Search for Standard Model Higgs boson production in association with top quark pairs at CMS

Dr. Ram Krishna Dewanjee

Assistant Professor, Dept. Of Physics,

Birla Institute Of Technology, Mesra, Ranchi

(Formerly at KBFI, Estonia)

On behalf of the CMS collaboration

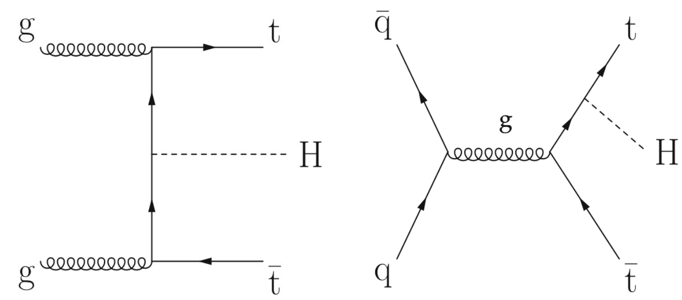
Motivation

- After 125 GeV Higgs discovery, LHC experiments focused on studying the properties of this new particle (spin, couplings etc.).
- Yukawa coupling b/w Higgs boson and top quark: important SM parameter that needs to be measured.
- Sensitive to BSM physics, may shed light on how top quark affects electroweak symmetry breaking.
- Although indirect constraints available (gluon fusion, H → $\gamma\gamma$ loop contribution), a direct measurement is possible only via studying top quark associated Higgs production ($t\bar{t}$H, tH).
- In addition, tH process is also sensitive to sign of top-Higgs Yukawa coupling.
- This overview covers all major Run-2 CMS tH/$t\bar{t}$H searches:
  $t\bar{t}$H (H → $b\bar{b}$),
  $t\bar{t}$H (H → $\gamma\gamma$),
  tH/$t\bar{t}$H (H → VV^***)** and **tH/$t\bar{t}$H (H → $\tau^+\tau^-$).**

(*V= W/Z boson)

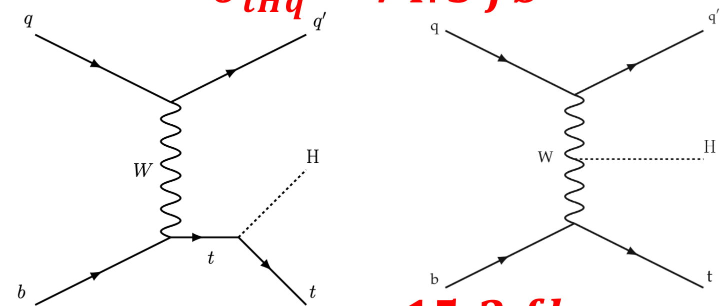
$\sigma_{ttH} = 506.5\ fb$

LO Feynman Diagrams for $t\bar{t}$H

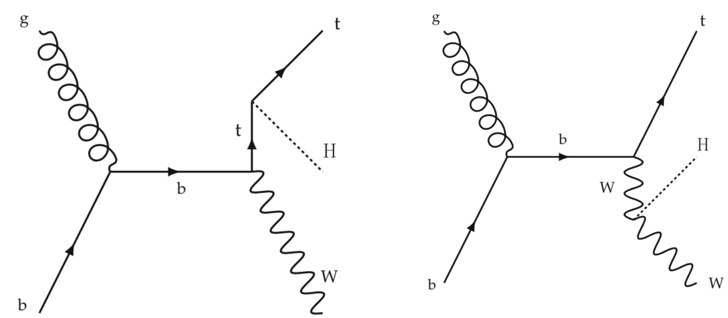


LO Feynman Diagrams for tH

$\sigma_{tHq} = 74.3\ fb$

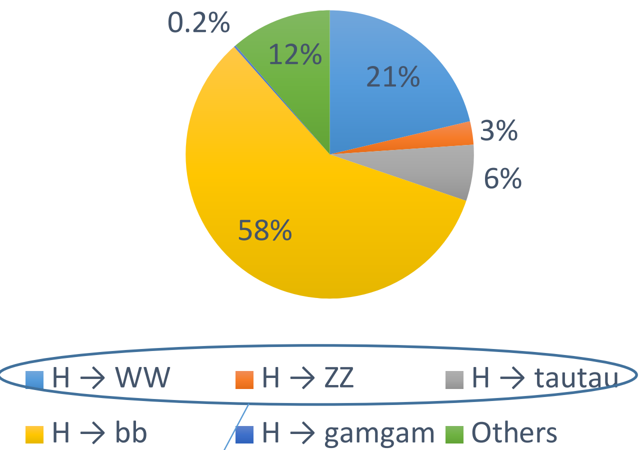


$\sigma_{tHW} = 15.2\ fb$



General strategy

Higgs Branching Fractions



Covered by tH/$t\bar{t}$H Multi-lepton analysis

- **$t\bar{t}$H (H→$b\bar{b}$)**
  - Large BR, sensitivity limited by systematic uncert. (on irreducible $t\bar{t} + b\bar{b}$ background).
  - Shape analysis with **DNN/MEM*** to achieve optimal S/B.
- **$t\bar{t}$H (H → $\gamma\gamma$)**
  - High signal purity, small BR => analysis sensitivity statistically limited.
  - Possible to probe CP structure of Higgs
- **tH/$t\bar{t}$H Multi-lepton**
  - Use of categories based on lepton and/or hadronic tau multiplicity.
  - Sufficient event yields, signal purities => most sensitive channel.
  - Mix of **H** → VV^* and **H** → $\tau^+\tau^-$ contributing to each category.

- Categorization used to separate events based on final states and event topology.
- Reducible backgrounds determined directly from data, irreducible backgrounds taken from MC and checked in data control regions.
- Signal-background separation in each category performed by shape analysis. Maximum Likelihood fit on the distribution of one "discriminating observable" (typically output of **BDT/DNN** @), performed for signal extraction.
- Analysis sensitivity improved by using discriminants computed with **MEM** , **BDT** or **DNN**#.

*** Matrix element method** **@ Boosted Decision tree** **# Deep Neural Network**

Analysis level objects

GENERIC GUIDELINES FOLLOWED IN MOST OF THE SEARCHES

- Particle Flow (**PF**) algorithm used for global event reconstruction: full use of sub-detector information.

  ***JINST* 12 P10003 (2017)**

- **Jets** reconstructed using anti-kT algorithm (R =0.4) using output of **PF** algorithm.

- **b-Jets** identified by DNN-based discriminator (**DeepCSV**) including impact parameter significance and track based lifetime information. Most analysis use working point corresponding to 70% (1%) b-tagging efficiency (light jet mis-tag rate).

- **Muons** identified by PF based selections designed to reduce muons fakes from pions/kaons and punch through hadrons.

- **Electrons** identified via multi-variate (**BDT**) based discriminators which include electron shower shape variables.

- **Hadronic Taus (τ_h)** identified via reconstructing the individual hadronic tau decay modes: 1Prong-0π_0, 1Prong-1π_0, 1Prong-2π_0 and 3Prong-0π_0 inside **PF** jets and qualifying strict isolation criteria. Dedicated **DNN** (DeepTau) based selections further reduce e $\rightarrow \tau_h$ and μ $\rightarrow \tau_h$ and quark/gluon jet $\rightarrow \tau_h$ fakes.

- **Photons (γ)** identified as ECAL energy clusters not linked to **PF** charged track. Dedicated shower shape-based clustering and **MVA** regression is used to recover the full energy of both converted and unconverted photons inside the detector.

$t\bar{t}$H (H→$b\bar{b}$): HIG-18-030

- Search performed in 2 distinct channels:

CMS-PAS-HIG-18-030

Hadronic channel

- Events divided into 6 categories depending on jet and b-jet multiplicity: **(7Jet, 3b-tag), (7Jet, ≥4b-tag), (8Jet, 3b-tag), (8Jet, ≥4b-tag), (≥ 9Jet, 3b-tag) and (≥ 9Jet, ≥ 4b-tag)**

- **MEM** employed in all categories to distinguish signal ($t\bar{t}$H) from background ($t\bar{t}$ + $b\bar{b}$) and used for signal extraction.

