Recent $t\bar{t}H$ and tH results from CMS

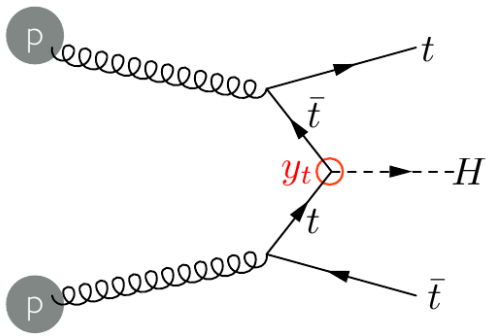
Carmen Diez Pardos (DESY)
28 May 2018

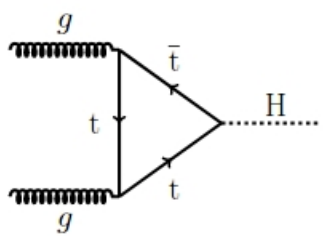
Higgs Toppings Workshop - Probing Top-Higgs Interactions at the LHC
27 May-2 Jun 2018, Benasque (Spain)

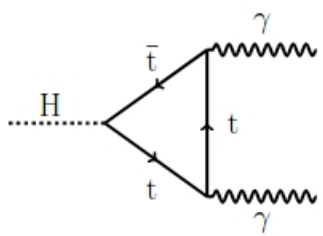
Top-Higgs coupling: the hunt for $t\bar{t}H$

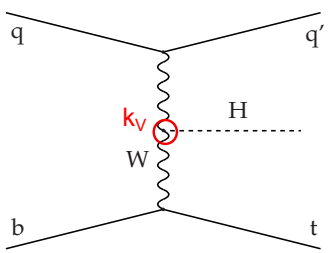
Best direct probe of the top-Higgs Yukawa coupling, vital step towards verifying the SM nature of the Higgs boson

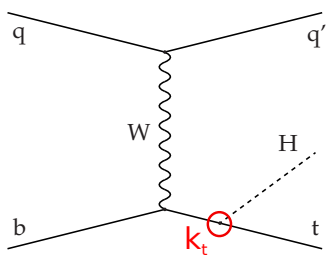
- Top quark is the most strongly-coupled SM fermion ($y_t \sim 1$)
- Direct measurement of y_t in $t\bar{t}H$ production:
  - gluon-gluon fusion: assumes no BSM coupling
  - Allows probing new physics in gg→H and H→ $\gamma\gamma$ effective vertices
- y_t in tH production: access to sign of the coupling





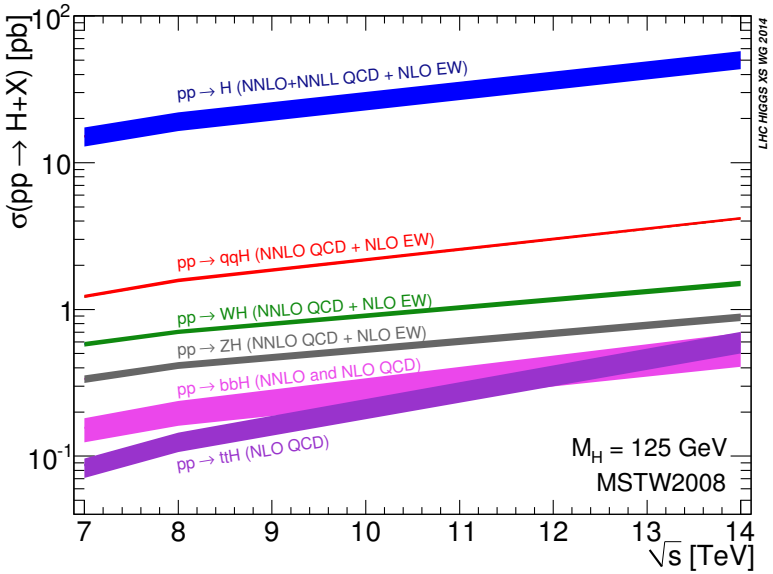






Challenges and analysis strategy

- Challenges
  - $\sigma_H \approx 0.5$ pb at $\sqrt{s}$=13TeV (m_H=125GeV), $\sigma_{t\bar{t}} \approx 830$ pb @13 TeV
    - → Larger increase in signal than backgrounds from 8 to 13 TeV
    - → By this year up to 6 times more data
  - Crucial to understand $t\bar{t}$+X (X = $b\bar{b}$, W, Z)
  - Large combinatorics of leptons and jets from top quark decays

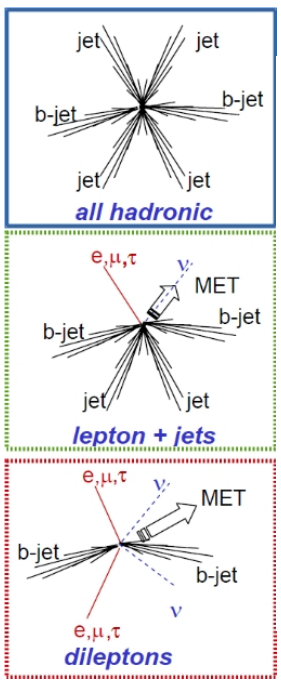


- tH: yet smaller $\sigma \approx 0.1$ pb!

- Sophisticated analysis strategies
  - $t\bar{t}$ like selections with additional searches for Higgs decay products
  - Event categorization based on top quark (W boson) and Higgs decay modes
  - MVA techniques, Matrix-Element-Methods used to extract signal

Top quark × Higgs decay channels

- 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

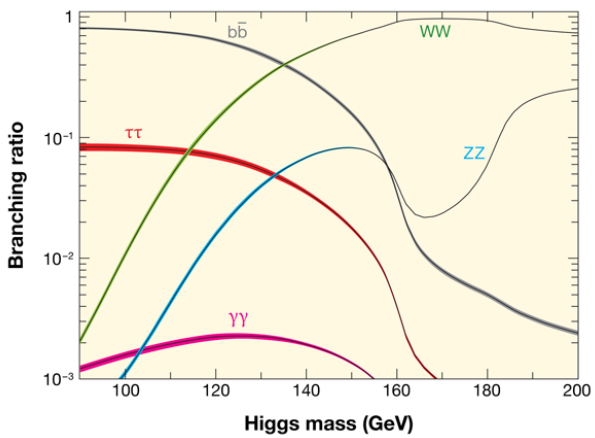


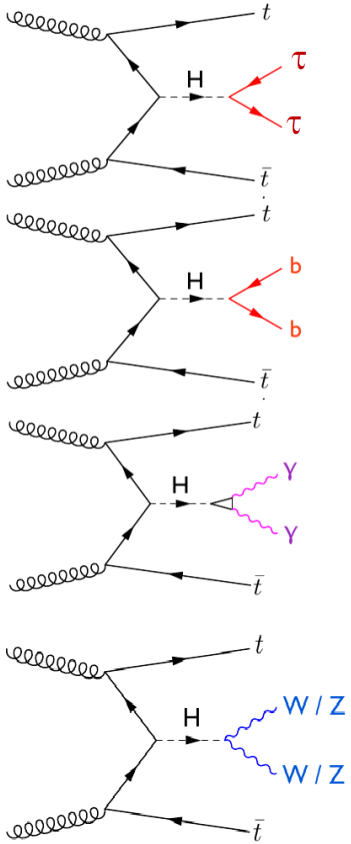
Top Pair Decay Channels

W decay	e^+	μ^+	τ^+	$u\bar{d}$	$c\bar{s}$
e^-	e^-e^+	$e^-\mu^+$	$e^-\tau^+$	electron+jets	
μ^-	μ^-e^+	$\mu^-\mu^+$	$\mu^-\tau^+$	muon+jets	
τ^-	τ^-e^+	$\tau^-\mu^+$	$\tau^-\tau^+$	tau+jets	
$\bar{u}d$	electron+jets	muon+jets	tau+jets	all-hadronic	
$\bar{c}s$					

dileptons

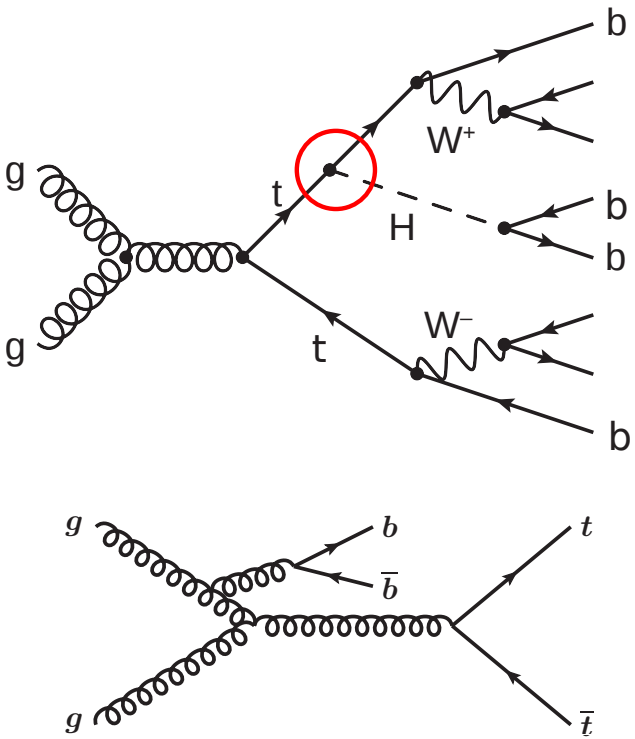
×





$t\bar{t}H(b\bar{b})$ Production

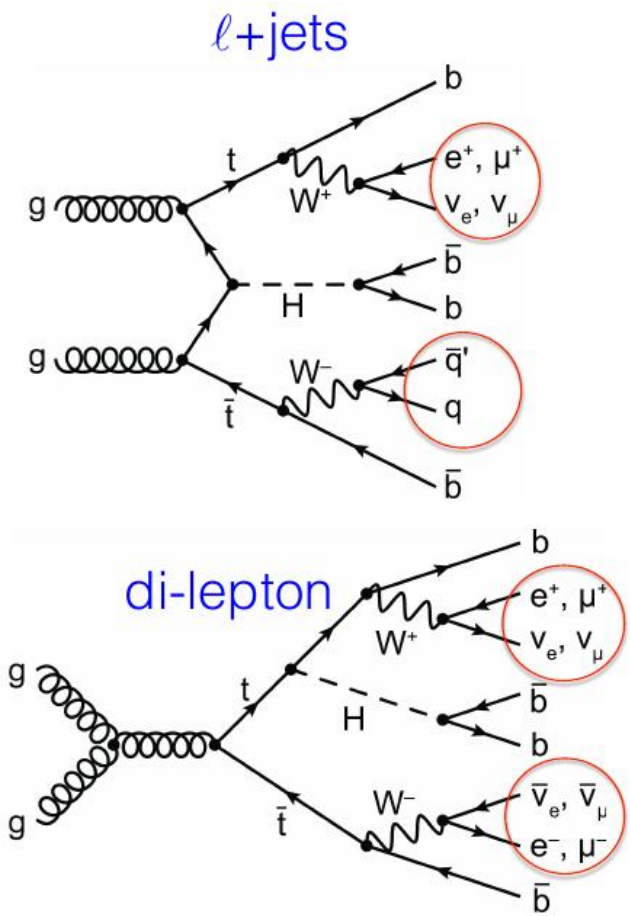
- Large $\mathcal{B}(H \to b\bar{b})$, access coupling 3rd generation quarks
- Challenging final state
  - Huge combinatorics in event reconstruction
  - Poor $H \to b\bar{b}$ mass resolution
  - Large $t\bar{t} + b\bar{b}$ background of $\mathcal{O}(10)$pb with associated large theory uncertainties: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



$t\bar{t}H(b\bar{b})$ Leptonic

arXiv:1804.03682

- Events with exactly 1 or 2 leptons (e, μ)
- At least 4 jets, with at least 3 b-tagged
- Backgrounds estimated from MC
- Exploiting Matrix-Element (ME) methods and MVA to discriminate signal from background

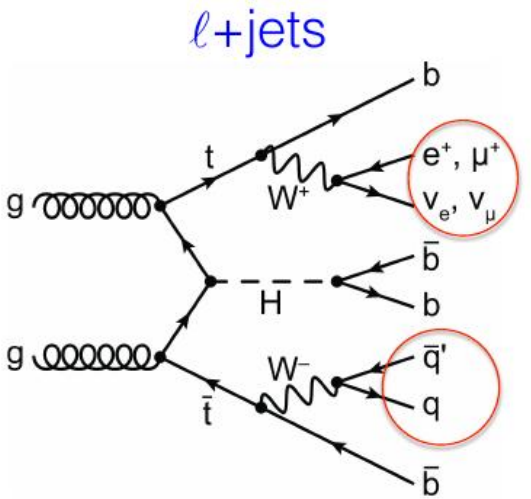


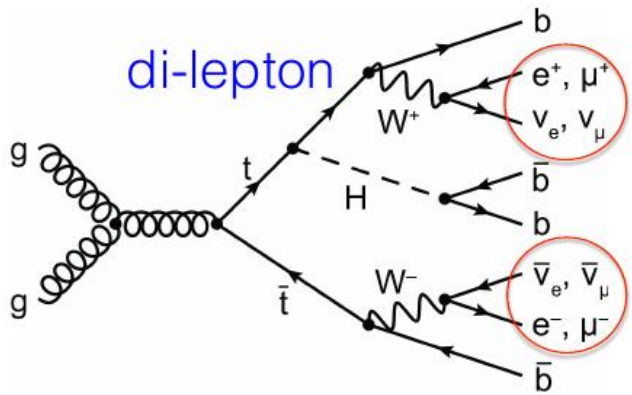
$t\bar{t}H(b\bar{b})$ Leptonic

arXiv:1804.03682

- Events with exactly 1 or 2 leptons (e, μ)
- At least 4 jets, with at least 3 b-tagged
- Backgrounds estimated from MC
- Exploiting Matrix-Element (ME) methods and MVA to discriminate signal from background
- Method chosen based on expected sensitivity

