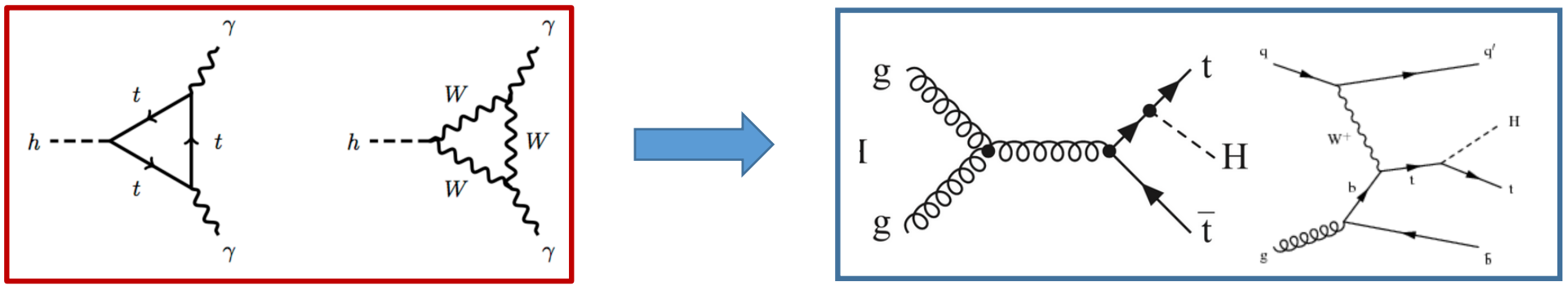
Top-Higgs associated production at CMS

Jeremy Andrea

Introduction

- Top-Higgs coupling measurements as a tool to probe the SM-EWK symmetry breaking.

- Top-Higgs coupling measured indirectly through the H→$\gamma\gamma$ production, via loops effects.



- Direct measurement of top-Higgs coupling through the Higgs associated production
  - **with a $t\bar{t}$ pair** : QCD production, largest cross section, sensitive to $|y_t|^2$.
  - **with a *t*-channel single top** : EWK production, lower cross section, sensitive to the sign of y_t.

- With the high luminosity at the LHC => direct measurements of top-Higgs couplings possible.

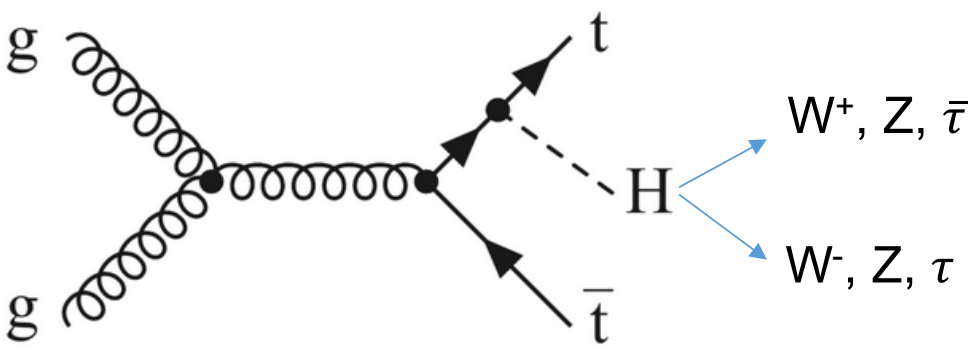
- Combinations of channels are presented here, with more emphasis put on multi-leptons.

Observation of ttH at CMS

HIG-17-035, PRL 120 (2018) 231801

ttH in multi-leptons

JHEP 08 (2018) 066



- ttH multi-leptons :
  - At least 1 electron or muon + hadronic taus,
  - At least 2 electrons or muons + no hadronic taus.
- Trigger selections :
  - One electron/muon (+ hadronic tau),
  - Presence of one, two or three electrons/muons.
- B-tagging : ≥1 tight bjet or ≥2 loose bjets.
- Signal categorisation based on jet multiplicities :
  - **2lss** : ≥4 jets,
  - 2lss + 1 τ_h: ≥3 jets,
  - **3l**, 3l + 1 τ_h: ≥2 jets,
  - 2l + 1 τ_h: ≥3 jets,
  - **4l** : ≥2 jets.

Lepton+τ channels already discussed yesterday
See Cristina's talk

- Z-mass veto on lepton pairs, low dilepton resonance veto.
- Minimum MET cut for some channels.

Analysis strategy

- Irreducible (non instrumental) backgrounds estimated from simulation, validated with Control Regions (CR).
- Instrumental backgrounds estimated from data.
- Photons conversion, minor, estimated from MC and validated in CR.
- Misidentified leptons : based on the Fake Factor method.
  - Estimations of fake rate in data (QCD multijet) and applied on a side band sample (tight lepton veto).
- Lepton charge mis-reconstruction: only relevant for 2lss. Similar technic, rate of charge mis-identification estimated from $Z/\gamma^* \rightarrow ll$ events.

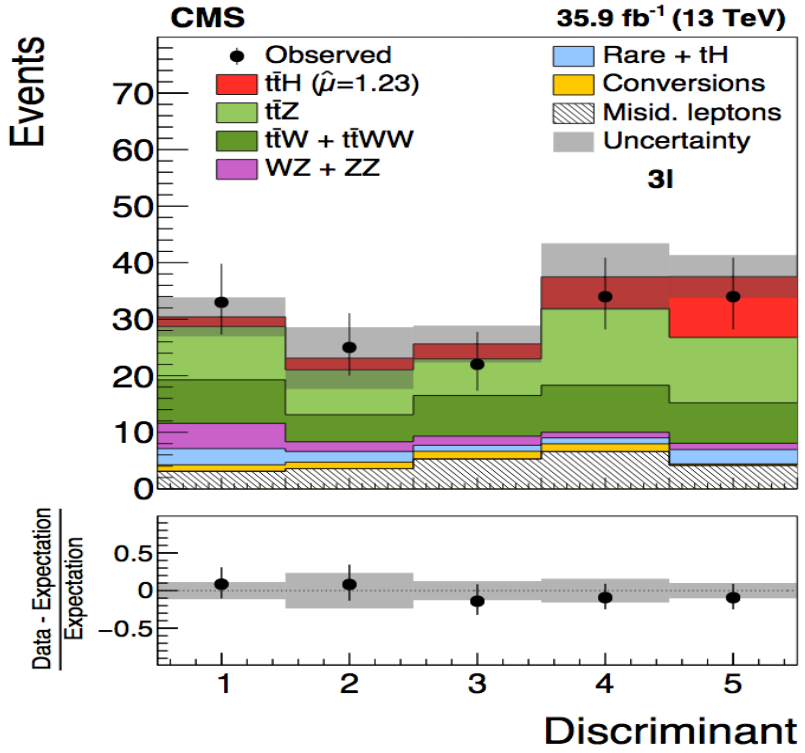
Process	$1\ell + 2\tau_h$	2ℓss	2ℓss $+ 1\tau_h$
$t\bar{t}H$	7.1 ± 2.4	66.3 ± 21.0	11.6 ± 3.5
$t\bar{t}Z/\gamma^*$	6.3 ± 1.1	80.9 ± 10.4	9.2 ± 1.2
$t\bar{t}W + t\bar{t}WW$	0.5 ± 0.1	150.0 ± 16.9	9.1 ± 1.0
$WZ + ZZ$	2.1 ± 1.6	16.5 ± 13.1	3.9 ± 3.0
tH	0.4 ± 0.1	2.7 ± 0.2	0.5 ± 0.04
Conversions	< 0.02	12.1 ± 5.8	1.4 ± 0.5
Sign flip	—	27.5 ± 8.0	0.5 ± 0.1
Misidentified leptons	195.7 ± 13.6	94.2 ± 21.2	8.6 ± 2.1
Rare backgrounds	1.4 ± 0.7	39.0 ± 21.2	3.1 ± 1.5
Total expected background	206.3 ± 14.0	423.0 ± 38.0	36.1 ± 4.2
Observed	212	507	49

