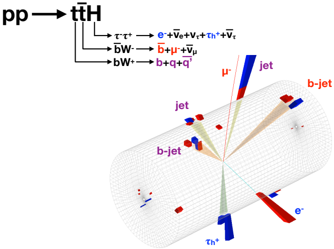
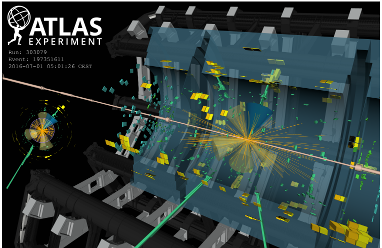
Observation of Top Quark Pair Production in Association with a Higgs Boson

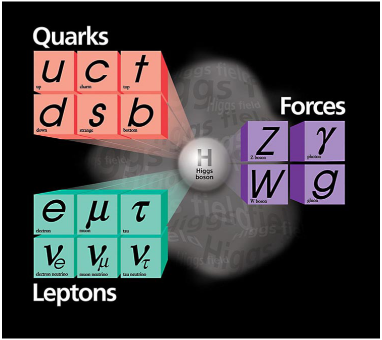
Carmen Diez Pardos (DESY)

6 June 2018





The fundamental building blocks of matter



- SM: Successful description of elementary particles and interactions
- LHC experiments discovered a new Higgs-like boson (m_H=125GeV)
- Candidate to close the long-standing puzzle of how elementary particles acquire mass in the SM
- But does it behave like the SM Higgs?

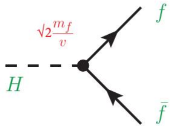
- Higgs boson: production and decay rates consistent with SM expectations
- Broad programme to measure properties
  - Confirm yet-unobserved processes
  - Search for deviations from SM expectation

The top quark and the Higgs boson

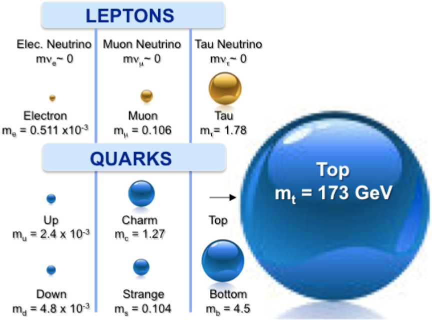
In the SM, elementary particles acquire mass via their interaction with the Higgs field

- Higgs coupling to the fermions (Yukawa coupling): proportional to fermion mass
- Top quark: most massive known particle → most strongly-coupled SM fermion (y_t ∼1)

→ Essential to study Higgs properties, measure the coupling



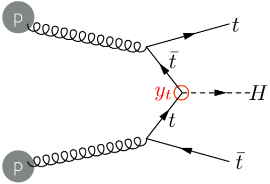
- Several open questions
    - Is the mass of the top quark generated by the Higgs mechanism?
    - Role in electroweak symmetry breaking?



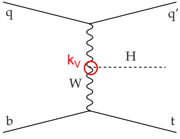
Top-Higgs coupling: the hunt for t$\bar{t}$H

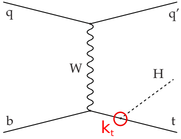
Best direct probe of the top-Higgs Yukawa coupling, vital step towards verifying the SM nature of the Higgs boson

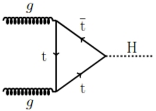
- Direct measurement of y_t in t$\bar{t}$H production:
  - gluon-gluon fusion: assumes no BSM coupling
  - Allows probing new physics in gg→H and H→ $\gamma\gamma$ effective vertices

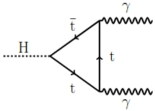


- y_t in tH production: access to sign of the coupling

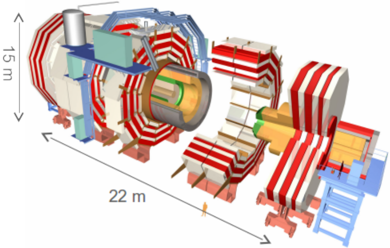


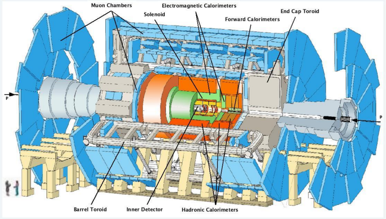


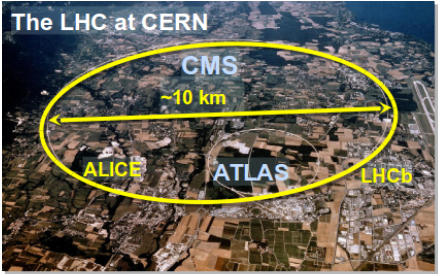




Where to look for it? The Large Hadron Collider







- proton-proton collision energies ($\sqrt{s}$)
- Run-1: 7 & 8 TeV, 25 fb^{-1}- stat. limited
- Run-2: 13 TeV, already $\sim$100 fb^{-1}

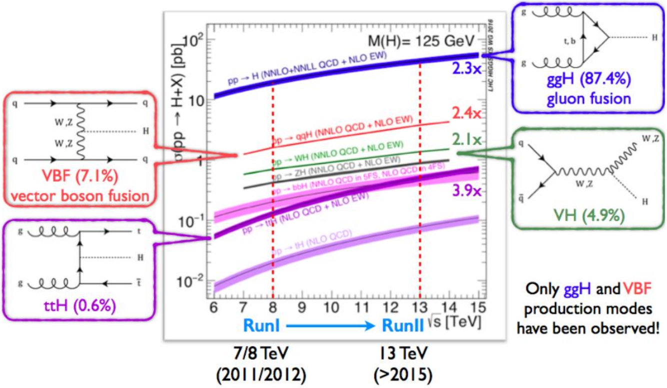
  Expected 75,000 $t\bar{t}H$ events at the end of this year

This presentation focuses on results with 13 TeV data (up to 80 fb^{-1}) + combination

$t\bar{t}H$ production

$\sigma_H \approx 0.5$ pb at $\sqrt{s}=13$TeV ($m_H=125$GeV)

- → Only 1% of total Higgs cross section
- → Larger increase in signal than backgrounds from 8 to 13 TeV
- → By this year up to 6 times more data

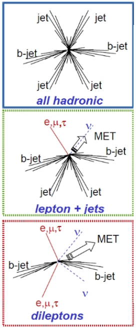


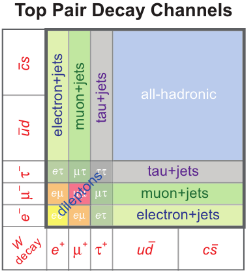
- $t\bar{t}H$ decay yields (very) complex final states, with many objects
- Crucial to understand the backgrounds (eg. $\sigma_{t\bar{t}} \approx 830$ pb @13 TeV)
- Large irreducible backgrounds: $t\bar{t}$+X
  (X = $b\bar{b}$, W, Z)

Top quark × Higgs decay channels

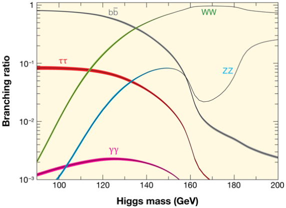
- Exploiting all $t\bar{t}$ decay channels and Higgs decays to
  - bottom quarks → Large BR, large background contributions
  - W, Z bosons, taus → smaller production rate, lower backgrounds
  - photons → clean final state, very small rate

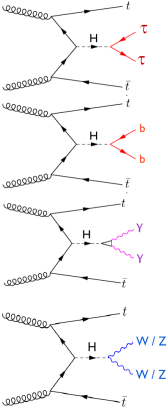
In the SM $\mathrm{t} \rightarrow \mathrm{Wb}$ almost 100%, W decay defines final state