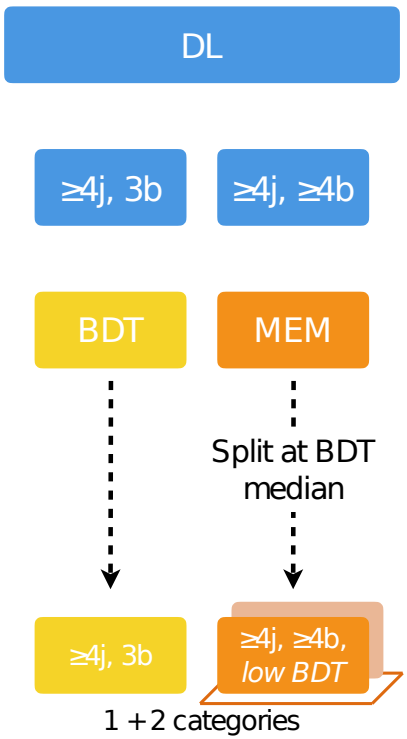
Channel	Method	Best-fit μ $\pm\text{tot}\,(\pm\text{stat}\ \pm\text{syst})$
Single-lepton	BDT+MEM	$1.0^{+0.69}_{-0.66}\left(^{+0.31\ +0.62}_{-0.30\ -0.59}\right)$
Single-lepton	DNN	$1.0^{+0.58}_{-0.55}\left(^{+0.30\ +0.50}_{-0.29\ -0.47}\right)$
Dilepton	BDT+MEM	$1.0^{+1.22}_{-1.12}\left(^{+0.65\ +1.04}_{-0.62\ -0.93}\right)$
Dilepton	DNN	$1.0^{+1.38}_{-1.36}\left(^{+0.71\ +1.18}_{-0.69\ -1.18}\right)$

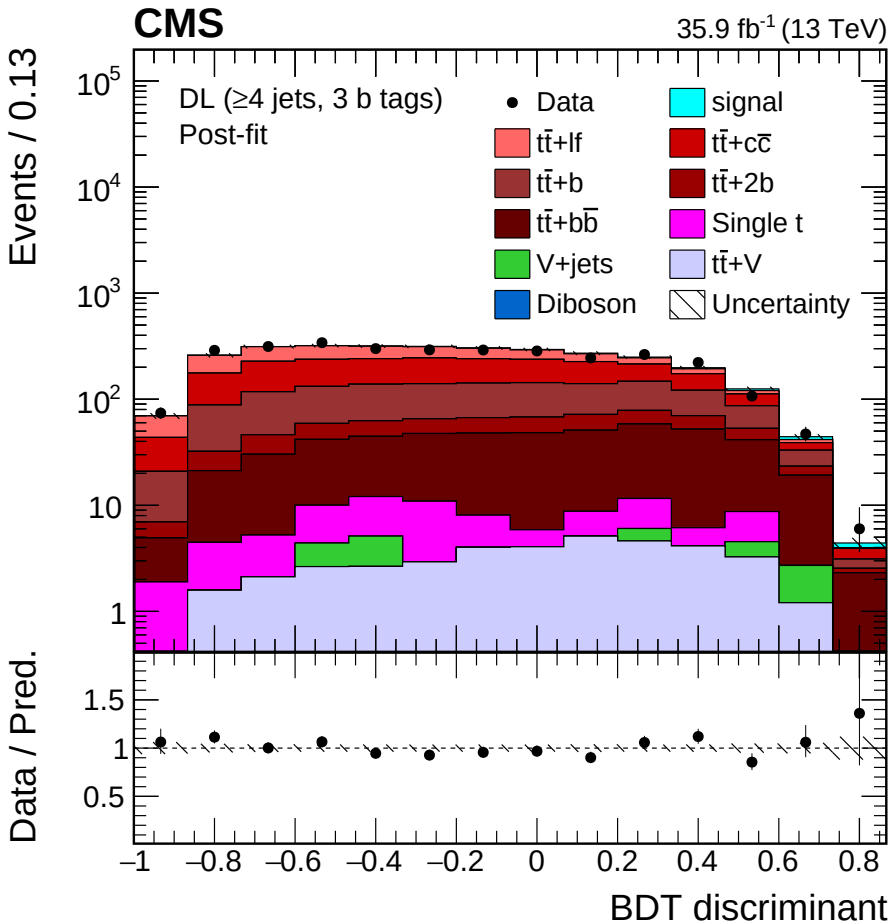




$t\bar{t}H(b\bar{b})$ Leptonic: **dilepton** $t\bar{t}$ channel

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

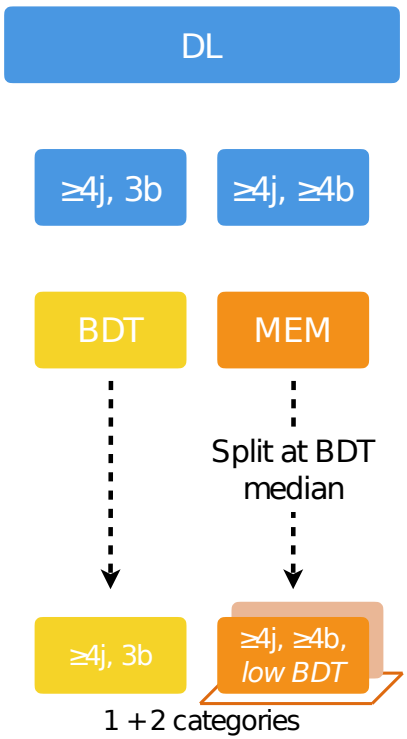


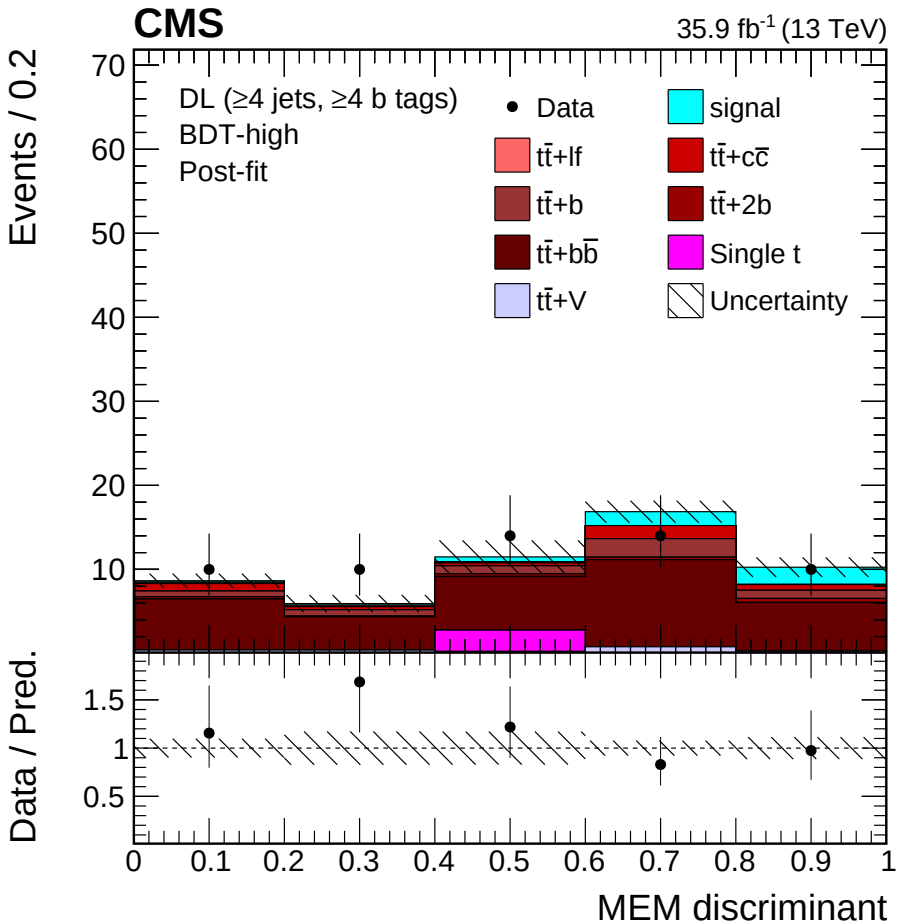


- $\geq$ 4j, 3b: BDT separating signal and inclusive $t\bar{t}$ + jets background as final discriminant

$t\bar{t}H(b\bar{b})$ Leptonic: **dilepton** $t\bar{t}$ channel

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

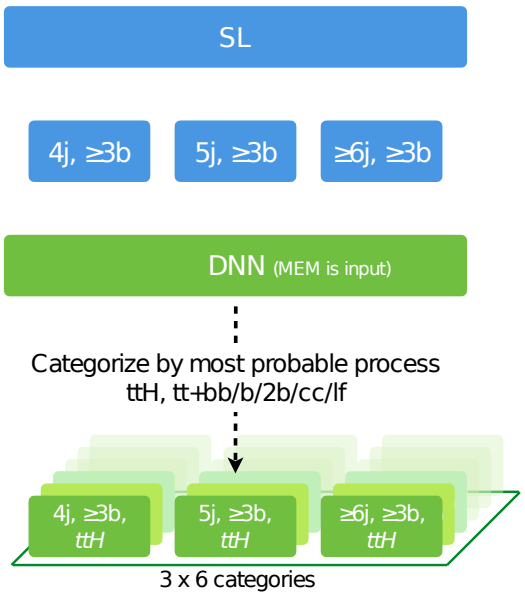


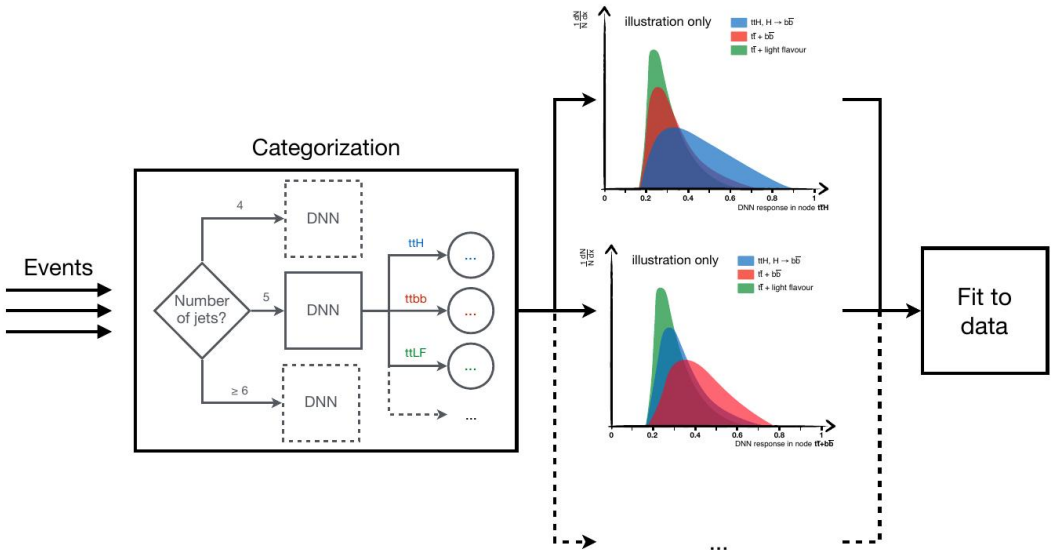


- $\geq$ 4j, $\geq$ 4b: low/high BDT sub-categories + Matrix Element Method (MEM) separating against $t\bar{t} + b\bar{b}$ background as final discriminant

$t\bar{t}H(b\bar{b})$ Leptonic: **lepton+jets** $t\bar{t}$ channel

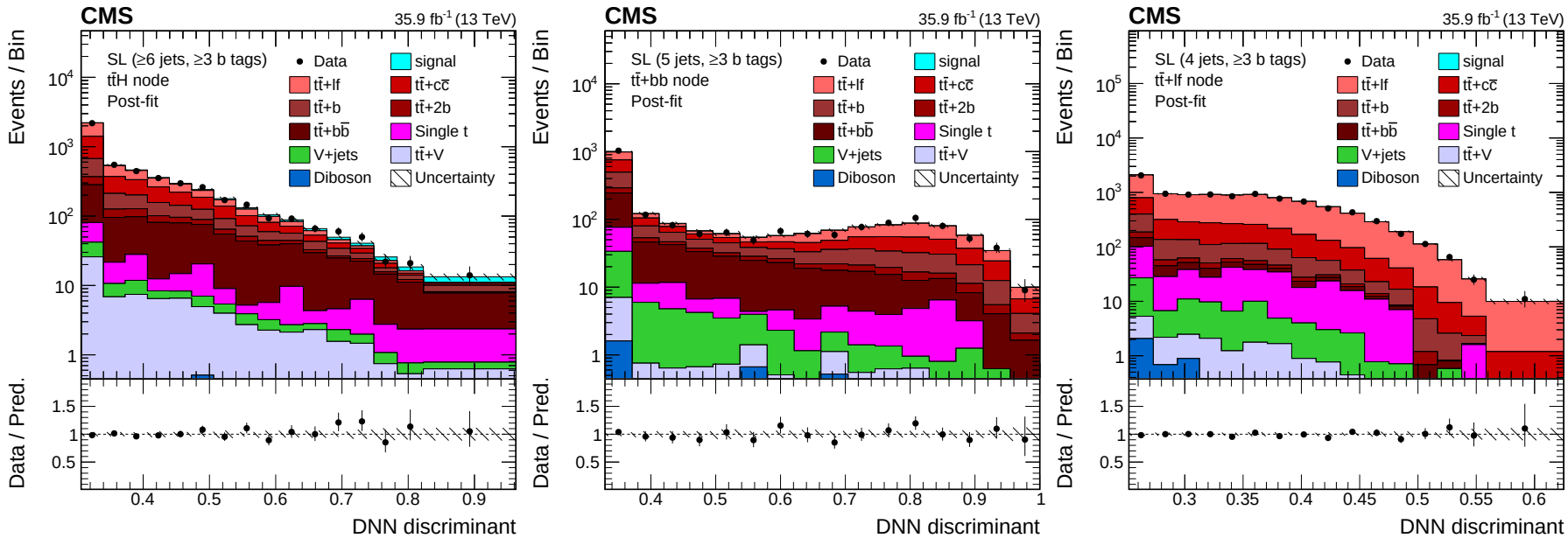
- Search in single-lepton $t\bar{t}$ channel
- Events categorised by number of jets: 4, 5, ≥ 6