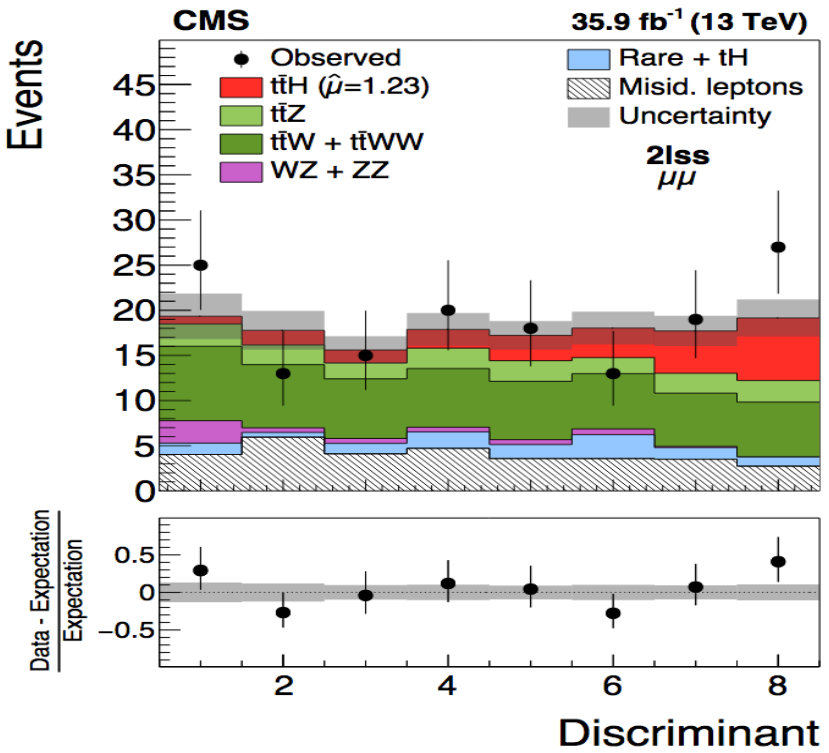
Process	3ℓ	$3\ell + 1\tau_h$	4ℓ
$t\bar{t}H$	22.8 ± 7.4	2.6 ± 0.9	1.1 ± 0.4
$t\bar{t}Z/\gamma^*$	49.0 ± 6.9	3.4 ± 0.5	2.1 ± 0.4
$t\bar{t}W + t\bar{t}WW$	35.2 ± 4.2	0.4 ± 0.04	$< 2 \times 10^{-3}$
$WZ + ZZ$	9.9 ± 2.4	0.3 ± 0.05	0.1 ± 0.1
tH	1.2 ± 0.2	0.1 ± 0.01	$< 4 \times 10^{-4}$
Conversions	5.3 ± 2.9	< 0.02	< 0.02
Misidentified leptons	22.7 ± 6.7	0.9 ± 0.2	< 0.04
Rare backgrounds	8.2 ± 13.8	0.2 ± 0.1	0.1 ± 0.2
Total expected background	131.4 ± 18.2	5.3 ± 0.5	2.4 ± 0.4
Observed	148	7	3

Systematic uncertainties

- Signal extraction from ML of all the signal regions, systematics as nuisance parameters,

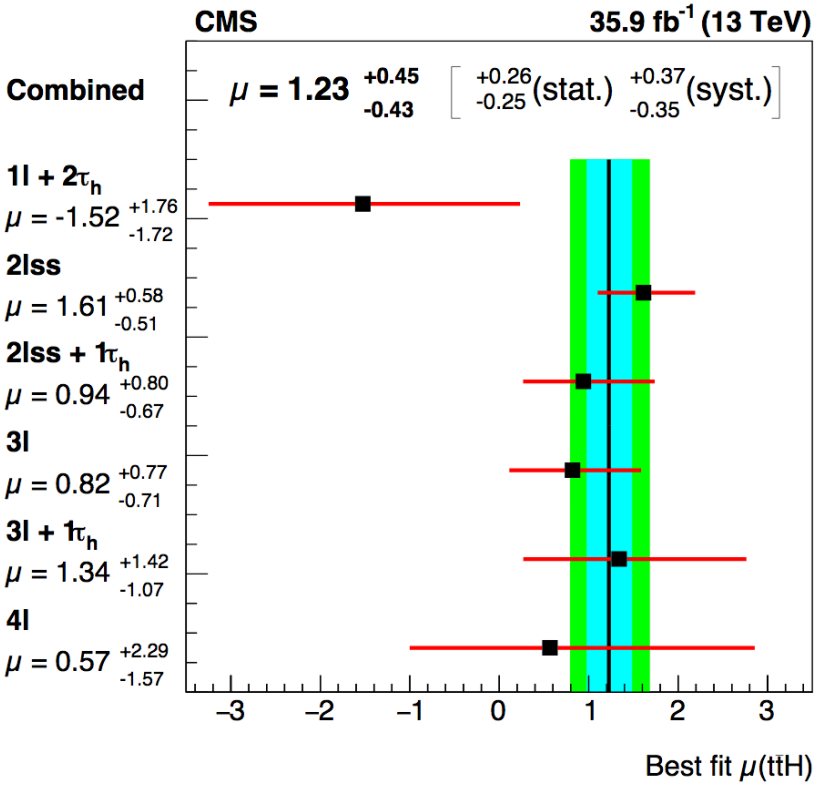
- Discriminating variables :
  - 2 BDTs for discriminating against ttV and $t\bar{t}$,
  - Variables : $\Delta R(l,j)$, N(b)Jets, jet-top association, leptons p_T…
  - BDTs joined into a single variable (~categorization based on S/B).





Results

- Observed (expected) significance of 3.2σ (2.8σ),
- Signal strength $\mu = 1.23^{+0.45}_{-0.43}$,
- Dominant systematics : Theory (ttV), same-sign/fakes, lepton ID.

