Karlsruhe Institute of Technology

Results for ttH (CMS)

A. Gilbert
on behalf of the CMS Collaboration

Higgs Coupling 2015, Lumley Castle | 13 October 2015

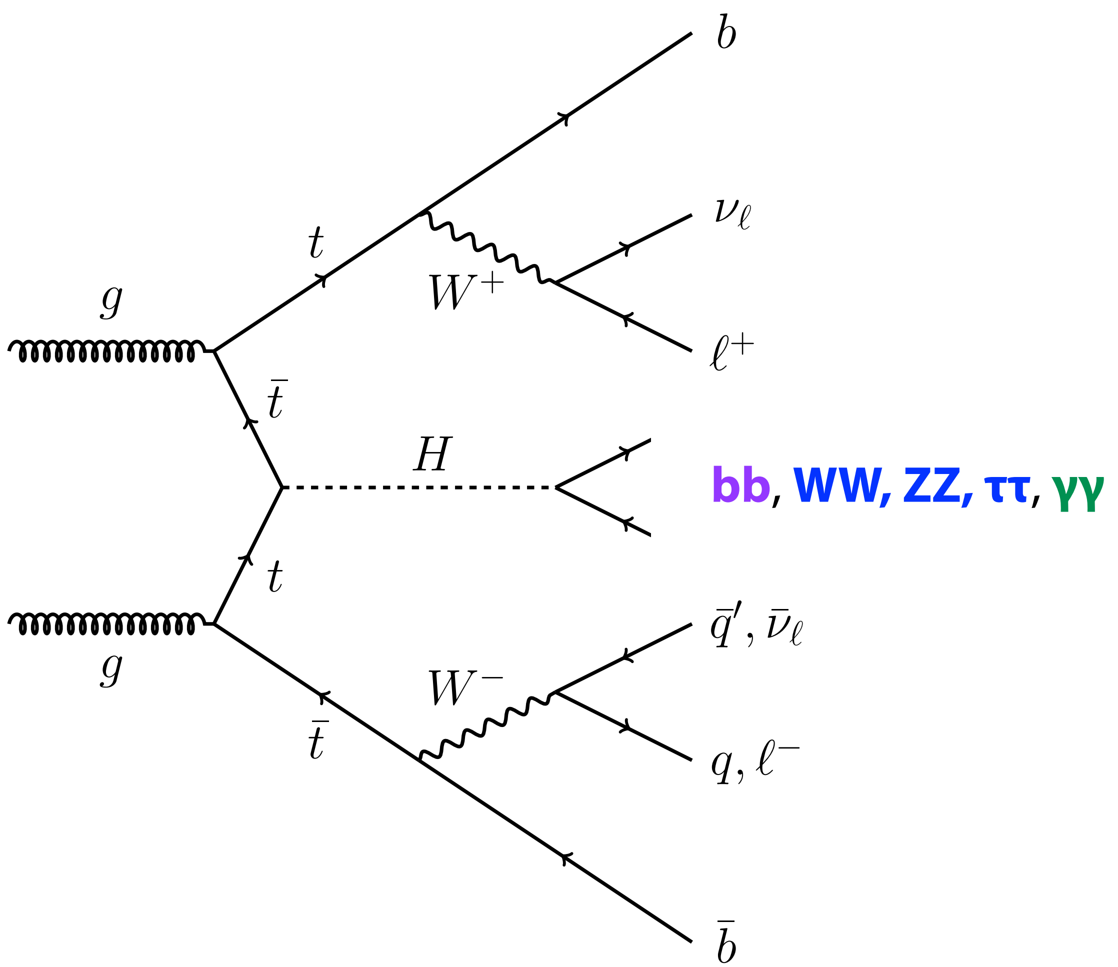
INSTITUTE OF EXPERIMENTAL PARTICLE PHYSICS (IEKP) – PHYSICS DEPARTMENT

Introduction

- Want to measure top-quark Yukawa coupling: **y_t ~ O(1) in the SM**

- Indirect evidence from ggH production and H→γγ decay via loop contribution
  - But possibly modified by BSM contributions
  - ttH production is the best handle for tree-level coupling, but σ ~ 130 fb

- **Strategy**:
  - Target as many combinations of the $t\bar{t}$ final state (0,1 or 2 leptons) and Higgs decay as possible
  - Target **hadrons**, **leptons** and **γγ** Higgs decays
  - Exploit **high jet** and **b-jet multiplicity**
  - Extract signal using MVA or matrix element methods



CMS Results

Combination of ttH searches

$H\to bb$ | $H\to \tau_h\tau_h$ | $H\to\gamma\gamma$ | $H\to WW/\tau\tau/ZZ$

5.1 fb^{-1} (7 TeV)
19.7 fb^{-1} (8TeV)

CMS-HIG-13-029
JHEP 09 (2014) 087

ttH→bb with a Matrix Element method

$H\to bb$

Alternative event categorisation and use of MEM to discriminate signal and background

19.5 fb^{-1} (8TeV)

CMS-HIG-14-010
EPJ C 75 (2015)

NEW

Combination of single-top + Higgs searches

$H\to bb$ | $H\to\gamma\gamma$ | $H\to WW/\tau_l\tau_l$ | $H\to \tau_h\tau_l$

19.7 fb^{-1} (8TeV)

CMS-HIG-14-027
arXiv:1509.08159
(submitted to JHEP)

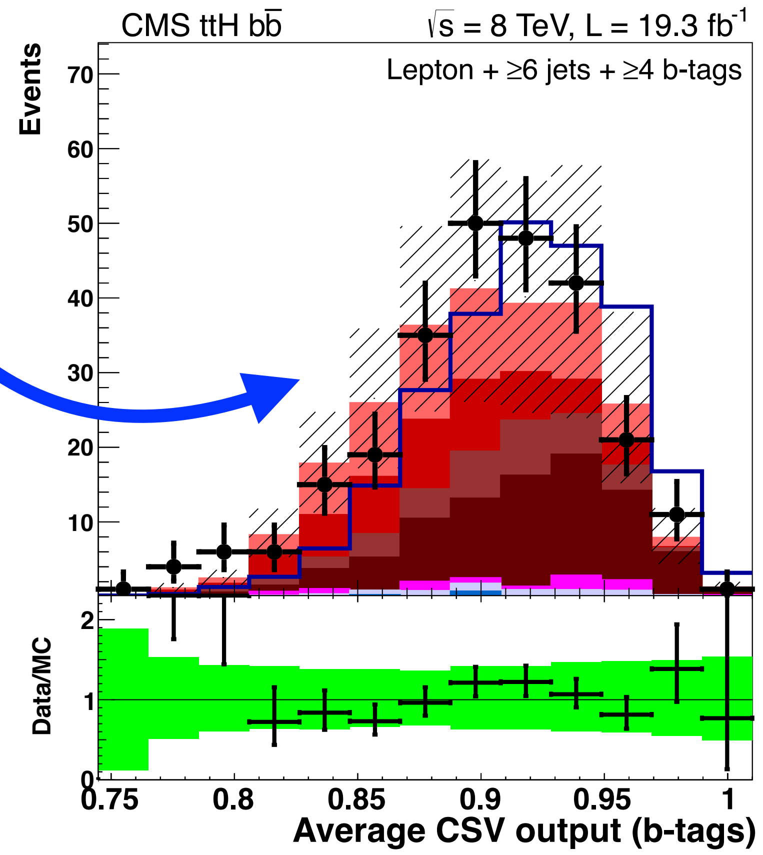
ttH - Analysis Overview

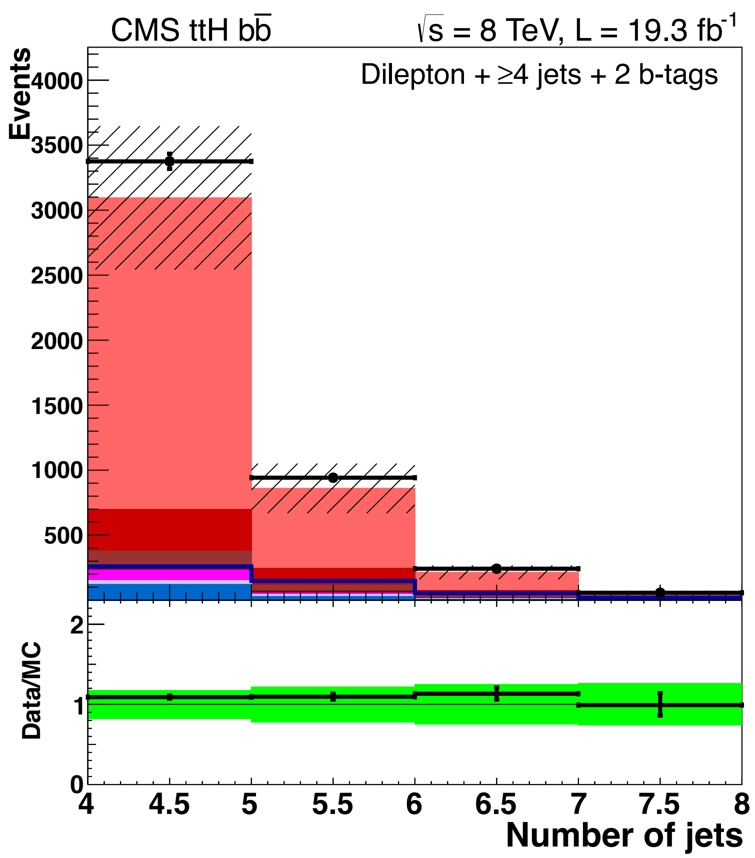
H→bb

- **Require at least one light lepton** to suppress large multi-jet background
  - **lepton+jets channel**: single isolated lepton + ≥ 4 jets of which ≥ 2 are b-tagged
  - **dilepton channel**: oppositely charged leptons + ≥ 3 jets of which ≥ 2 are b-tagged
- **Main backgrounds** from **tt+bb**, mis-tagged **tt+light** and **tt+cc, tt+V and single t**
- Categorise on jet and b-tagged jet multiplicity, multivariate discriminator for signal extraction
- 10-15 input variables including object kinematics, event shape, b-tagging discriminant