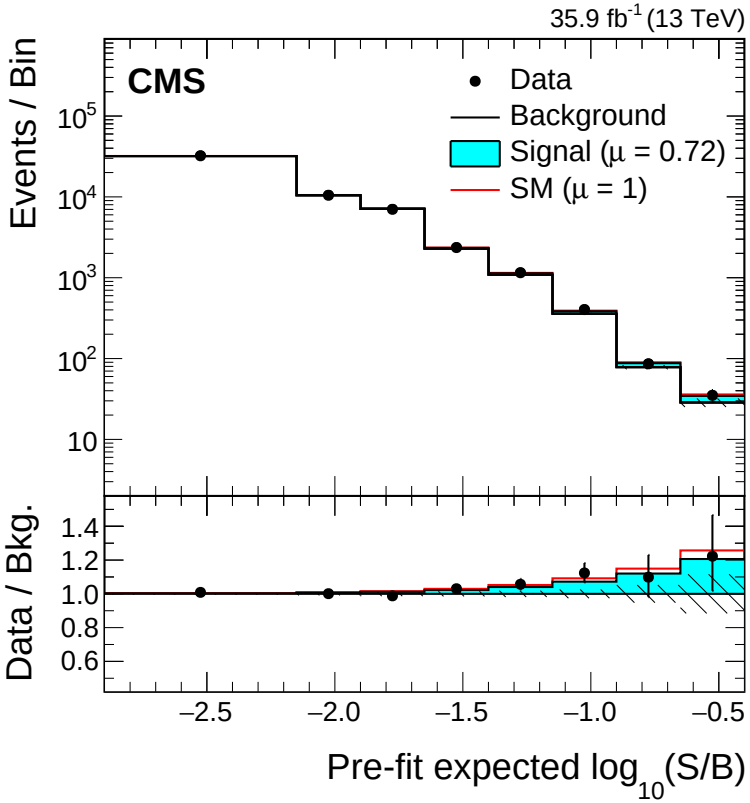
- Deep Neural Network (DNN) per jet category: multi-classification as signal or any of 5 $t\bar{t}$ + jets bkgs. ($t\bar{t}+b\bar{b}$, $t\bar{t}+2b$, $t\bar{t}+b$, $t\bar{t}+c\bar{c}$, $t\bar{t}+LF$)
- Final discriminant: DNN output of chosen process node

$t\bar{t}H(b\bar{b})$ Leptonic: **lepton+jets** $t\bar{t}$ channel



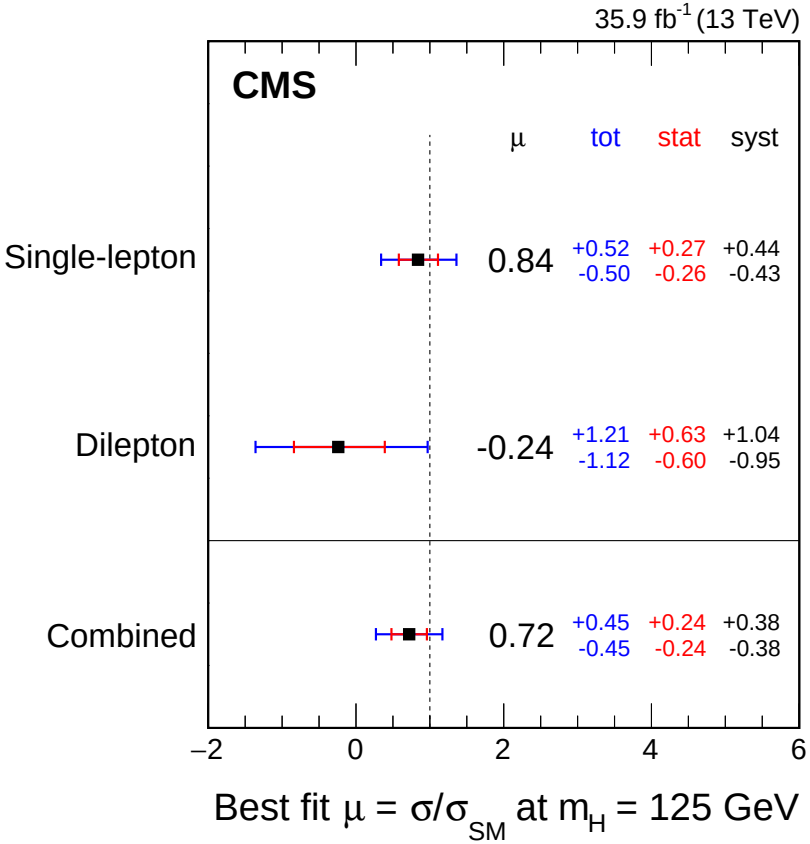
- Deep Neural Network (DNN) per jet category: multi-classification as signal or any of 5 $t\bar{t}$ + jets bkgs. ($t\bar{t}+b\bar{b}$, $t\bar{t}+2b$, $t\bar{t}+b$, $t\bar{t}+c\bar{c}$, $t\bar{t}+LF$)
- Final discriminant: DNN output of chosen process node

$t\bar{t}H(b\bar{b})$ Leptonic: Results



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

$t\bar{t}H(b\bar{b})$ Leptonic: Results



Uncertainty source	$\pm\sigma_\mu$
total experimental	+0.15/-0.16
b tagging	+0.11/-0.14
jet energy scale	+0.06/-0.07
total theory	+0.28/-0.29
$t\bar{t}$ + HF cross-section and PS	+0.24/-0.28
size of MC samples	+0.14/-0.15
total systematic	+0.38/-0.38
statistical	+0.24/-0.24
total	+0.45/-0.45

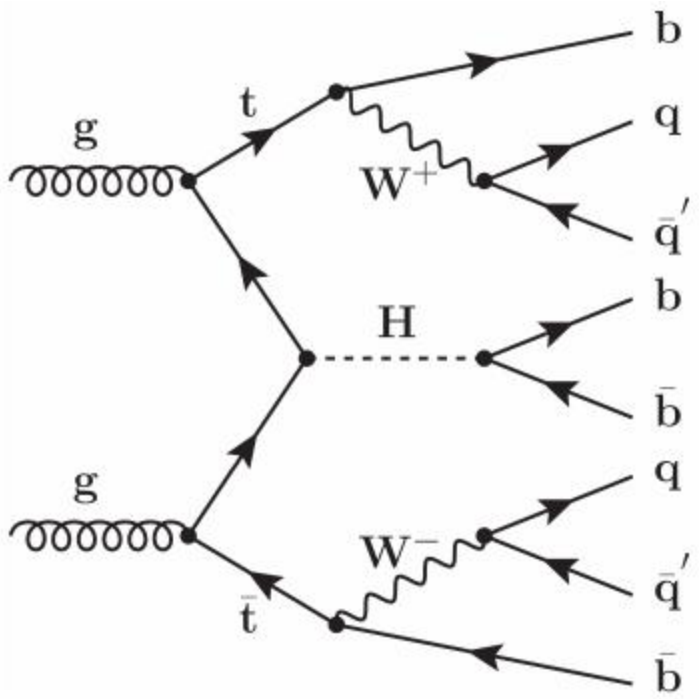
Best-fit $\mu = 0.72^{+0.45}_{-0.45}$, at 1.6 (2.2) σ obs. (exp.) significance

- Limited by $t\bar{t}$ + HF and b-tagging uncertainties

$t\bar{t}H(b\bar{b})$ **Hadronic**

arXiv:1803.06986

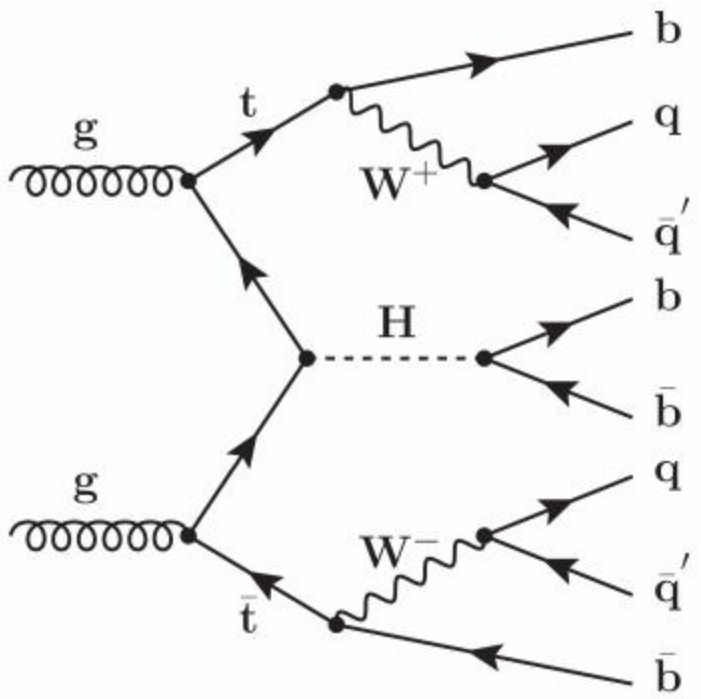
- Trigger: $\geq$ 6 jets, large HT, $\geq$1 or 2 b-tagged jets
- Challenge:
  - Large backgrounds from QCD multi-jets, $t\bar{t}$ + jets, and the irreducible $t\bar{t} + b\bar{b}$
- Larger signal contribution
- Possibility to fully reconstruct the event



$t\bar{t}H(b\bar{b})$ Hadronic

arXiv:1803.06986

- Trigger: $\geq$ 6 jets, large HT, $\geq$1 or 2 b-tagged jets
- Challenge:
  - Large backgrounds from QCD multi-jets, $t\bar{t}$ + jets, and the irreducible $t\bar{t} + b\bar{b}$
- Larger signal contribution
- Possibility to fully reconstruct the event

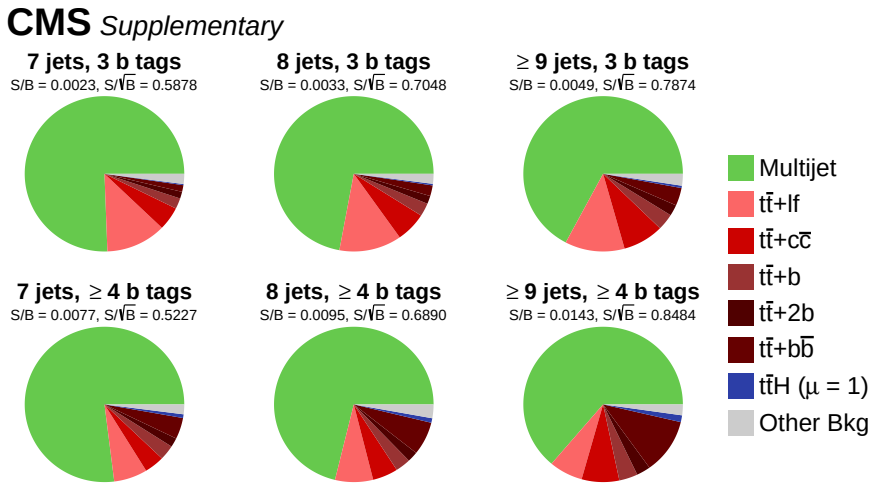


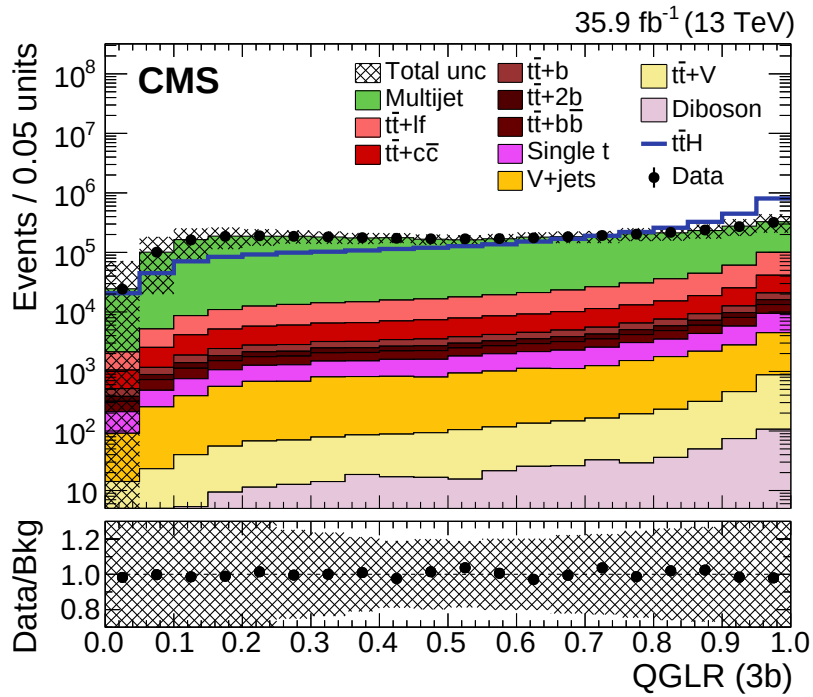
- A quark-gluon discriminant is used to differentiate quarks jets from gluon jets
  - Discrimination against QCD multijets
- Dedicated MEM to discriminate signal against $t\bar{t}$ + jets and $t\bar{t} + b\bar{b}$