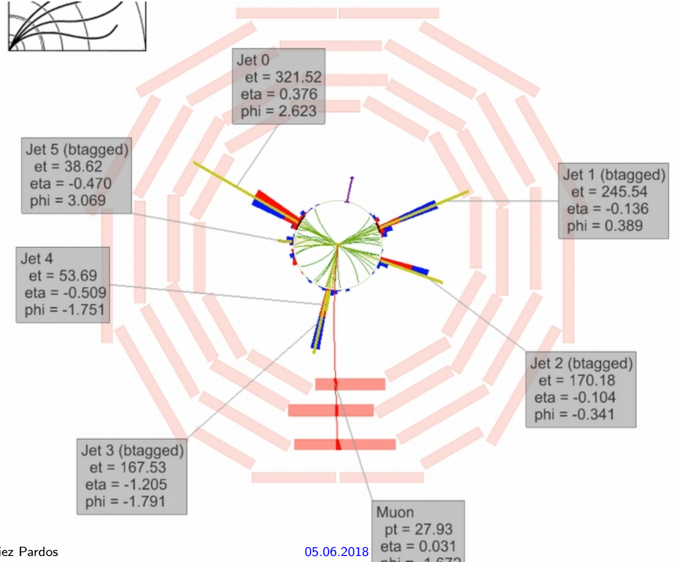
×



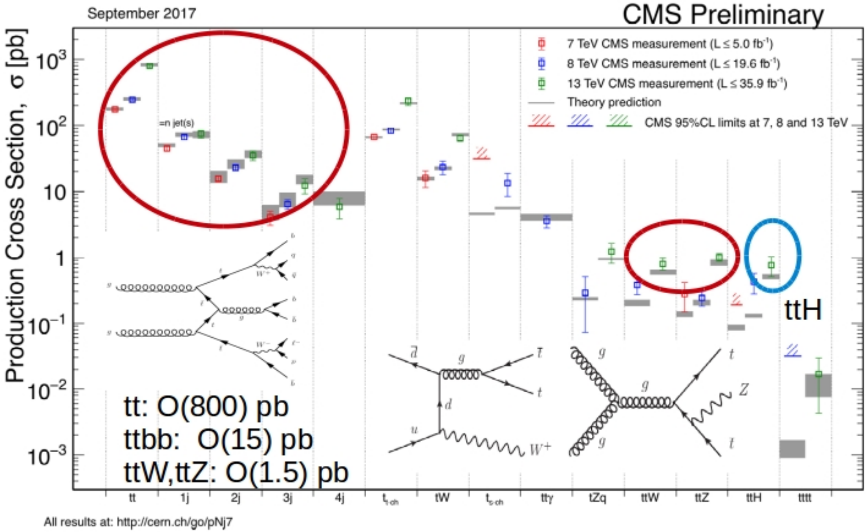


Complex final states

- Complex final states, with many objects: leptons, jets, taus
- Large combinatorics of leptons and jets from top quark decays



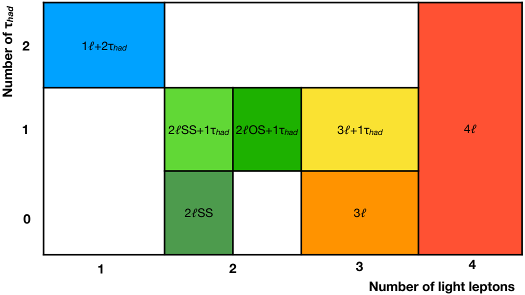
Challenging backgrounds

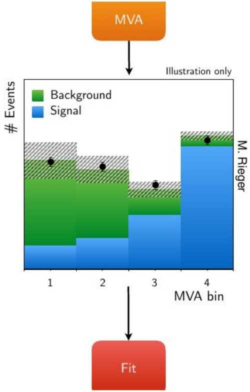


+ diboson production (WW, ZZ), QCD multijets...

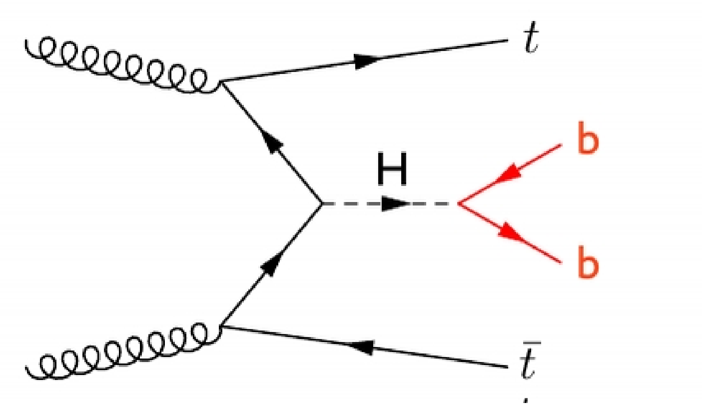
Sophisticated analysis strategies

- $t\bar{t}$ like selections with additional searches for Higgs decay products
- Event categorization based on top quark (W boson) and Higgs decay modes
- Multivariate analysis (MVA) techniques, eg. boosted decision trees (BDT) or deep neural networks (DNN), Matrix-Element-Methods (MEM) used to extract signal, boosted-object reconstruction
- Profile likelihood fits across all categories to extract the signal



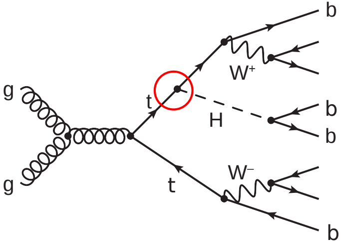


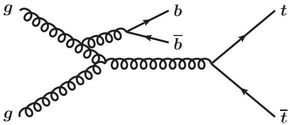
$t\bar{t}H(b\bar{b})$



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

- Large $\mathcal{B}(H \rightarrow b\bar{b})$, access coupling 3rd generation quarks
- Challenging final state
  - Huge combinatorics in event reconstruction
  - Poor $H \rightarrow b\bar{b}$ 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

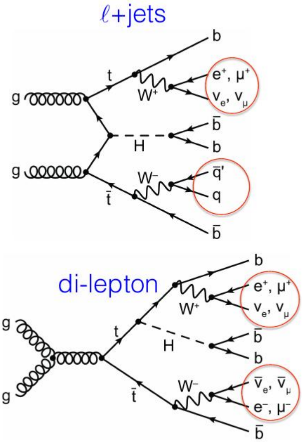




$t\bar{t}H(b\bar{b})$ Leptonic

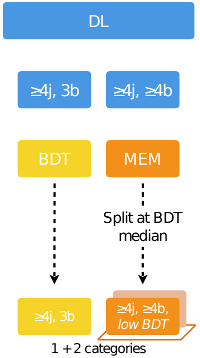
arXiv:1804.03682, PhysRevD.97.072016

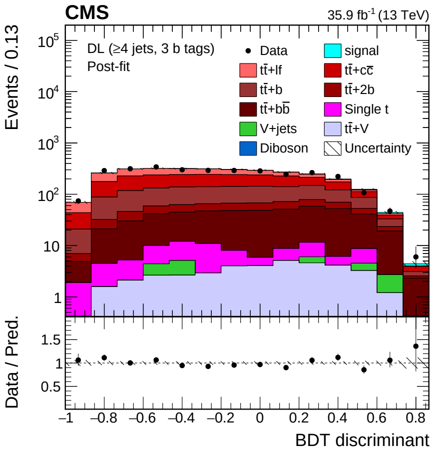
- Events with exactly 1 (2) leptons (e, μ)
- At least 3 (4) jets, with at least 1 (3) b-tagged
- Create categories enriched in signal and background events
- Exploiting MEM and MVA and boosted topologies to discriminate signal from background



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

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

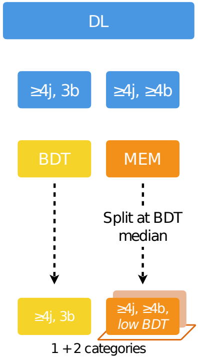


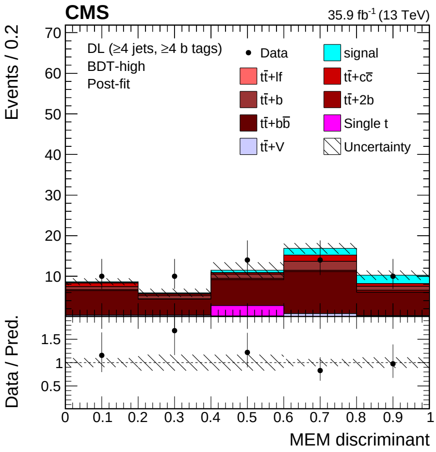


- $\geq$ 4j, 3b: BDT separating signal and inclusive $t\bar{t}$ + jets background as final discriminant