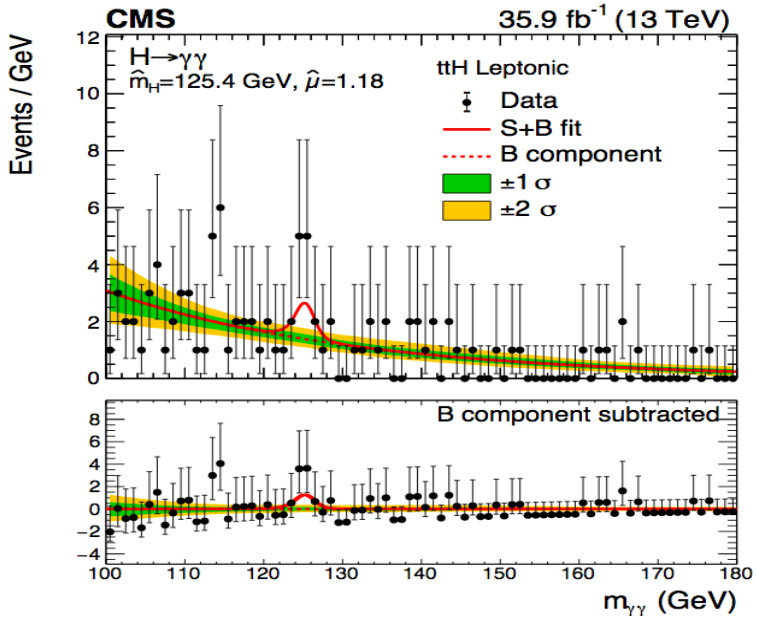


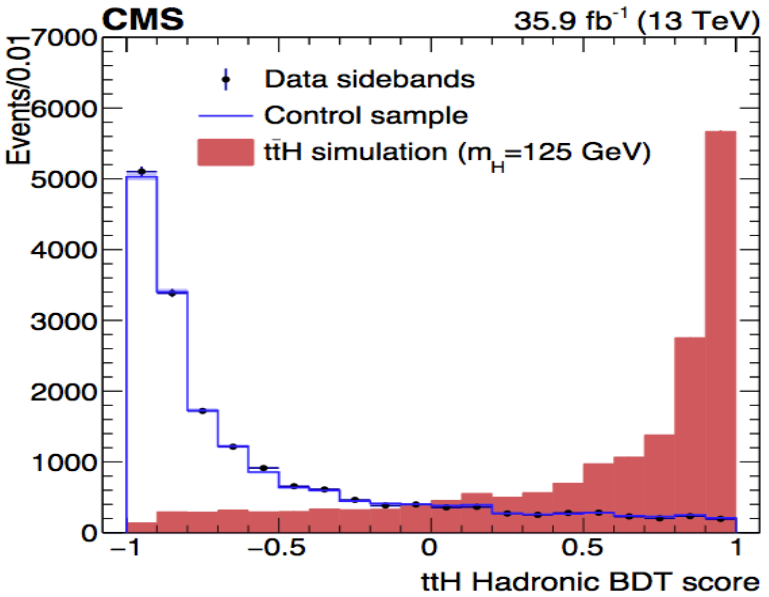
Source	Uncertainty [%]	$\Delta\mu/\mu$ [%]
e, μ selection efficiency	2–4	11
τ_h selection efficiency	5	4.5
b tagging efficiency	2–15 [57]	6
Reducible background estimate	10–40	11
Jet energy calibration	2–15 [65]	5
τ_h energy calibration	3	1
Theoretical sources	≈10	12
Integrated luminosity	2.5	5

Other channels, H→$\gamma\gamma$

JHEP 11 (2018) 185

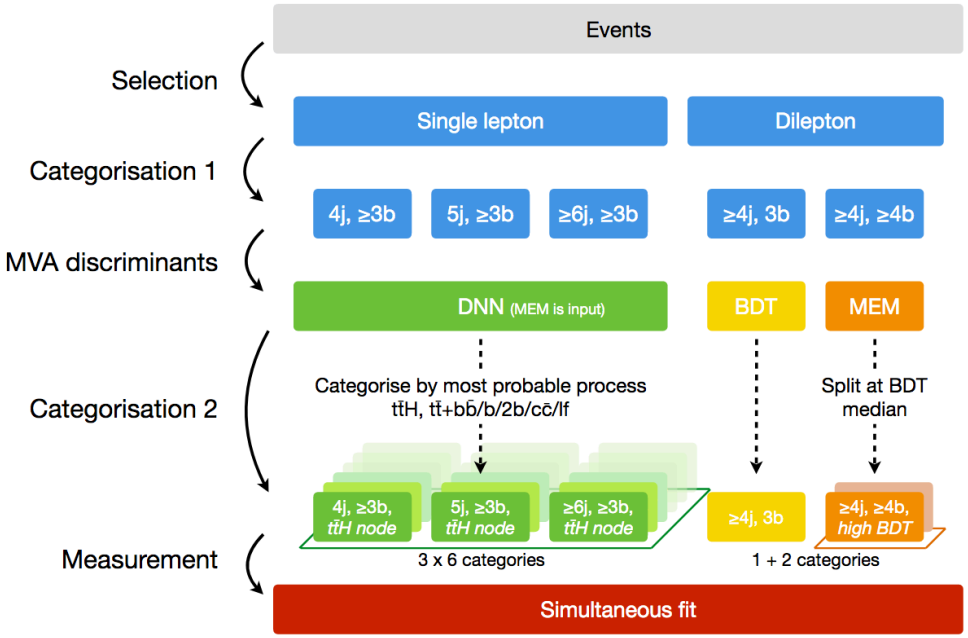
- ttH, H→$\gamma\gamma$, part of the H→$\gamma\gamma$ "grand" analysis (several different production modes).
- Semi-leptonic and fully hadronic channels accounted.
- "Usual" l+jet selection for single lepton channel.
- Full hadronic channels, signal event selection based on BDT.
- Fit of the di-photon invariant mass.