$t\bar{t}H(b\bar{b})$ Hadronic: Analysis strategy

- ≥ 7 jets, ≥ 3 b-tagged jets, $H_T > 500$ 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

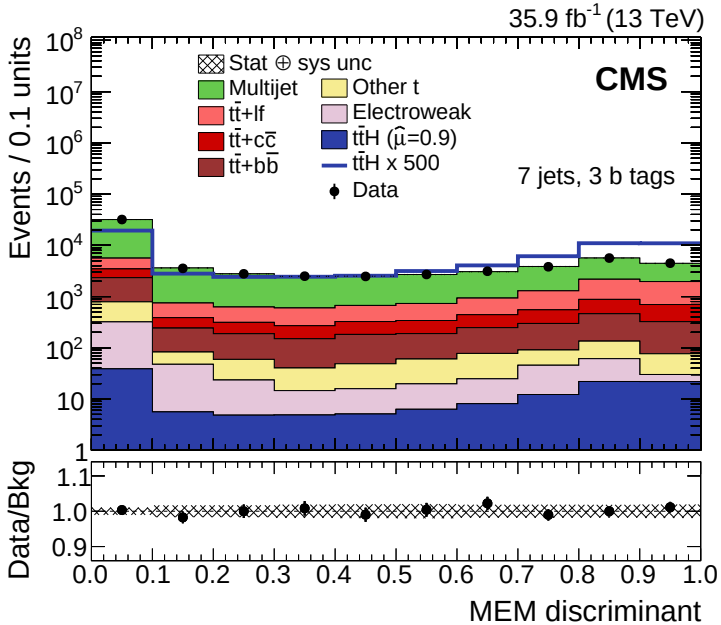
	$N_{CSVM} = 2$ $N_{CSVL} \geq 3$	$N_{CSVM} \geq 3$
QGLR > 0.5	CR (to extract distribution)	SR (final analysis)
QGLR < 0.5	Validation CR (to validate distribution)	VR (comparison with data)

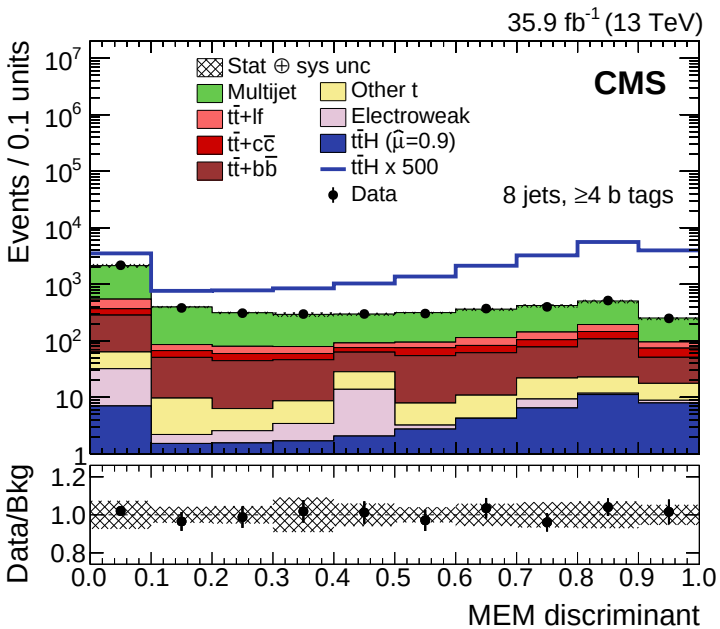


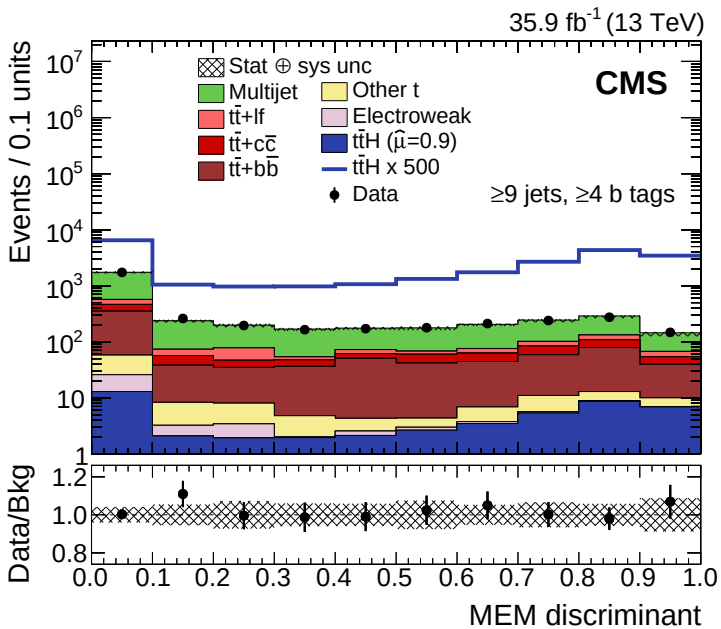


$t\bar{t}H(b\bar{b})$ Hadronic: Analysis strategy

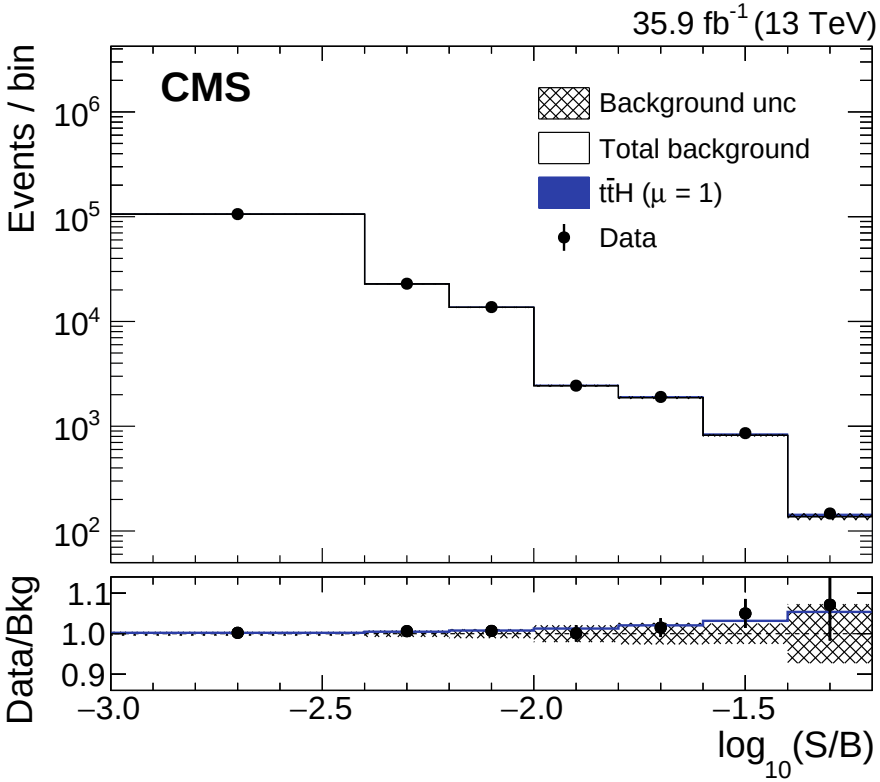
- Final discriminant: MEM
- Constructed from LO matrix elements for the $t\bar{t}H$ signal and $t\bar{t} + b\bar{b}$ backgrounds
- Also performs well against the $t\bar{t}$ + LF jets and QCD multijets backgrounds

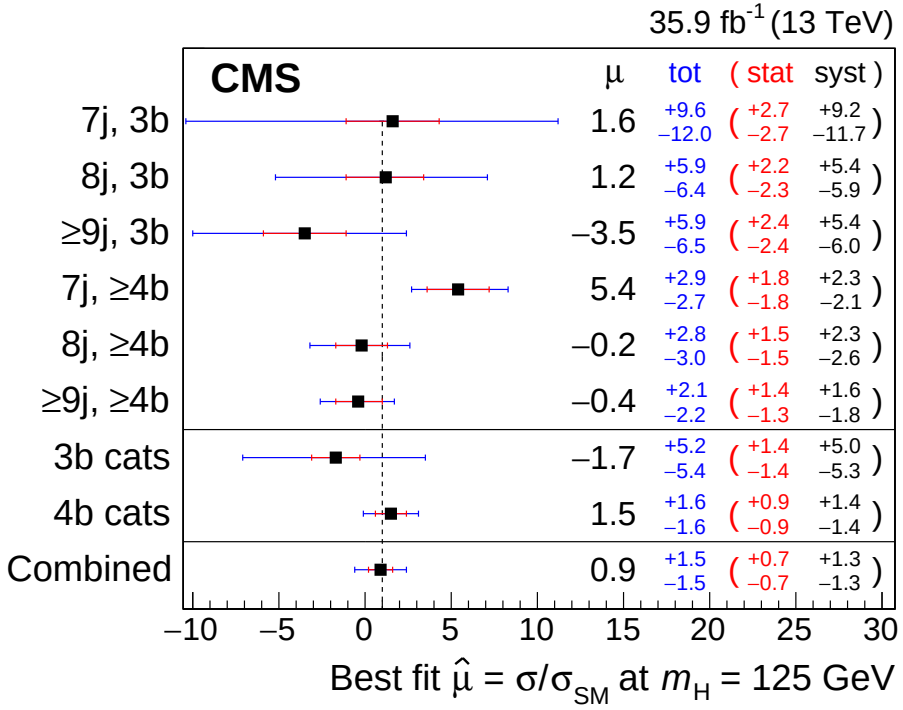






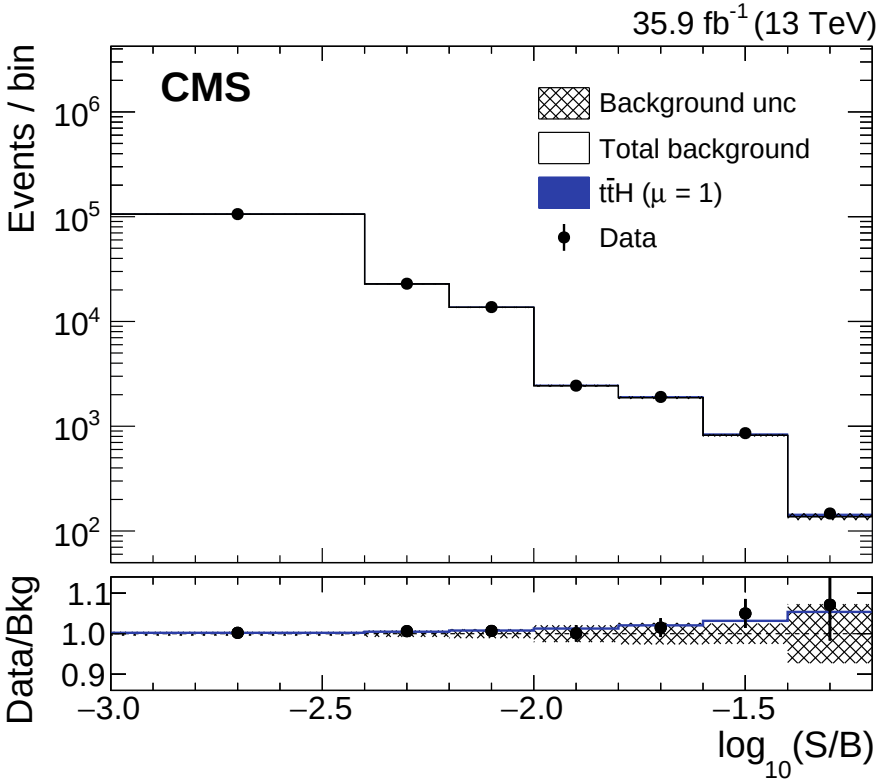
$t\bar{t}H(b\bar{b})$ Hadronic: Results

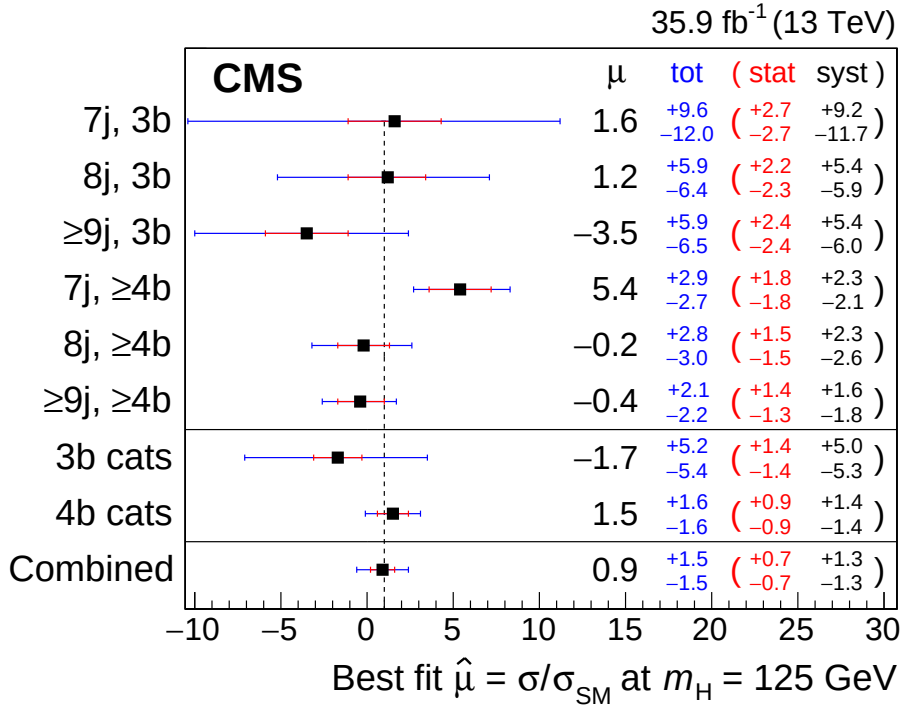




Best-fit $\mu = 0.9^{+1.5}_{-1.5}$, upper 95% C.L. limit 3.8 (3.1) obs. (exp.) $\times$ SM