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

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

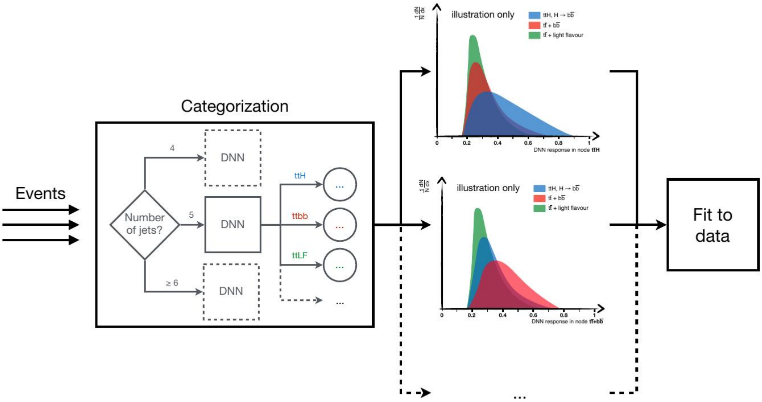




- $\geq$ 4j, $\geq$ 4b: low/high BDT sub-categories + MEM separating against $t\bar{t} + b\bar{b}$ background as final discriminant

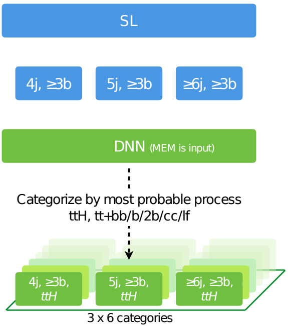
$t\bar{t}H(b\bar{b})$ Leptonic: **lepton+jets** $t\bar{t}$ channel (CMS)

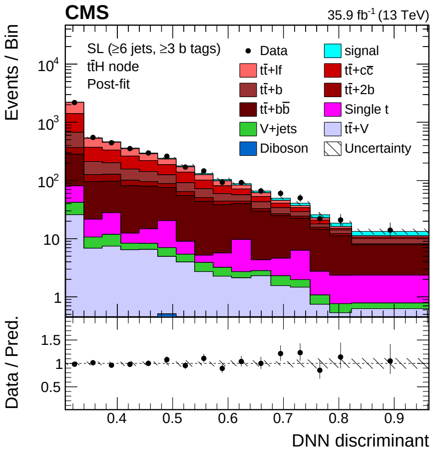
- Search in single-lepton $t\bar{t}$ channel
- Deep Neural Network per jet category & most probable process: multi-classification as signal or any of 5 $t\bar{t}$ + jets bkgs. ($t\bar{t} + b\bar{b}$, $t\bar{t} + 2b$, $t\bar{t} + b$, $t\bar{t} + c\bar{c}$, $t\bar{t} + LF$)
- Output of categorization yields powerful discriminators
  → One for each process vs all other processes



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

- Events categorised by number of jets: 4, 5, ≥ 6

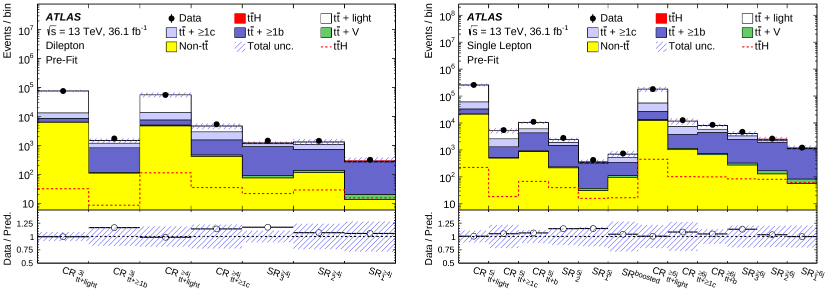


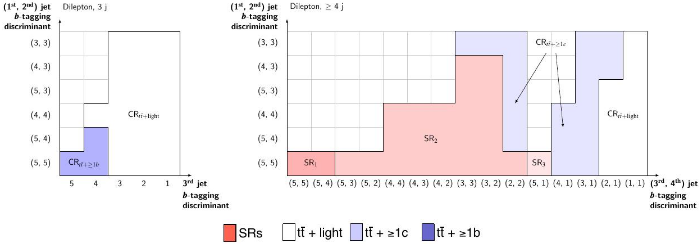


- Final discriminant: DNN output of chosen process node

tt̄H(bb̄) Leptonic: analysis strategy (ATLAS)

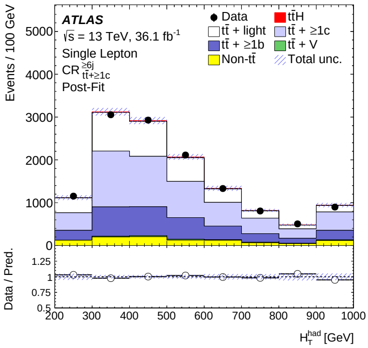
- Events categorised by number of leptons, jets, and b-tagging discriminant





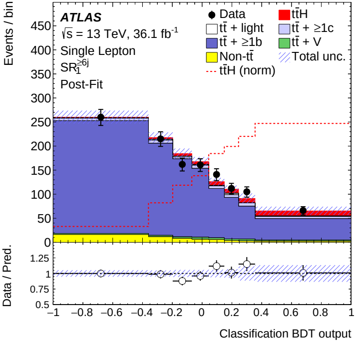
$t\bar{t}H(b\bar{b})$ Leptonic: analysis strategy (ATLAS)

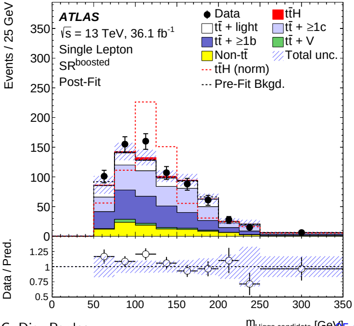
- 10 control regions to constrain different backgrounds: H_T distribution or yields
- 9 signal regions: BDT as final discriminant, with inputs



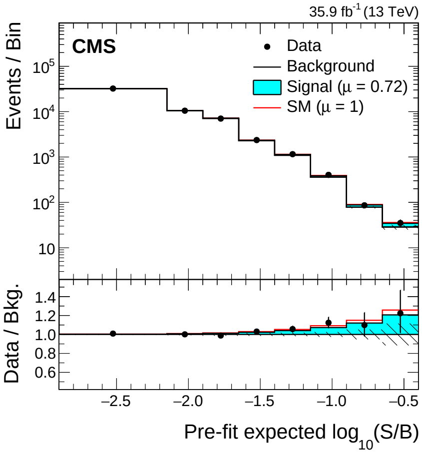
tt̄H(bb̄) Leptonic: analysis strategy (ATLAS)

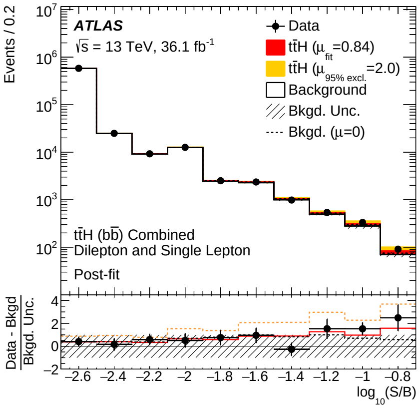
- 10 control regions to constrain different backgrounds: H_T distribution or yields
- 9 signal regions: BDT as final discriminant, with inputs
  - MEM
  - Likelihood discriminant: tt̄H against tt̄ + bb̄
  - Event reconstruction techniques: BDT to reconstruct tt̄H system, Boosted-object techniques





