Observation of Top Quark Pair Production in Association with a Higgs Boson at CMS

Carmen Diez Pardos (DESY)

7 July 2018

ICHEP2018 SEOUL

XXXIX INTERNATIONAL CONFERENCE ON *high energy* PHYSICS

JULY 4 - 11, 2018 COEX, SEOUL

# Top-Higgs coupling: the hunt for t$\bar{\text{t}}$H

Best direct probe of the top-Higgs Yukawa coupling, vital step towards verifying the SM nature of the Higgs boson

- Top quark is the most strongly-coupled SM fermion ($y_t \sim 1$)

- Direct measurement of y_t in t$\bar{\text{t}}$H production:
  - gluon-gluon fusion: assumes no BSM coupling
  - Allows probing new physics in gg→H and H→ $\gamma\gamma$ effective vertices

- y_t in tH production: access to sign of the coupling
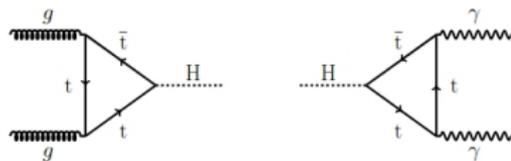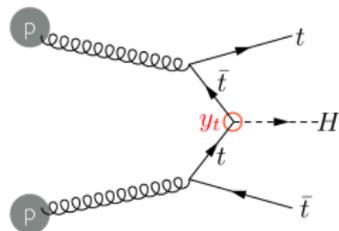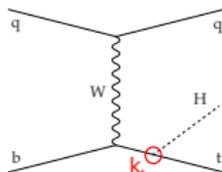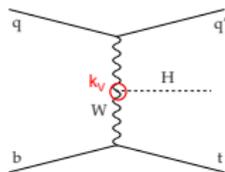
Top quark × Higgs decay channels

- Challenges: $\sigma_{t\bar{t}H} \approx 0.5$ pb, $\sigma_{t\bar{t}} \approx 830$ pb @13 TeV
  - Crucial to understand $t\bar{t}$+X (X = $b\bar{b}$, W, Z)
  - Large combinatorics of leptons and jets from top quark decays
- Exploiting all $t\bar{t}$ decay channels and Higgs decays to
  - bottom quarks → Large BR, large background contributions
  - W, Z bosons, taus → smaller production rate, lower backgrounds
 - photons → clean final state, very small rate
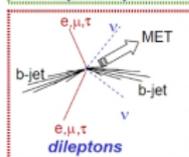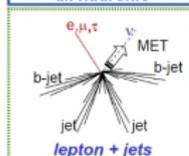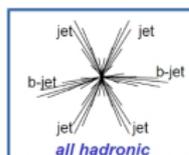
Top Pair Decay Channels
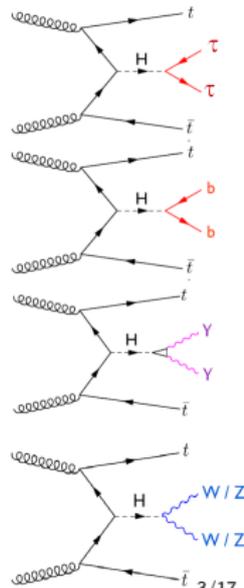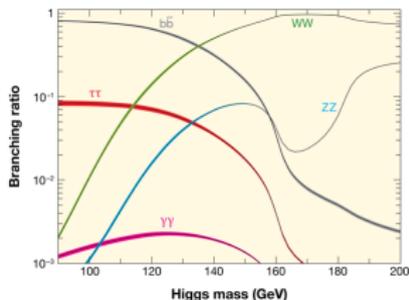
t̄tH(bb̄) Production

- Large $\mathcal{B}(H \to b\bar{b})$, access to coupling 3rd generation quarks
- Challenging final state
  - Huge combinatorics in event reconstruction
  - Poor $H \to b\bar{b}$ mass resolution
  - Large $t\bar{t} + b\bar{b}$ background of $\mathcal{O}(10)$pb with associated large theory uncertainties: from simulation
- Search channels
  - Leptonic $t\bar{t}$: higher purity
  - Fully-hadronic $t\bar{t}$: higher rate
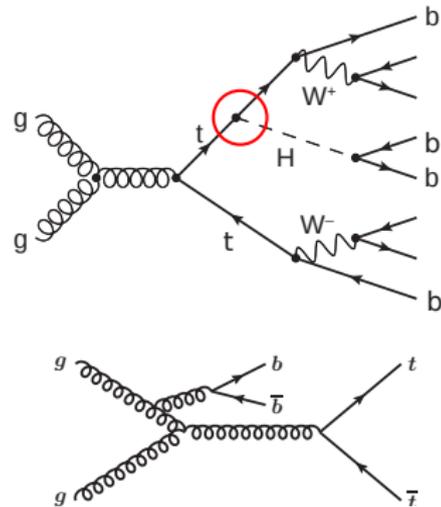
# tt̄H(bb̄): **dilepton** tt̄ channel arXiv:1804.03682, sub. to JHEP