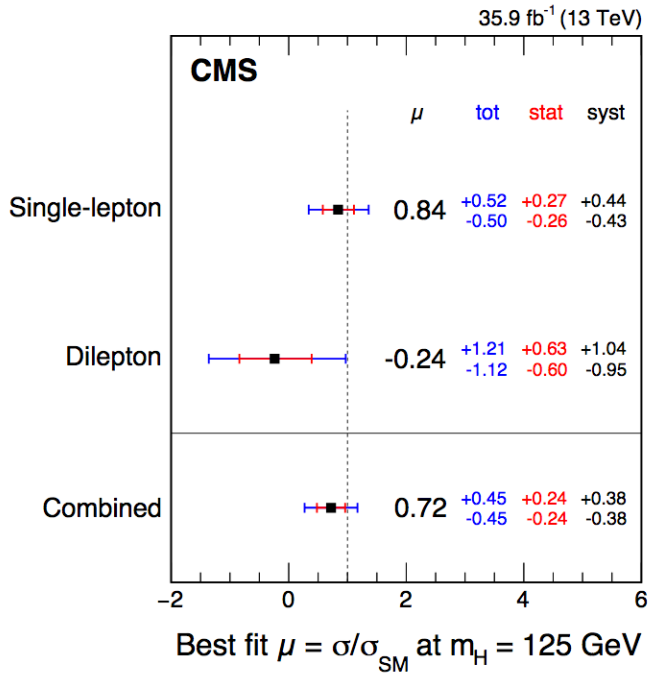


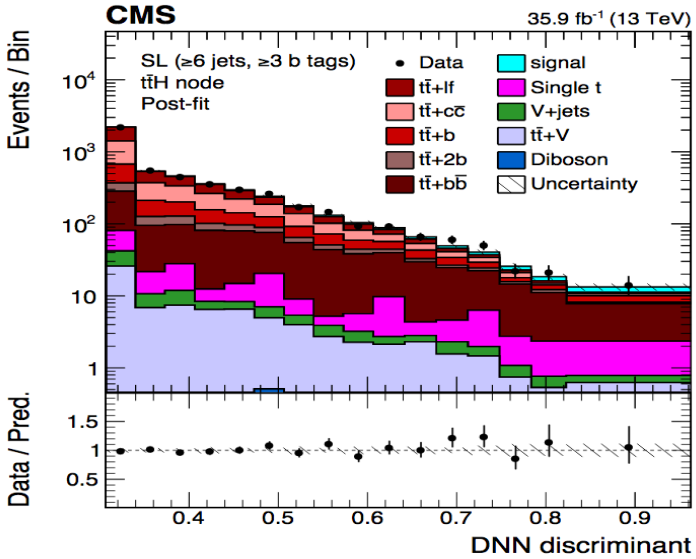


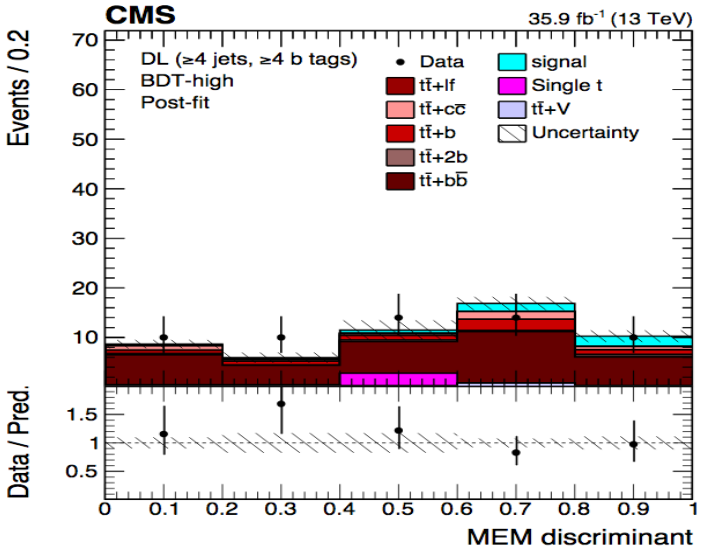
Other channels, H→$b\bar{b}$

JHEP 03 (2019) 026

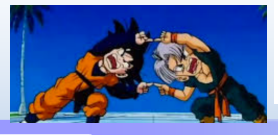




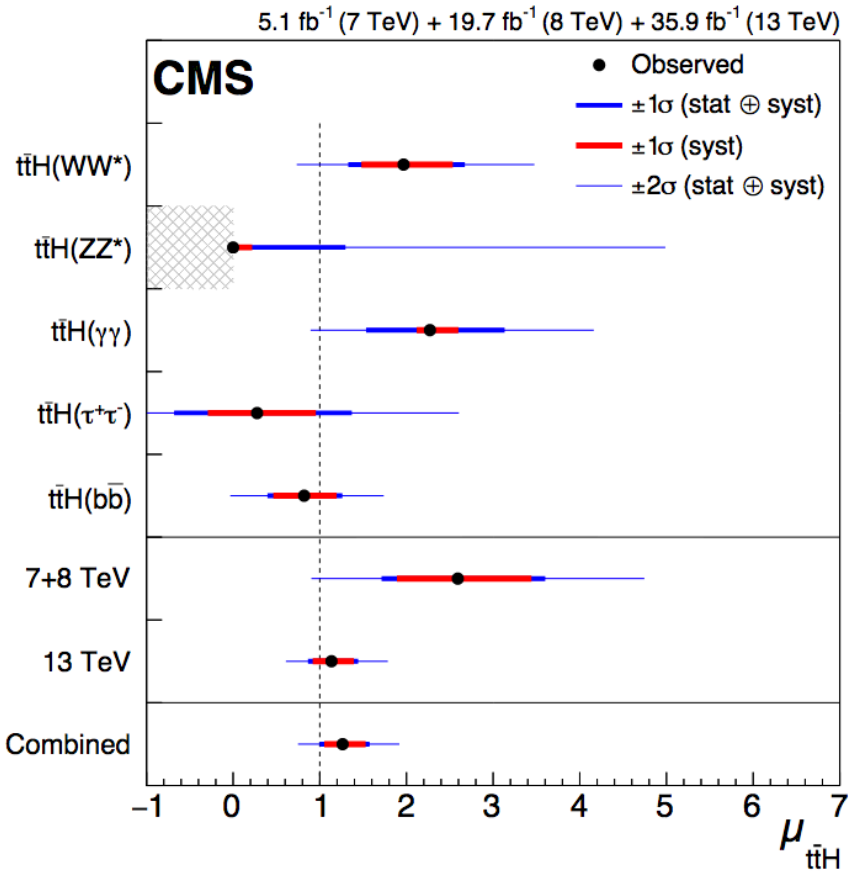


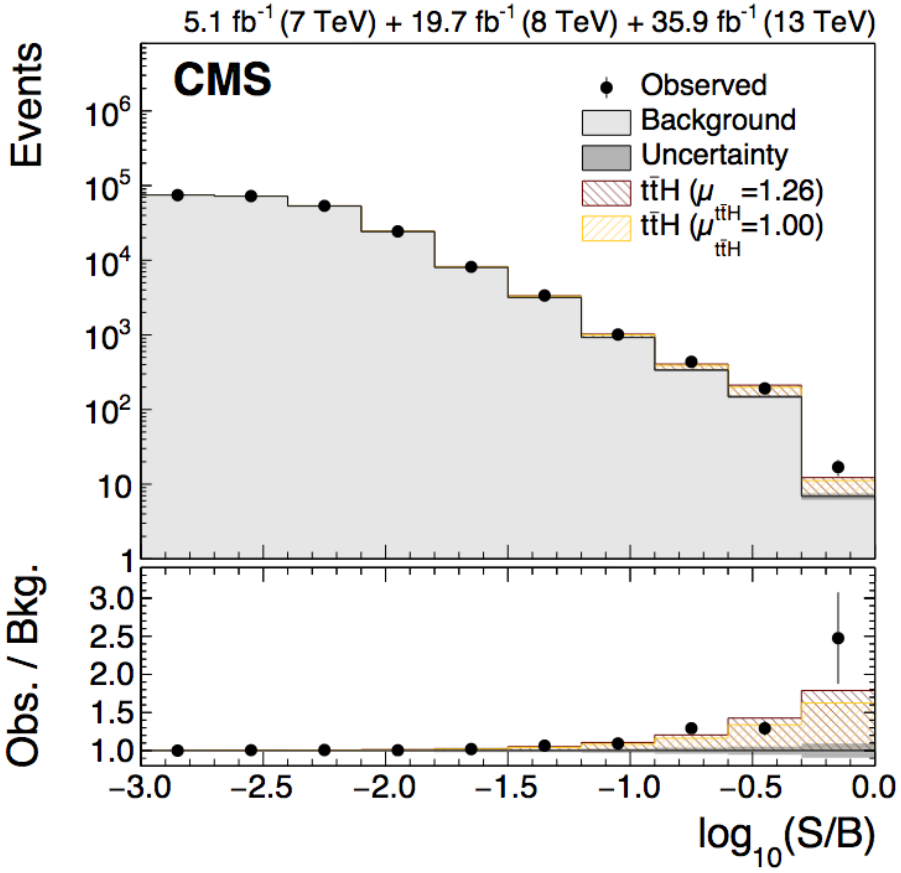


ttH combination



- Combination of multi-leptons, $\tau\tau$, $\gamma\gamma$ and $b\bar{b}$, for various samples, taken at different energies.
- Observed (expected) significance of 5.2 (4.2).
- Signal strength : $1.26^{+0.31}_{-0.26}$.