Leptonic channel

- Events divided into single-lepton and di-lepton (split by lepton flavor) channels.

- Di-lepton events divided into 5 categories: **(3Jets, 2b-tag), (3Jets, 3b-tag), (≥4Jets, 2b-tag), (≥4Jets, 3b-tag)** and **(≥4Jets, ≥4b-tag)** categories.

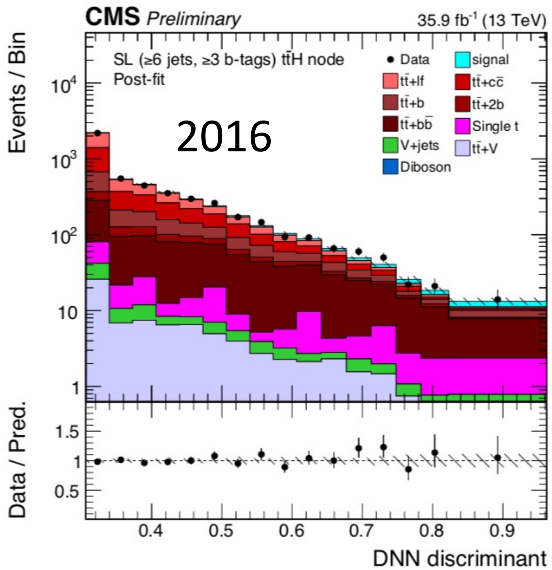
- **BDTs** used for signal extraction in all categories. For categories: **(≥4Jets, 3b-tag), (≥4Jets, ≥4b-tag)** the **MEM** discriminator was used as an input variable to the BDT.

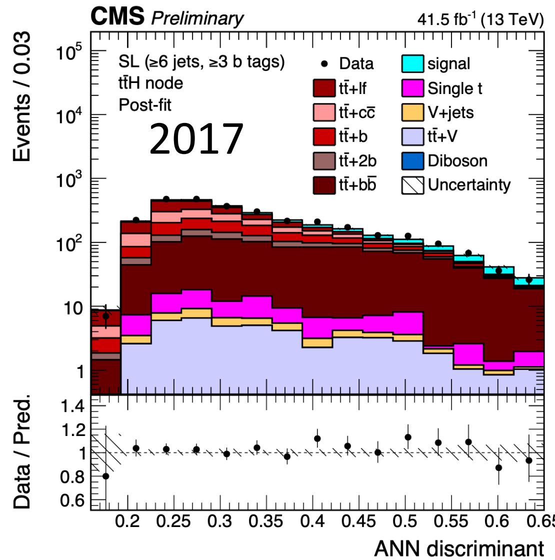
- Single lepton events split into: **(4Jets, ≥3b-tag), (5Jets, ≥3b-tag)** and **(≥6Jets, ≥3b-tag)** categories. Multi-Class **NN**s trained to distinguish signal ($t\bar{t}$H) from backgrounds ($t\bar{t}$+**X**) as well as for signal extraction.

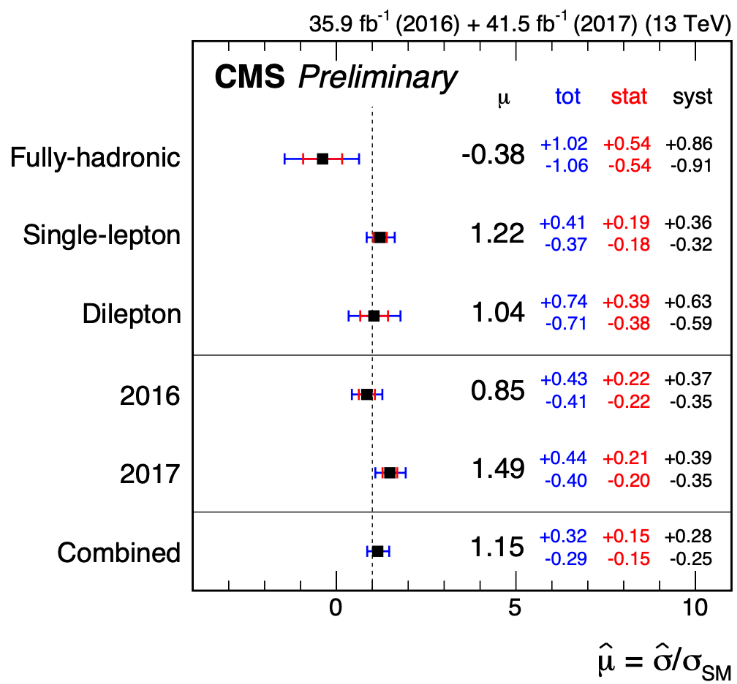
Full 2016+2017 data

X = b$\bar{b}$, c$\bar{c}$, 2b, b, light flavors.

$t\bar{t}$H (H→$b\bar{b}$): Results







Post-Fit Event Yields & Systematics in Backup (Slide-18,19)

Full Run-2 Legacy Result soon !!

$t\bar{t}$H (H → $\gamma\gamma$): HIG-19-013

Phys. Rev. Lett. 125, 061801

- Events triggered by asymmetric di-photon triggers with E_T thresholds of 30 and 18 GeV and loose ECAL based ID.
- Photon energy scale corrected, and resolution smeared to match data using Z→ e^+e^- events ($e^\pm$ reconstructed as γ).
- Photons entering analysis required to satisfy pre-selection criteria and **BDT** (trained to separate prompt photons from mis-identified jet fragments).
- Additional **BDT**s are used for di-photon vertex assignment and vertex probability computation.
- Selected events further split into **Leptonic** and **Hadronic** channels.
- Dedicated **BDT**s ("BDT-bkg") used per channel to distinguish b/w $t\bar{t}$**H (H** → $\gamma\gamma$) signal vs backgrounds: γ+Jets, $\gamma\gamma$+Jets, $t\bar{t}$+Jets, $t\bar{t}$+γ, $t\bar{t}$+$\gamma\gamma$, W/Z+γ. **All the non-$t\bar{t}$H Higgs production modes also treated as background.**
- Each channel specific "BDT-bkg" takes as input the following variables:
  - Photon kinematics
  - Lepton kinematics
  - Jet Kinematics
  - Output of the "top-tagger **BDT**" (used to separate top quarks decaying into 3 jets from non-top Bkg.)
  - Output of DNN (used to separate $t\bar{t}$H from **channel specific dominant Bkg.s***)

Full Run-2 data 137 fb-1

*** $\gamma\gamma$+Jets, $t\bar{t}$+$\gamma\gamma$ (Hadronic) / $t\bar{t}$+$\gamma\gamma$ (Leptonic)**

$t\bar{t}$H (H $\rightarrow \gamma\gamma$): Analysis strategy

- The main category-wise selections of this analysis are:
  - ***$t\bar{t}$*H Leptonic**: targeting semi-Leptonic top decays.
    - $\boldsymbol{p_T^{\gamma 1} > \frac{m_{\gamma\gamma}}{3}, p_T^{\gamma 2} > \frac{m_{\gamma\gamma}}{4}}$ (where 1/2 refer to leading/sub-leading photons)
    - 100 GeV < $m_{\gamma\gamma}$ < 180 GeV
    - Photon ID **BDT** score $\geq$ -0.7 (for both photons)
    - Both photons must pass the conversion safe electron veto (for e sub-channel only)
    - $\geq$ 1 isolated e (μ) satisfying p_T > 10 (5) GeV and $|\eta|$ < 2.4
    - $\geq$ 1 Jet satisfying p_T > 25 GeV and $|\eta|$ < 2.4
  - ***$t\bar{t}$*H Hadronic**: targeting hadronic top decays.
    - $\boldsymbol{p_T^{\gamma 1} > \frac{m_{\gamma\gamma}}{3}, p_T^{\gamma 2} > \frac{m_{\gamma\gamma}}{4}}$
    - 100 GeV < $m_{\gamma\gamma}$ < 180 GeV
    - Photon ID **BDT** score $\geq$ -0.7 (for both photons)
    - $\geq$ 3 Jets satisfying p_T > 25 GeV and $|\eta|$ < 2.4
    - $\geq$ 1 loose b-tagged Jet
    - Lepton veto (i.e., 0 leptons (e/μ) in the event)
- Events are further divided into 8 categories (based on BDT-bkg output) to maximize expected significance.
- To measure CP structure, selected events split into 12 categories to maximize sensitivity to CP structure of $t\bar{t}$H amplitude.
- Simultaneous binned maximum likelihood fit performed across all categories done for signal extraction.