Combined Secondary Vertex Discriminator

Medium WP:
70% b-jet eff. / 2% fakes
Loose WP:
85% b-jet eff. / 10% fakes





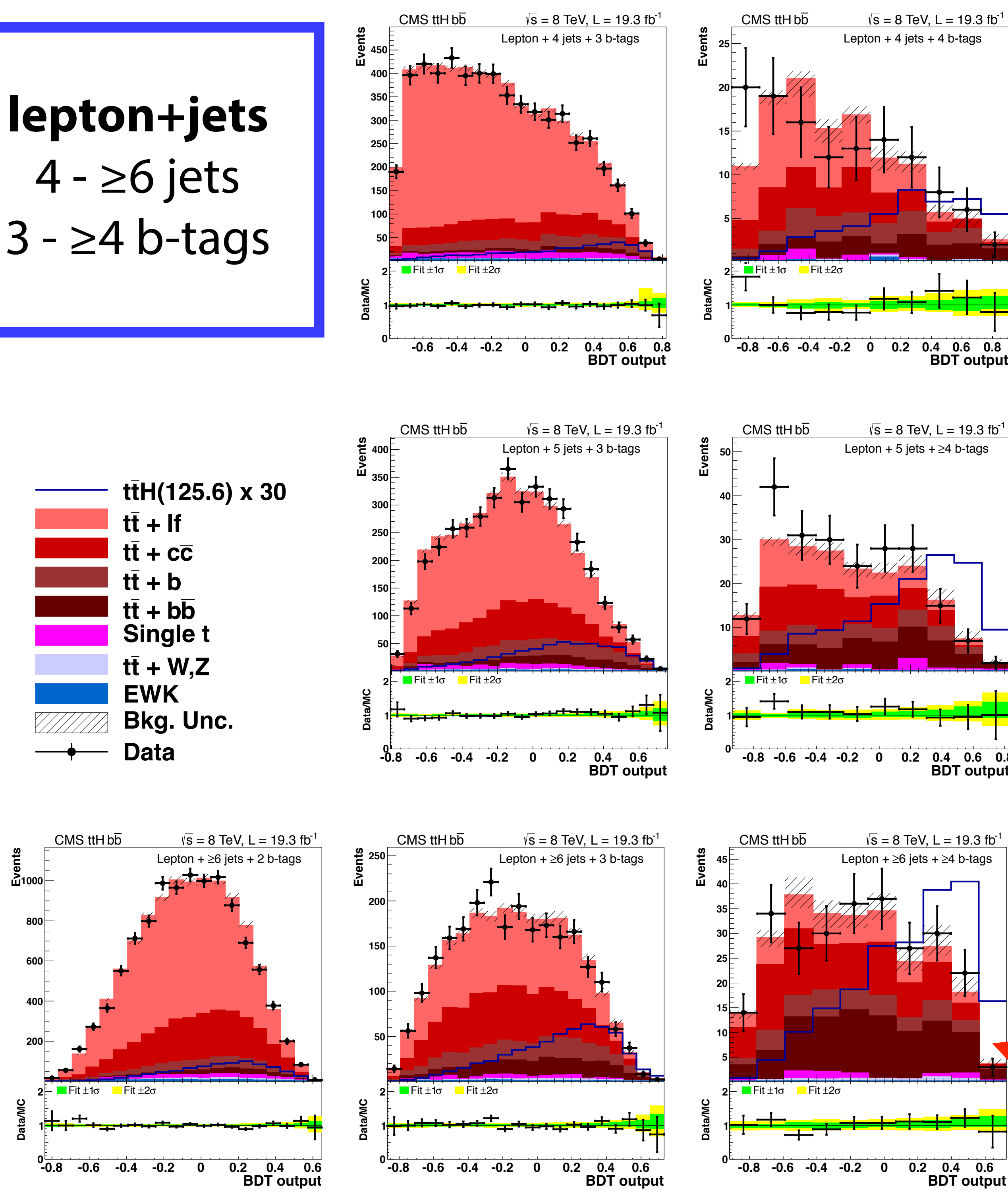
ttH - Event Categories

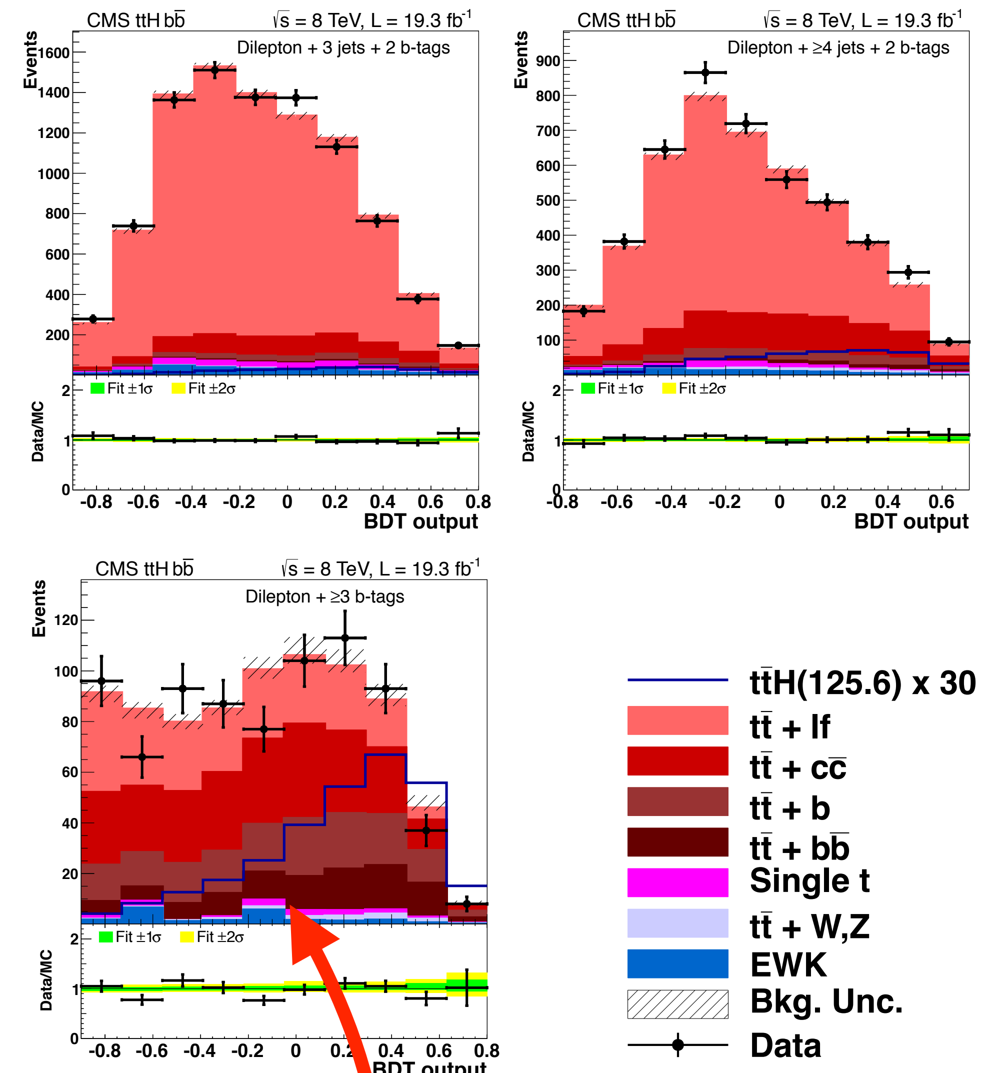
H→bb

b-tagged jets

jets

lepton+jets
4 - ≥6 jets
3 - ≥4 b-tags





jets

b-tags

Most sensitive categories

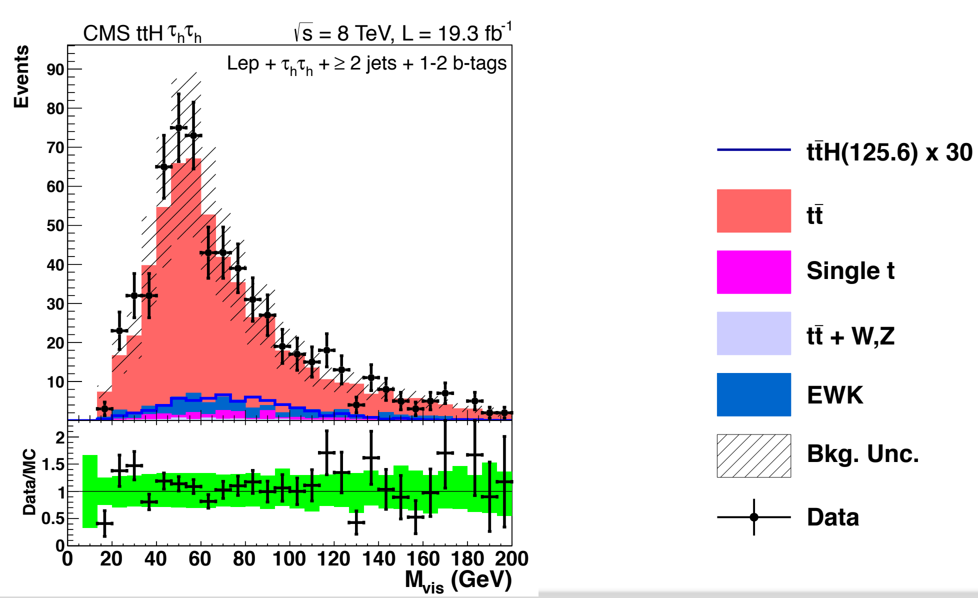
dilepton
3 - ≥4 jets
2 - ≥3 b-tags

ttH - Event Categories

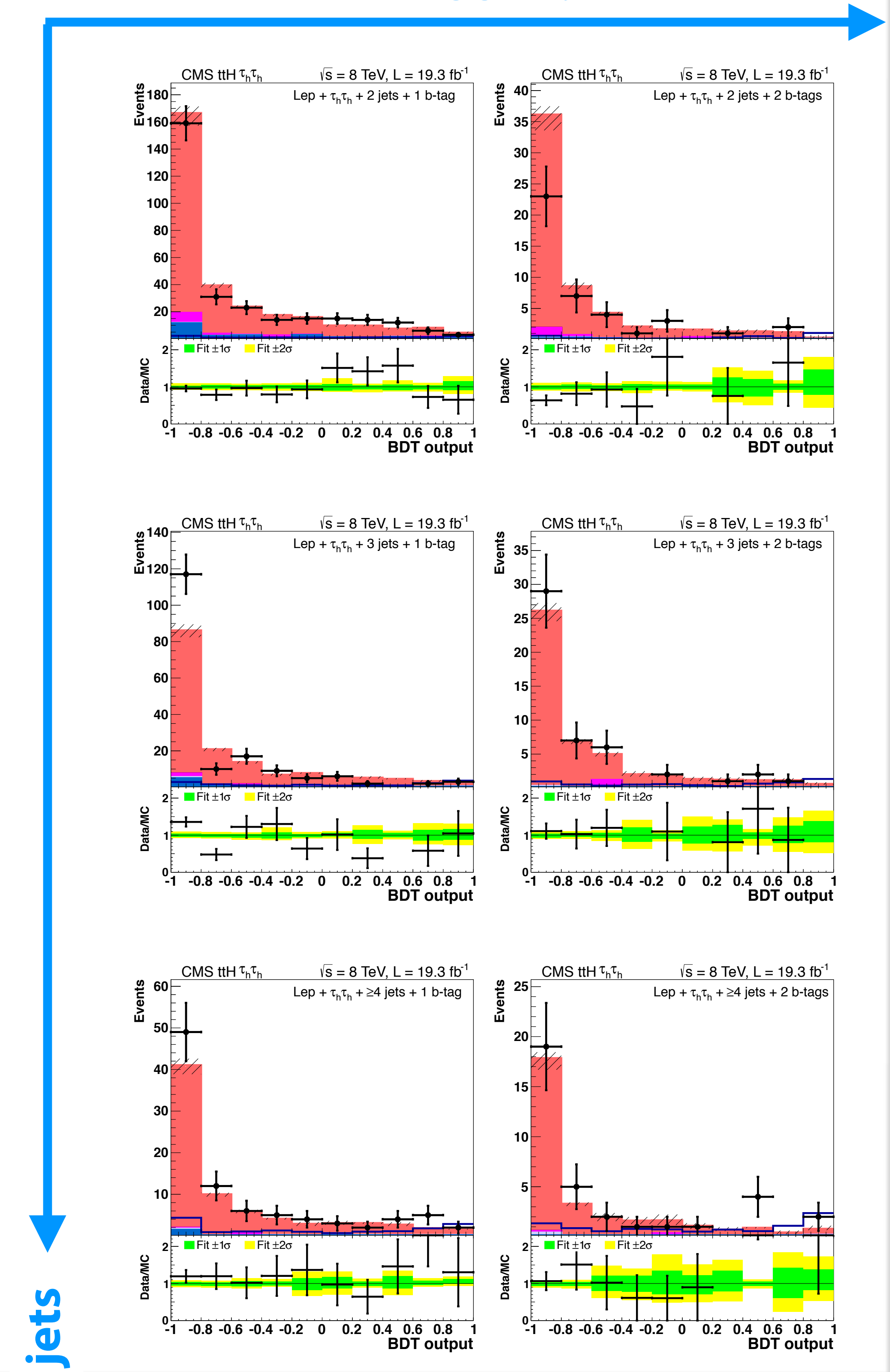
$H \to \tau_h \tau_h$

- A third channel targets Higgs decays to a pair of τ leptons which decay hadronically (τ_h)
  - Requires one e/μ from $t\bar{t}$ decay, ≥ 2 additional jets of which ≥ 1 b-tagged
  - τ_h candidates must contain a **single charged hadron**, pass **decay mode identification** and an **MVA-based isolation** discriminator
- Similar categorisation scheme with 2-4 jets and 1-2 b-tags
- One BDT trained for all categories
  - Most input variables related to the di-τ_h system

Visible mass of the di-tau decay products



b-tagged jets

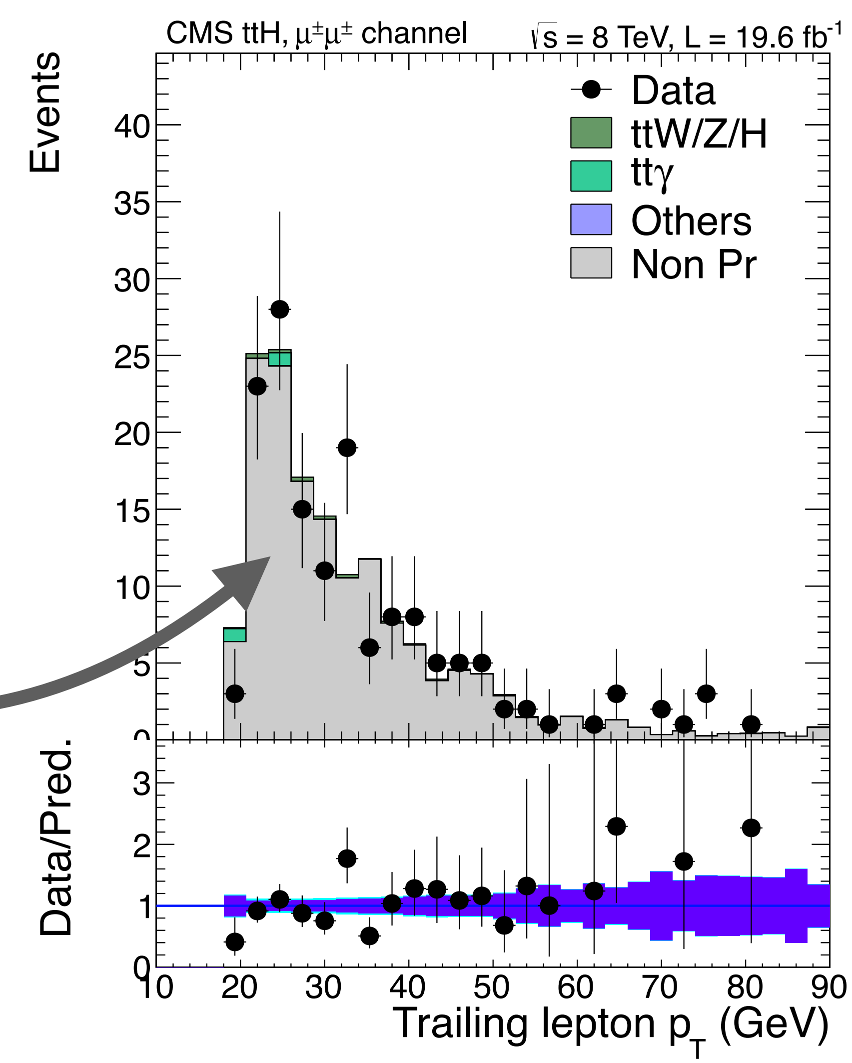


ttH - Analysis Overview

H→WW/ττ/ZZ

- Consider **three multi-lepton final states**, where at least one lepton originates in top decay