- Events categorised
  by number
  of jets and
  b-tagged jets
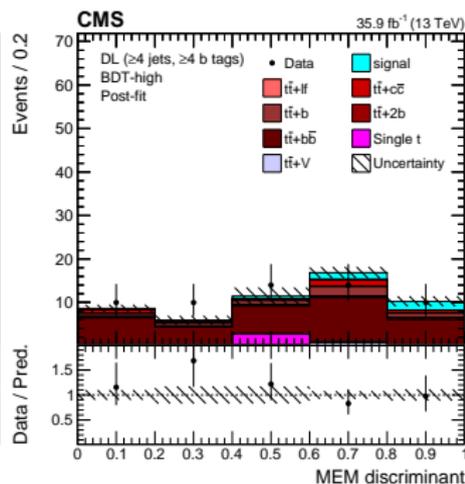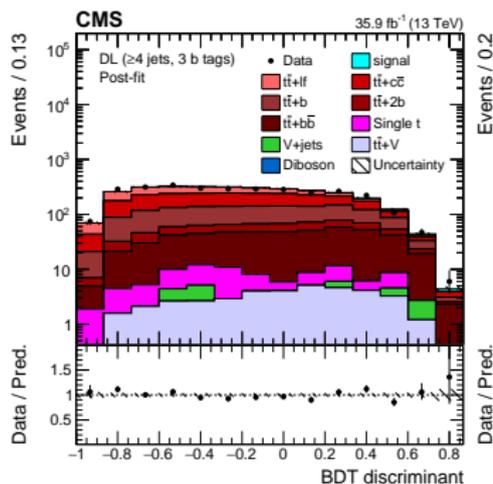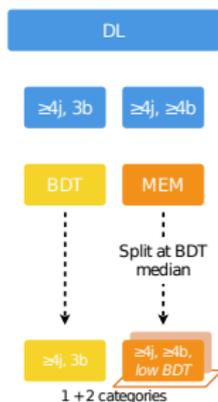

- ≥ 4j, 3b: BDT separating signal and inclusive tt̄ + jets background
- ≥ 4j, ≥ 4b: low/high BDT sub-categories + Matrix Element Method
  (MEM) separating against tt̄ + bb̄ background

tt̄H(bb̄): **lepton+jets** tt̄ channel arXiv:1804.03682, sub. to JHEP

- Search in single-lepton tt̄ channel
- Events categorised by number of jets: 4, 5, ≥ 6
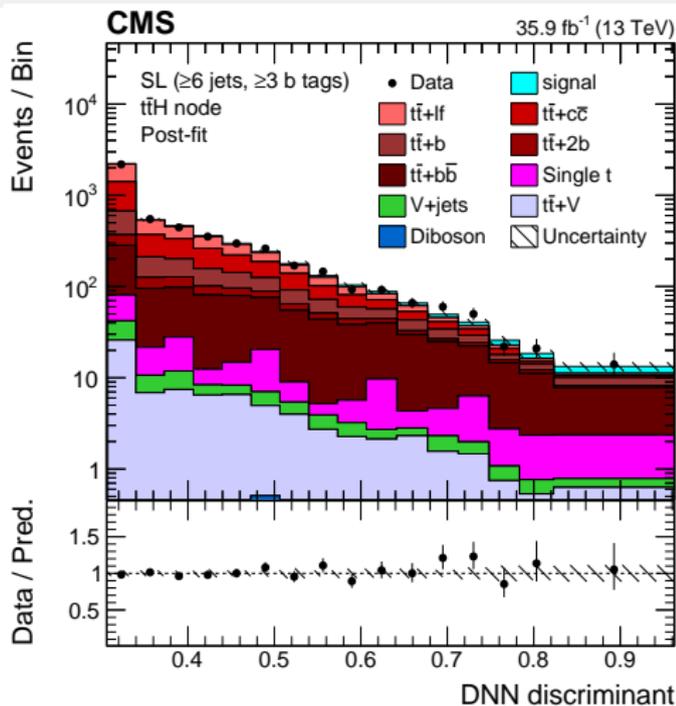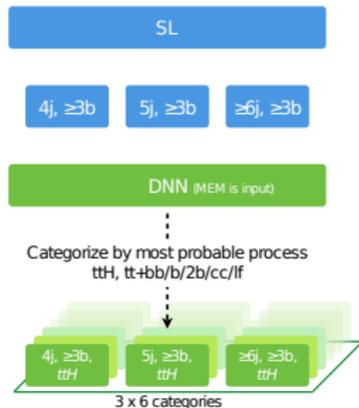
- Deep Neural Network (DNN) per jet category: multi-classification as signal or any of 5 tt̄ + jets bkgs. (tt̄ + bb̄, tt̄ + 2b, tt̄ + b, tt̄ + cc̄, tt̄ + LF)
- Final discriminant: DNN output of chosen process node