$t\bar{t}H(b\bar{b})$ Hadronic: Results





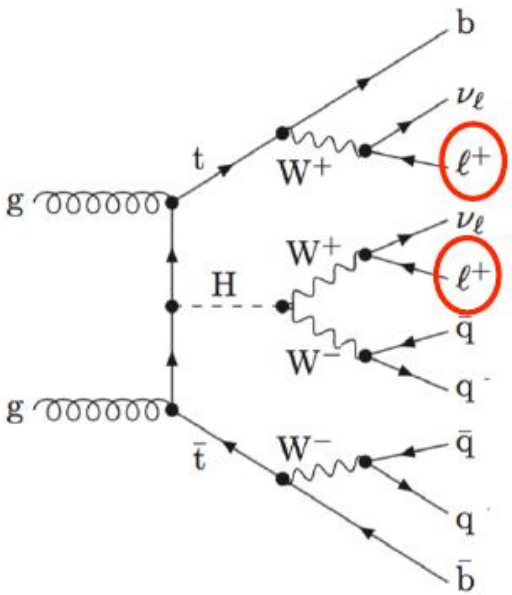
Best-fit $\mu = 0.9^{+1.5}_{-1.5}$, upper 95% C.L. limit 3.8 (3.1) obs. (exp.) $\times$ SM

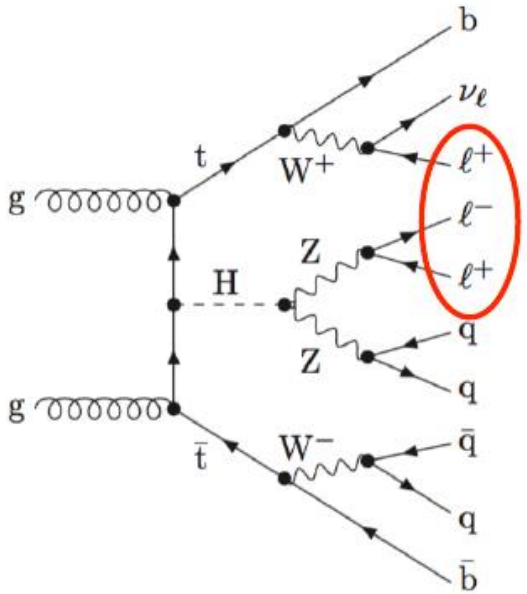
- Major systematic uncertainties: Multijet estimation, tt+HF prediction, b-tagging and JES etc.

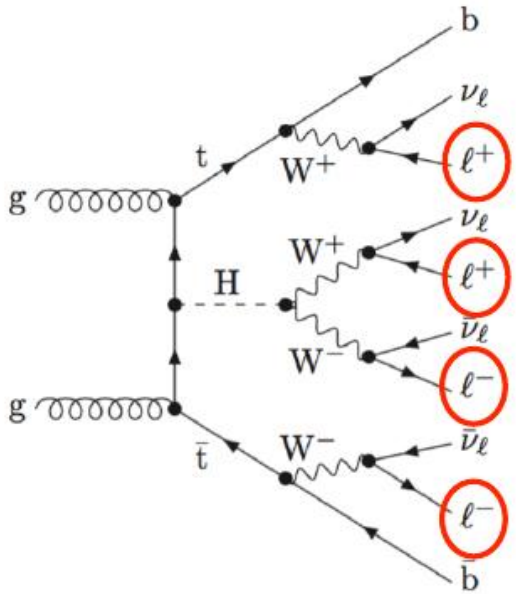
$t\bar{t}H$ multilepton

arXiv:1803.05485

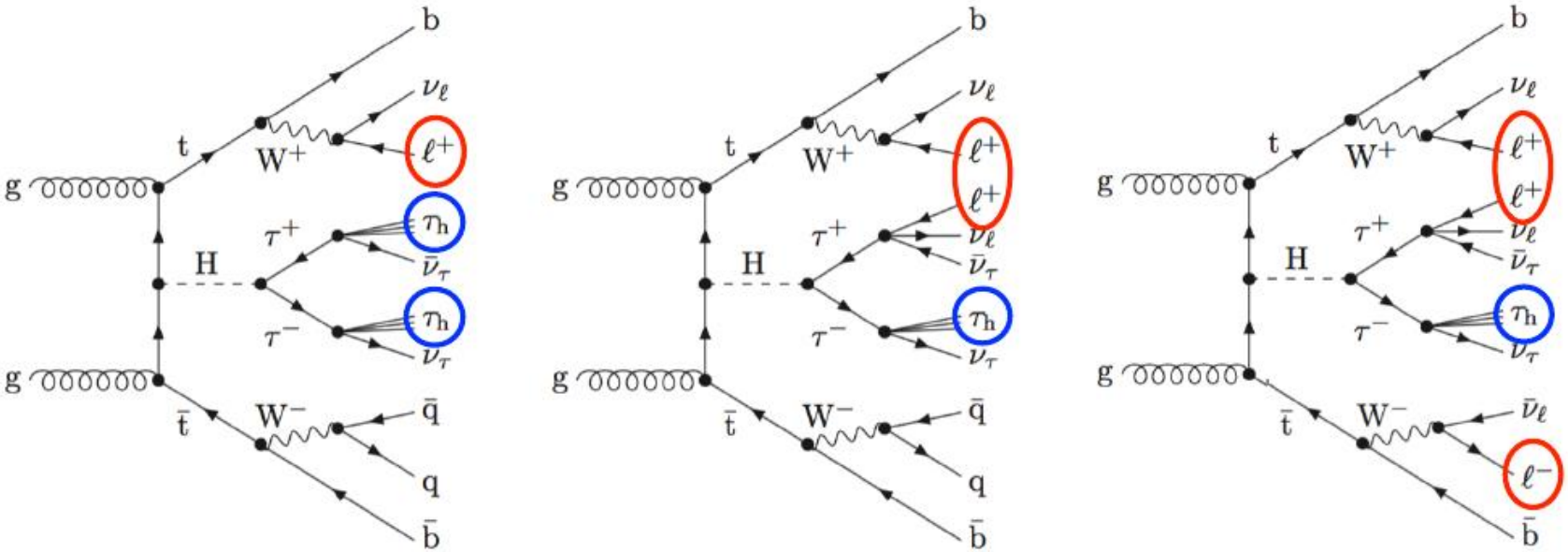
- Multilepton final states: Higgs decay to W^+W^-, ZZ, and $\tau\tau$
- Events categorized based on number of leptons and τ_h candidates







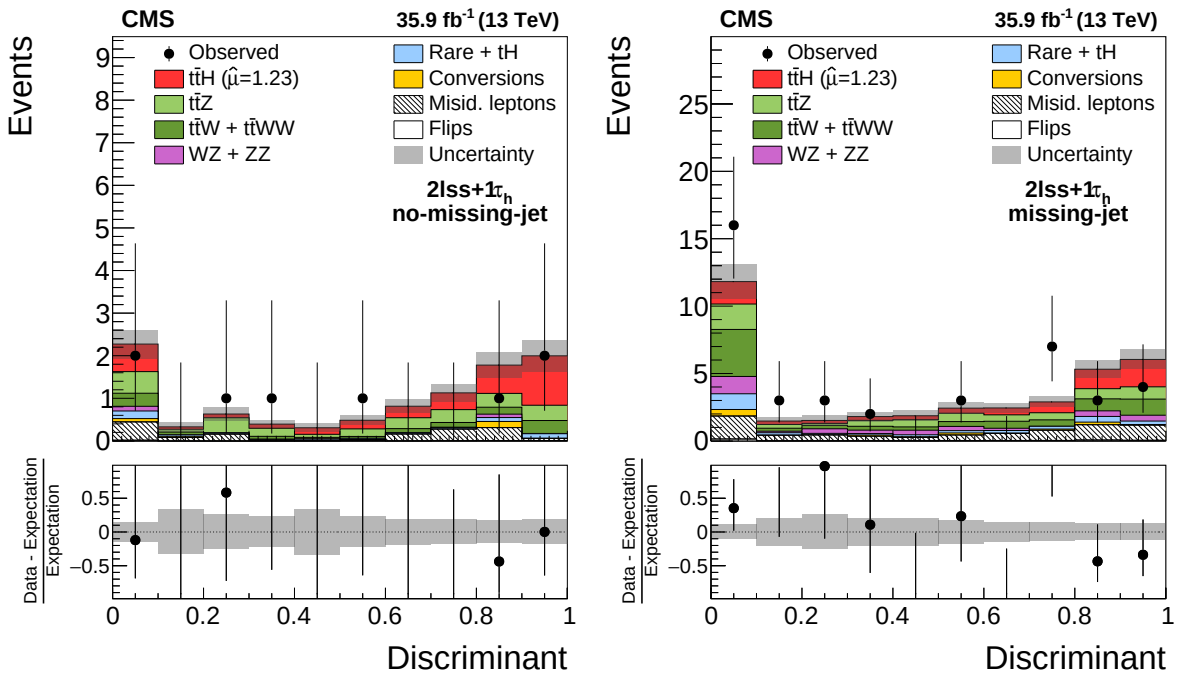
$t\bar{t}H$ multilepton: analysis strategy



- Additional requirements on jets
  - At least 2 loose or 1 medium b-tagged jets
  - At least 2 to 4 jets, depending on event category
- 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 discriminate signal from backgrounds

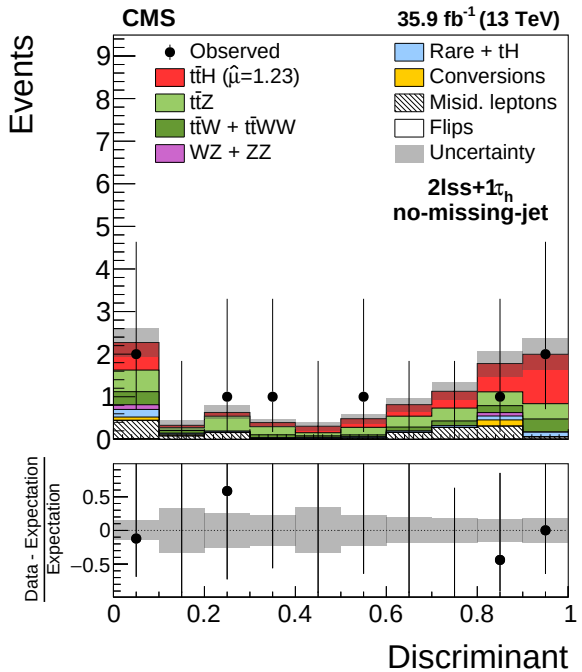
ttH multilepton: analysis strategy

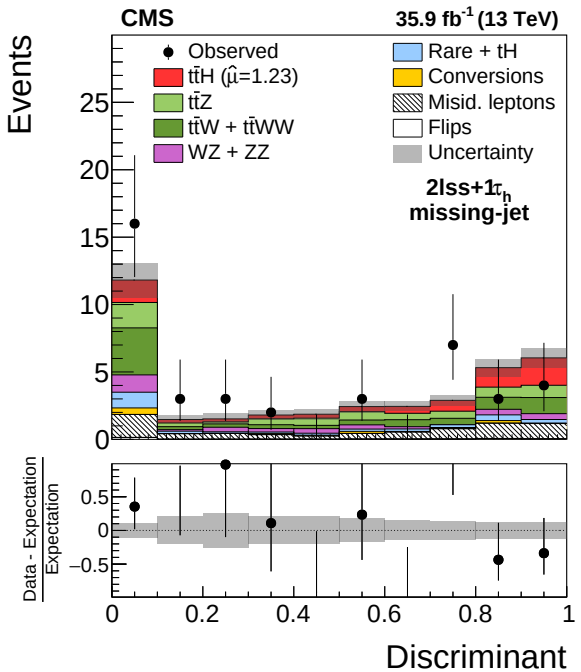
- Event categorization in lepton flavor, and b-jet multiplicity
- Discriminating variables
  - MEM against ttZ (2 leptons same-sign + 1 τ_h)