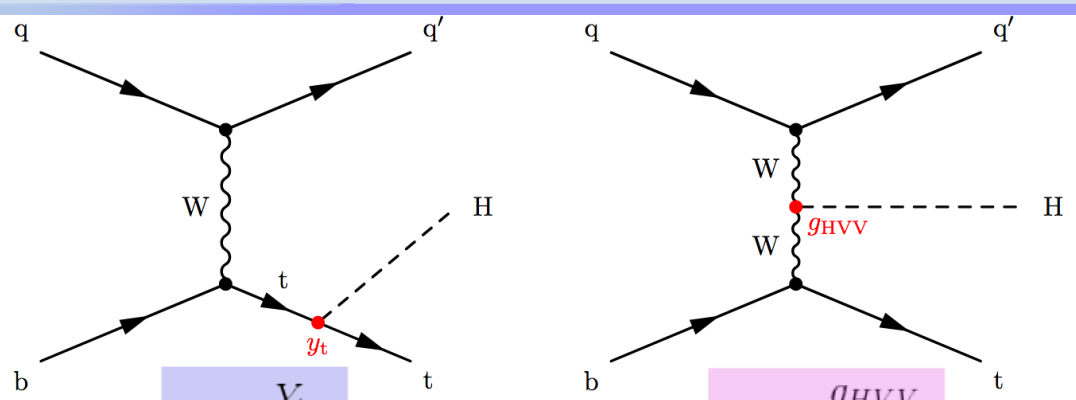




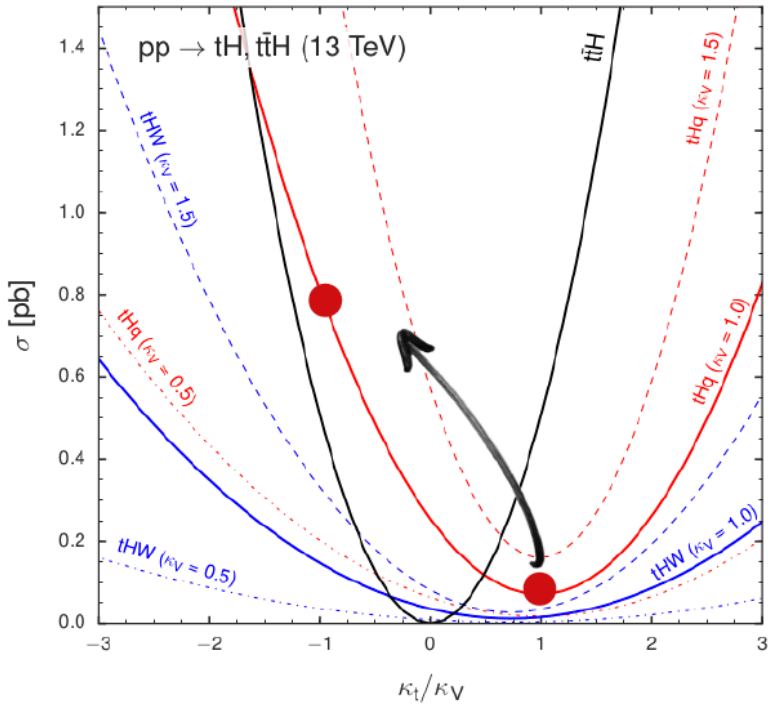
Search for tHq

HIG-18-009, Submitted to Phys. Rev. D.

tHq motivation



$$\kappa_t = \frac{Y_t}{Y_t^{SM}} \qquad \mathcal{A} \propto (\kappa_V - \kappa_t) \qquad \kappa_V = \frac{g_{HVV}}{g_{HVV}^{SM}}$$



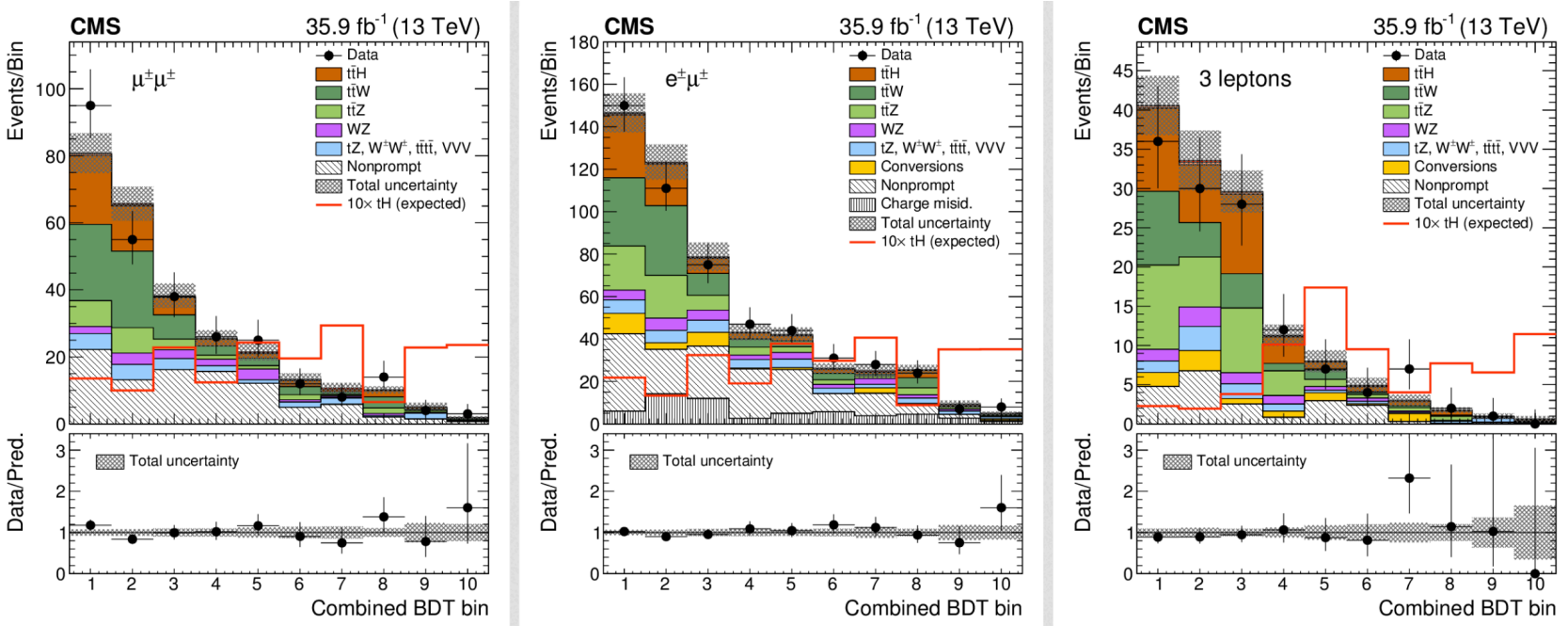
- ttH cross section proportional to $|y_t|^2$.
- tHq cross section sensitive to the sign of y_t.
- In SM, destructive interferences between top Higgs-Strahlung and W boson "fusion" diagrams.
- Large increase of tHq cross section in case $y_t < 0$.

tHq multilepton

- Extension of the ttH multilepton analysis.
  - Same object selections,
  - Similar data-driven techniques for fake lepton and charge mis-reconstruction backgrounds.

- Adapted signal event selection :
  - ≥ 1 b-tag jet, ≥ 1 untagged jet, potentially forward ($|\eta| < 4.5$),
  - Channels : same sign dileptons $\mu^{\pm}\mu^{\pm}$, $e^{\pm}\mu^{\pm}$ and tri-leptons (di-electron neglected).