$t\bar{t}$H (H → $\gamma\gamma$): Results

PROBING $t\bar{t}$H CP STRUCTURE

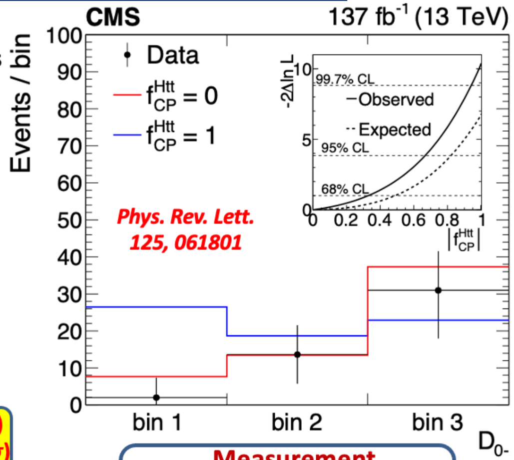
- BDT (D_{0-}) trained to distinguish CP-even and CP-odd contributions to $t\bar{t}$H process.
- $m_{\gamma\gamma}$ distribution fitted simultaneously across 12 categories to measure:

$$f_{CP}^{Htt} = \frac{|\tilde{\kappa}_t|^2}{|\kappa_t|^2 + |\tilde{\kappa}_t|^2}\,\mathrm{sign}(\tilde{\kappa}_t/\kappa_t)$$

CP-Even top-Higgs Yukawa coupling

CP-Odd top-Higgs Yukawa coupling

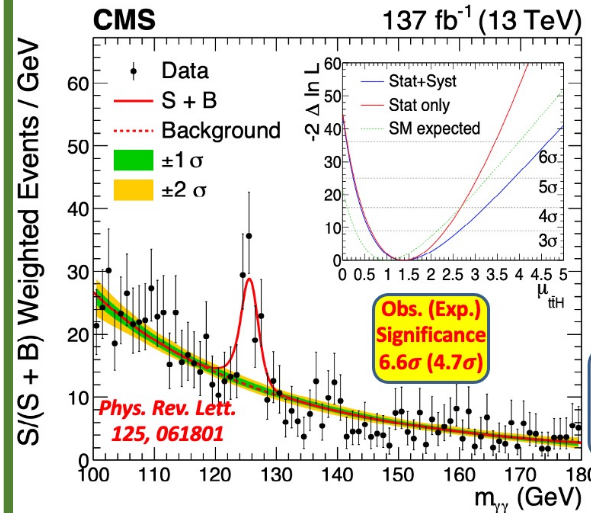
Pure Pseudoscalar Model ($f_{CP}^{Htt} = 1$) excluded at 3.2σ (exp. exclusion 2.6σ)



Measurement

$$f_{CP}^{Htt} = 0.00 \pm 0.49\ (@\ 68\%\ CL)$$
$$|f_{CP}^{Htt}| < 0.82\ (@\ 95\%\ CL)$$

$t\bar{t}$H CROSS-SECTION MEASUREMENT



Obs. (Exp.) Significance 6.6σ (4.7σ)

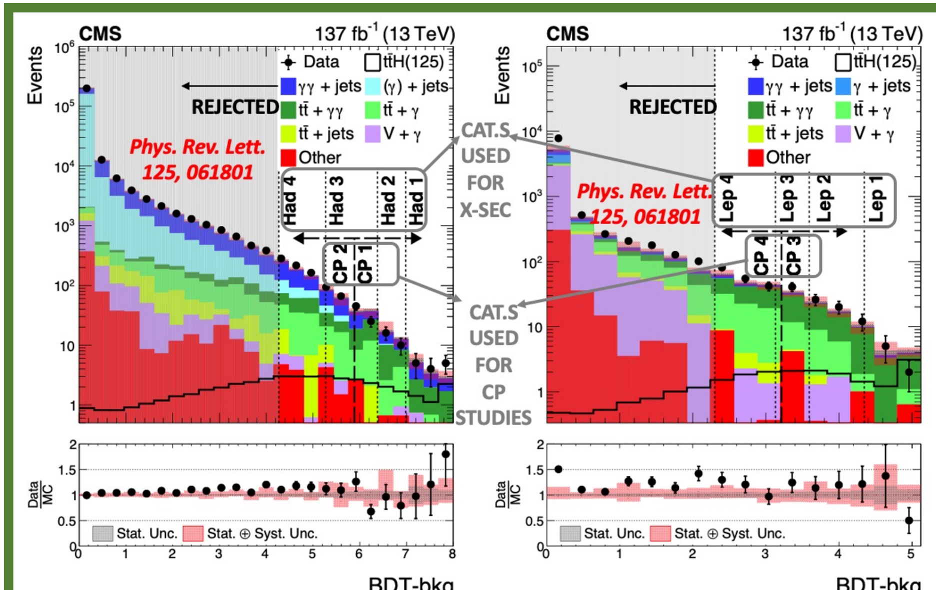
- **Signal modelling:** *Gaussian + Crystal Ball function*
- **Background modelling:** *Discrete Profiling Method*

Measured Signal Strength

$$\mu_{t\bar{t}H} = 1.38^{+0.36}_{-0.29}(\mathrm{stat.})^{+0.21}_{-0.11}(\mathrm{syst})$$

(Compatible with SM)

Post-Fit Event Yields & Systematics in Backup (Slide-20,21)



BDT Distributions for Hadronic (Left) and Leptonic (Right) channels

$t\bar{t}$H Multi-lepton (H → VV^*/(H → $\tau^+\tau^-$): HIG-19-008

EPJC 81:378 (2021)

- Events are triggered based on the presence of electrons/muons/τ_h using lepton or lepton + τ_h triggers.
- Selected events required to have at least 2 "loose" b-tagged jets[$] out of which at least one is medium b-tagged.
- Special **BDT** used to distinguish "prompt" leptons (produced by W/Z/leptonic τ decays) from "non-prompt" ones (produced in b-hadron decays, decays-in-flight and photon conversions).
- This **BDT** is trained on simulated $t\bar{t}$H ($t\bar{t}$) events as signal (background). Leptons passing (failing) it are called tight (loose) leptons.
- Events with pair of loose leptons having $m_{ll} < 12$ GeV rejected due to mis-modelling by simulation.
- Events categorized into 2 major categories and split further depending on lepton and/or τ_h multiplicity:
  - **Leptonic:** 8 channels targeting ≥ 1 leptonic $t/\bar{t}$ decays in the tH/$t\bar{t}$H event topology.
    $1l + 2\tau_h$, $2l$ss $+1\tau_h$, $2l$os $+ 1\tau_h$, $2l + 2\tau_h$, $3l + 1\tau_h$, $2l$ss $+0\tau_h$, $3l + 0\tau_h$, $4l + 0\tau_h$.
  - **Hadronic:** 2 channels targeting hadronic $t/\bar{t}$ decays in the tH/$t\bar{t}$H event topology.
    $0l + 2\tau_h$, $1l + 1\tau_h$

[$] $p_T > 25$ GeV, $|\eta| < 2.4$, $\Delta R(lep, Jet) > 0.4$
$l = e/\mu$
τ_h = hadronic tau decay
OS/SS = Opposite/Same-Sign

Full Run-2 data
137 fb-1

$t\bar{t}$H Multi-Lepton: Background Estimation

- Dominant reducible backgrounds arise from "non-prompt leptons"* passing "tight" selections and from lepton charge mis-identification.