  - **2l (**same-sign + at least 4 jets), **3l, 4l**
  - All categories require ≥ 2 loose or ≥ 1 medium b-tagged jets
- Main backgrounds have ≥ 1 non-prompt lepton, e.g. from b hadron decay. Require leptons to pass BDT "prompt vs. non-prompt" discriminator using IP, isolation and nearest-jet properties

- **Three main classes of backgrounds:**
  - **$t\bar{t}$+V**: estimated from simulation normalised to NLO cross sections
  - **VV**: mainly WW and WZ, normalised in a signal-depleted control region with ≥ 2 jets and a b-tag veto, or with inverted m_{ll} veto
  - **$t\bar{t}$/W/Z+jets** with non-prompt leptons: data-driven using a "fake-rate" method applied to events in control region where at least one lepton fails lepton MVA

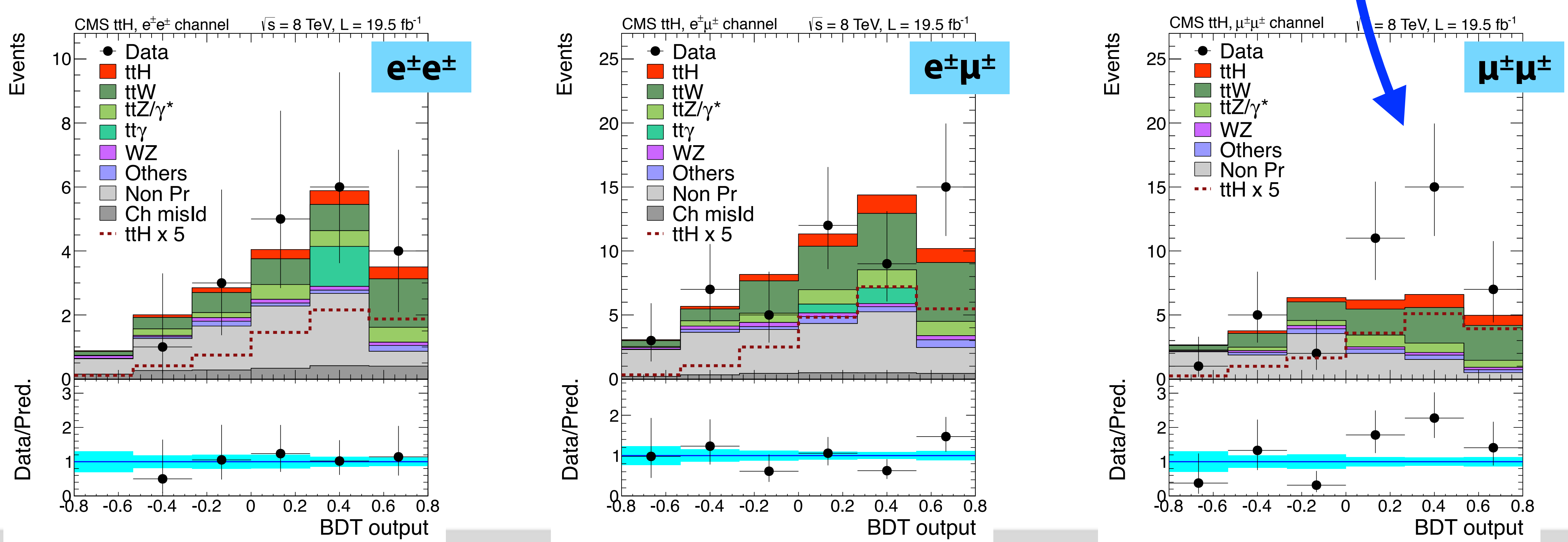


ttH - Signal Extraction

H→WW/ττ/ZZ

- **2l categories:** fit output of a **BDT discriminator** trained to separate ttH vs $t\bar{t}$+jets
- Input variables include the p_T and |η| of the trailing lepton, transverse mass of the leading lepton and E_T^{miss}, event topology and energy sum variables
- Divide events into two categories based on lepton charge sum
  - ~ 5% improvement in sensitivity due to charge asymmetry in SM backgrounds