tt̄H(bb̄) Leptonic: Results

arXiv:1804.03682, sub. to JHEP
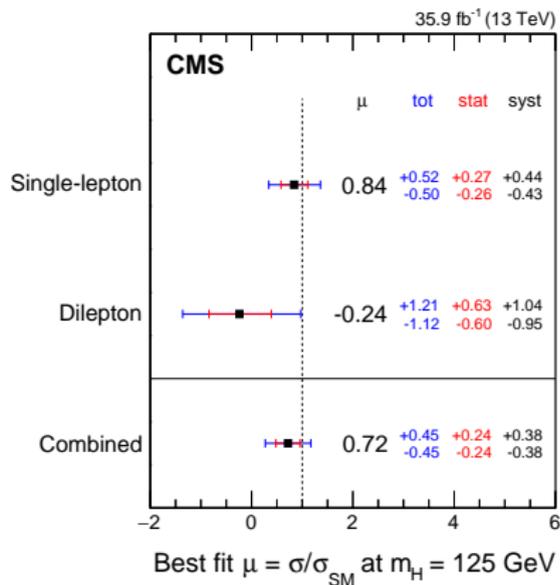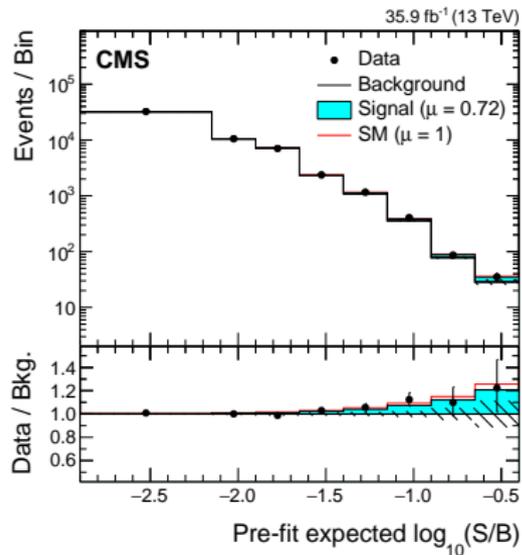
Best-fit $\mu = 0.72^{+0.45}_{-0.45}$, at 1.6 (2.2) σ obs. (exp.) significance

tt̄H(bb̄) Leptonic: Results

arXiv:1804.03682, sub. to JHEP
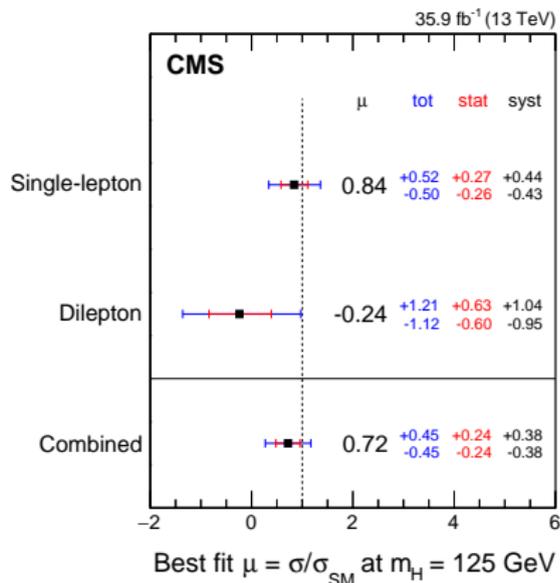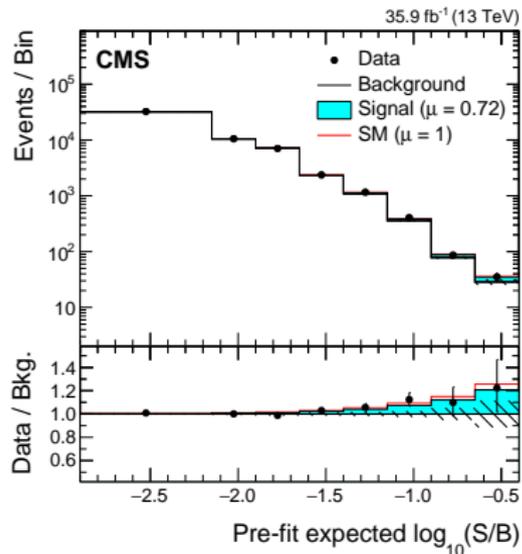
Best-fit $\mu = 0.72^{+0.45}_{-0.45}$, at 1.6 (2.2) σ obs. (exp.) significance