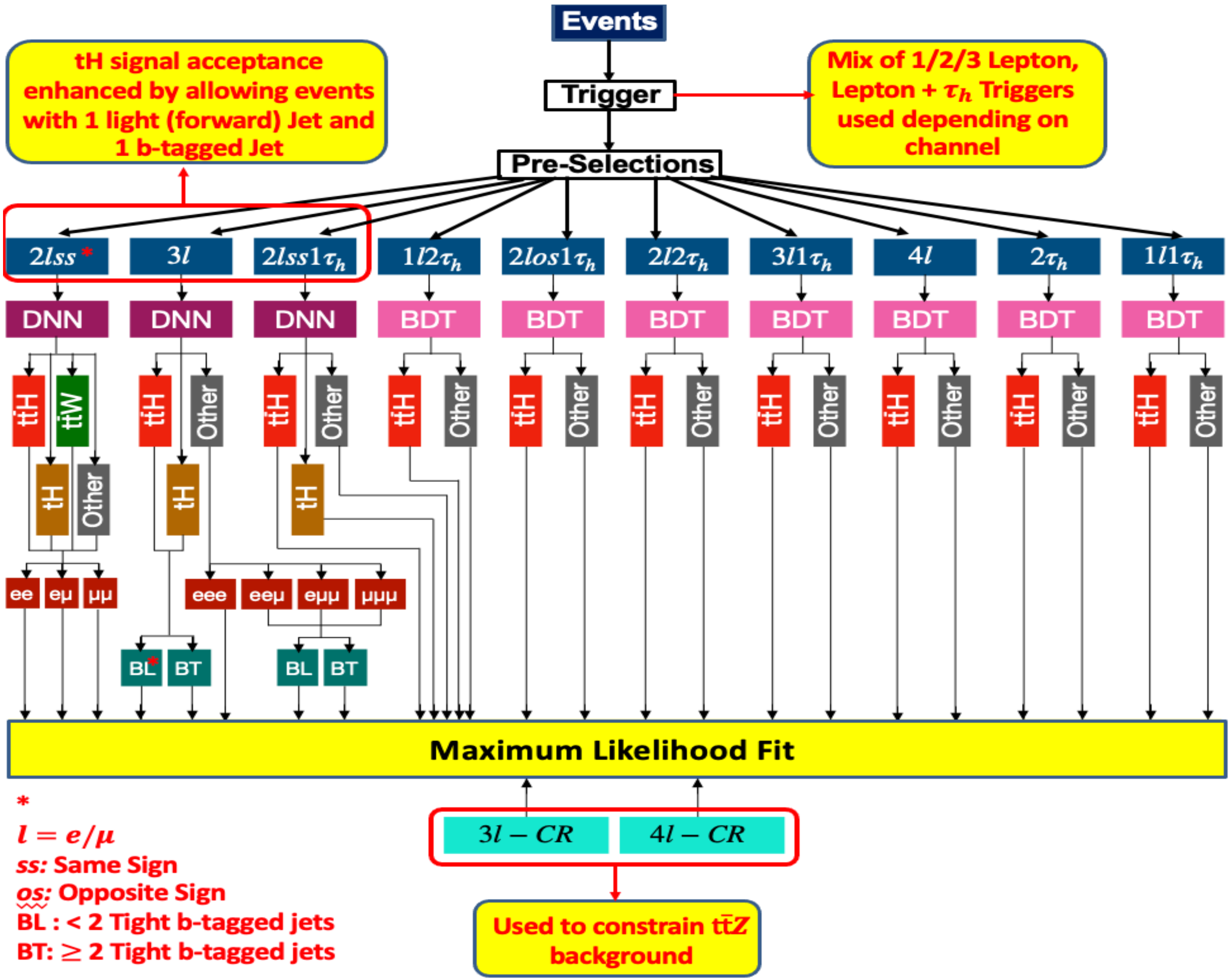
- Non-prompt lepton contribution estimated from data via fake rate method:
  - Measure "Probability(non-prompt lepton→ tight lepton)" in QCD enriched regions in data collected by lepton, lepton+jet triggers as function of lepton (p_T, η).
  - Weigh events in sidebands defined by relaxed lepton ID criteria (w.r.t. signal region) with the above probability to get the final background contribution.

- Lepton charge mis-id estimated from Z→ ll events in data:
  - Measure the "charge mis-id probability" in sample of SS di-lepton events compatible with Z-boson decay.
  - Measurement done in bins of lepton (p_T, η).
  - Weigh events with OS leptons in signal region with the above measured probability to get the background estimate.

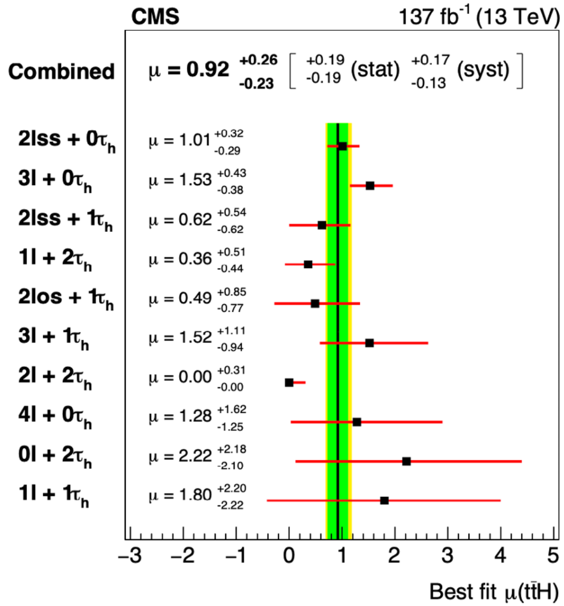
- Irreducible $t\bar{t}V$ and WZ/ZZ backgrounds modelled by MC and validated in control regions in data.

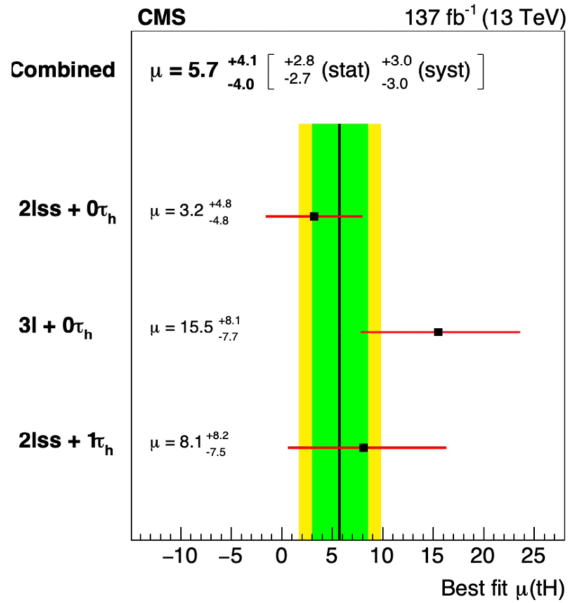
* Leptons not produced by W/Z decays e.g. leptons in semi-leptonic decays of b-hadrons

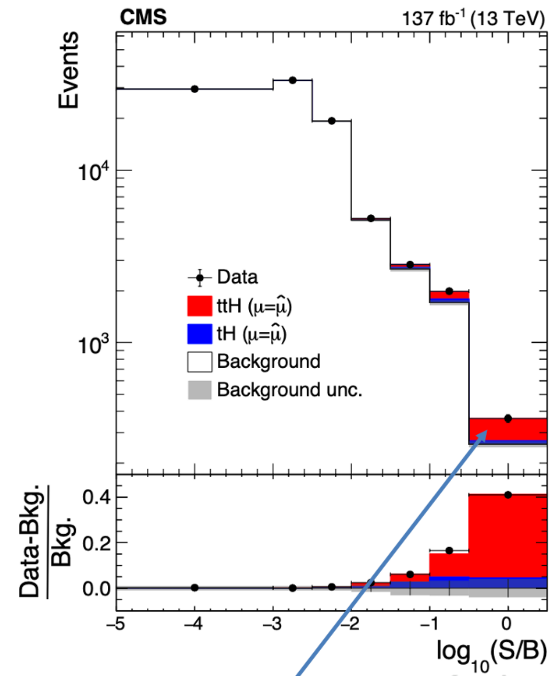
tH/$t\bar{t}$H Multi-Lepton: Signal Extraction