Visible excess of events in μ±μ± final state

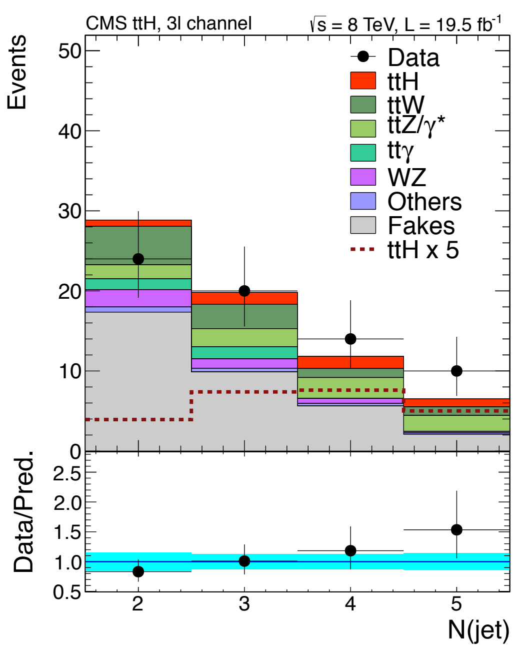


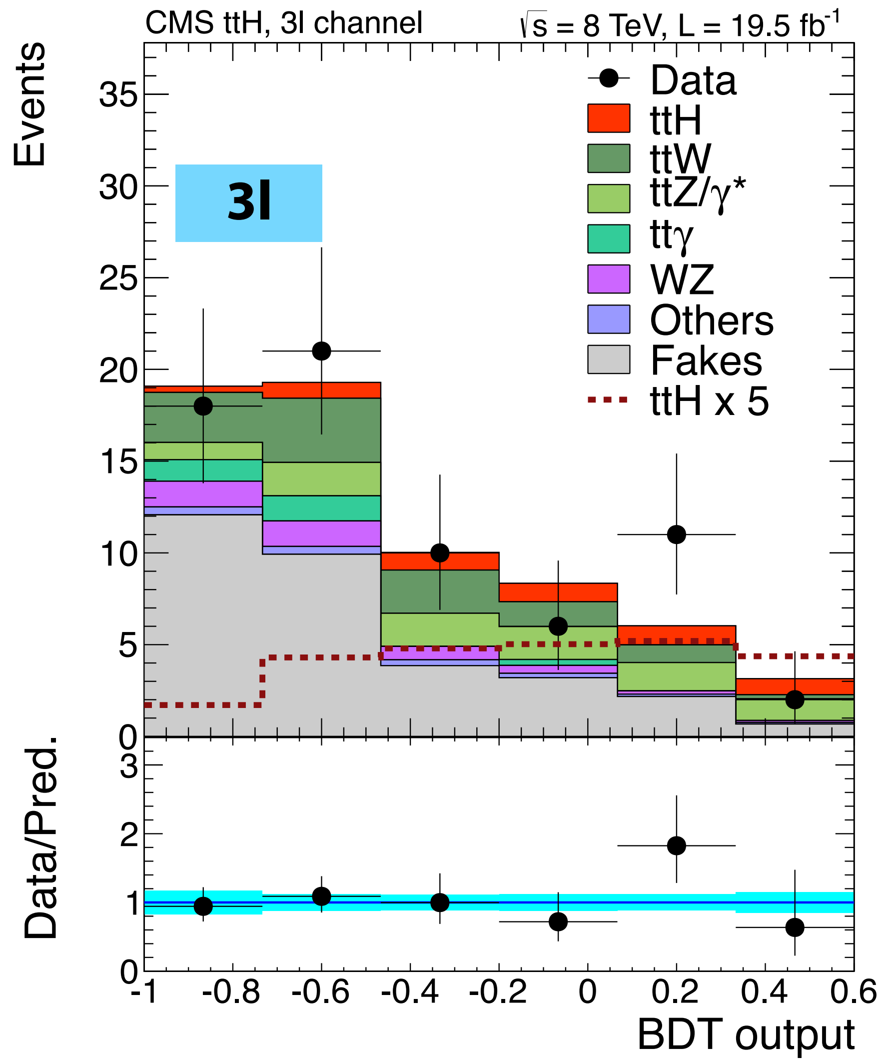
ttH - Signal Extraction

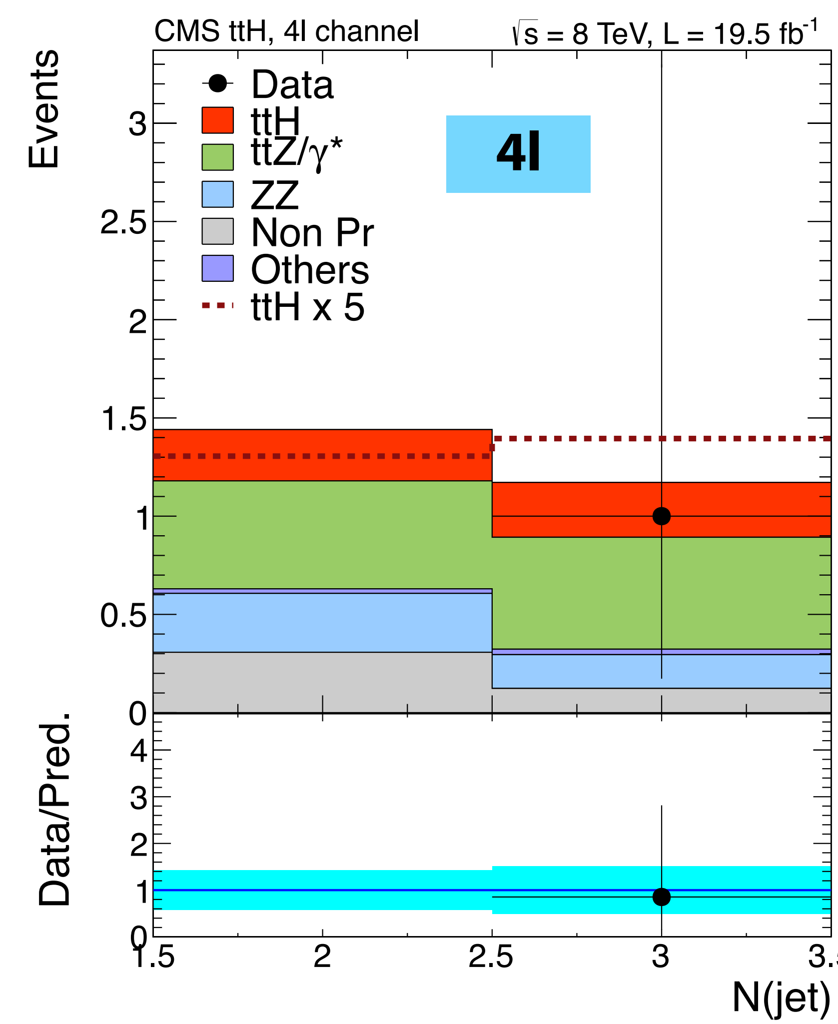
H→WW/ττ/ZZ

- **3l category** uses a similar **BDT discriminator** trained to separate ttH vs backgrounds
- **4l category** uses the jet multiplicity as discriminating variable due to low event statistics

BDT gives ~10% improvement in 2l & 3l sensitivity compared to fitting nJets distribution







ttH - Analysis Overview

H→γγ

- Despite **small H→γγ branching ratio** (~2x10^{-3}) exploit clean signature and excellent photon energy resolution
- Photon identification and energy measurement closely follows inclusive CMS H→γγ analysis
- Loose selection on $t\bar{t}$ final state objects to accept as many signal events as possible
- Categorise events based on **hadronic** and **leptonic** $t\bar{t}$ final states:
  - **Hadronic:** ≥4 jets, of which ≥ 1 b-tagged, no high-p_T e/μ
  - **Leptonic:** ≥ 2 jets, of which ≥ 1 b-tagged, ≥ 1 e/μ with p_T > 20 GeV
- Extract signal from fit to the **di-photon invariant mass** $m_{\gamma\gamma}$

Analysis selections aim to maximise ttH acceptance and minimise acceptance of other signal modes which peak at same $m_{\gamma\gamma}$