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

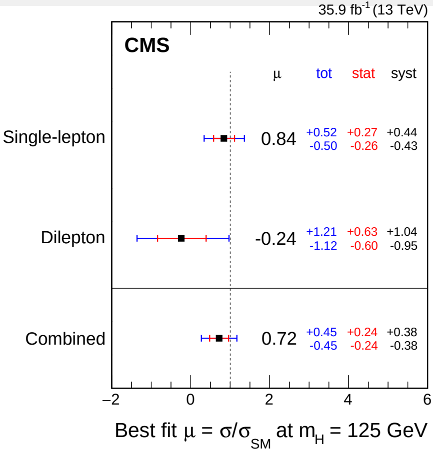


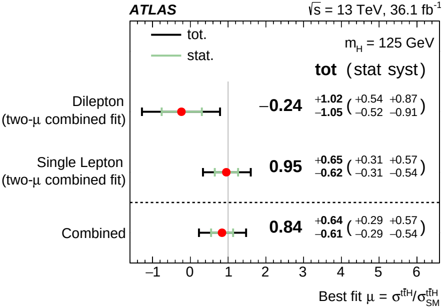


Best-fit $\mu = 0.72^{+0.45}_{-0.45}$, at 1.6 (2.2) σ obs. (exp.) significance

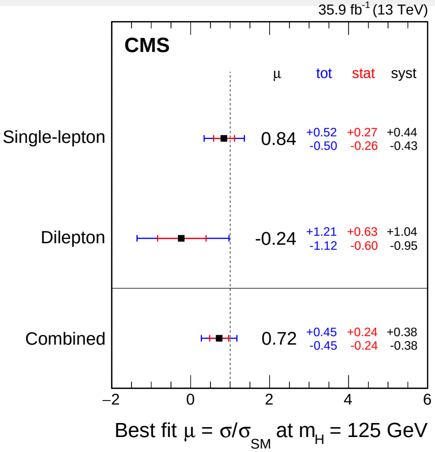
Best-fit $\mu = 0.84^{+0.64}_{-0.61}$, at 1.4 (1.6) σ obs. (exp.) significance

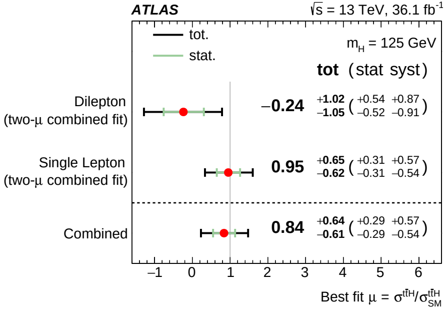
$t\bar{t}H(b\bar{b})$ Leptonic: Results





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



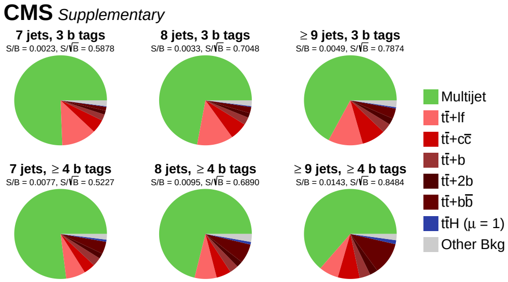


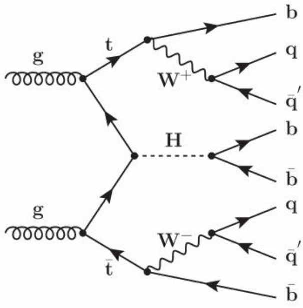
- Uncertainty on $t\bar{t}$ + heavy flavour largest impact
- Statistical uncertainty of MC
- Experimentally limited by b-tagging uncertainties

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

arXiv:1803.06986

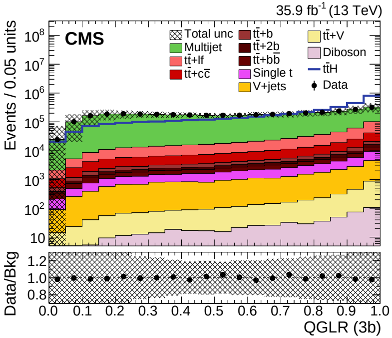
- Challenge:
  - Large backgrounds from QCD multijets, $t\bar{t}$ + jets, and the irreducible $t\bar{t} + b\bar{b}$
- Larger signal contribution
- Possibility to fully reconstruct the event





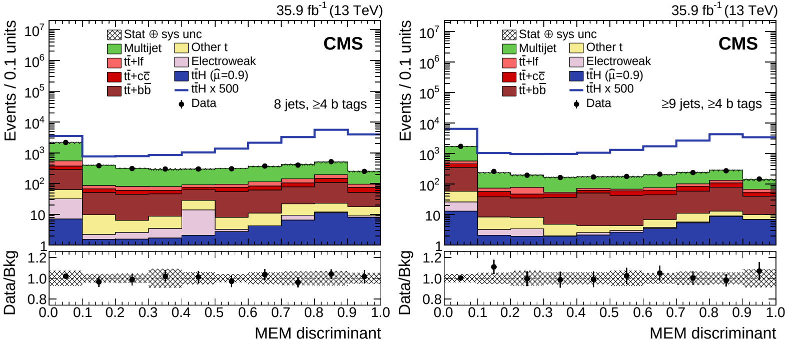
tt̄H(bb̄) Hadronic: Analysis strategy

- ≥ 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
- A quark-gluon discriminant is used to differentiate quarks jets from gluon jets
    - Shape from low b-tag multiplicity control region in data
    - Rate from final fit to data

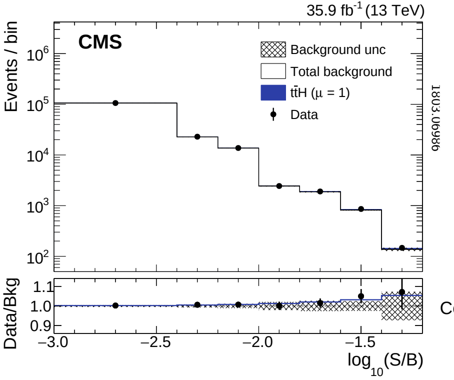


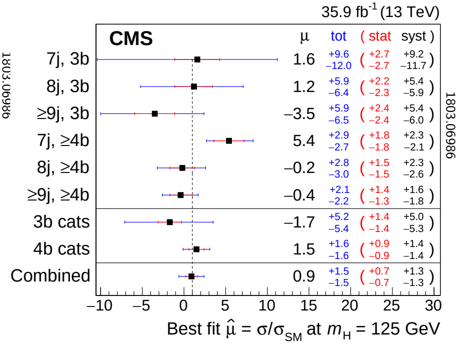
$t\bar{t}H(b\bar{b})$ Hadronic: Analysis strategy

- 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



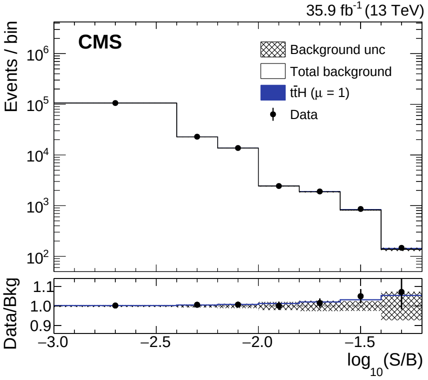
$t\bar{t}H(b\bar{b})$ Hadronic: Results

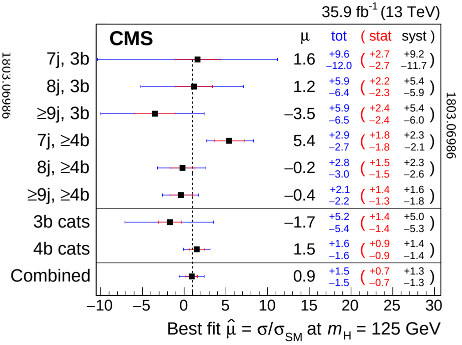




Best-fit $\mu = 0.9^{+1.5}_{-1.5}$, upper 95% C.L. limit 3.8 (3.1) obs. (exp.) × SM

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

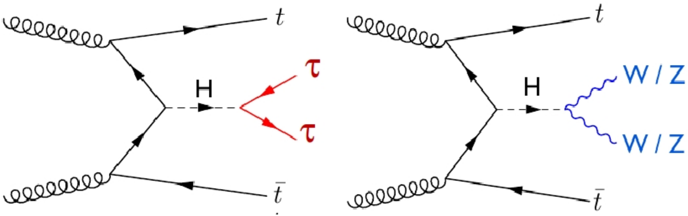




Best-fit $\mu = 0.9^{+1.5}_{-1.5}$, upper 95% C.L. limit 3.8 (3.1) obs. (exp.) $\times$ SM

- Major systematic uncertainties: Multijet estimation, $t\bar{t}$ + HF prediction, b-tagging and JES etc.