- Limited by tt̄ + HF and b-tagging uncertainties

tt̄H(bb̄) Hadronic

arXiv:1803.06986, sub. to JHEP

- ≥ 7 jets, ≥ 3 b-tagged jets, $H_T > 500\,\text{GeV}$, no leptons
- Events categorised by number of jets and b-tagged jets
- Dominant background: QCD-multijet production
  - Shape from low b-tag multiplicity control region in data
  - Rate from final fit to data
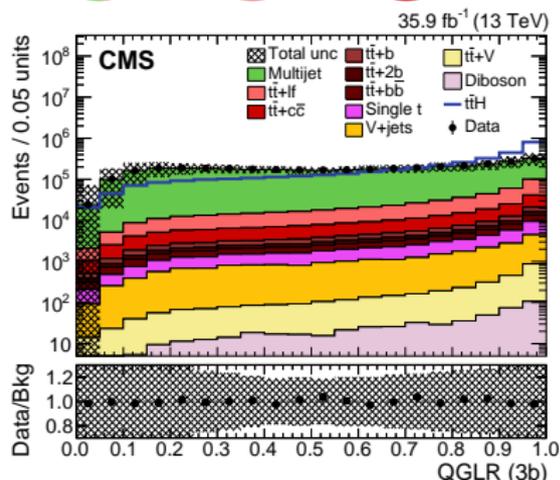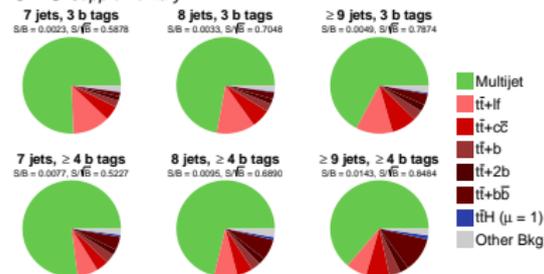
# tt̄H(bb̄) hadronic   arXiv:1803.06986, sub. to JHEP

- Final discriminant: MEM (discriminate tt̄H signal and $t\bar{t} + b\bar{b}$)
- Also performs well against the $t\bar{t} + LF$ jets and QCD multijets backgrounds
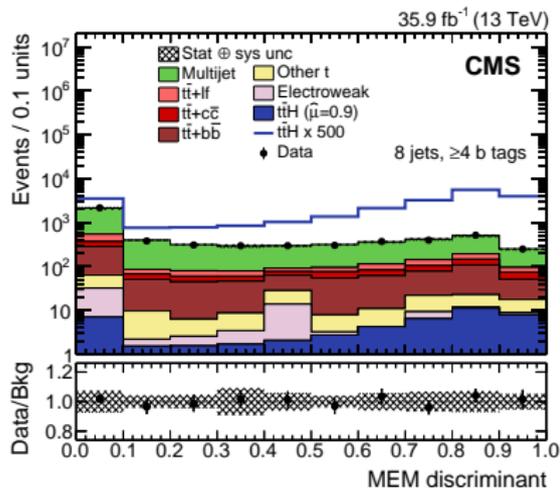
tt̄H(bb̄) hadronic   arXiv:1803.06986, sub. to JHEP

- Final discriminant: MEM (discriminate tt̄H signal and tt̄ + bb̄)
- Also performs well against the tt̄ + LF jets and QCD multijets backgrounds
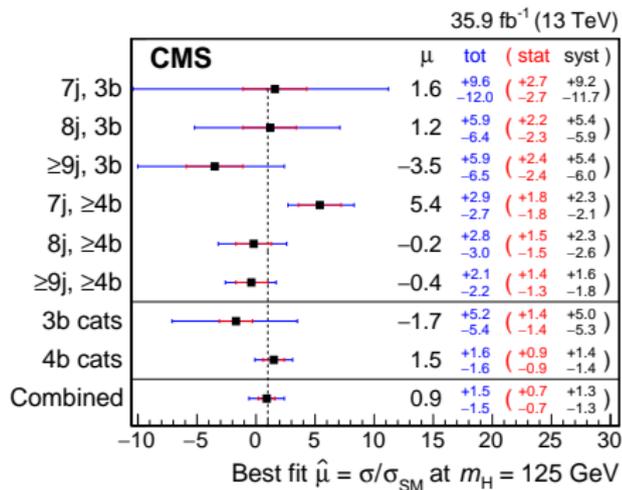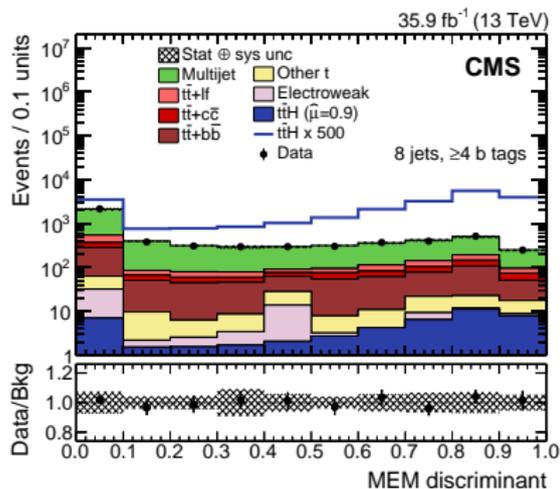