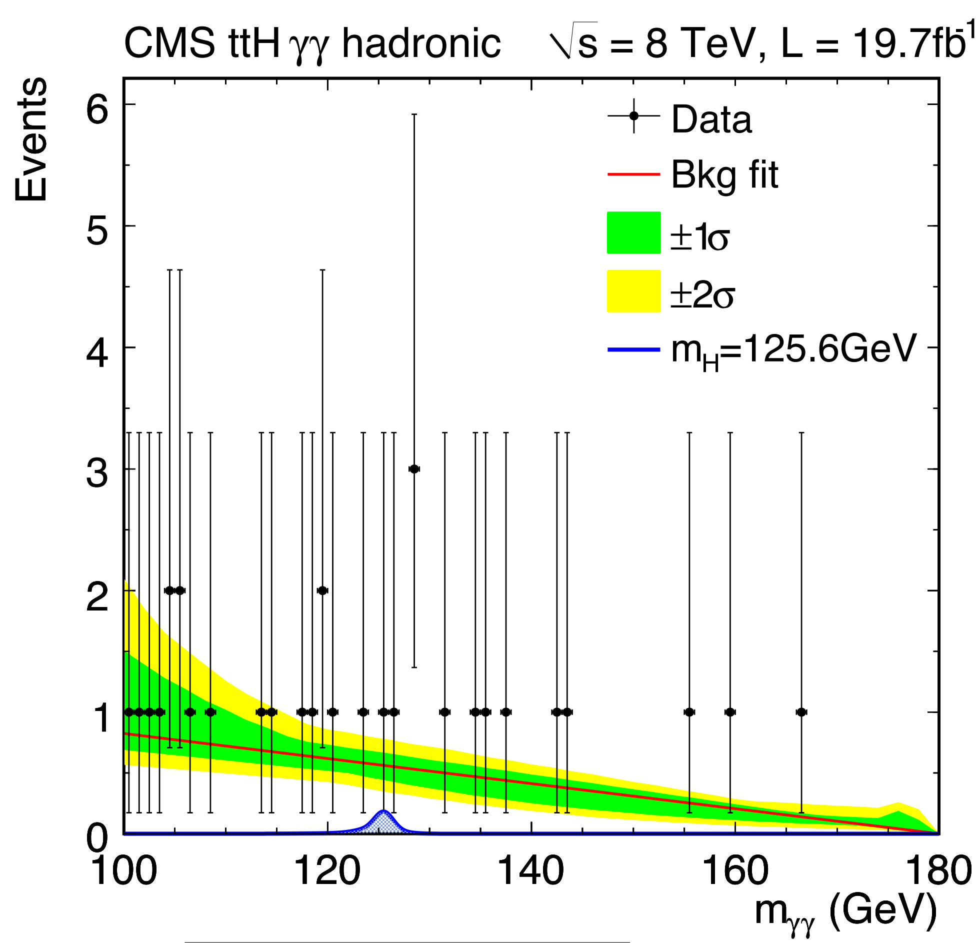
Expected signal events

	7 TeV	8 TeV	
	All decays	Hadronic channel	Leptonic channel
ttH	0.21	0.51	0.45
gg → H	0.01	0.02	0
VBF H	0	0	0
WH/ZH	0.01	0.01	0.01
Total H	0.23	0.54	0.46
Data	9	32	11

ttH - Signal extraction

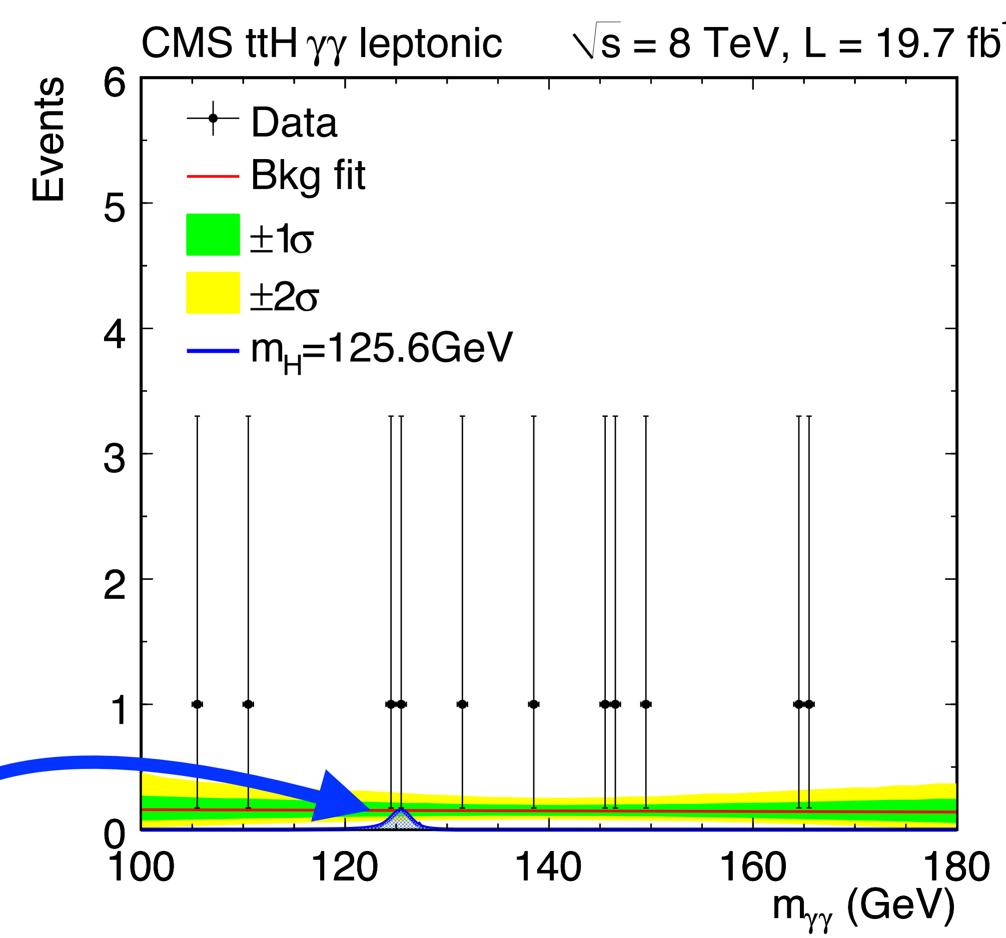
H→γγ

- **Main backgrounds**: top quarks + real/fake photons, high p_T photons + multi-jet
- Modelled by a functional form and determined by fit to $m_{\gamma\gamma}$ in the range 100-180 GeV
  - Exact form treated as a **discrete nuisance parameter** including exponential, power-law and polynomial functions