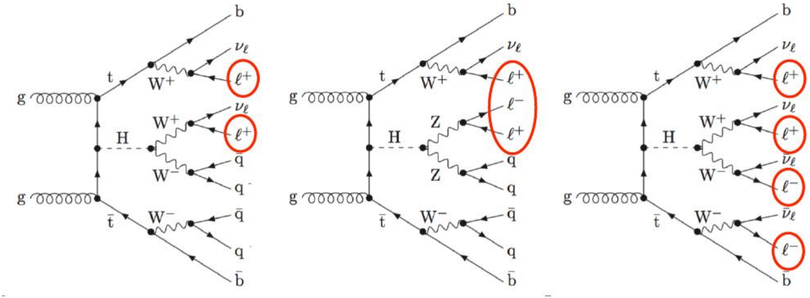
$t\bar{t}H \to \tau\tau$, $t\bar{t}H \to ZZ^*$, $t\bar{t}H \to WW^*$



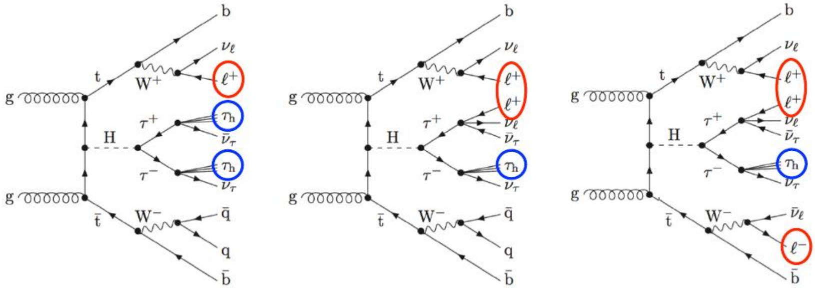
$t\bar{t}H$ multilepton

arXiv:1803.05485, PRD 97 (2018) 072003

- 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

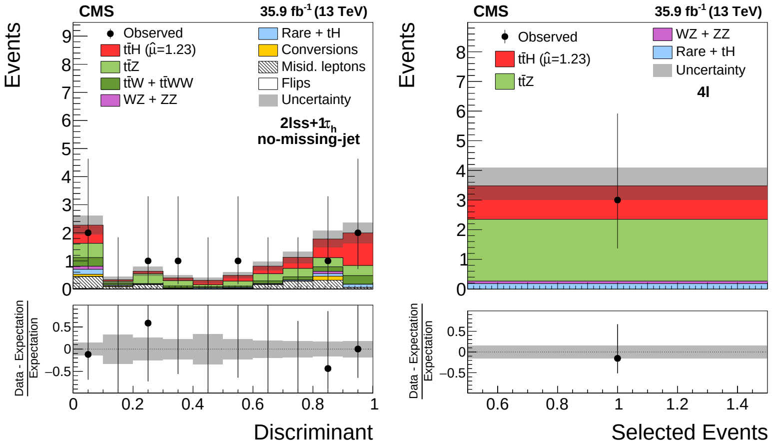


ATLAS: also 2 leptons OS + 1 τ_h

- Additional requirements on jets and 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

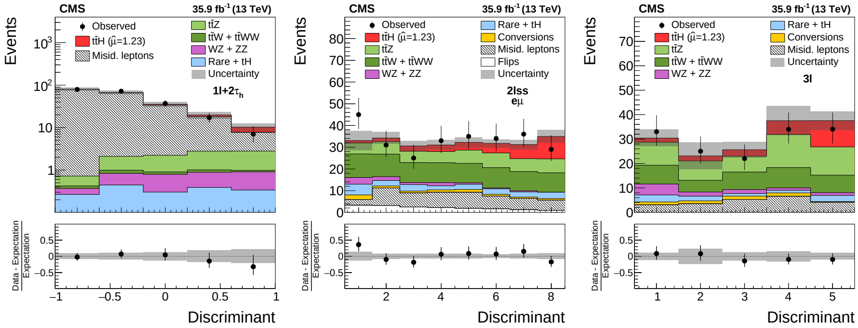
tt̄H multilepton: analysis strategy (CMS)

- Event categorization 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.)



$t\bar{t}H$ multilepton: analysis strategy (CMS)

- 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)

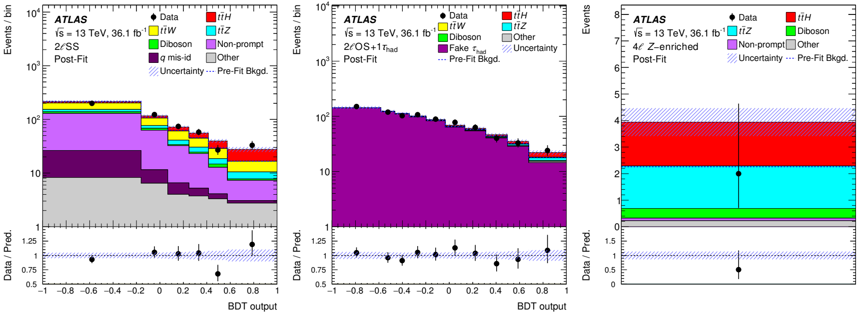


$t\bar{t}H$ multilepton: analysis strategy (ATLAS)

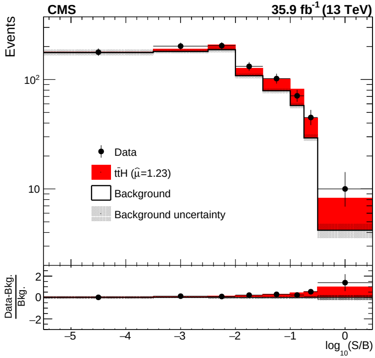
- MVA discriminant trained against main backgrounds
  - 2lSS: $t\bar{t}H$ vs $t\bar{t}$ + jets and $t\bar{t}H$ vs $t\bar{t}$ + V
  - 3l: 5-dimensional multinomial BDT: $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}$ + jets, VV
  - τ channels: $t\bar{t}H$ vs $t\bar{t}$ + jets
  - 4l: $t\bar{t}Z$

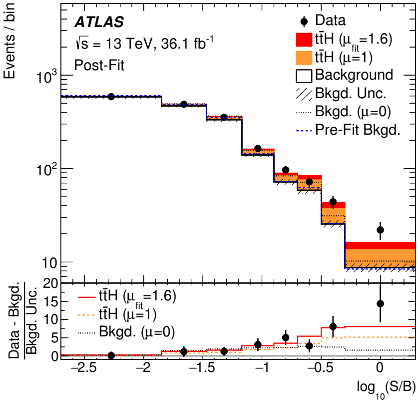
$t\bar{t}H$ multilepton: analysis strategy (ATLAS)