$tH/t\bar{t}H$ Multi-Lepton: Results

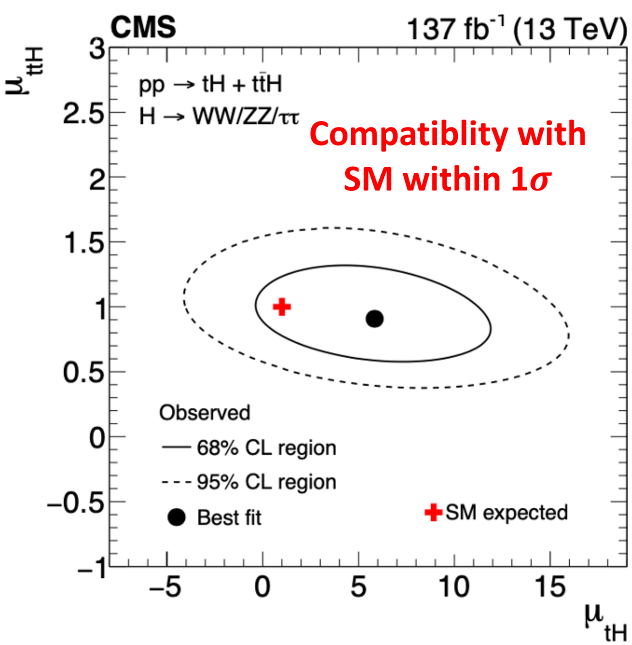


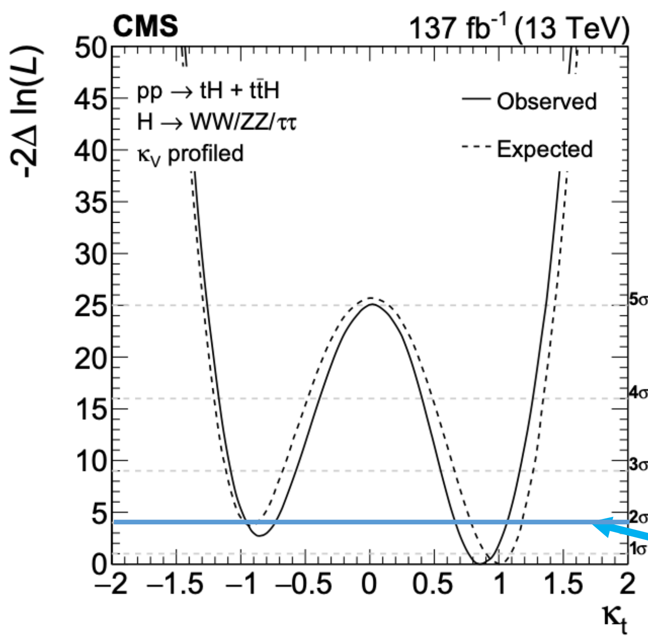




$t\bar{t}H$ **Obs. (Exp.) Significance: 4.7σ (5.2σ)**

tH **Obs. (Exp.) Significance: 1.4σ (0.3σ)**





Post-Fit Event Yields & Systematics in Backup (Slide-22,23)

κ_t constrained to be:

$-0.9 < \kappa_t < -0.7$ or $0.7 < \kappa_t < 1.1$ (@ 95% CL)

Summary

- Comprehensive overview of all CMS Run-2 $\mathrm{tH}/t\bar{t}$H searches till date was presented.

- First joint measurement of $t\bar{t}$H and tH cross-sections performed by CMS in the multi-lepton channel using full Run-2 data. Anomalous top-Higgs Yukawa coupling (κ_t= -1) disfavored within 2 sigma.

- Higgs CP measurements performed in the $t\bar{t}$**H** (**H** $\rightarrow \gamma\gamma$) channel which exclude the pseudo-scalar (CP = 1) hypothesis by 3.2 sigma.

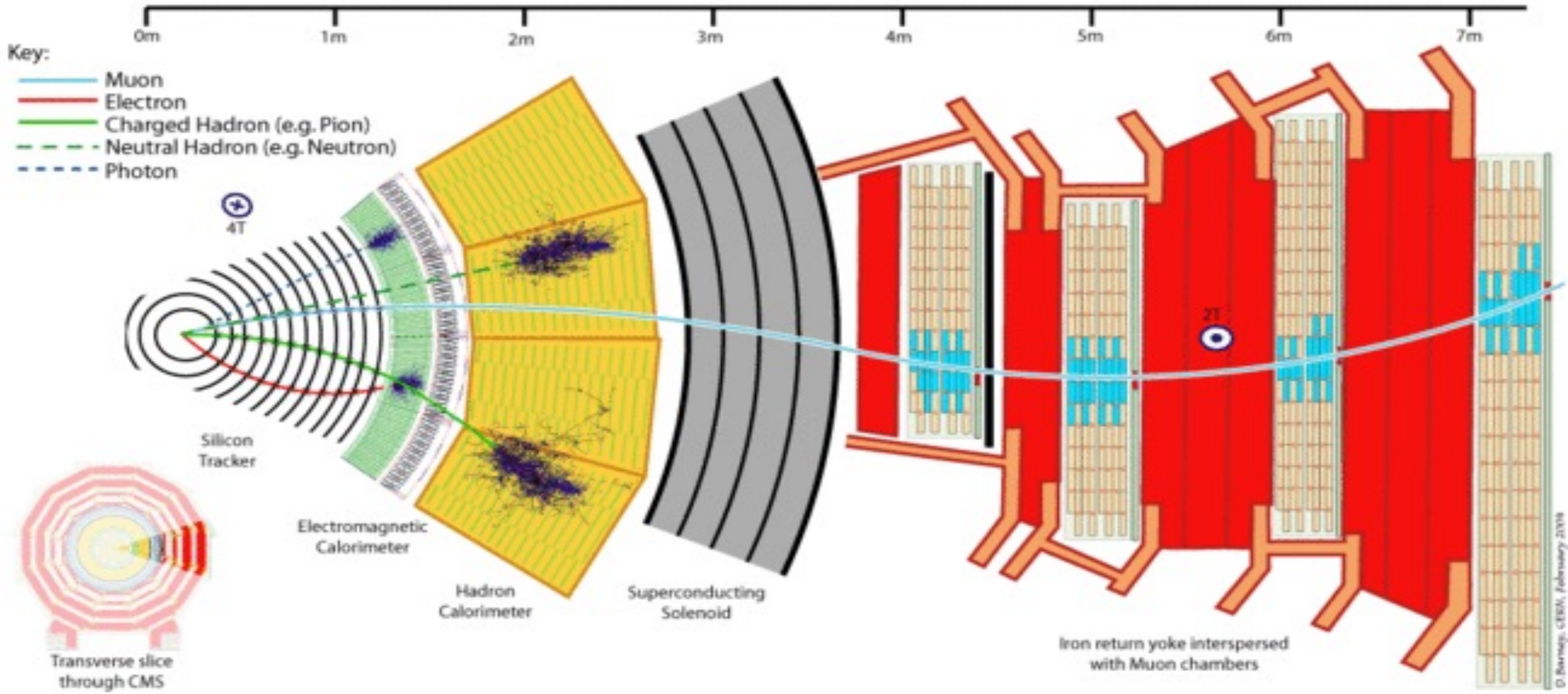
- Work currently ongoing for $t\bar{t}$**H (H→**$b\bar{b}$**)** Run-2 legacy result with more data and improved experimental methods to further increase analysis sensitivity in this channel.

- With 2022 data taking (at $\sqrt{s}$ = 13.6 TeV) recently concluded, looking forward to improve upon these results and to throw more light on the top-Higgs interaction in the future.