Best-fit $\mu = 0.9^{+1.5}_{-1.5}$, upper 95% C.L. limit 3.8 (3.1) obs. (exp.) × SM

- Major systematic uncertainties: Multijet estimation, tt̄ + HF prediction, b-tagging and JES etc.

tt̄H multilepton $\qquad$ arXiv:1803.05485, sub. to JHEP

- Multilepton final states: Higgs decay to W^+W^-, ZZ, and $\tau\tau$
- Events categorised based on number of leptons and τ_h candidates
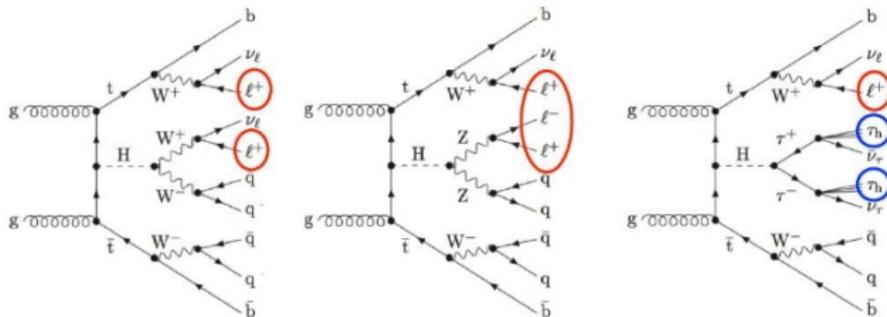
- 1 ℓ + 2 τ_h, 2 same-sign ℓ + 0, 1 τ_h, 3 ℓ + 0, 1 τ_h, 4 ℓ
- Additional requirements on jets, b-tagged jets
- Major backgrounds
    - Irreducible: $t\bar{t} + V$ and diboson, predicted from simulation and control regions
    - Reducible: non-prompt leptons in $t\bar{t} + $ jets events, estimated from data
    - Large $t\bar{t} + $ fake τ_h for 1 lepton + 2 τ_h
- BDT and MEM discriminants to separate signal from backgrounds

tt̄H multilepton: analysis strategy arXiv:1803.05485, sub. to JHEP

- Event categorisation in lepton flavor, and b-jet multiplicity
- Discriminating variables
  - MEM against tt̄Z (2 leptons same-sign + 1 τ_h)
  - Yield in 4-leptons (low stats.)