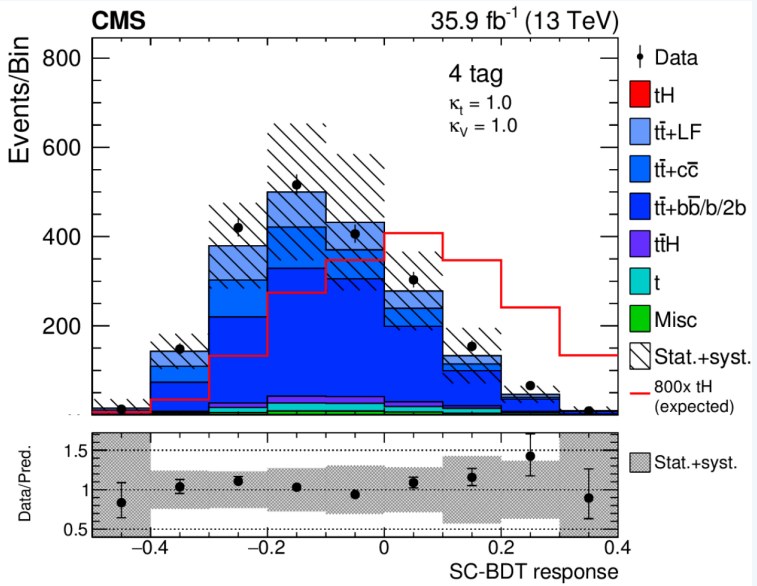
- 2 different BDT for discriminations against *ttW/Z* and $t\bar{t}$.

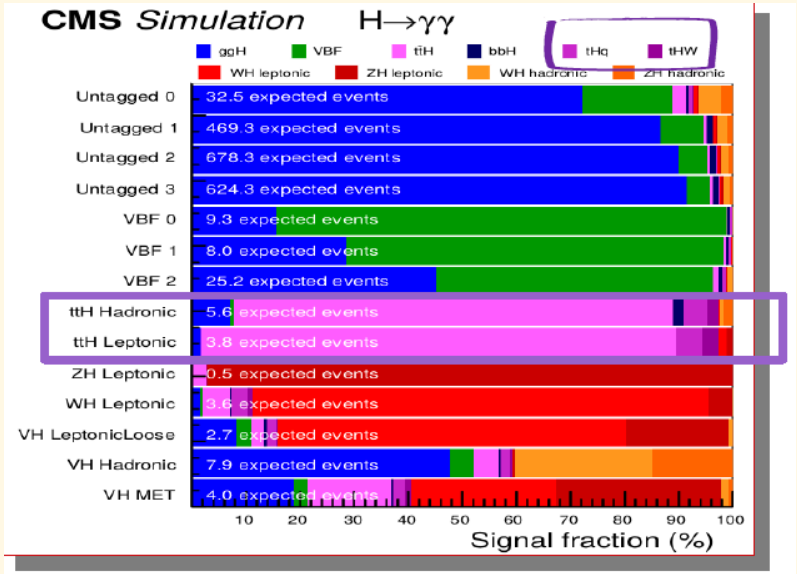


other channels, H→bb, H→$\gamma\gamma$

- H→bb channel:
  - Largest stat,
  - Large ttbar+jets backgrounds,
  - Large b-to-top/higgs combinatorics.

- Analysis strategy :
  - 1 CR to constrains tt+light w.r.t. tt+jets,
  - Jets assignments from a BDT score,
  - 1 BDT to separate signal from ttbar in SR.

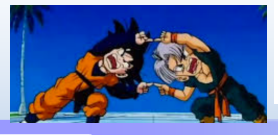




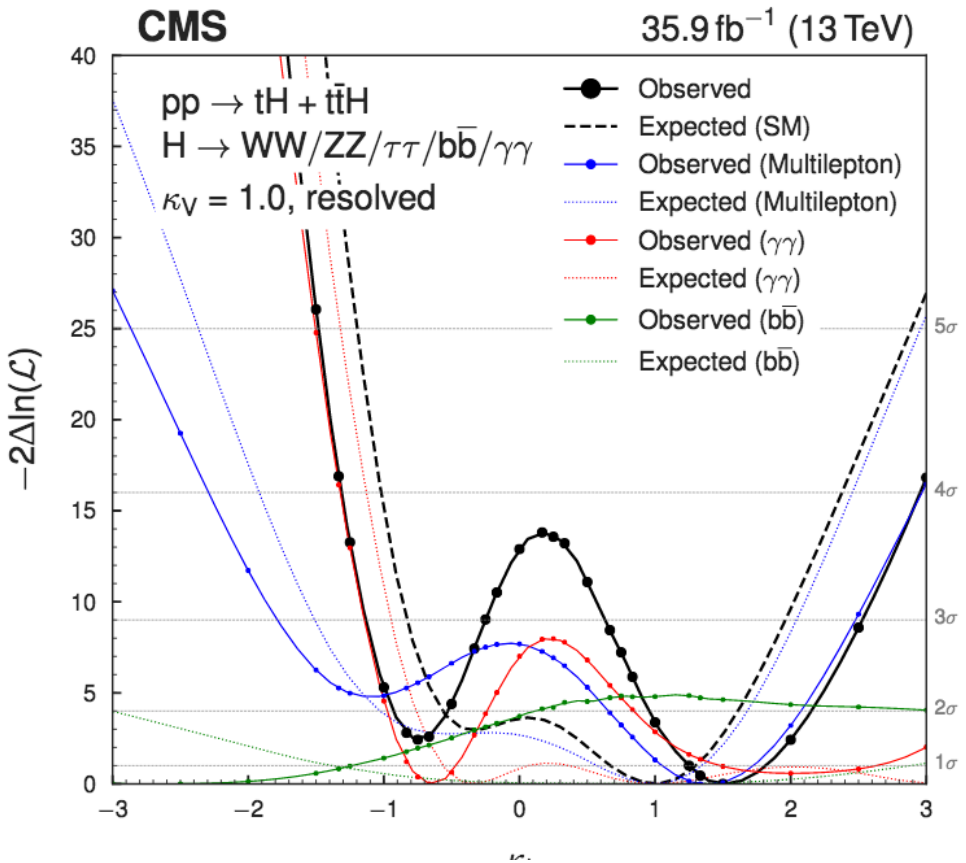
- H→$\gamma\gamma$ channel.
  - Based on a re-interpretation of H→$\gamma\gamma$ combination, using $\gamma\gamma$ invariant mass distribution.
  - tHq yields extracted from ttH categories.

- Acceptance corrected to account for different κ_t.

Combinations



- Combinations of multi-leptons, $\gamma\gamma$ and $b\bar{b}$ channels.
- Data favour a positive value of κ_t.
- Exclude κ_t values outside the ranges [−0.9,−0.5] and [1.0, 2.1] at 95% CL.