THANK YOU

Backup Slides

EVENT RECONSTRUCTION (PARTICLE FLOW)



- Muons: Tracker hits, Calo Energy deposits (ECAL + HCAL), Muon chamber hits
- Charged Hadrons: Tracker hits, Calo Energy deposits (ECAL + HCAL)
- Electron/Photon (Converted): Tracker hits, Calo Energy deposits (ECAL)
- Neutral Hadron: Calo Energy deposits (ECAL + HCAL)
- Photon (Unconverted): Calo Energy deposits (ECAL)

$t\bar{t}$H (H→$b\bar{b}$): Event Yields

CMS-PAS-HIG-18-030

Process	FH channel	SL channel	DL channel
QCD	2938305 ± 301286	—	—
$t\bar{t}$+lf	357488 ± 52694	718341 ± 83944	275407 ± 19610
$t\bar{t}$+$c\bar{c}$	93674 ± 11860	96581 ± 13795	24721 ± 3145
$t\bar{t}$+b	23737 ± 2892	27222 ± 3749	5613 ± 824
$t\bar{t}$+2b	14039 ± 2183	10537 ± 2206	1697 ± 351
$t\bar{t}$+$b\bar{b}$	19730 ± 2413	12770 ± 2050	1813 ± 262
Single t	24117 ± 1847	38170 ± 3720	14044 ± 1133
V+jets	31154 ± 2319	14491 ± 1754	2199 ± 264
$t\bar{t}$+V	2924 ± 228	2963 ± 286	1028 ± 99
Diboson	354 ± 40	503 ± 61	420 ± 55
Total bkg.	3505523 ± 339615	921576 ± 97714	326942 ± 23458
$t\bar{t}$H	2556 ± 164	1747 ± 167	363 ± 22
Data	3508079	923936	331055

$t\bar{t}$H (H→$b\bar{b}$): Systematics

Table 6: Contributions of different sources of uncertainties to the result for the combined fit to the 2016 and 2017 datasets. The quoted uncertainties $\Delta\hat{\mu}$ in $\hat{\mu}$ are obtained by fixing the listed sources of uncertainties to their post-fit values in the fit and subtracting the obtained result in quadrature from the result of the full fit. The statistical uncertainty is evaluated by fixing all nuisance parameters to their post-fit values. The quadratic sum of the contributions is different from the total uncertainty because of correlations between the nuisance parameters.

Uncertainty source	$\Delta\hat{\mu}$
Total experimental	$+0.15/-0.13$
b tagging	$+0.08/-0.07$
jet energy scale and resolution	$+0.05/-0.04$
Total theory	$+0.23/-0.19$
signal	$+0.15/-0.06$
$t\bar{t}$+hf modelling	$+0.14/-0.15$
QCD background prediction	$+0.10/-0.08$
Size of simulated samples	$+0.10/-0.10$
Total systematic	$+0.28/-0.25$
Statistical	$+0.15/-0.15$
Total	$+0.32/-0.29$

CMS-PAS-HIG-18-030

$t\bar{t}$H (H $\rightarrow \gamma\gamma$): Event Yields

ttH Hadronic: Signal Regions

Process	Yield	$\mathcal{F}$ of bkg
$\gamma\gamma$ + Jets	343.67 ± 5.29	0.47
(γ) + Jets (Data)	96.38 ± 12.26	0.13
$t\bar{t}+\gamma\gamma$	133.16 ± 13.48	0.18
$t\bar{t}+\gamma$ + Jets	113.19 ± 15.12	0.15
$t\bar{t}$ + Jets	19.30 ± 9.25	0.03
Drell-Yan	-2.52 ± 2.52	-0.00
V + γ	6.70 ± 2.68	0.01
$t+\gamma$	10.26 ± 4.05	0.01
$t\bar{t}+V$	1.18 ± 0.09	0.00
VV	0.00 ± 0.00	0.00
tV	8.93 ± 1.92	0.01
ggH	4.58 ± 0.28	0.01
VBF	0.16 ± 0.02	0.00
VH	1.42 ± 0.06	0.00
tHq	1.57 ± 0.00	0.00
tHW	0.78 ± 0.00	0.00
All bkg.	738.77 ± 26.60	1.00
Data	800.00 ± 28.28	1.08
ttH	27.59 ± 0.24	0.04

ttH Leptonic: Signal Regions

Process	Yield	$\mathcal{F}$ of bkg
$\gamma\gamma$ + Jets	20.19 ± 1.56	0.08
γ + Jets	0.00 ± 0.00	0.00
$t\bar{t}+\gamma\gamma$	114.66 ± 5.08	0.48
$t\bar{t}+\gamma$ + Jets	74.50 ± 4.74	0.31
$t\bar{t}$ + Jets	5.85 ± 1.17	0.02
Drell-Yan	-0.35 ± 0.35	-0.00
V + γ	9.49 ± 7.56	0.04
$t+\gamma$	1.95 ± 2.76	0.01
$t\bar{t}+V$	2.44 ± 0.13	0.01
VV	2.04 ± 0.75	0.01
tV	6.86 ± 1.38	0.03
ggH	0.11 ± 0.05	0.00
VBF	0.00 ± 0.01	0.00
VH	1.15 ± 0.06	0.00
tHq	0.81 ± 0.00	0.00
tHW	0.59 ± 0.00	0.00
All bkg.	240.28 ± 10.93	1.00
Data	271.00 ± 16.46	1.13
ttH	17.13 ± 0.18	0.07

Table 12: Background composition in the signal regions for the Hadronic channel (left) and the Leptonic channel (right).

TABLE I. The expected number of H events in the hadronic and leptonic channels per category and the fractional contribution per H production mode.

	Total	$t\bar{t}H$ (%)	tH (%)	ggH (%)	VH (%)	VBF (%)	$b\bar{b}H$ (%)
Had1	5.8	89.1	6.8	3.3	0.8	< 0.1	0.1
Had2	4.2	82.9	6.8	8.7	1.4	0.2	0.1
Had3	11.6	78.6	7.2	10.3	3.5	0.3	0.1
Had4	13.6	65.4	7.7	19.3	6.9	0.7	0.1
Lep1	5.8	90.6	7.9	0.5	1.0	< 0.1	< 0.1
Lep2	4.9	90.0	6.7	0.4	2.9	< 0.1	< 0.1
Lep3	3.5	86.2	7.4	0.4	6.0	< 0.1	< 0.1
Lep4	5.7	78.1	8.2	1.1	12.7	< 0.1	< 0.1
Total	55.1	79.5	7.4	8.2	4.7	0.3	< 0.1

$t\bar{t}$H (H $\rightarrow \gamma\gamma$): Systematics

Name	Type	Value/Comment
QCD Scale uncertainty	Theoretical	8%
PDF uncertainty	Theoretical	Category migrations due to this < 1%
α_S uncertainty	Theoretical	2.6%
BR(H $\rightarrow \gamma\gamma$) uncertainty	Theoretical	2%
ggH contamination	Theoretical	2%
b-tagging	Experimental	4%
Photon ID uncertainty	Experimental	1%
JES/JER Uncertainty	Experimental	2%
Luminosity uncertainty	Experimental	1.8%
L1 Pre-fire probability uncertainty	Experimental	20%
Top tagger & DNN O/p uncertainty	Experimental	Not applied since good Data/MC agreement in Input Var.s & discrepancies (if any) covered by above uncert.s

$tH/t\bar{t}H$ Multi-Lepton: Event Yields-1

Process	$2\ell SS + 0\tau_h$	$3\ell + 0\tau_h$	$2\ell SS + 1\tau_h$
$t\bar{t}H$	222 ± 51	61 ± 15	28.9 ± 6.4
tH	119 ± 85	20 ± 14	12.7 ± 9.0
$t\bar{t}Z + t\bar{t}\gamma^*$	322 ± 25	145 ± 11	29.6 ± 3.3
$t\bar{t}W + t\bar{t}WW$	1153 ± 64	171.1 ± 9.5	47.4 ± 6.5
WZ	296 ± 31	89.7 ± 9.7	19.4 ± 2.9
ZZ	31.2 ± 3.3	16.2 ± 1.6	1.6 ± 0.3
Misidentified leptons	1217 ± 91	140 ± 11	52.0 ± 9.6
Flips	121 ± 19	–	–
Rare backgrounds	222 ± 48	41.0 ± 8.9	13.3 ± 3.1
Conversion	42 ± 12	5.6 ± 1.6	–
ggH + qqH + VH + $t\bar{t}$VH	35.3 ± 4.0	3.4 ± 0.3	1.8 ± 0.3
Total expected background	3517 ± 85	627 ± 20	179 ± 13
Data	3738	744	201