tt̄H multilepton: analysis strategy arXiv:1803.05485, sub. to JHEP
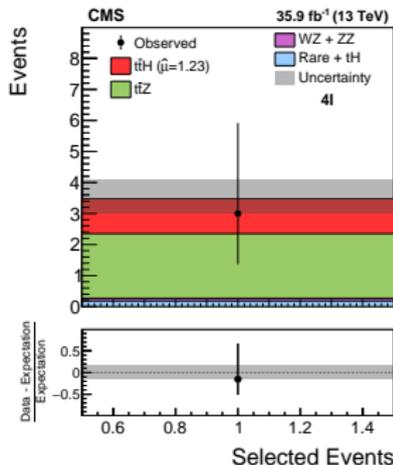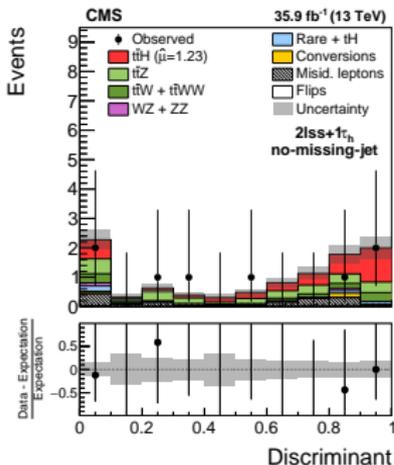
- Event categorisation in lepton flavor, and b-jet multiplicity
- Discriminating variables
  - MEM against tt̄Z (2 leptons same-sign + 1 τ_h)
  - Yield in 4-leptons (low stats.)
 - BDTs against tt̄ + jets (1l+2 τ_h) and tt̄ + jets, tt̄ + V (2 leptons same-sign, 3 leptons)
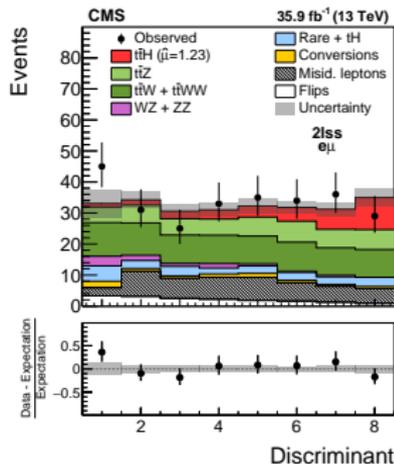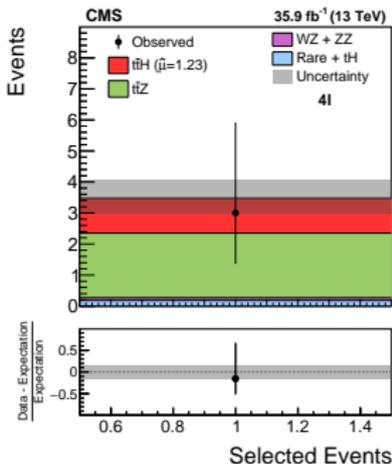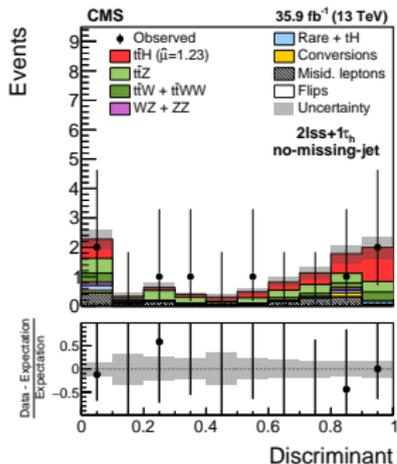
tt̄H multilepton results   arXiv:1803.05485, sub. to JHEP
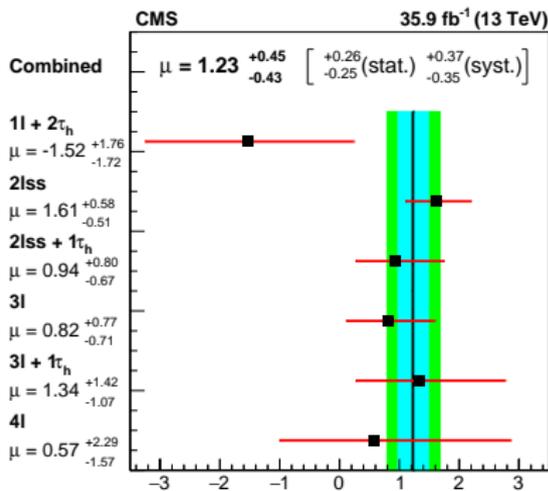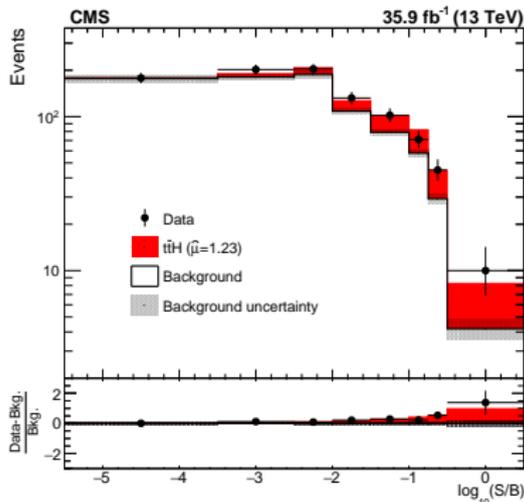

Best-fit $\mu = 1.23^{+0.45}_{-0.43}$, at 3.2 (2.8) σ obs. (exp.) significance

tt̄H multilepton results arXiv:1803.05485, sub. to JHEP
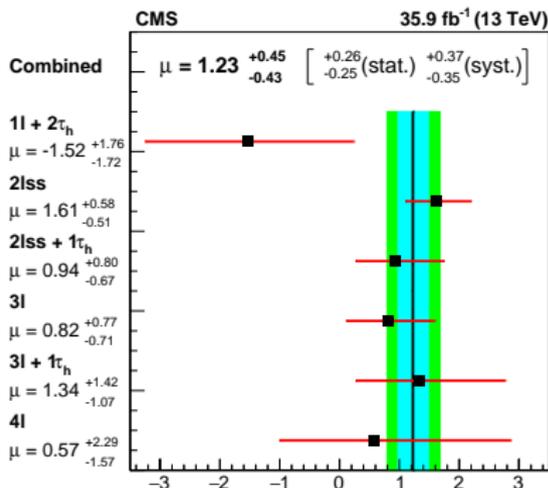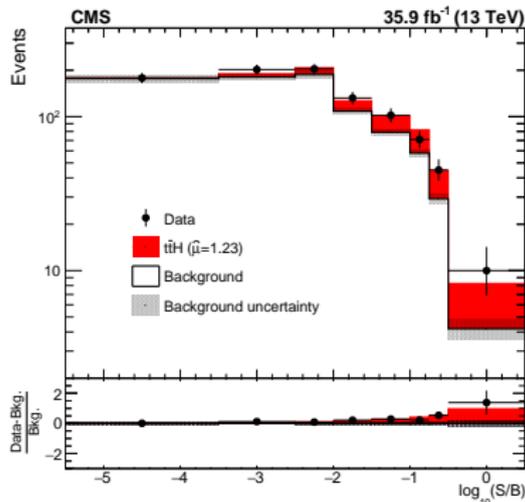
Best-fit $\mu = 1.23^{+0.45}_{-0.43}$, at 3.2 (2.8) σ obs. (exp.) significance

