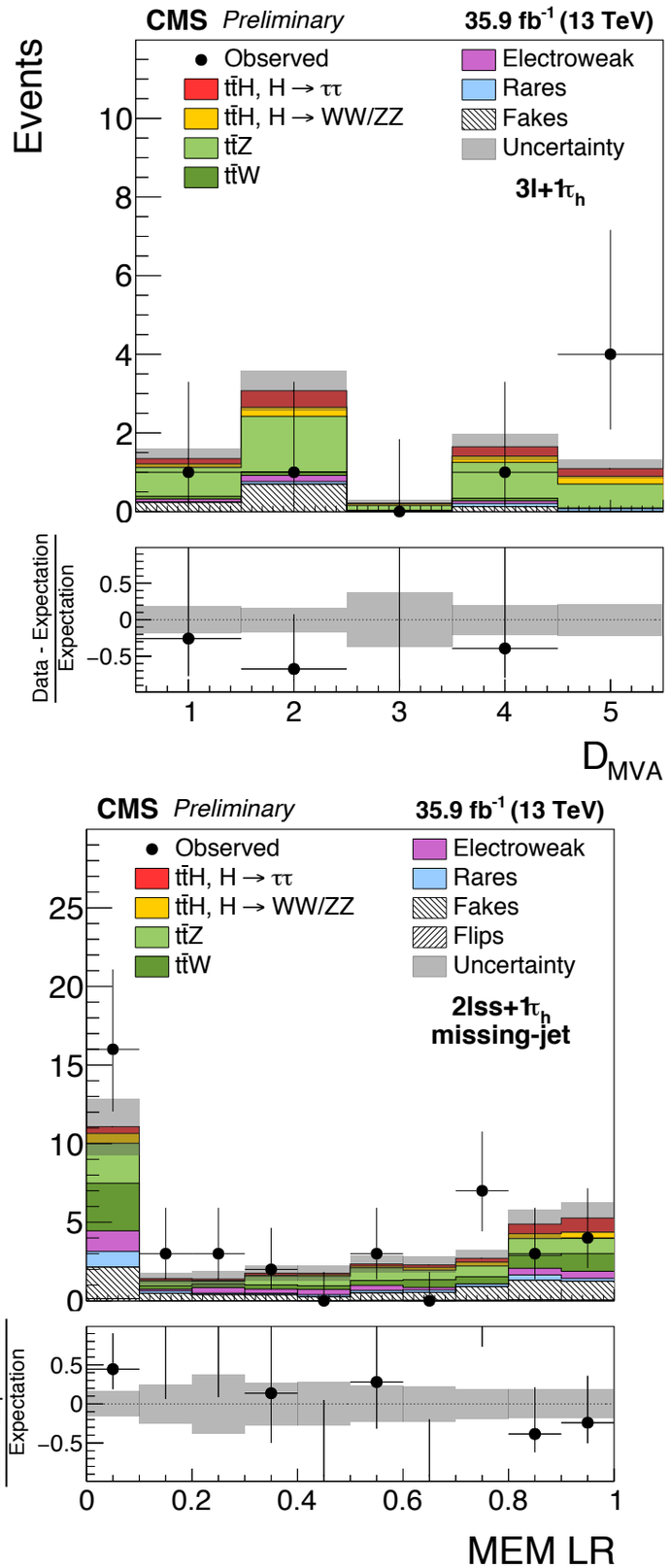
  - **$H\to\gamma\gamma$:** 35.9 fb^{-1} (2016)

    CMS PAS HIG-16-040

    - **$\mu_{t\bar{t}H} = 2.2^{+0.9}_{-0.8}$** (stat. dominated)
    - **observed** (expected) **significance** of **3.3σ** (1.5σ)

  - **$H\to b\bar{b}$:** 12.9 fb^{-1} (2016)

    CMS PAS HIG-16-038

    - **$\mu_{t\bar{t}H} = -0.19^{+0.8}_{-0.81}$** ($^{+0.45}_{-0.44}$ stat. $^{+0.66}_{-0.68}$ syst.)
    - **observed** (expected) **limits** on **$\mu_{t\bar{t}H}$ = 1.5**(1.7)

$t\bar{t}H$ TAU-HADRONIC (τ_H) : SIGNAL EXTRACTION