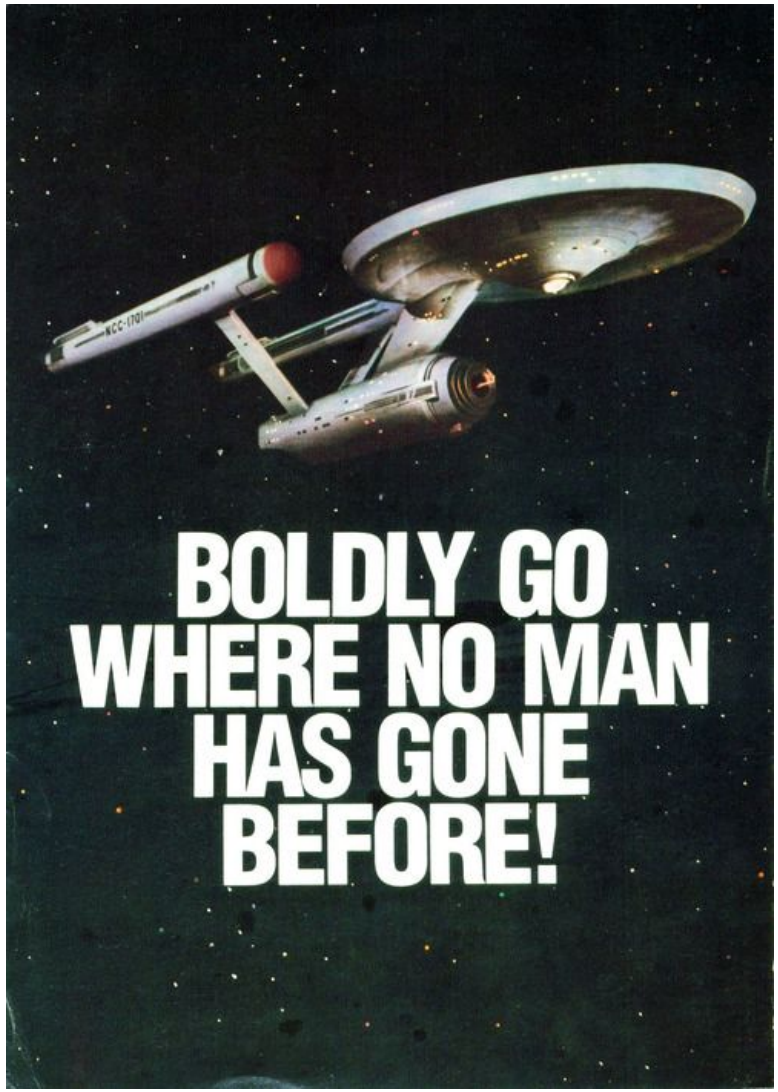


Conclusion

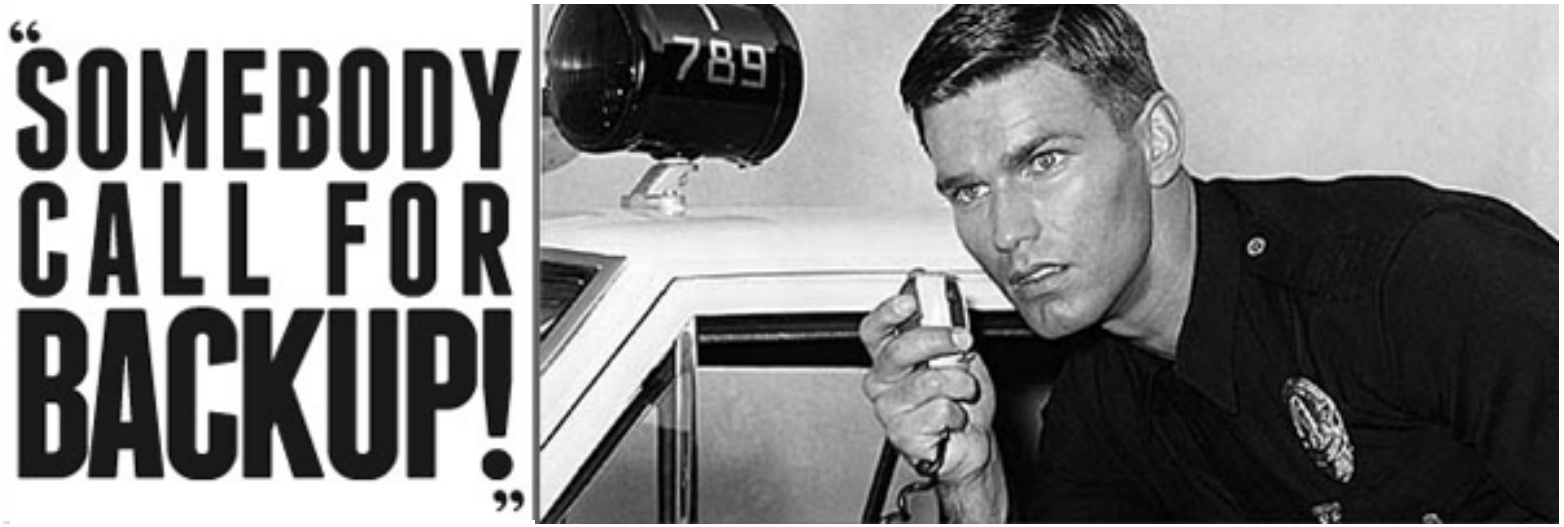
- With the high luminosity available at the LHC, we can start exploring direct top-Higgs couplings.

- Observation of ttH process preformed by combining 7+8+13 TeV data and using several decay channels.

- Sensitivity on tHq is not yet enough to reach the SM predictions, **but limits on BSM models can be set !** Promising channel for searching for new physics.



Backup



Categories definition

Selection	2ℓss	$2\ell\text{ss} + 1\tau_h$
Targeted $t\bar{t}H$ decay	$t \to b\ell\nu$, $t \to bqq$, $H \to WW \to \ell\nu qq$	$t \to b\ell\nu$, $t \to bqq$, $H \to \tau\tau \to \ell\tau_h + \nu$'s
Trigger	Single- and double-lepton triggers	
Lepton p_T	$p_T > 25$ / 15 GeV	$p_T > 25$ / 15 (e) or 10 GeV (μ)
Lepton η	$\lvert\eta\rvert < 2.5$ (e) or 2.4 (μ)	
τ_h p_T	—	$p_T > 20$ GeV
τ_h η	—	$\lvert\eta\rvert < 2.3$
Charge requirements	2 same-sign leptons and charge quality requirements	2 same-sign leptons and charge quality requirements $\sum_{\ell,\tau_h} q = \pm 1$
Jet multiplicity	$\geq$4 jets	$\geq$3 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	
Missing transverse momentum	$L_D > 30$ GeV	$L_D > 30$ GeV *
Dilepton mass	$m_{\ell\ell} > 12$ GeV and $\lvert m_{ee} - m_Z\rvert > 10$ GeV *	

Categories definition

Selection	3ℓ	$3\ell + 1\tau_h$
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu$, $t \to b\ell\nu$, $H \to WW \to \ell\nu qq$ $t \to b\ell\nu$, $t \to bqq$, $H \to WW \to \ell\nu\ell\nu$ $t \to b\ell\nu$, $t \to bqq$, $H \to ZZ \to \ell\ell qq$ or $\ell\ell\nu\nu$	$t \to b\ell\nu$, $t \to b\ell\nu$, $H \to \tau\tau \to \ell\tau_h + \nu$'s
Trigger	Single-, double- and triple-lepton triggers	
Lepton p_T	$p_T > 25$ / 15 / 15 GeV	$p_T > 20$ / 10 / 10 GeV
Lepton η	$\lvert\eta\rvert < 2.5$ (e) or 2.4 (μ)	
τ_h p_T	—	$p_T > 20$ GeV
τ_h η	—	$\lvert\eta\rvert < 2.3$
Charge requirements	$\sum_\ell q = \pm 1$	$\sum_{\ell,\tau_h} q = 0$
Jet multiplicity	≥ 2 jets	
b tagging requirements	≥ 1 tight b-tagged jet or ≥ 2 loose b-tagged jets	
Missing transverse momentum	No requirement if $N_j \geq 4$ $L_D > 45$ GeV [†] $L_D > 30$ GeV otherwise	
Dilepton mass	$m_{\ell\ell} > 12$ GeV and $\lvert m_{\ell\ell} - m_Z\rvert > 10$ GeV [‡]	
Four-lepton mass	$m_{4\ell} > 140$ GeV [§]	—