Process	$1\ell + 1\tau_h$	$0\ell + 2\tau_h$	$2\ell OS + 1\tau_h$	$1\ell + 2\tau_h$
$t\bar{t}H$	183 ± 41	24.4 ± 6.0	19.1 ± 4.3	19.3 ± 4.2
tH	65 ± 46	16 ± 12	4.8 ± 3.4	2.6 ± 1.9
$t\bar{t}Z + t\bar{t}\gamma^*$	203 ± 24	27.1 ± 3.8	25.5 ± 2.9	20.3 ± 2.1
$t\bar{t}W + t\bar{t}WW$	254 ± 34	3.8 ± 0.5	17.4 ± 2.4	2.6 ± 0.4
WZ	198 ± 37	42.5 ± 8.7	8.4 ± 1.6	11.8 ± 2.2
ZZ	98 ± 13	34.2 ± 4.8	1.9 ± 0.3	1.8 ± 0.3
DY	4480 ± 460	1430.0 ± 220	519 ± 28	250 ± 16
$t\bar{t}$+jets	41900 ± 1900	861 ± 98	–	–
Misidentified leptons	25300 ± 1900	3790 ± 220	–	–
Rare backgrounds	1930 ± 420	60 ± 14	5.9 ± 1.3	5.6 ± 1.3
Conversion	–	–	0.5 ± 0.2	–
ggH + qqH + VH + $t\bar{t}$VH	38.5 ± 3.6	26.7 ± 3.6	0.8 ± 0.1	–
Total expected background	73550 ± 610	6290 ± 130	584 ± 27	295 ± 16
Data	73736	6310	603	307

Process	$4\ell + 0\tau_h$	$3\ell + 1\tau_h$	$2\ell + 2\tau_h$
$t\bar{t}H$	2.0 ± 0.5	4.0 ± 0.9	2.2 ± 0.5
tH	0.2 ± 0.2	0.8 ± 0.6	0.3 ± 0.2
$t\bar{t}Z + t\bar{t}\gamma^*$	5.9 ± 0.4	6.6 ± 0.7	2.5 ± 0.3
$t\bar{t}W + t\bar{t}WW$	0.2 ± 0.0	1.1 ± 0.2	–
ZZ	0.6 ± 0.2	0.3 ± 0.1	0.2 ± 0.0
Misidentified leptons	–	1.5 ± 0.9	3.4 ± 0.9
Rare backgrounds	0.6 ± 0.1	1.0 ± 0.3	0.3 ± 0.1
Conversion	–	–	–
Total expected background	7.4 ± 0.5	11.5 ± 1.3	6.8 ± 1.0
Data	12	18	3

$tH/t\bar{t}H$ Multi-Lepton: Event Yields-2

Process	3ℓ-CR	4ℓ-CR	$t\bar{t}W(W)$ CR
$t\bar{t}H$	15.9 ± 4.4	1.4 ± 0.4	62 ± 14
tH	4.4 ± 3.0	–	22 ± 18
$t\bar{t}Z + t\bar{t}\gamma^*$	550 ± 43	41.5 ± 3.0	100.3 ± 8.1
$t\bar{t}W + t\bar{t}WW$	26.8 ± 1.7	–	588 ± 35
WZ	4320 ± 120	–	51.6 ± 7.5
ZZ	298 ± 18	1030 ± 32	0.2 ± 0.1
Nonprompt leptons	210 ± 20	–	102 ± 14
Flips	–	–	24.9 ± 4.0
Rare backgrounds	311 ± 61	17.0 ± 3.4	58 ± 13
Conversions	1.0 ± 0.3	0.1 ± 0.1	1.4 ± 0.6
$ggH + qqH + VH + t\bar{t}VH$	42.8 ± 3.1	5.8 ± 0.4	1.6 ± 0.3
Total expected background	5761 ± 99	1094 ± 33	949 ± 33
Data	5778	1089	986

$tH/t\bar{t}H$ Multi-Lepton: Systematics

Source	$\Delta\mu_{t\bar{t}H}/\mu_{t\bar{t}H}$ (%)	$\Delta\mu_{tH}/\mu_{tH}$ (%)	$\Delta\mu_{t\bar{t}W}/\mu_{t\bar{t}W}$ (%)	$\Delta\mu_{t\bar{t}Z}/\mu_{t\bar{t}Z}$ (%)
Trigger efficiency	2.3	8.1	1.2	1.9
e, μ reconstruction and identification efficiency	2.9	7.1	1.7	3.2
τ_h identification efficiency	4.6	9.1	1.7	1.3
b tagging efficiency and mistag rate	3.6	13.6	1.3	2.9
Misidentified leptons and flips	6.0	36.8	2.6	1.4
Jet energy scale and resolution	3.4	8.3	1.1	1.2
MC sample and sideband statistical uncertainty	7.1	27.2	2.4	2.3
Theory-related sources affecting acceptance and shape of distributions	4.6	18.2	2.0	4.2
Normalization of MC-estimated processes	13.3	12.3	13.9	11.3
Integrated luminosity	2.2	4.6	1.8	3.1
Statistical uncertainty	20.9	48.0	5.9	5.8

$tH/t\bar{t}H$ Multi-Lepton: Pre-selections

Selection step	$2\ell SS + 0\tau_h$	$2\ell SS + 1\tau_h$
Targeted $t\bar{t}H$ decay	$t \to b\ell\nu, t \to bqq'$ with $H \to WW \to \ell\nu qq'$	$t \to b\ell\nu, t \to bqq'$ with $H \to \tau\tau \to \ell\nu\nu\tau_h\nu$
Targeted tH decays	$t \to b\ell\nu,$ $H \to WW \to \ell\nu qq'$	$t \to b\ell\nu,$ $H \to \tau\tau \to \ell\tau_h + \nu$'s
Trigger	Single- and double-lepton triggers	Single- and double-lepton triggers
Lepton p_T	$p_T > 25 / 15$ GeV	$p_T > 25 / 15$ GeV (e) or 10 GeV (μ)
Lepton η	$\|\eta\| < 2.5$ (e) or 2.4 (μ)	$\|\eta\| < 2.5$ (e) or 2.4 (μ)
τ_h p_T	–	$p_T > 20$ GeV
τ_h η	–	$\|\eta\| < 2.3$
τ_h identification	–	Very loose
Charge requirements	2 SS leptons and charge quality requirements	2 SS leptons and charge quality requirements $\sum_{\ell,\tau_h} q = \pm 1$
Multiplicity of central jets	$\geq$3 jets	$\geq$3 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets
Missing transverse momentum	$L_D > 30\,\text{GeV}^{\dagger}$	$L_D > 30\,\text{GeV}^{\dagger}$
Dilepton invariant mass	$\|m_{\ell\ell} - m_Z\| > 10\,\text{GeV}^{\ddagger}$ and $m_{\ell\ell} > 12$ GeV	

Selection step	$3\ell + 0\tau_h$	$3\ell + 1\tau_h$
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu, t \to b\ell\nu$ with $H \to WW \to \ell\nu qq'$ $t \to b\ell\nu$, $t \to bqq'$ with $H \to WW \to \ell\nu\ell\nu$ $t \to b\ell\nu, t \to bqq'$ with $H \to ZZ \to \ell\ell qq'$ or $\ell\ell\nu\nu$	$t \to b\ell\nu, t \to b\ell\nu$ with $H \to \tau\tau \to \ell\nu\nu\tau_h\nu$
Targeted tH decays	$t \to b\ell\nu, H \to WW \to \ell\nu\ell\nu$	–
Trigger	Single-, double- and triple-lepton triggers	Single-, double- and triple-lepton triggers
Lepton p_T	$p_T > 25 / 15 / 10$ GeV	$p_T > 25 / 15 / 10$ GeV
Lepton η	$\|\eta\| < 2.5$ (e) or 2.4 (μ)	$\|\eta\| < 2.5$ (e) or 2.4 (μ)
τ_h p_T	–	$p_T > 20$ GeV
τ_h η	–	$\|\eta\| < 2.3$
τ_h identification	–	Very loose
Charge requirements	$\sum_{\ell} q = \pm 1$	$\sum_{\ell,\tau_h} q = 0$
Multiplicity of central jets	$\geq$2 jets	$\geq$2 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets
Missing transverse momentum	$L_D > 0/30/45\,\text{GeV}^{\ddagger}$	$L_D > 0/30/45\,\text{GeV}^{\ddagger}$
Dilepton invariant mass	$m_{\ell\ell} > 12$ GeV and $\|m_{\ell\ell} - m_Z\| > 10\,\text{GeV}^{\S}$	$m_{\ell\ell} > 12$ GeV and $\|m_{\ell\ell} - m_Z\| > 10\,\text{GeV}^{\S}$
Four-lepton invariant mass	$m_{4\ell} > 140\,\text{GeV}^{\P}$	–