- MVA discriminant trained against main backgrounds
  - 2lSS: $t\bar{t}H$ vs $t\bar{t}+$ jets and $t\bar{t}H$ vs $t\bar{t}+V$
  - 3l: 5-dimensional multinomial BDT: $t\bar{t}H$, $t\bar{t}W$, $t\bar{t}Z$, $t\bar{t}+$ jets, VV
  - τ channels: $t\bar{t}H$ vs $t\bar{t}+$ jets
  - 4l: $t\bar{t}Z$
- Discriminating variables: BDT in all regions, except 4 leptons and 3 leptons + 1 τ_h



tt̄H multilepton results

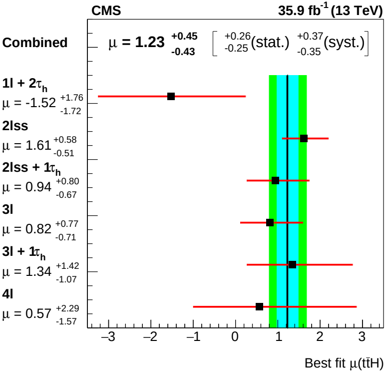


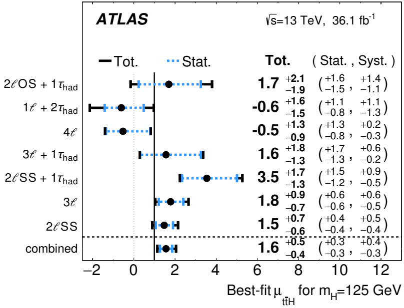


Best-fit $\mu = 1.23^{+0.45}_{-0.43}$, at 3.2 (2.8) σ obs. (exp.) significance

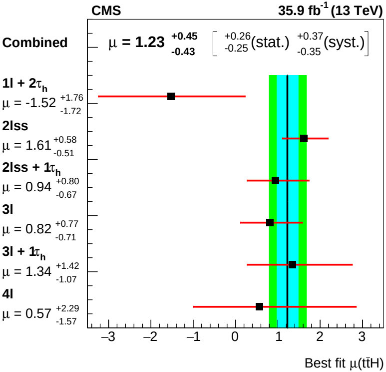
Best-fit $\mu = 1.6^{+0.5}_{-0.4}$, at 4.1 (2.8) σ obs. (exp.) significance

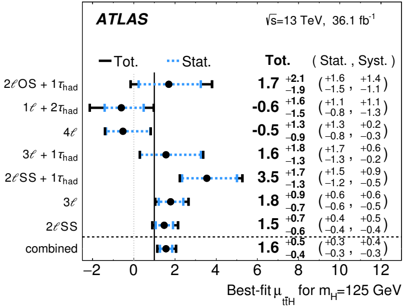
tt̄H multilepton results





$t\bar{t}H$ multilepton results

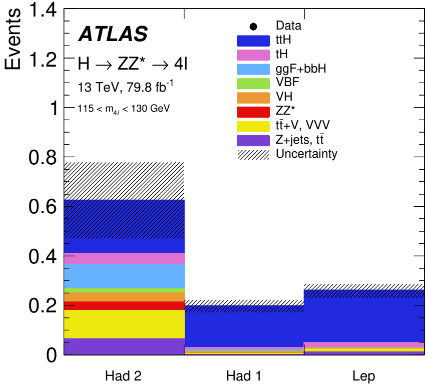




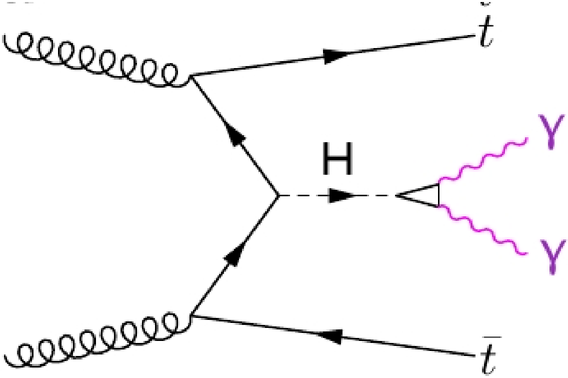
- Limited by non-prompt lepton estimation and τ identification, jet energy scale and resolution, $t\bar{t}H$ and $t\bar{t}+V$ modelling
- Several channels limited by statistics

ATLAS $t\bar{t}H(ZZ^* \to 4l)$, **80 fb^{-1}** arXiv:1806.00425, sub. PLB

- Improved sensitivity: separate leptonic and hadronic categories with BDT (for hadronic)
- No event was observed (0.45 expected) $\to$ Very statistically limited!
- 1.2 σ expected



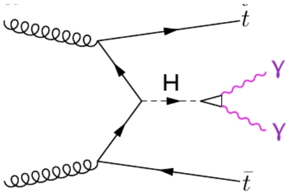
t$\bar{t}$H($\gamma\gamma$)

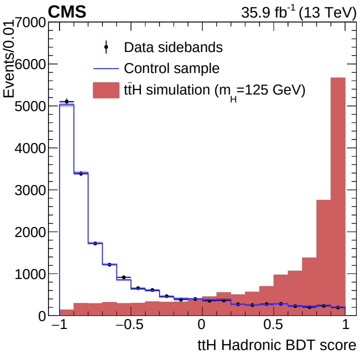


$t\bar{t}H(\gamma\gamma)$

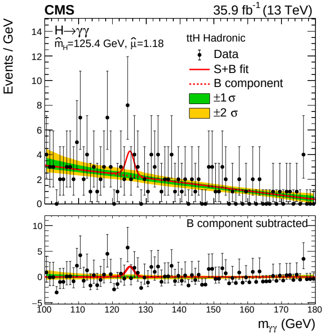
arXiv:1804.02610

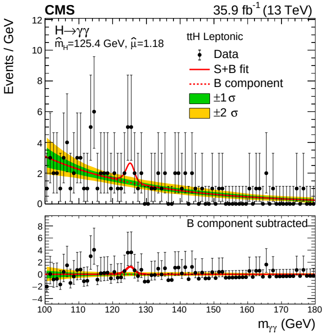
- Clear signature coming from the photons
- Higgs boson can be reconstructed as a narrow peak
- Backgrounds estimated from sideband regions
- Dedicated $t\bar{t}H$ channel part of the global $H \to \gamma\gamma$ analysis
- $t\bar{t}$ hadronic and leptonic channels
  - Hadronic $t\bar{t}$ decay: MVA is used for background rejection
- Signal extracted from fit to $m_{\gamma\gamma}$





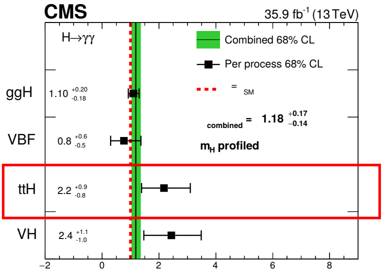
CMS $t\bar{t}H(\gamma\gamma)$ results





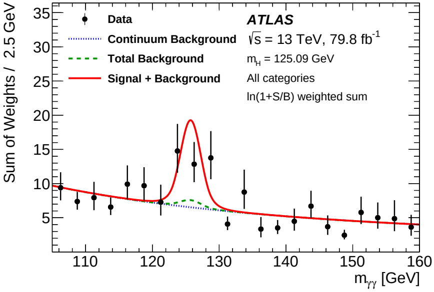
- Statistically limited
- Leading systematic uncertainties:

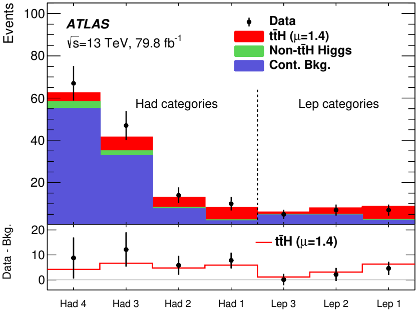
  Photon shower shape and energy scale



ATLAS $t\bar{t}H(\gamma\gamma)$ results, 80 fb^{-1} arXiv:1802.04146

- Analysis strategy: categorisation in 3 leptonic and 4 hadronic categories
- Increased sensitivity (50% for the same luminosity) by analysis improvements e.g: MVA utilizing γ and jet kinematic properties





Best-fit $\mu = 1.39^{+0.48}_{-0.42}$, at 4.1 (3.7) σ obs. (exp.) significance

$t\bar{t}H$ Combination(s)

CMS $t\bar{t}H$ combination

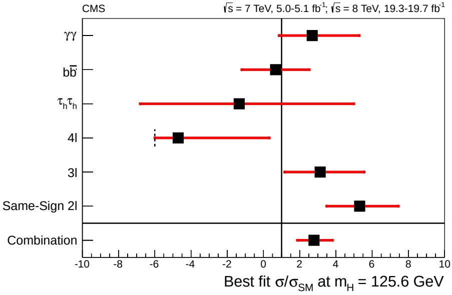
Phys.Rev.Lett. 120 (2018) 231801

Contributing analyses

- All of the presented $t\bar{t}H$ 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