Categories definition

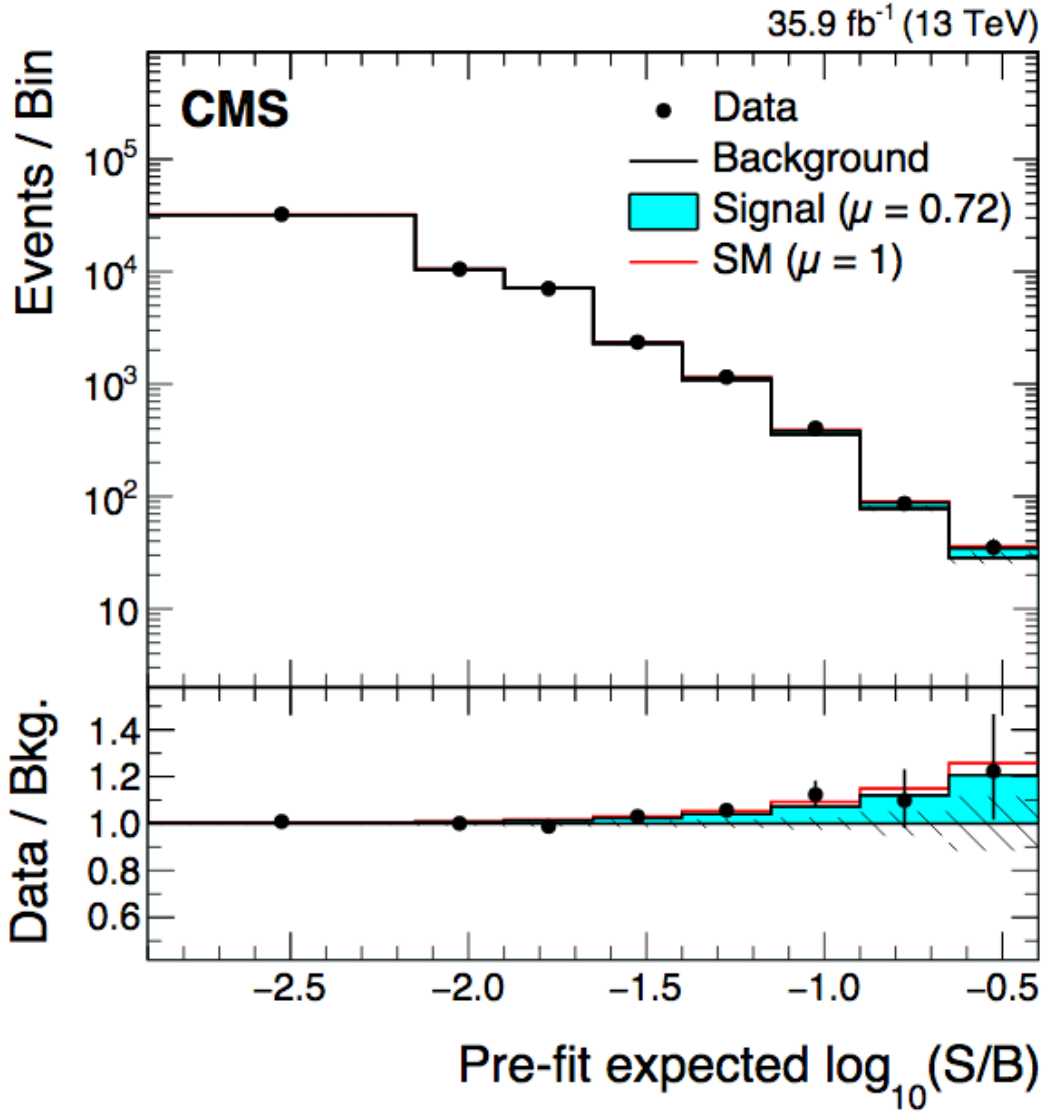
Selection	$1\ell + 2\tau_h$	4ℓ
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu$, $t \to bqq$, $H \to \tau\tau \to \tau_h\tau_h + \nu$'s	$t \to b\ell\nu$, $t \to b\ell\nu$, $H \to WW \to \ell\nu\ell\nu$
		$t \to b\ell\nu$, $t \to b\ell\nu$, $H \to ZZ \to \ell\ell qq$ or $\ell\ell\nu\nu$
Trigger	Single=lepton and lepton+τ_h triggers	Single-, double- and triple-lepton triggers
Lepton p_T	$p_T > 25$ (e) or 20 GeV (μ)	$p_T > 25$ / 15 / 15 / 10 GeV
Lepton η	$\lvert\eta\rvert < 2.1$	$\lvert\eta\rvert < 2.5$ (e) or 2.4 (μ)
τ_h p_T	$p_T > 30$ / 20 GeV	—
τ_h η	$\lvert\eta\rvert < 2.3$	—
Charge requirements	$\sum_{\tau_h} q = 0$ and $\sum_{\ell,\tau_h} q = \pm 1$	—
Jet multiplicity	$\geq$3 jets	$\geq$2 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	
Dilepton mass	$m_{\ell\ell} > 12$ GeV	$m_{\ell\ell} > 12$ GeV and $\lvert m_{\ell\ell} - m_Z\rvert > 10$ GeV ‡
Four-lepton mass	—	$m_{4\ell} > 140$ GeV §

BTD variables

Table 3: Observables used as input to the BDTs that separate the $t\bar{t}H$ signal from the $t\bar{t}V$ and $t\bar{t}$+jets backgrounds in the $1\ell + 2\tau_h$, 2ℓss, 3ℓ, and $3\ell + 1\tau_h$ categories.

Observable	$1\ell + 2\tau_h$	2ℓss	3ℓ	$3\ell + 1\tau_h$
$\Delta R(\ell_1, j)$	—	√	√	√
$\Delta R(\ell_2, j)$	—	√	√	√
$\langle \Delta R_{jj} \rangle$	√	—	—	√[2]
$\Delta R_{\tau\tau}$	√	—	—	—
$\max(\lvert\eta^{\ell 1}\rvert, \lvert\eta^{\ell 2}\rvert)$	—	√	√	√
H_T^{miss}	√	—	—	√[2]
N_j	√	√	√	√
N_b	√	—	—	—
$m_{\tau\tau}^{vis}$	√	—	—	—
$m_T^{\ell 1}$	—	√	√	√
$p_T^{\ell 1}$	—	√[1]	√[1]	√[1]
$p_T^{\ell 2}$	—	√[1]	-	-
$p_T^{\ell 3}$	—	—	√[1]	√[1]
$p_T^{\tau 1}$	√	—	—	—
$p_T^{\tau 2}$	√	—	—	—
$LR(3\ell)$	—	—	√[1]	—
MVA_{thad}^{max}	—	√[2]	—	—
MVA_{Hj}^{max}	—	√[1]	—	—

[1] Used only in BDT that separates $t\bar{t}H$ signal from $t\bar{t}V$ background.
[2] Used only in BDT that separates $t\bar{t}H$ signal from $t\bar{t}$+jets background.


35.9 fb-1 (13 TeV)
CMS
Data
Background
Signal (μ = 0.72)
SM (μ = 1)
Events / Bin
10^5
10^4
10^3
10^2
10
Data / Bkg.
1.4
1.2
1.0
0.8
0.6
−2.5
−2.0
−1.5
−1.0
−0.5
Pre-fit expected log10(S/B)

Source	Uncertainty [%]	$\Delta\mu/\mu$ [%]
e, μ selection efficiency	2–4	17
b tagging efficiency	2–15	6
Jet energy calibration	2–15	3
Forward jet modeling	10–35	3
Integrated luminosity	2.5	10
Reducible background estimate	10–40	14
Theoretical sources	≈10	14
$t\bar{t}$+HF normalization	≈50	7
PDFs	2–6	8