‡ Applied to all SFOS lepton pairs and to pairs of electrons of SS charge
§ Applied to all SFOS lepton pairs
¶ If the event contains two SFOS pairs of leptons that pass the loose lepton selection criteria

Selection step	$0\ell + 2\tau_h$	$1\ell + 1\tau_h$
Targeted $t\bar{t}H$ decays	$t \to bqq', t \to bqq'$ with $H \to \tau\tau \to \tau_h\nu\tau_h\nu$	$t \to bqq', t \to bqq'$ with $H \to \tau\tau \to \ell\nu\nu\tau_h\nu$
Trigger	Double-τ_h trigger	Single-lepton and lepton+τ_h triggers
Lepton p_T	–	$p_T > 30$ (e) or 25 GeV (μ)
Lepton η	–	$\|\eta\| < 2.1$
τ_h p_T	$p_T > 40$ GeV	$p_T > 30$ GeV
τ_h η	$\|\eta\| < 2.1$	$\|\eta\| < 2.1$
τ_h identification	Loose	Medium
Charge requirements	$\sum_{\tau_h} q = 0$	$\sum_{\ell,\tau_h} q = 0$
Multiplicity of central jets	$\geq$4 jets	$\geq$4 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets
Dilepton invariant mass	$m_{\ell\ell} > 12$ GeV	$m_{\ell\ell} > 12$ GeV

Selection step	$1\ell + 2\tau_h$	$2\ell + 2\tau_h$
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu, t \to bqq'$ with $H \to \tau^+\tau^- \to \tau_h\nu\tau_h\nu$	$t \to b\ell\nu, t \to b\ell\nu$ with $H \to \tau^+\tau^- \to \tau_h\nu\tau_h\nu$
Trigger	Single-lepton and lepton+τ_h triggers	Single- and double-lepton triggers
Lepton p_T	$p_T > 30$ (e) or 25 GeV (μ)	$p_T > 25 / 10$ (15) GeV (e)
Lepton η	$\|\eta\| < 2.1$	$\|\eta\| < 2.5$ (e) or 2.4 (μ)
τ_h p_T	$p_T > 30 / 20$ GeV	$p_T > 20$ GeV
τ_h η	$\|\eta\| < 2.1$	$\|\eta\| < 2.3$
τ_h identification	medium	medium
Charge requirements	$\sum_{\ell,\tau_h} q = \pm 1$	$\sum_{\ell,\tau_h} q = 0$
Multiplicity of central jets	$\geq$3 jets	$\geq$2 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets
Missing transverse momentum	–	$L_D > 0 / 30 / 45\,\text{GeV}^{\dagger}$
Dilepton invariant mass	$m_{\ell\ell} > 12$ GeV	$m_{\ell\ell} > 12$ GeV

Selection step	$2\ell OS + 1\tau_h$	$4\ell + 0\tau_h$
Targeted $t\bar{t}H$ decays	$t \to b\ell\nu, t \to bqq'$ with $H \to \tau^+\tau^- \to \ell\nu\nu\tau_h\nu$	$t \to b\ell\nu, t \to b\ell\nu$ with $H \to WW \to \ell\nu\ell\nu$ $t \to b\ell\nu, t \to b\ell\nu$ with $H \to ZZ \to \ell\ell qq'$ or $\ell\ell\nu\nu$
Trigger	Single- and double-lepton triggers	Single-, double- and triple-lepton triggers
Lepton p_T	$p_T > 25 / 15$ GeV (e) or 10 GeV (μ)	$p_T > 25 / 15 / 15 / 10$ GeV
Lepton η	$\|\eta\| < 2.5$ (e) or 2.4 (μ)	$\|\eta\| < 2.5$ (e) or 2.4 (μ)
τ_h p_T	$p_T > 20$ GeV	–
τ_h η	$\|\eta\| < 2.3$	–
τ_h identification	Tight	–
Charge requirements	$\sum_{\ell} q = 0$ and $\sum_{\ell,\tau_h} q = \pm 1$	$\sum_{\ell} q = 0$
Multiplicity of central jets	$\geq$3 jets	$\geq$2 jets
b tagging requirements	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets	$\geq$1 tight b-tagged jet or $\geq$2 loose b-tagged jets
Missing transverse momentum	$L_D > 30\,\text{GeV}^{\dagger}$	$L_D > 0 / 30 / 45\,\text{GeV}^{\ddagger}$
Dilepton invariant mass	$m_{\ell\ell} > 12$ GeV	$\|m_{\ell\ell} - m_Z\| > 10\,\text{GeV}^{\S}$ and $m_{\ell\ell} > 12$ GeV
Four-lepton invariant mass	–	$m_{4\ell} > 140\,\text{GeV}^{\P}$

$tH/t\bar{t}H$ Multi-Lepton: DNN Input Variables

	$2\ell SS + 0\tau_h$	$2\ell SS + 1\tau_h$	$3\ell + 0\tau_h$	$1\ell + 1\tau_h$	$0\ell + 2\tau_h$	$2\ell OS + 1\tau_h$	$1\ell + 2\tau_h$	$4\ell + 0\tau_h$	$3\ell + 1\tau_h$	$2\ell + 2\tau$
Electron multiplicity	✓	✓	✓	–	–	–	–	–	–	–
Three-momenta of leptons and/or τ_hs	✓	✓	✓	✓	✓	✓	✓	–	✓	✓
p_T of leptons and/or τ_hs	–	–	–	–	–	–	–	✓	–	–
Transverse mass of leptons and/or τ_hs	✓	✓	–	✓	✓	✓	✓	–	–	–
Invariant mass of leptons and/or τ_hs	✓	–	–	✓	✓	✓	✓	✓	✓	✓
SVFit mass of leptons and/or τ_hs	–	–	–	✓	✓	–	–	–	–	–
ΔR between leptons and/or τ_hs	✓	✓	✓	✓	✓	✓	✓	–	–	✓
$\cos\theta^*$ of leptons and τ_hs	–	–	–	✓	✓	–	✓	–	–	✓
Charge of leptons and/or τ_hs	✓	✓	✓	✓	–	–	–	–	–	–
Has SFOS lepton pairs	–	–	✓	–	–	–	–	✓	✓	–
Jet multiplicity	✓	✓	✓	–	–	–	–	–	–	–
Jets three-momenta	✓	✓	✓	–	–	–	–	–	–	–
Average ΔR between jets	✓	✓	✓	✓	✓	✓	✓	–	–	✓
Forward jet multiplicity	✓	✓	✓	–	–	–	–	–	–	–
Leading forward jet three-momenta	✓	✓	✓	–	–	–	–	–	–	–
Minimum $\lvert\Delta\eta\rvert$ between leading forward jet and jets	–	✓	✓	–	–	–	–	–	–	–
b jet multiplicity	✓	✓	✓	–	–	–	–	–	–	–
Invariant mass of b jets	✓	✓	✓	✓	✓	✓	✓	–	–	✓
Linear discriminant L_D	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hadronic top quark tagger	✓	✓	✓	✓	✓	✓	✓	–	–	–
Hadronic top p_T	–	✓	✓	–	–	✓	✓	–	–	–
Higgs boson jet tagger	✓	–	–	–	–	–	–	–	–	–
Number of variables	36	41	37	16	15	18	17	7	9	9