Hadronic

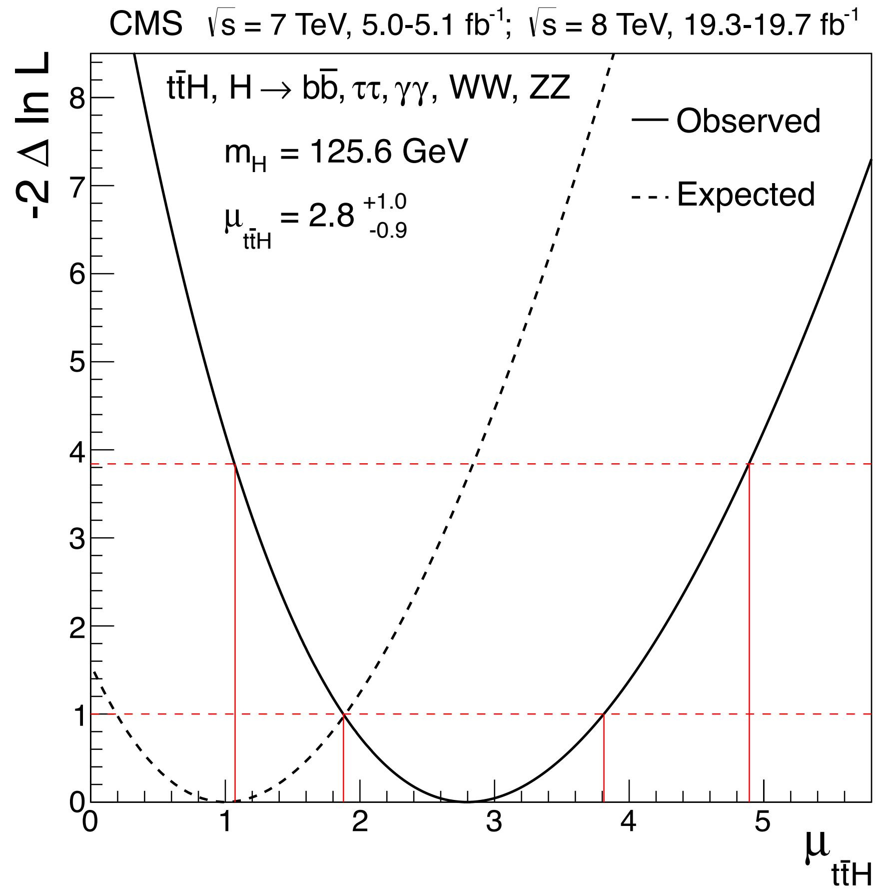
Expected signal contribution

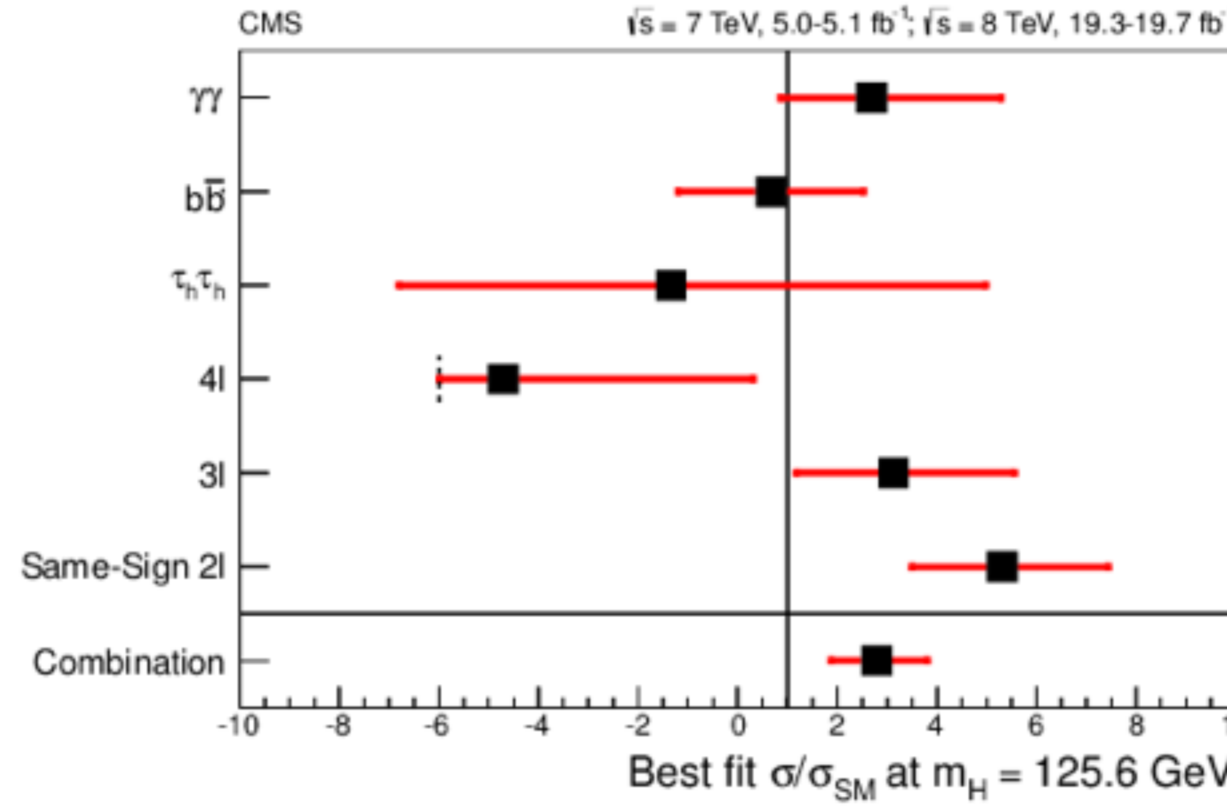


Leptonic

ttH - Results

p-value for compatibility of separate channel results with common μ = 29%





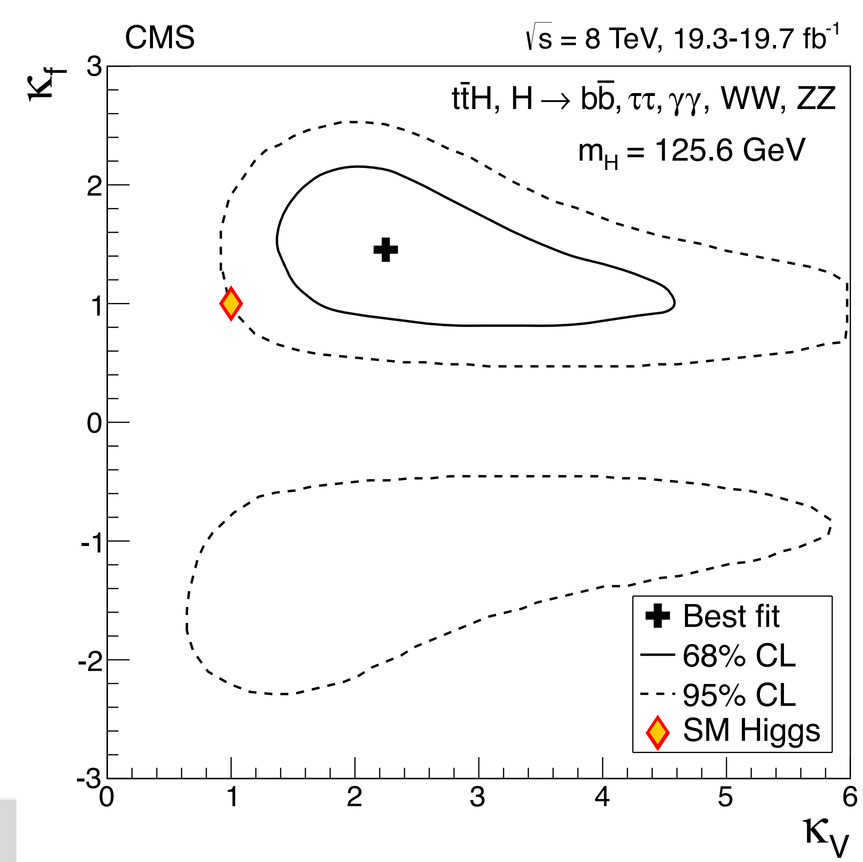
$\hat{\mu} = 2.8^{+1.0}_{-0.9}$ **@ 125.6 GeV**

~ 2.1σ excess compared to SM (μ = 1)

2D coupling fit:

$\kappa_V, \kappa_F = (2.2, 1.5)$

Compatible with SM (1,1) at 95% CL

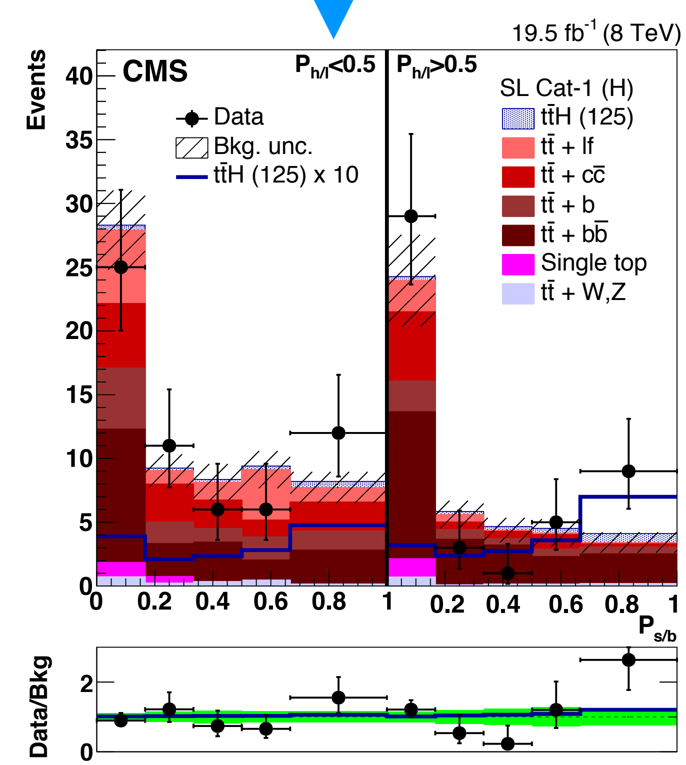


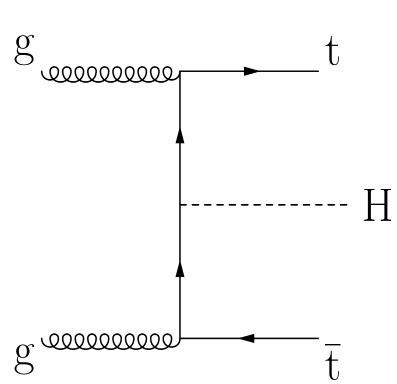
Matrix Element Method

H→bb

- Alternative H→bb analysis using a Matrix Element method instead of a BDT discriminator
- Fit 2D distribution of $\mathbf{P_{s/b}}$ and $\mathbf{P_{h/l}}$
  - $\mathbf{P_{s/b}}$**:** Ratio of signal **(ttH) and** bkg (**tt+bb**) likelihoods computed from **LO matrix elements** with transfer functions to model experimental resolution
  - $\mathbf{P_{h/l}}$**:** likelihood of b-tagging observables
- **~30% improved sensitivity** compared to BDT analysis

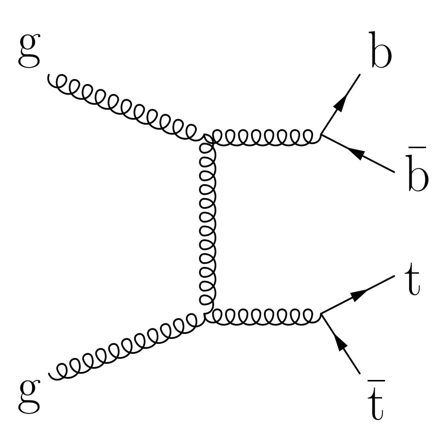
$$P_{\mathrm{h/l}} = \frac{f(\vec{\xi}|\mathrm{t\bar{t}+hf})}{f(\vec{\xi}|\mathrm{t\bar{t}+hf}) + k_{\mathrm{h/l}} f(\vec{\xi}|\mathrm{t\bar{t}+lf})}$$