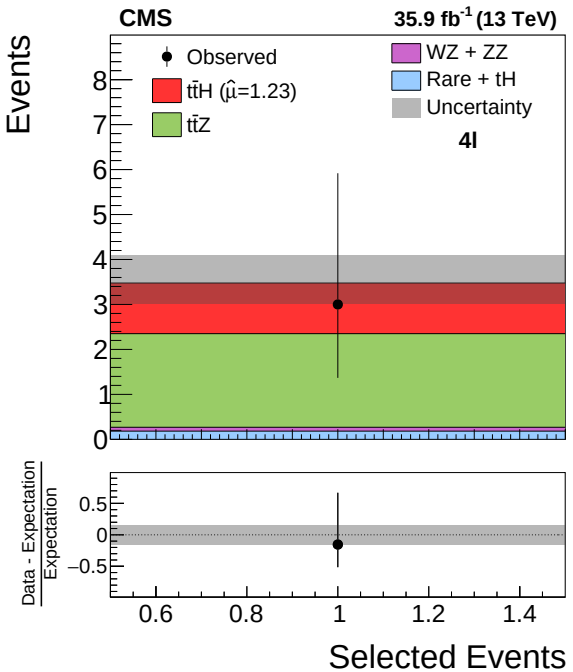


ttH multilepton: analysis strategy

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

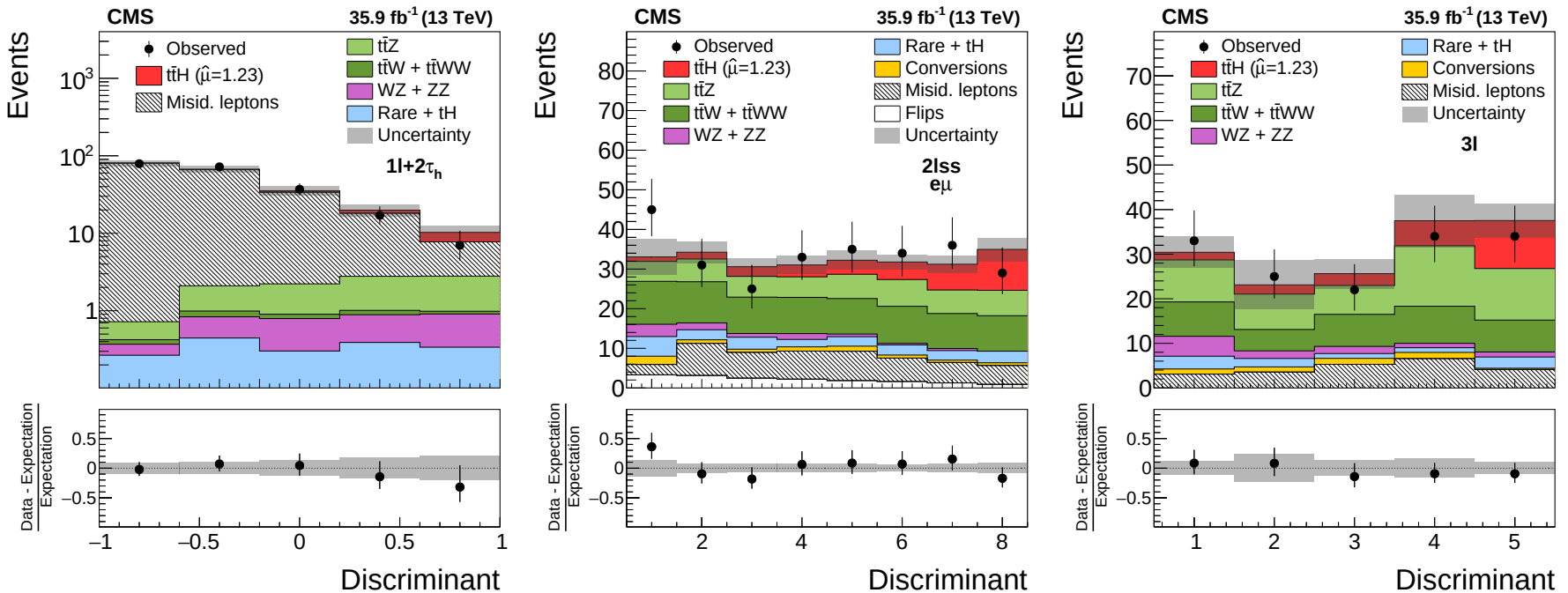




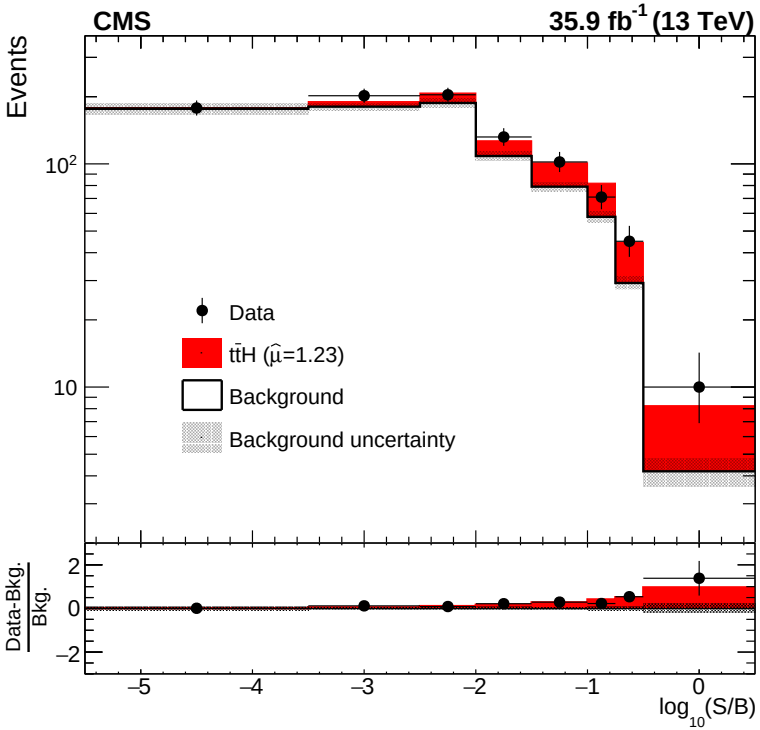


$t\bar{t}H$ multilepton: analysis strategy

- Event categorization in lepton flavor, and b-jet multiplicity
- Discriminating variables
  - MEM against $t\bar{t}Z$ (2 leptons same-sign + 1 τ_h)
  - Yield in 4-leptons (low stats.)
  - BDTs against $t\bar{t}$ + jets (1l+2 τ_h) and $t\bar{t}$ + jets, $t\bar{t}$ + V (2 leptons same-sign, 3 leptons has MEM as input)



$t\bar{t}H$ multilepton results

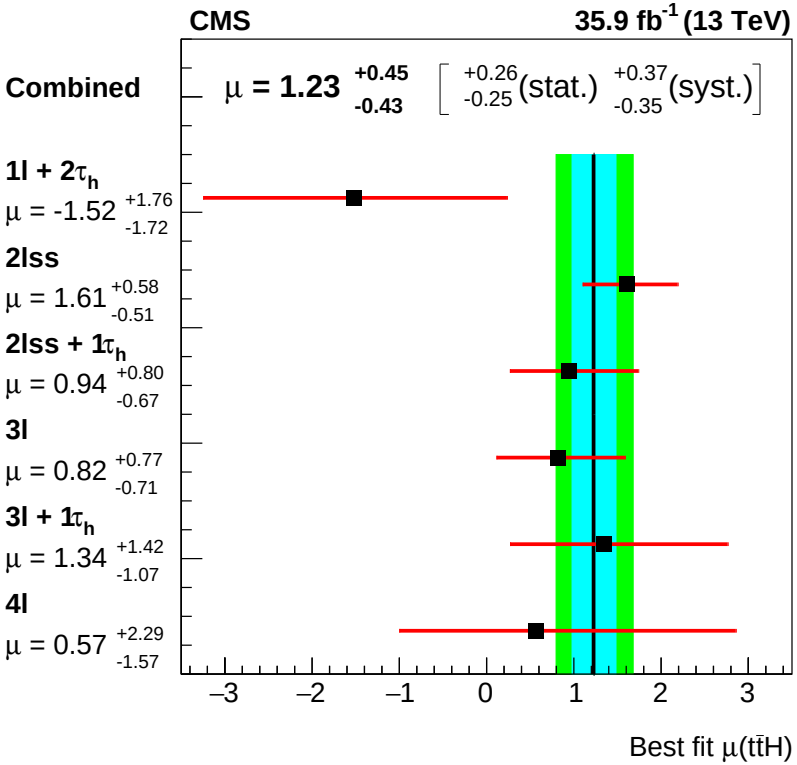


Best-fit $\mu = 1.23^{+0.45}_{-0.43}$, at 3.2 (2.8) σ obs. (exp.) significance

Cross check analysis with $t\bar{t} + V$ freely floating, normalization constrained using control regions:

$\mu = 1.04^{+0.50}_{-0.36}$, 2.7σ sig.

$t\bar{t}H$ multilepton results



Source	Unc. [%]	$\Delta\mu/\mu$ [%]
Lepton selection efficiency	2–4	11
τ_h selection efficiency	5	4.5
b tagging efficiency	2–15	6
Reducible background	10–40	11
Jet energy calibration	2–15	5
τ_h energy calibration	3	1
Theoretical sources	≈10	12
Integrated luminosity	2.5	5

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

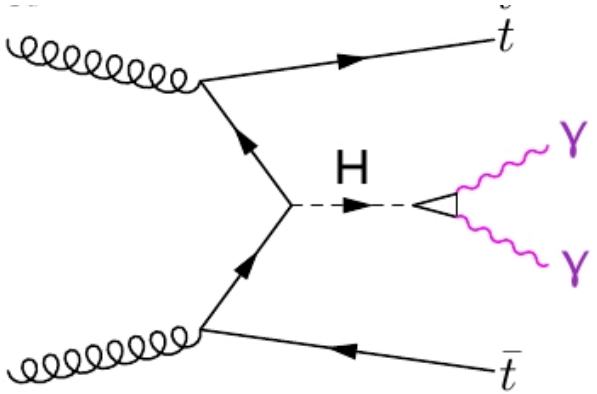
Cross check analysis with $t\bar{t} + V$ freely floating, normalization constrained using control regions:

$\mu = 1.04^{+0.50}_{-0.36}$, 2.7σ sig.

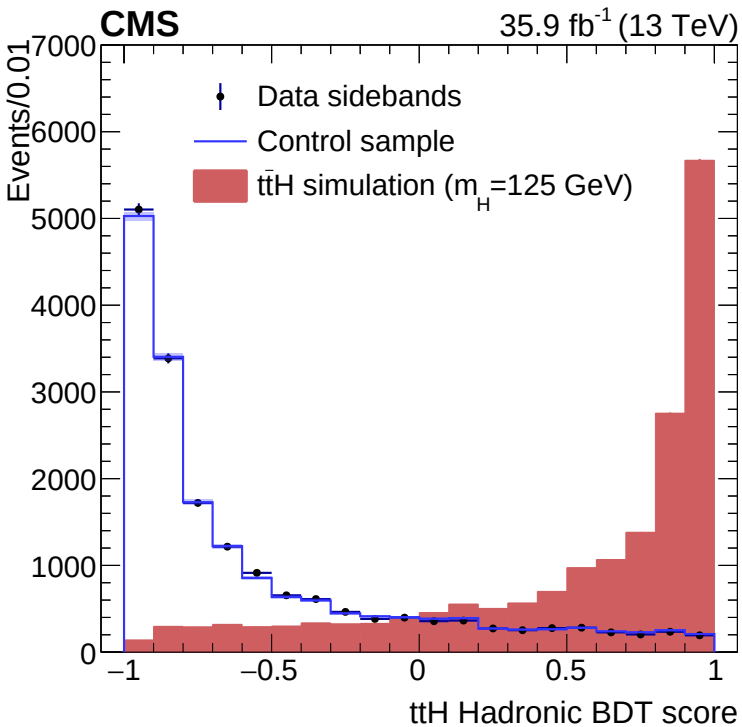
- Limited by non-prompt lepton estimation and tau identification, JES and JER, $t\bar{t}H$ and $t\bar{t} + V$ modelling
- Several channels limited by statistics

$t\bar{t}H(\gamma\gamma)$

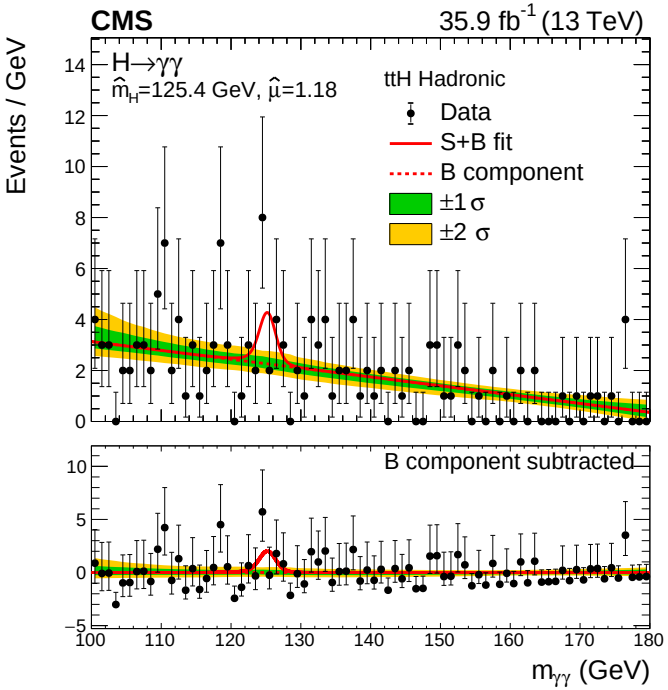
arXiv:1804.02610

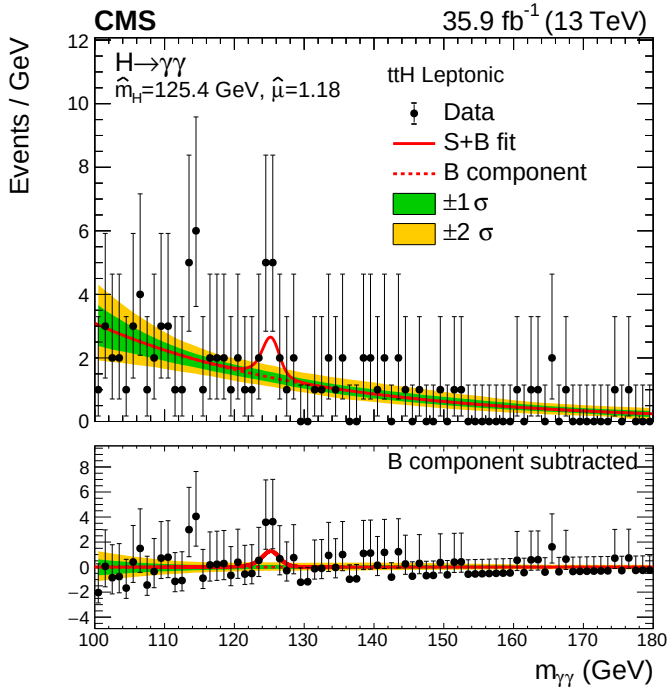


- Clear signature coming from the photons
- Dedicated $t\bar{t}H$ channel part of the global $H \to \gamma\gamma$ analysis
- $t\bar{t}$ hadronic and leptonic channels
  - Hadronic $t\bar{t}$ decay: MVA is used for background rejection
- Signal extracted from fit to $m_{\gamma\gamma}$