- Limited by non-prompt lepton estimation and tau identification, JES, tt̄H and tt̄ + V modelling
- Several channels limited by statistics

tt̄H → ZZ* → 4ℓ   PAS-HIG-18-001

- Analysis with full 2017 dataset
- Very clean final state, but tiny branching fraction
- Dedicated tt̄H channel part of the global H→ ZZ* analysis
- tt̄ hadronic and leptonic channels
  - ≥ 4 jets, ≥ 1 b-tagged jet and additional 0/1 leptons
- Combined fit (relying on $m_{4\ell}$ and a kinematic discriminant) with analysis of 2016 data (doi:10.1007/JHEP11(2017)047)

tt̄H($\gamma\gamma$)

arXiv:1804.02610, sub. to JHEP
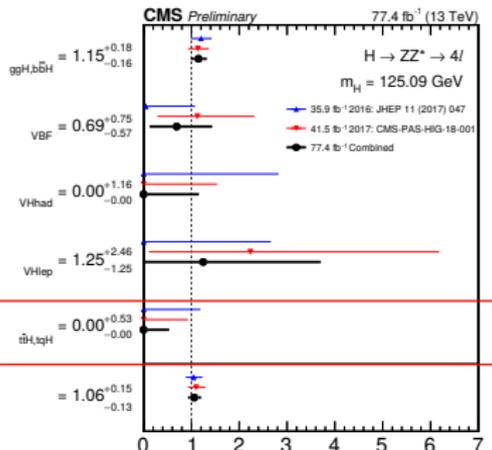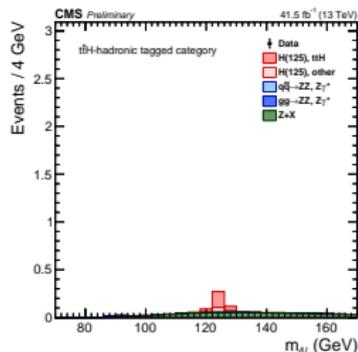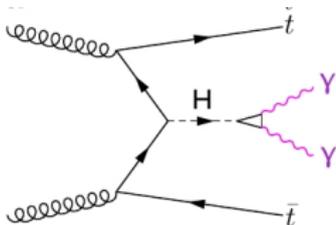
- Clear signature coming from the photons
- Higgs boson can be reconstructed as a narrow peak
- Dedicated tt̄H channel part of the global H $\rightarrow \gamma\gamma$ analysis
- tt̄ hadronic and leptonic channels
- Signal extracted from fit to $m_{\gamma\gamma}$
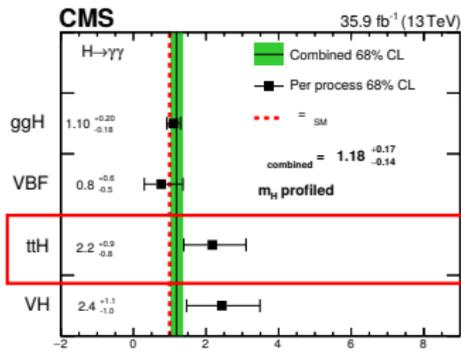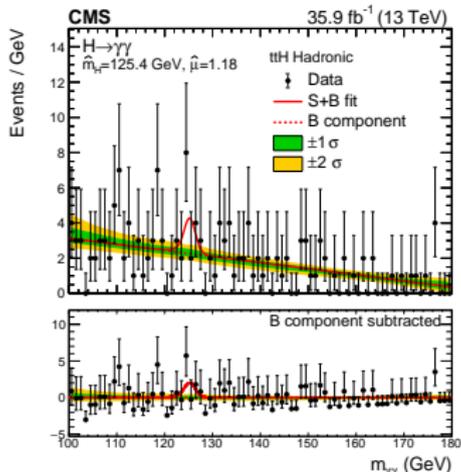
Combination: First observation of tt̄H   PRL 120 (2018) 231801

Contributing analyses

- All tt̄H analyses with 2016 data
- 7 TeV (up to 5.1 fb^{-1}) + 8 TeV (up to 19.7 fb^{-1}):