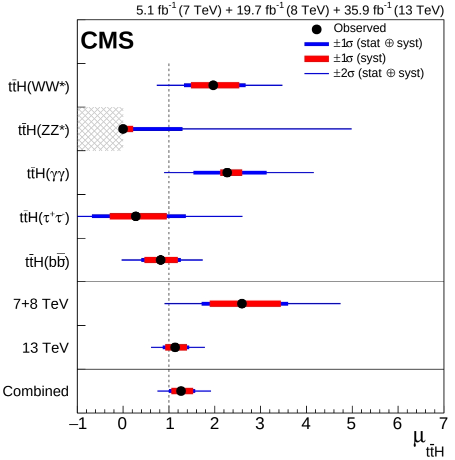


Correlations between Run-1 and Run-2 analyses

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

CMS tt̄H combination

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



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

CMS $t\bar{t}H$ combination

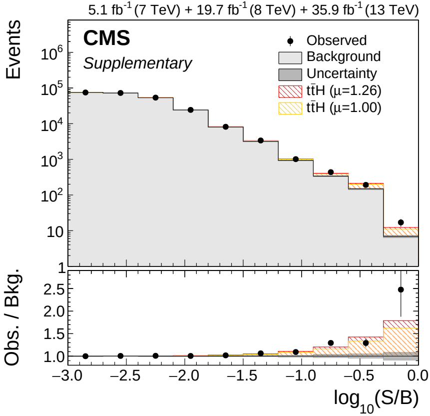
- $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ$ channels still limited by statistics
- Other channels dominated by systematics
- Signal theory mainly from inclusive $t\bar{t}H$ prediction
- Background theory mainly from $t\bar{t} + HF$ prediction in $t\bar{t}H(b\bar{b})$
- Experimental: lepton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

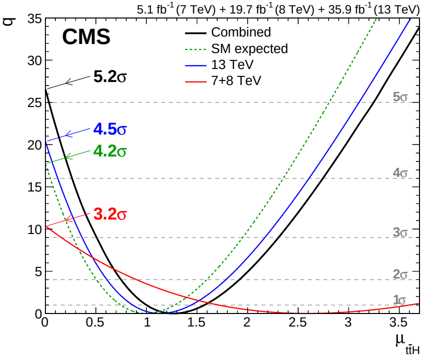
Uncertainty source	$\Delta\mu$	
Signal theory	+0.15	−0.07
Inclusive ttH normalisation (cross section and BR)	+0.15	−0.07
ttH acceptance (scale, pdf, PS and UE)	+0.004	−0.004
Other Higgs boson production modes	+0.002	−0.003
Background theory	+0.14	−0.13
tt + bb/cc prediction	+0.13	−0.11
tt + V(V) prediction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

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

CMS tt̄H combination

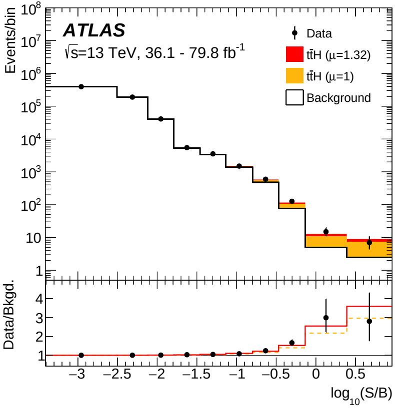
- First observation of the tt̄H production process (10 April 2018)
- Observed significance is 5.2σ (4.2σ exp.) with respect to the $\mu_{t\bar{t}H} = 0$ hypothesis





ATLAS tt̄H combination arXiv:1806.00425

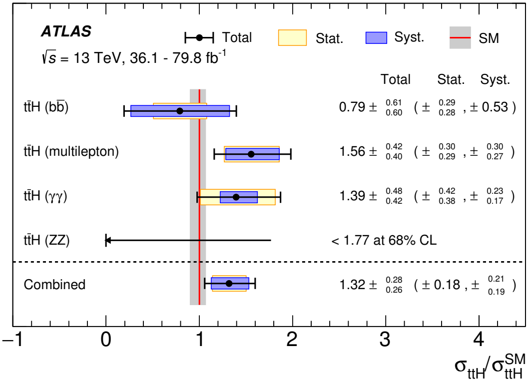
- 79.8 fb^{-1} tt̄H($\gamma\gamma$), tt̄H →4l results combined with 36.1 $^{-1}$ tt̄H(bb̄), multilepton, as well as with the Run-1 result



Uncertainty source	$\Delta\sigma_{t\bar{t}H}/\sigma_{t\bar{t}H}$ [%]
Theory uncertainties (modelling)	11.9
$t\bar{t}$ + heavy flavour	9.9
$t\bar{t}H$	6.0
Non-$t\bar{t}H$ Higgs boson production modes	1.5
Other background processes	2.2
Experimental uncertainties	9.3
Fake leptons	5.2
Jets, $E_{\mathrm{T}}^{\mathrm{miss}}$	4.9
Electrons, photons	3.2
Luminosity	3.0
τ-lepton	2.5
Flavour tagging	1.8
MC statistical uncertainties	4.4

$$\mu_{t\bar{t}H} = 1.32^{+0.28}_{-0.26} = 1.32^{+0.18}_{-0.18}(\text{stat})^{+0.21}_{-0.19}(\text{syst})$$

ATLAS tt̄H combination



- Observation of tt̄H production with 5.8 σ (4.9 σ) sign. (Run-2) and 6.3 σ (5.1 σ) sign. including Run-1 (4 June 2018)

Observation of $t\bar{t}H$ production!

- Results presented for $t\bar{t}H$ searches with 36-80 fb^{-1} of pp collision data @ 13 TeV (2016-17 data)
  - Improvements in analysis techniques compared to Run 1 (e.g. DNN, multivariate analysis $t\bar{t}H(\gamma\gamma)$, etc)
  - Addition of new challenging final states: fully hadronic mode, final states with hadronic decaying τ leptons
  - Several channels already systematic limited: Working on further improvements
- Combination resulted in the first observation of $t\bar{t}H$ production by CMS just published in PRL, ATLAS just submitted results including more 13 TeV data to PLB with larger significance
- New data being analyzed as we speak
  - More statistics helpful for developing more sophisticated strategies
  - Statistic limited channels will become more and more relevant

ATLAS: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HiggsPublicResults
CMS: http://cms-results.web.cern.ch/cms-results/public-results/publications/HIG/index.html

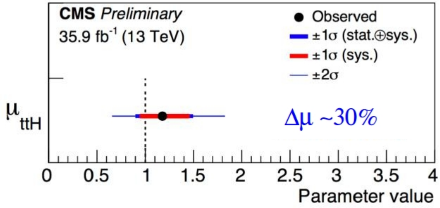
BACKUP

$t\bar{t}H$ combination + other Higgs measurements

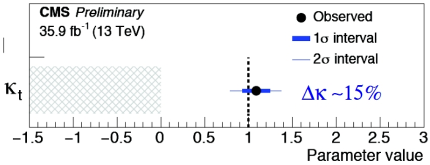
CMS-HIG-17-031

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.