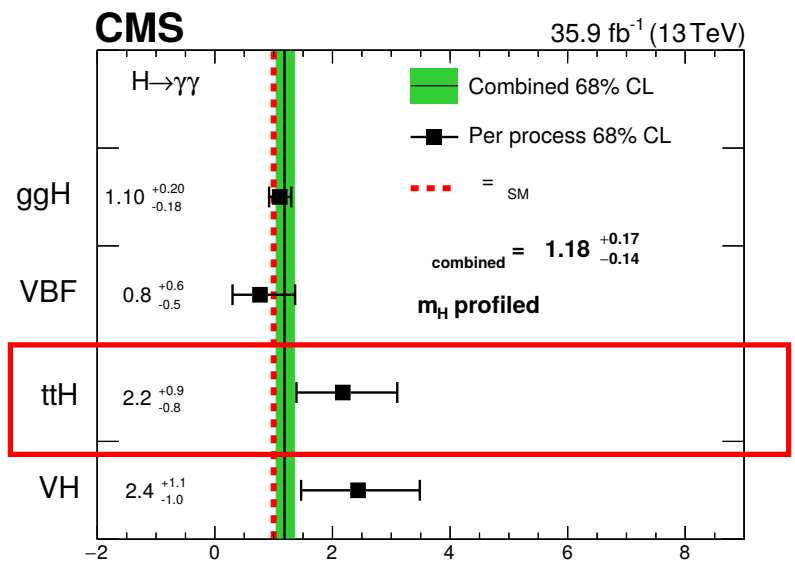
$t\bar{t}H(\gamma\gamma)$ results





- Statistically limited
- Leading systematic uncertainties:

  Photon shower shape and energy scale



$t\bar{t}H$ combination

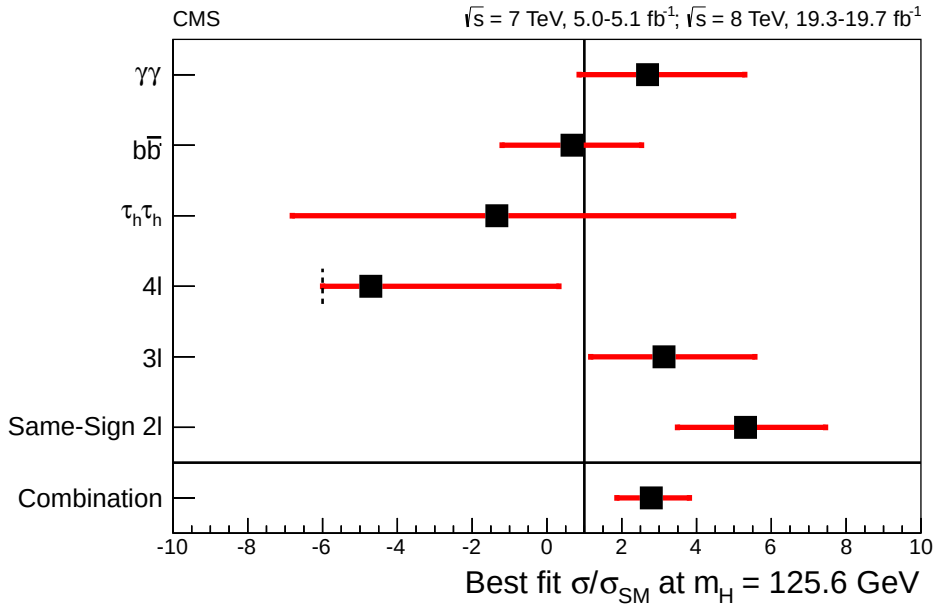
arXiv:1804.02610

Contributing analyses

- All $t\bar{t}H$ 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 $t\bar{t}H$ categories of the $H\rightarrow\gamma\gamma$ analysis

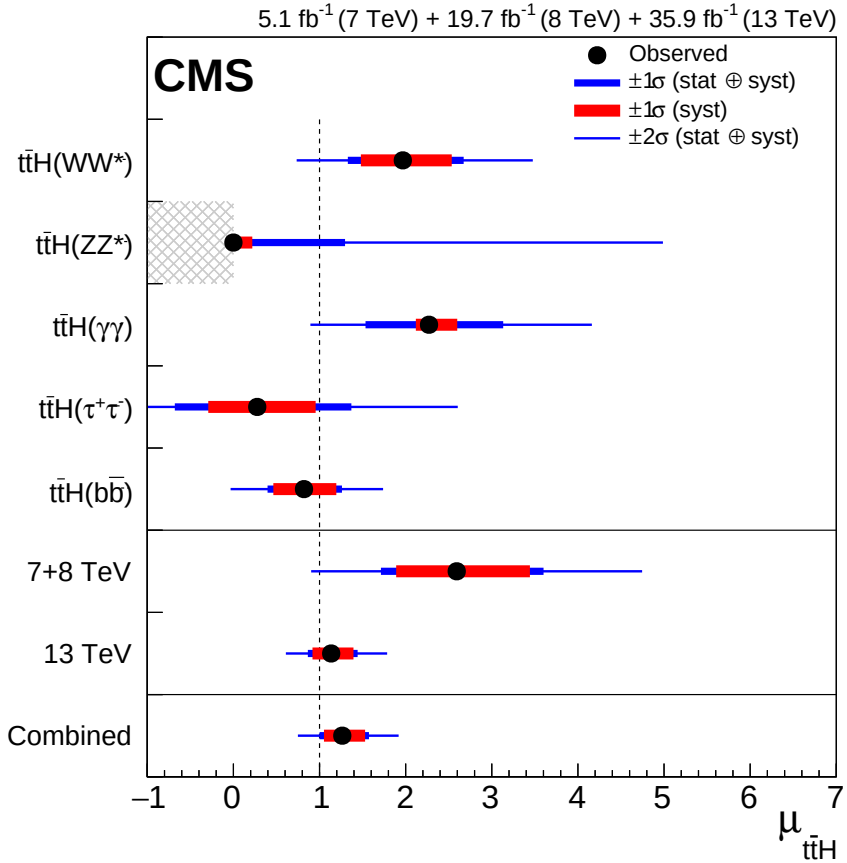


Correlations between Run-1 and Run-2 analyses

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

$t\bar{t}H$ combination

- $H \to \gamma\gamma$ and $H \to ZZ$ channels still limited by statistics
- Other channels dominated by systematics
- Signal theory mainly from inclusive $t\bar{t}H$ prediction
- Background theory mainly from $t\bar{t} + HF$ flavour prediction in $t\bar{t}H(b\bar{b})$
- Experimental: lepton efficiencies, lepton mis-id, b-tagging and MC stats all important



$$\mu_{t\bar{t}H} = 1.26^{+0.31}_{-0.26} = 1.26^{+0.16}_{-0.16}(\text{stat})^{+0.17}_{-0.15}(\text{expt})^{+0.14}_{-0.13}(\text{Th. bkg})^{+0.15}_{-0.07}(\text{Th. sig})$$

$t\bar{t}H$ combination

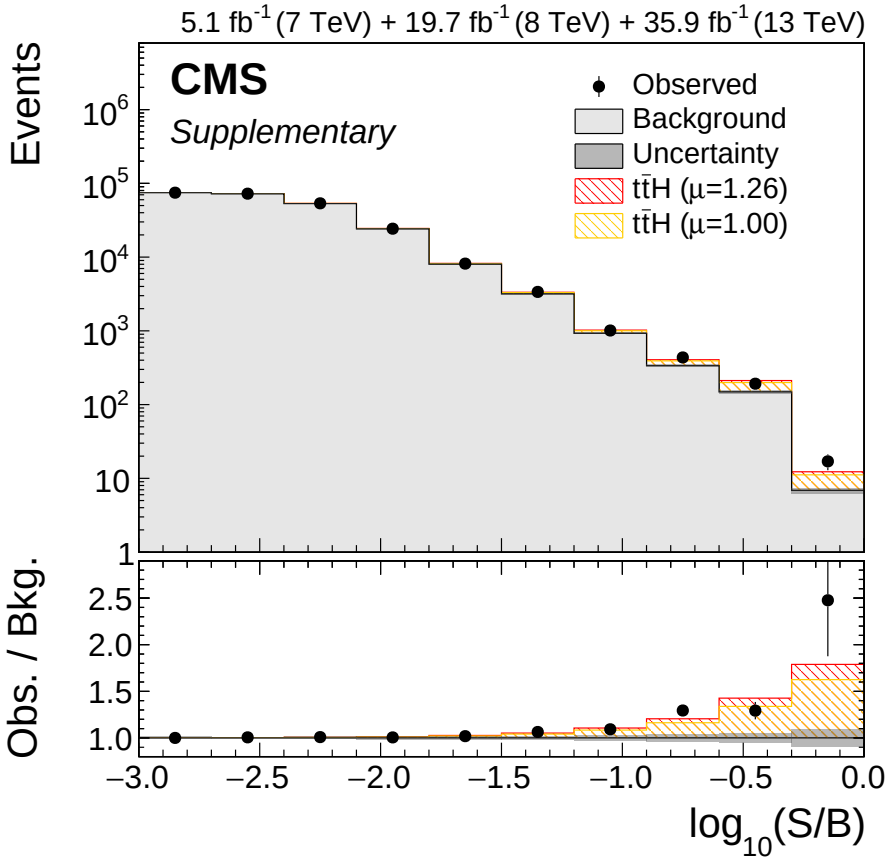
- $H \to \gamma\gamma$ and $H \to ZZ$ channels still limited by statistics
- Other channels dominated by systematics
- Signal theory mainly from inclusive $t\bar{t}H$ prediction
- Background theory mainly from $t\bar{t} + HF$ flavour prediction in $t\bar{t}H(b\bar{b})$
- Experimental: lepton efficiencies, lepton mis-id, b-tagging and MC stats all important

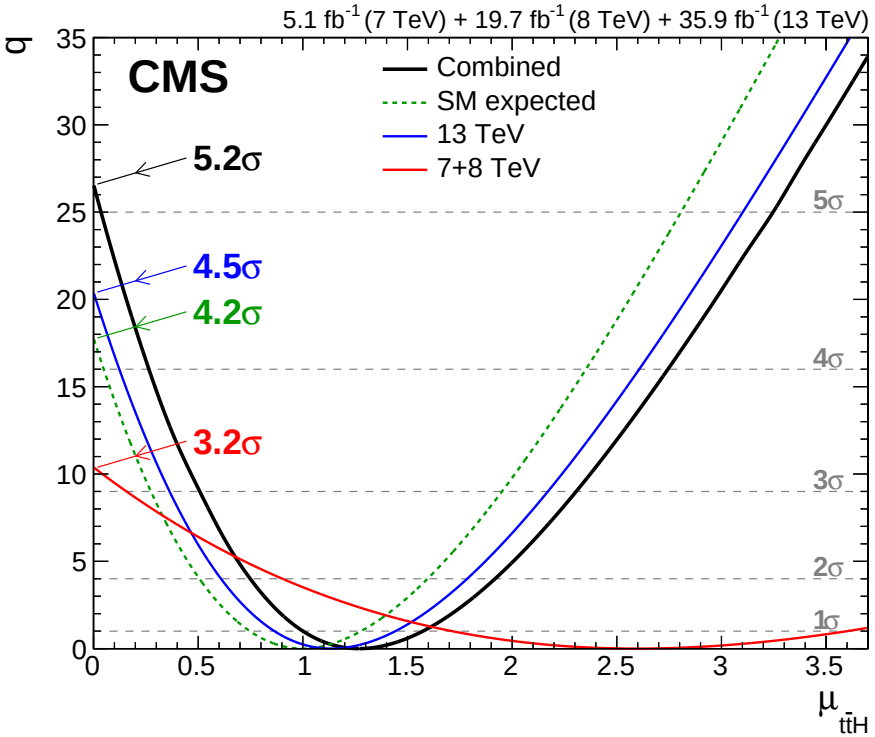
Uncertainty source	$\Delta\mu$	
Signal theory	+0.15	−0.07
Inclusive ttH normalisation (cross section and BR)	+0.15	−0.07
ttH acceptance (scale, pdf, PS and UE)	+0.004	−0.004
Other Higgs boson production modes	+0.002	−0.003
Background theory	+0.14	−0.13
tt + bb/cc prediction	+0.13	−0.11
tt + V(V) prediction	+0.06	−0.06
Other background uncertainties	+0.03	−0.03
Experimental	+0.17	−0.15
Lepton (inc. τ_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 τ_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

$$\mu_{t\bar{t}H} = 1.26^{+0.31}_{-0.26} = 1.26^{+0.16}_{-0.16}(\text{stat})^{+0.17}_{-0.15}(\text{expt})^{+0.14}_{-0.13}(\text{Th. bkg})^{+0.15}_{-0.07}(\text{Th. sig})$$