ttH→bb

vs



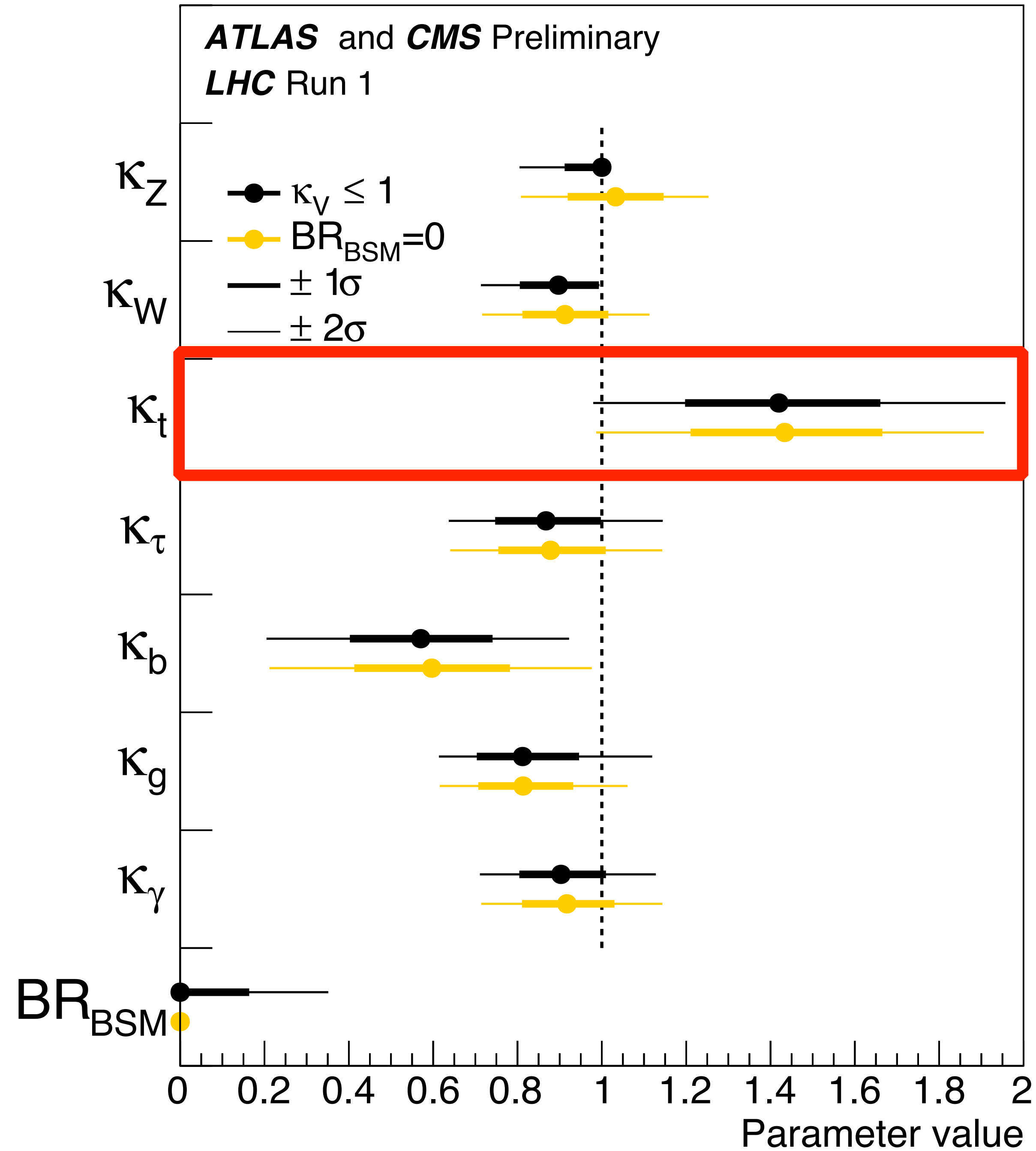
tt + bb

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

$\hat{\mu} = 1.2^{+1.6}_{-1.5}$ **@ 125.6 GeV**

Role of ttH in the Coupling Combination

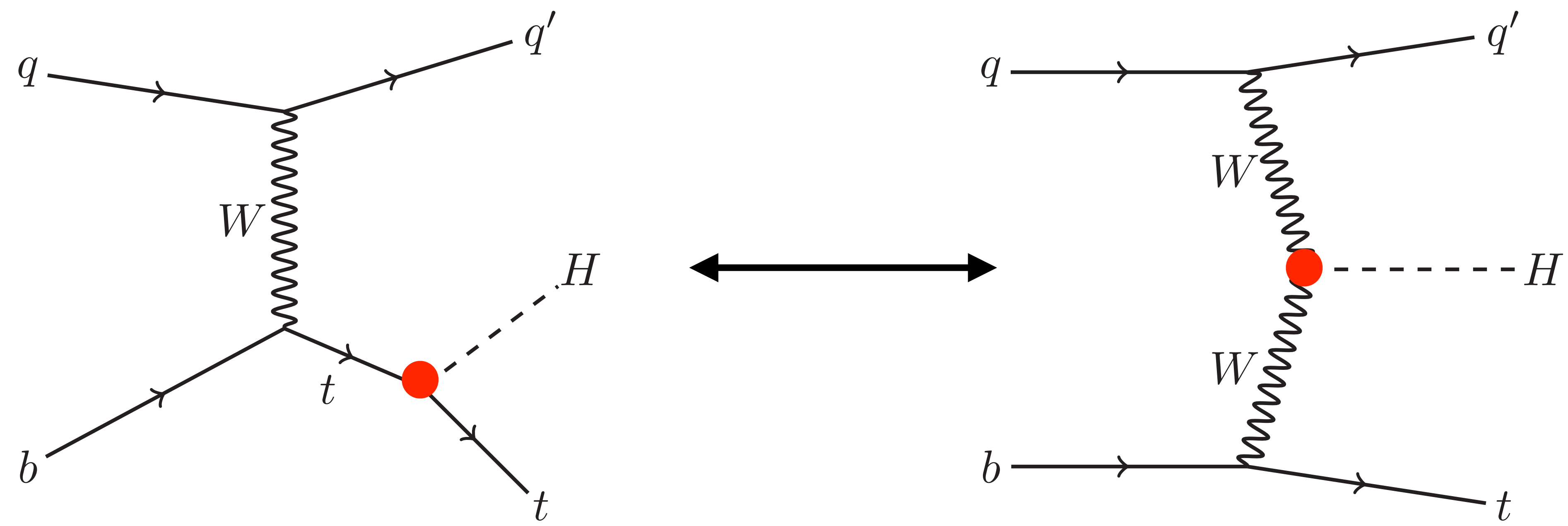
- Results of a combination of CMS and ATLAS Higgs analyses has been released
- ttH analyses play an important role
- In particular: constraint on κ_t in coupling modifier model where loop processes are not resolved
  - E.g. gluon-gluon fusion production assumed to scale with κ_g, not as a function of κ_t and κ_b
- Combined result of $\kappa_t = 1.43^{+0.23}_{-0.22}$



ATLAS-CONF-2015-044/ CMS-PAS-HIG-15-002

Single-top + Higgs Production

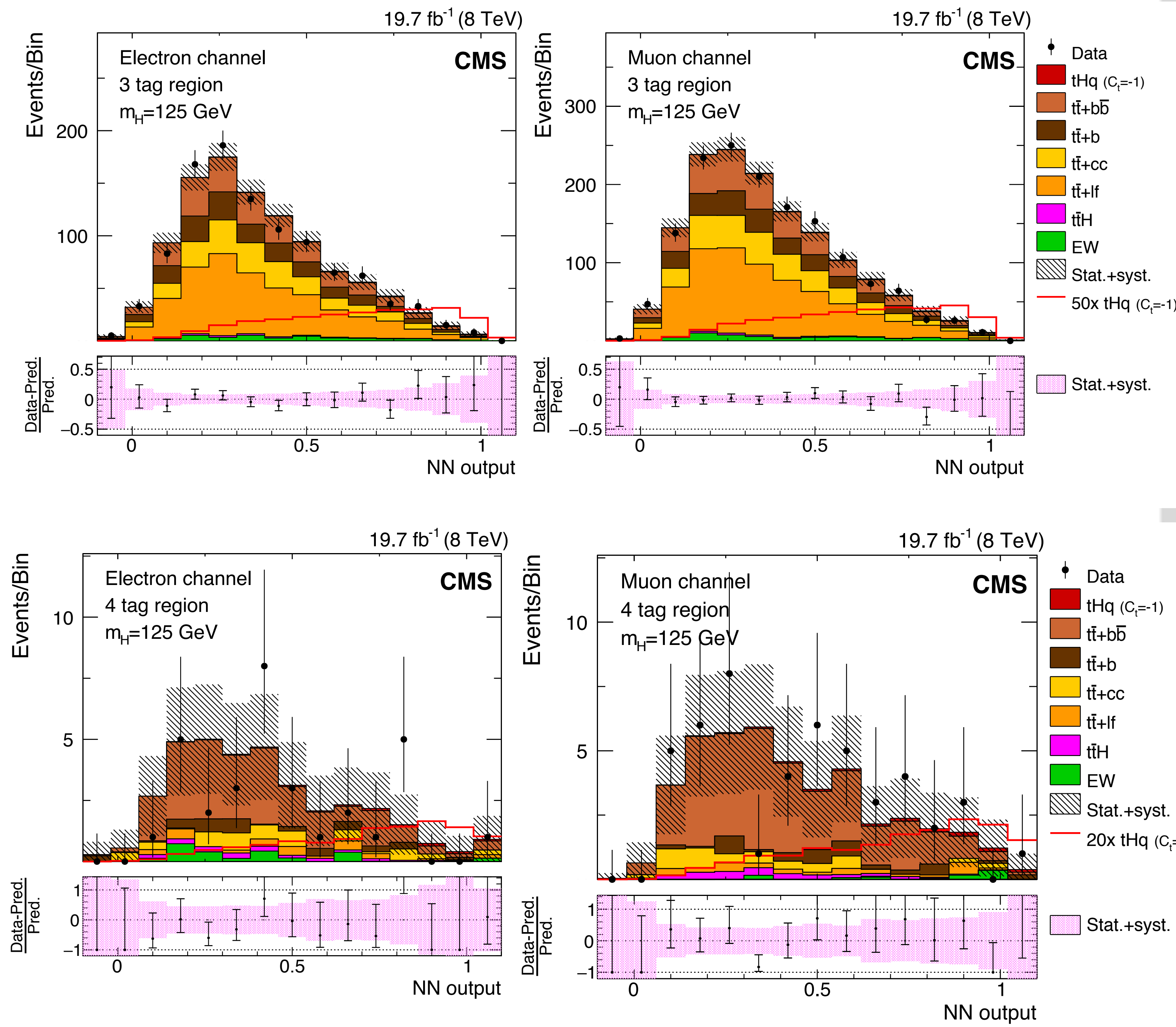
- Novel channel to probe the sign of C_t and search for new physics: **single top + Higgs production**
  - Proceeds mainly via the t-channel tHq process:



- t and W couple to H with opposite sign: destructive interference of main diagrams
- SM cross section of ~ 18 fb (cf. 130 fb for ttH)**, but enhanced by factor of 15 if C_t = -1**
- **C_t = -1 can be excluded given constraint from inclusive H→γγ search, but assumes no new particles in loop for that decay**
- Searches exploit t→lνb decay and presence of a typically forward hadronic jet in the final sate

Single-top + Higgs Production **H→bb**

- Expect **at least five quarks** in the final state: H→bb (2), t→lvb (1), forward q (1), b from strong interaction in tHq process (1)
- Preselects events with a single e/μ candidate and ≥ 4 jets
- 2 event categories:
  - ≥ 4 jets + ≥ 3 b-tags
  - ≥ 5 jets + ≥ 4 b-tags
- Significant tt+jets background reduced with artificial neural networks (NN)
  - 1st stage: use NNs trained to identify correct association of jets → quarks for either tHq or tt+jets hypothesis
  - 2nd stage: signal vs background NN including lepton & jet kinematics, b-tagging and event topology variables under both hypotheses
- Signal extracted by fit to NN output distribution