  Dedicated analyses targeting the bb and multilepton final states

  The tt̄H categories of the H$\to \gamma\gamma$ analysis

5.1 fb^{-1} (7 TeV) + 19.7 fb^{-1} (8 TeV) + 35.9 fb^{-1} (13 TeV)
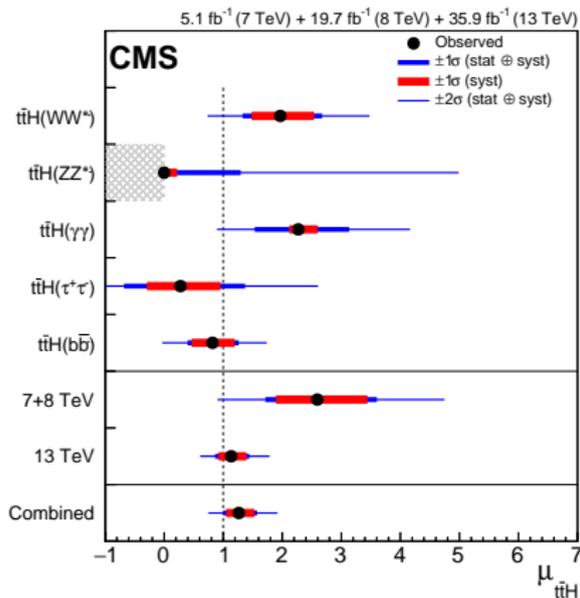
Correlations between Run-1 and Run-2 analyses

- Inclusive signal theory and some background theory uncertainties correlated
- Experimental uncertainties largely uncorrelated

Combination: First observation of tt̄H   PRL 120 (2018) 231801

- Observed significance is 5.2σ (4.2σ exp.) with respect to the $\mu_{t\bar{t}H} = 0$ hypothesis
- First observation of the tt̄H production process
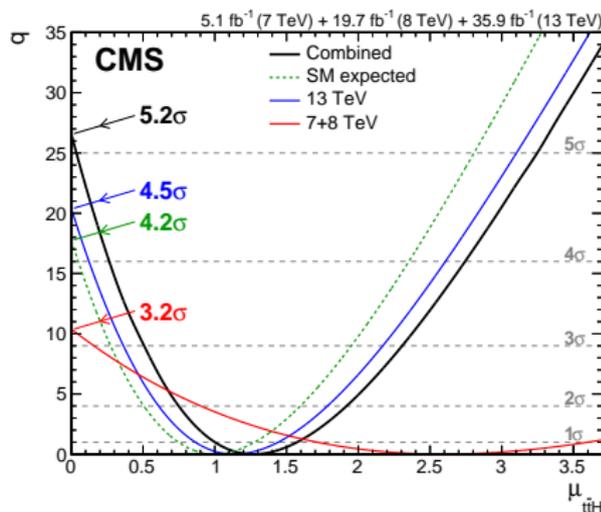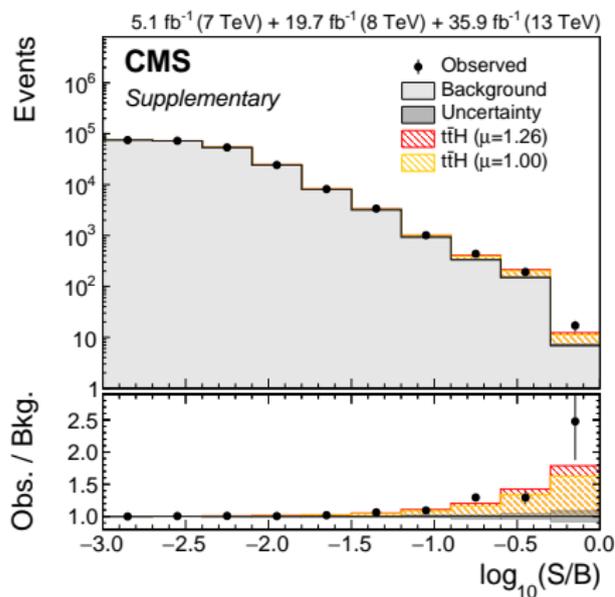
$$\mu_{t\bar{t}H} = 1.26^{+0.31}_{-0.26} = 1.26^{+0.16}_{-0.16}(\text{stat})^{+0.17}_{-0.15}(\text{expt})^{+0.14}_{-0.13}(\text{Th. bkg})^{+0.15}_{-0.07}(\text{Th. sig})$$

Summary and outlook

- Results presented for t$\bar{\text{t}}$H searches with 36 fb^{-1} of pp collision data @ 13 TeV (2016 data)
  - Improvements in analysis techniques compared to Run 1 (e.g. DNN)
  - Addition of new challenging final states: fully hadronic mode, final states with hadronically decaying τ leptons
  - Work ongoing to reduce limitations of systematic uncertainties of results with full Run-2 lumi.(signal and background modeling (t$\bar{\text{t}}$ + b$\bar{\text{b}}$, t$\bar{\text{t}}$ + V), improve non-prompt lepton estimation, jet flavor tagging...)
- First observation of t$\bar{\text{t}}$H production, combining 7, 8, and 13 TeV analyses
- New data...
  - More statistics helpful for developing more sophisticated strategy for background control
  - Statistic limited channels will become more and more relevant