$t\bar{t}H$ combination: observation

- Observed significance is 5.2σ (4.2σ exp.) with respect to the $\mu_{t\bar{t}H} = 0$ hypothesis
- First observation of the $t\bar{t}H$ production process

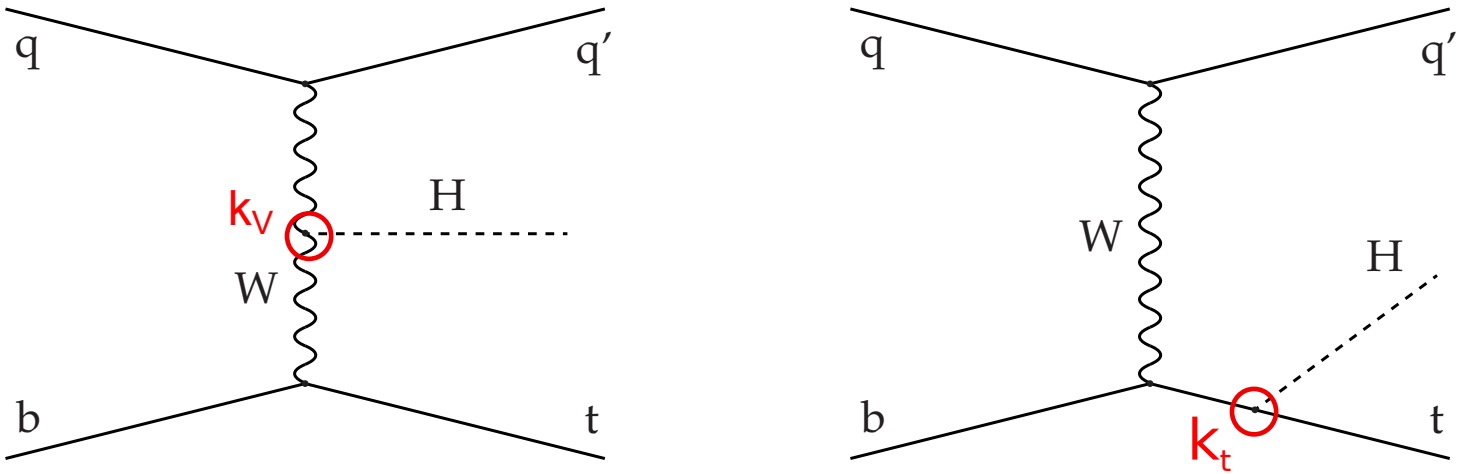




Search for tHq

CMS-HIG-17-005

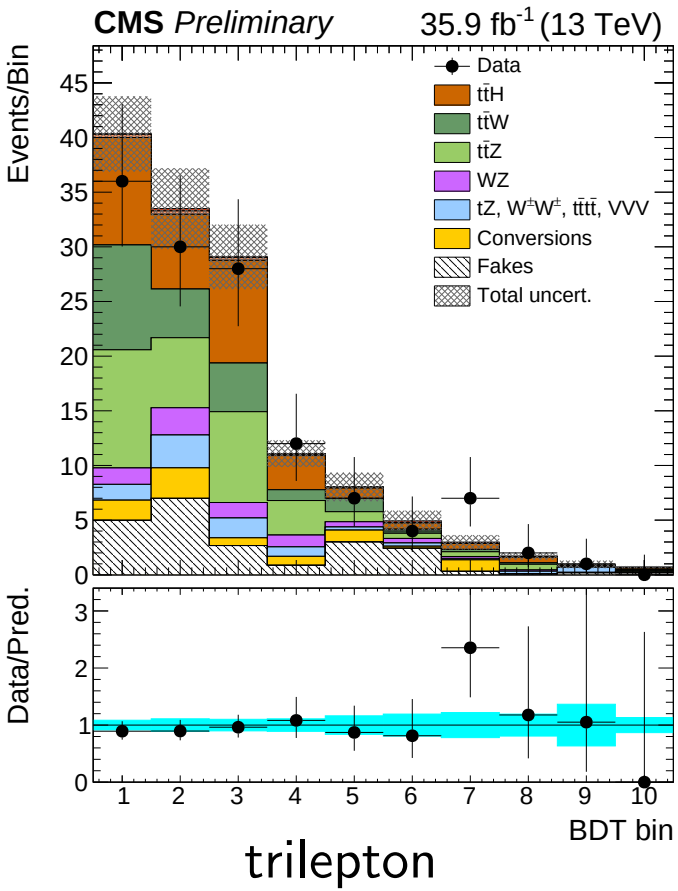
- Study of tHq process exposes the relative sign of top-Higgs and W-Higgs couplings via interference ($\sigma_{tHq}^{SM} \approx 71fb$, $\sigma_{tHq}^{\kappa_t=-1} \approx 790fb$)



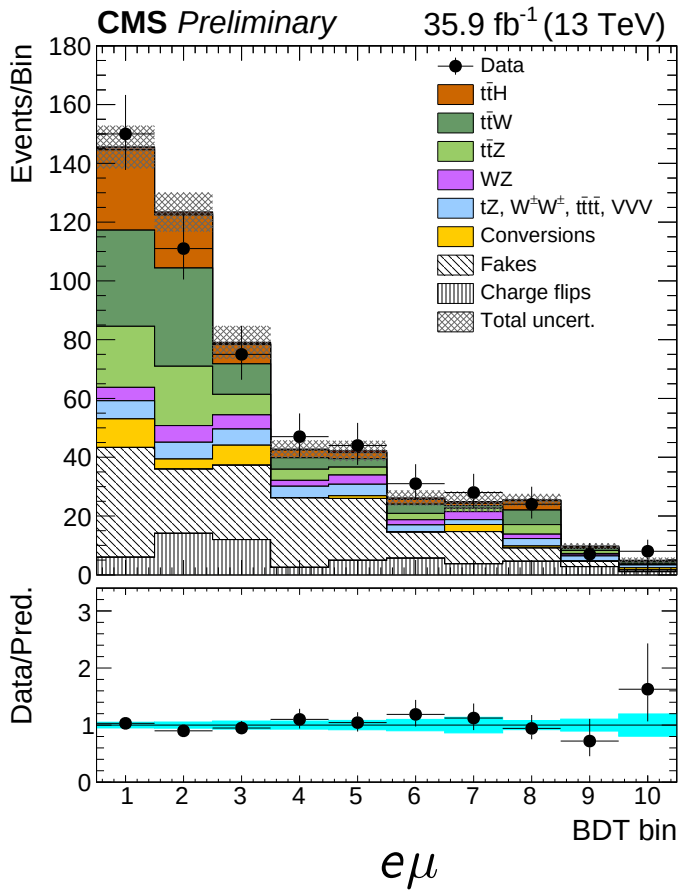
- Studied using similar strategies as for $t\bar{t}H$
- Channels:
  - Same-sign dilepton ($\mu\mu$, $e\mu$): one W from Higgs decays hadronically, others decay leptonically.
  - Trilepton: All three Ws decay leptonically
- Background sources
  - Irreducible ($t\bar{t}$ +X, diboson, rare) from MC
  - Reducible (fakes) estimated from data

Search for tHq: analysis strategy

- Two separate BDT trainings using MC samples for signal (tHq with $\kappa_t = -1$, $\kappa_V = -1$) and backgrounds
  - Against $t\bar{t}$: non-prompt lepton type background
  - Against combined $t\bar{t}Z$ and $t\bar{t}W$: prompt lepton type background



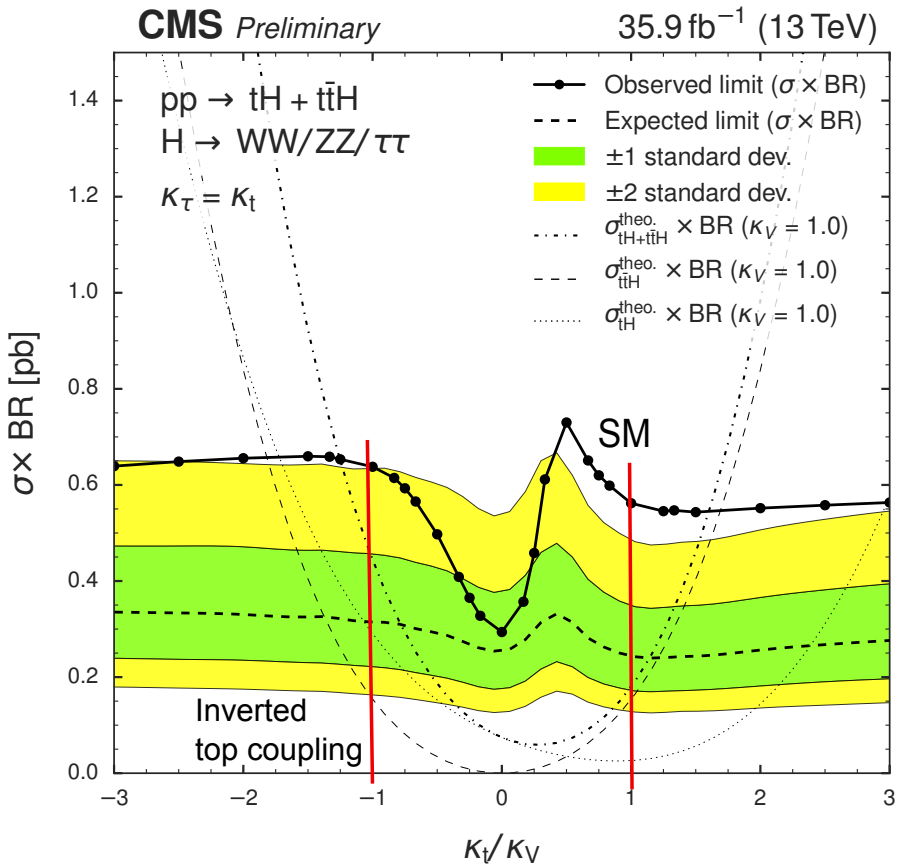
trilepton



$e\mu$

Search for tHq: results

- Limit on common signal strength for (tHq/W+t$\bar{t}$H) as function of κ_t/κ_V
  t$\bar{t}$H is included as signal since its cross section varies as κ_t^2
- Enhanced production cross section in the case of anomalous top-Higgs couplings



κ_t values outside (-1.25, 1.6) are excluded at 95% C.L. for $\kappa_V = 1$

Summary

- Results presented for $t\bar{t}H$ 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 hadronic decaying τ leptons
  - Several channels already systematic limited (with $\sim$30% of the expected Run-2 lumi)
  - Work ongoing to reduce limitations of these analyses:
    Reduce impact of systematic uncertainties like signal and background modeling ($t\bar{t} + b\bar{b}$, $t\bar{t} + V$), improve non-prompt lepton estimation, jet flavor tagging...
- Observation of $t\bar{t}H$ 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

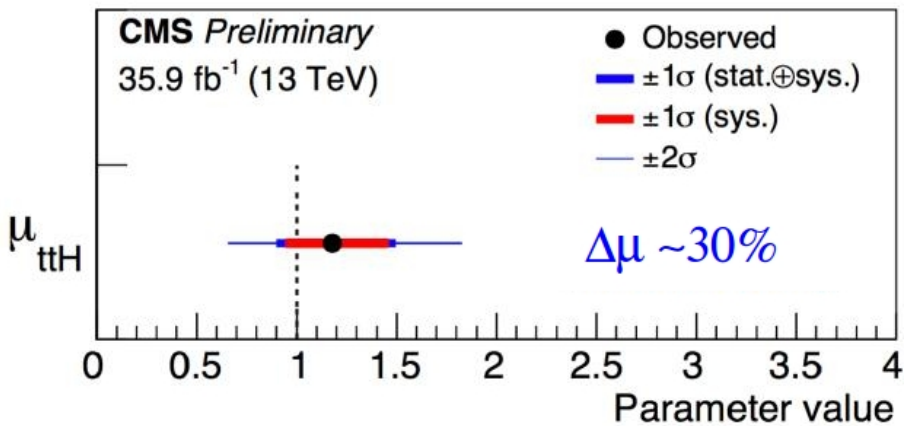
Extras

Combination with other Higgs meas.

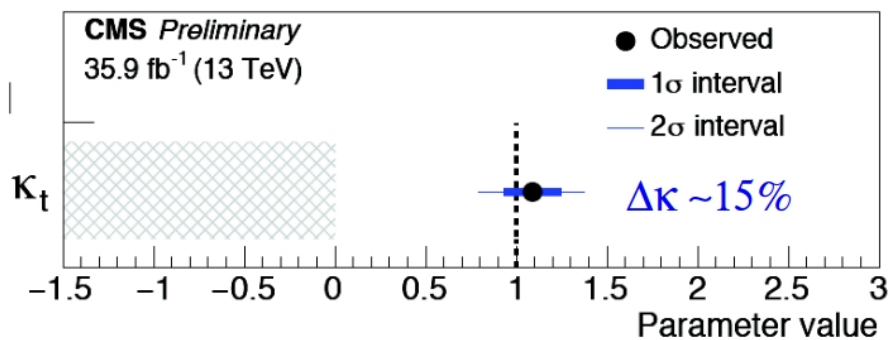
CMS-HIG-17-031

Combination of $t\bar{t}H$ analyses, along with other Higgs measurements, for 13 TeV data

- $t\bar{t}H$ +tH production cross section modifier from per-production mode fit (other production modes floating)
- Top coupling modifier from κ-framework fit with effective loops



ttH			
Best fit value		Uncertainty Stat.	Syst.
1.18	$^{+0.31}_{-0.27}$	$^{+0.16}_{-0.16}$	$^{+0.26}_{-0.21}$



		Uncertainty	
Best fit		Stat.	Syst.
1.09	$^{+0.14}_{-0.14}$	$^{+0.08}_{-0.08}$	$^{+0.12}_{-0.12}$