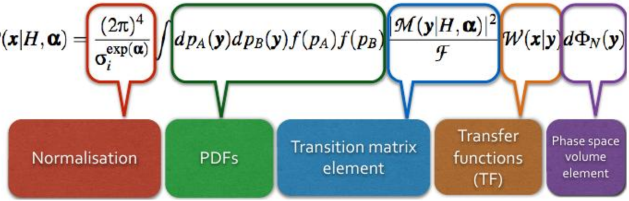
Best fit		Uncertainty Stat.	Syst.
1.09	$^{+0.14}_{-0.14}$	$^{+0.08}_{-0.08}$	$^{+0.12}_{-0.12}$

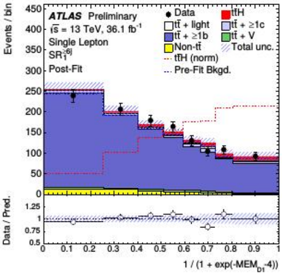
MEM

$t\bar{t}H(b\bar{b})$: Matrix Element Method

Taken from these slides by I. Connelly

$$P(\mathbf{x}|H,\boldsymbol{\alpha}) = \frac{(2\pi)^4}{\sigma_i^{\exp(\boldsymbol{\alpha})}} \int dp_A(\mathbf{y})dp_B(\mathbf{y})f(p_A)f(p_B)\frac{|\mathcal{M}(\mathbf{y}|H,\boldsymbol{\alpha})|^2}{\mathcal{F}} W(\mathbf{x}|\mathbf{y})d\Phi_N(\mathbf{y})$$





- MEM linking a set of measured quantities (x, e.g. b-jet energy) with a set of unobservable partonic objects (y, e.g. b-quark energy) associated to a hypothesised process
- Test two hypotheses (H): $t\bar{t}H$ (signal) and $t\bar{t}b\bar{b}$ (background)
- Transfer function $W(x|y)$: likelihood that partonic configuration y is measured as x (from MC)
- Discriminant defined as the difference between the logarithms of the signal and background likelihoods

MEM

- Signal extraction via Matrix Element Methods (MEM):
  - Event-by-event discriminator build upon matrix elements, combined with reconstruction-level information

Numerical integration — Momentum conservation — Resolution function (allow ISR)

$$w(\vec{y}|\mathcal{H}) = \sum_{i=1}^{N_C} \int \frac{dx_a dx_b}{2x_a x_b s} \int \prod_{k=1}^{8} \left(\frac{d^3\vec{p}_k}{(2\pi)^3 2E_k} \right) (2\pi)^4 \delta^{(E,z)} \left(p_a + p_b - \sum_{k=1}^{8} p_k \right) \mathcal{R}^{(x,y)} \left(\vec{\rho}_T, \sum_{k=1}^{8} p_k \right)$$
$$\times g(x_a, \mu_F) g(x_b, \mu_F) |\mathcal{M}(p_a, p_b, p_1, \ldots, p_8)|^2 W(\vec{y}, \vec{p})$$

Parton density functions — LO scattering amplitude (Open Loops) — Detector transfer function

- Construct per-event signal/background probabilities using full kinematic information in an analytic approach

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

- $t\bar{t}$+bb take as background hypothesis, permuting overall jet assignments
- Works best for final states with multiple reconstructed jets

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

ATLAS

Uncertainty source	$\Delta\mu$	
$t\bar{t} + \geq 1b$ modeling	+0.46	−0.46
Background-model stat. unc.	+0.29	−0.31
b-tagging efficiency and mis-tag rates	+0.16	−0.16
Jet energy scale and resolution	+0.14	−0.14
$t\bar{t}H$ modeling	+0.22	−0.05
$t\bar{t} + \geq 1c$ modeling	+0.09	−0.11
JVT, pileup modeling	+0.03	−0.05
Other background modeling	+0.08	−0.08
$t\bar{t}$ + light modeling	+0.06	−0.03
Luminosity	+0.03	−0.02
Light lepton (e, μ) id., isolation, trigger	+0.03	−0.04
Total systematic uncertainty	+0.57	−0.54
$t\bar{t} + \geq 1b$ normalization	+0.09	−0.10
$t\bar{t} + \geq 1c$ normalization	+0.02	−0.03
Intrinsic statistical uncertainty	+0.21	−0.20
Total statistical uncertainty	+0.29	−0.29
Total uncertainty	+0.64	−0.61

(PhysRevD.97.072016)

CMS

Uncertainty source	$\pm\Delta\mu$ (observed)	$\pm\Delta\mu$ (expected)
Total experimental	+0.15/−0.16	+0.19/−0.17
b tagging	+0.11/−0.14	+0.12/−0.11
jet energy scale and resolution	+0.06/−0.07	+0.13/−0.11
Total theory	+0.28/−0.29	+0.32/−0.29
$t\bar{t}$+hf cross section and parton shower	+0.24/−0.28	+0.28/−0.28
Size of the simulated samples	+0.14/−0.15	+0.16/−0.16
Total systematic	+0.38/−0.38	+0.45/−0.42
Statistical	+0.24/−0.24	+0.27/−0.27
Total	+0.45/−0.45	+0.53/−0.49

(HIG-17-026)

$t\bar{t}H$ multilepton results

Uncertainty Source	$\Delta\mu$	
$t\bar{t}H$ modelling (cross section)	+0.20	−0.09
Jet energy scale and resolution	+0.18	−0.15
Non-prompt light-lepton estimates	+0.15	−0.13
Jet flavour tagging and τ_{had} identification	+0.11	−0.09
$t\bar{t}W$ modelling	+0.10	−0.09
$t\bar{t}Z$ modelling	+0.08	−0.07
Other background modelling	+0.08	−0.07
Luminosity	+0.08	−0.06
$t\bar{t}H$ modelling (acceptance)	+0.08	−0.04
Fake τ_{had} estimates	+0.07	−0.07
Other experimental uncertainties	+0.05	−0.04
Simulation statistics	+0.04	−0.04
Charge misassignment	+0.01	−0.01
Total systematic uncertainty	+0.39	−0.30

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

Statistical methodology

- Results calculated using the profile likelihood (**L**) ratio, **q**

$\vec{\alpha}$ = Set of POIs at some fixed values to be tested

$\vec{\theta}$ = Nuisance parameters

$$q(\vec{\alpha}) = -2\ln\left(\frac{L(\vec{\alpha}\,,\,\hat{\hat{\vec{\theta}}}_{\vec{\alpha}})}{L(\hat{\vec{\alpha}},\hat{\vec{\theta}})}\right)$$

Values of $\vec{\theta}$ that maximise the likelihood given the fixed values of $\vec{\alpha}$ being tested (conditional estimate)

Values of $\vec{\alpha}$ and $\vec{\theta}$ that globally maximise the likelihood (unconditional estimate)

- Exploit the asymptotic limit:
  - Test statistic $q(\vec{\alpha})$ is assumed to follow a χ2 distribution with $\vec{\alpha}$ degrees of freedom
  - ⇒ To determine a confidence-level (CL) interval for a single parameter α, we only need to find the values of α where $q(\vec{\alpha})$ = the χ2 critical value for that CL, e.g.
    - 1D 68% CL at q(α) = 1.00

ATLAS $t\bar{t}H(b\bar{b})$ selection

- *b*-tagging:
  - Considering 4 working points: *loose*, *medium*, *tight*, *very-tight*
  - Efficiency for *b*-jets: 85% → 60%
  - Rejection factor for *c*-jets [light jets]: 3→35 [30→1500]
  - *b*-tagging discriminant built as:

	none	*loose*	*medium*	*tight*	*very-tight*
Efficiency	-	85%	77%	70%	60%
Discriminant value	1	2	3	4	5

Channel classification:

- Two separate channels depending on the number of light leptons (ℓ=e, μ): 1ℓ, 2ℓ
- 2ℓ opposite-sign (OS) with $p_T > 27, 15$ GeV (veto $m_{\ell\ell} \sim m_Z$, and events with τ_{had})
  - Require ≥3 jets and ≥2 *medium* *b*-tagged jets
- 1ℓ with $p_T > 27$ GeV (veto events with ≥2 τ_{had}'s)
  - High-p_T category:
    - 'Boosted' event: boosted Higgs and top candidates (large-R jets, reclustered from $R = 0.4$ jets), plus a *loose* *b*-tagged jet
    - Higgs boson candidate ($p_T > 200$ GeV): two *loose* *b*-tagged jets
    - Top candidate ($p_T > 250$ GeV): one *loose* *b*-tagged + ≥1 non-*b*-tagged jets
  - If failing the 'boosted' selection → 'Resolved' event:
    - Require ≥5 jets and ≥2 *very-tight* *b*-tagged jets or ≥3 *medium* *b*-tagged jets