Single-top + Higgs Production

H→γγ

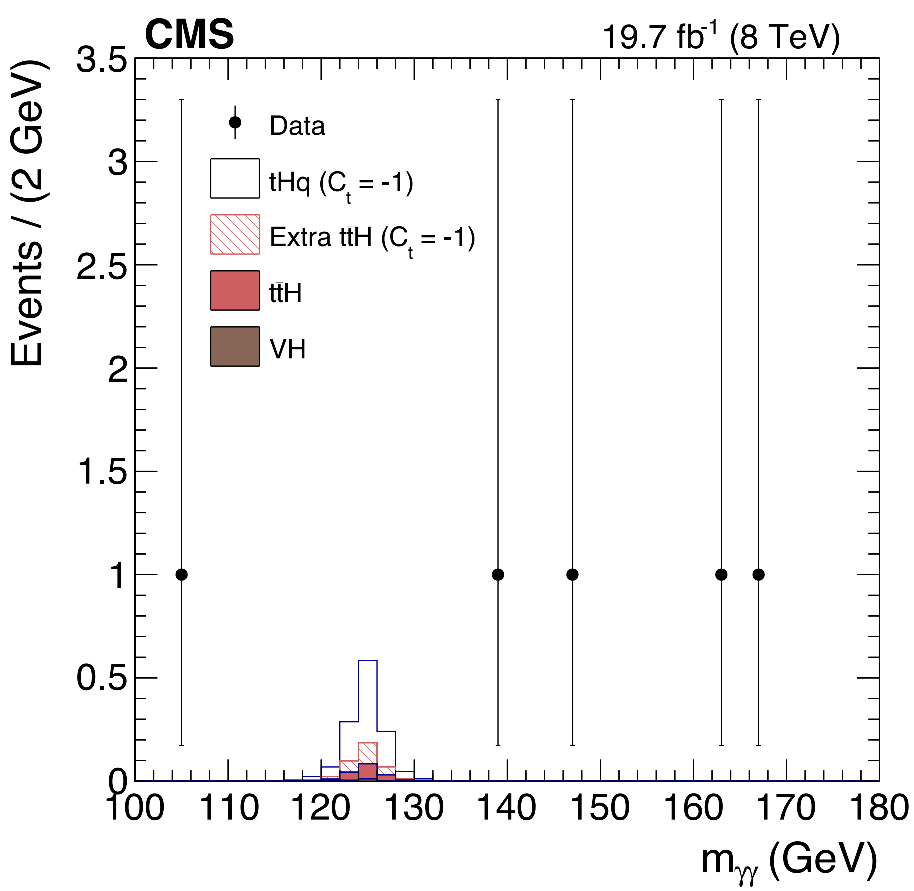
- With C_t = -1 H→γγ BR also enhanced by ~ x2
- Initial selection:
  - 2 high pT photons, 1 isolated lepton, ≥ 1 b-tagged jet, forward jet with |η| > 1
- Then apply cut on likelihood classifier using jet multiplicity, top quark m_T, light jet η, $\Delta\eta^{lep\text{-}jet}$, lepton charge
- Zero events observed in signal region

Bayes classifier:
Ratio of s/ s+b likelihoods

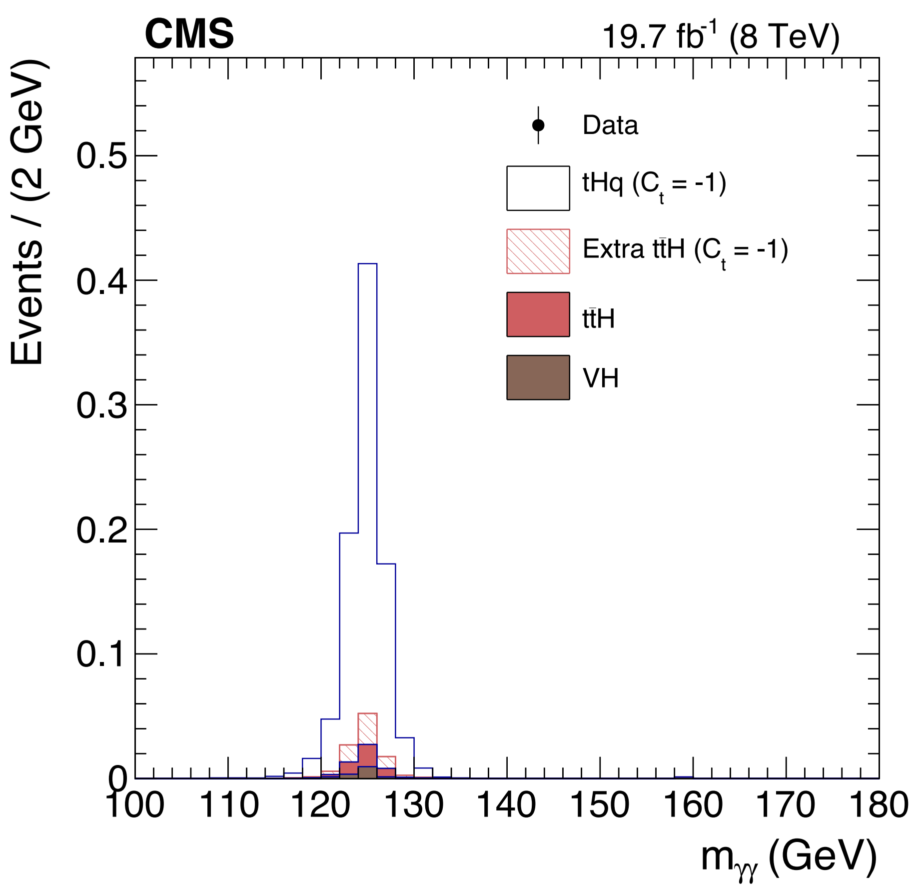
$$\mathrm{L}(\mathrm{x}) = \frac{\mathrm{L_S}(\mathrm{x})}{\mathrm{L_S}(\mathrm{x}) + \mathrm{L_B}(\mathrm{x})}$$

$$\mathrm{L}^i(\mathrm{x}) = \prod_j p^i_j(\mathrm{x}^j),$$

i = signal or bkg. process
$p_j(x_j)$ pdf of variable j evaluated at x



Initial selection



Full selection

Single-top + Higgs Production

H→WW

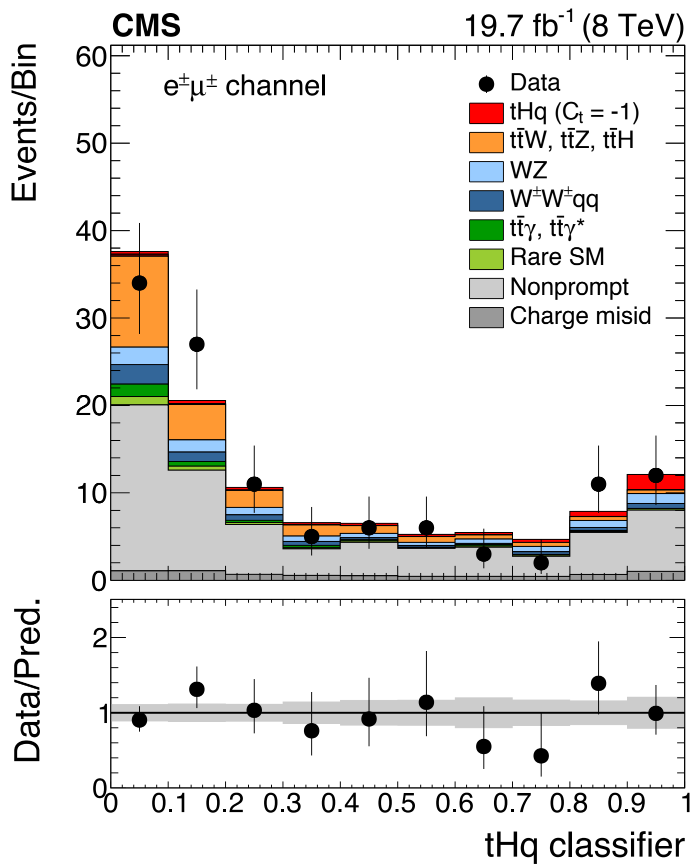
- **Event categories:**
  - Same sign dilepton (eμ or μμ)
  - Three lepton (eee, μμμ, eeμ, eμμ)
  - All categories require a central b-tagged jet and an additional forward jet
- Largest background from tt+jets production (**Nonprompt**)
- Fit for signal with multi-variate likelihood classifier, inputs include jet multiplicity and kinematic variables

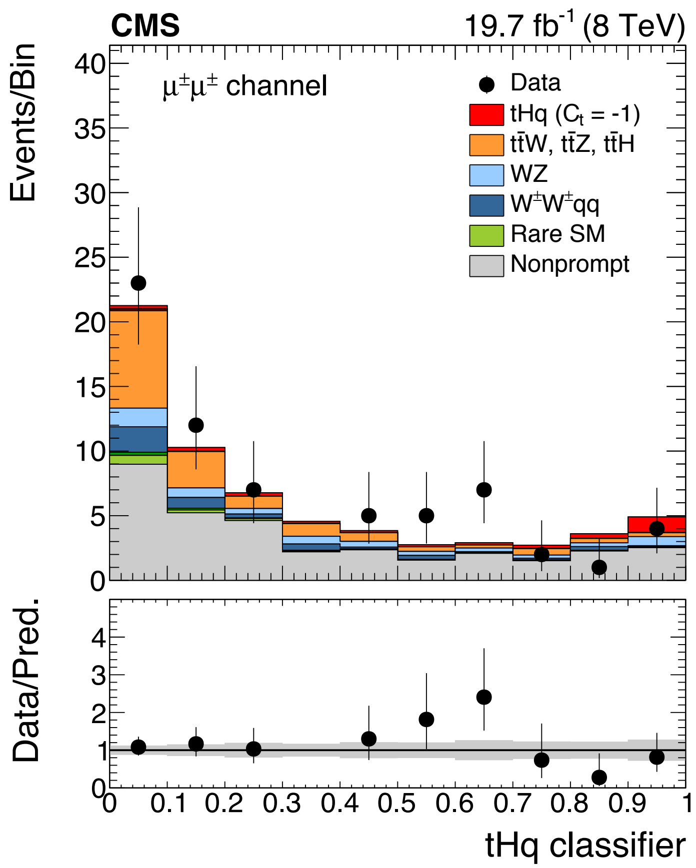
Bayes classifier:
Ratio of s/ s+b likelihoods

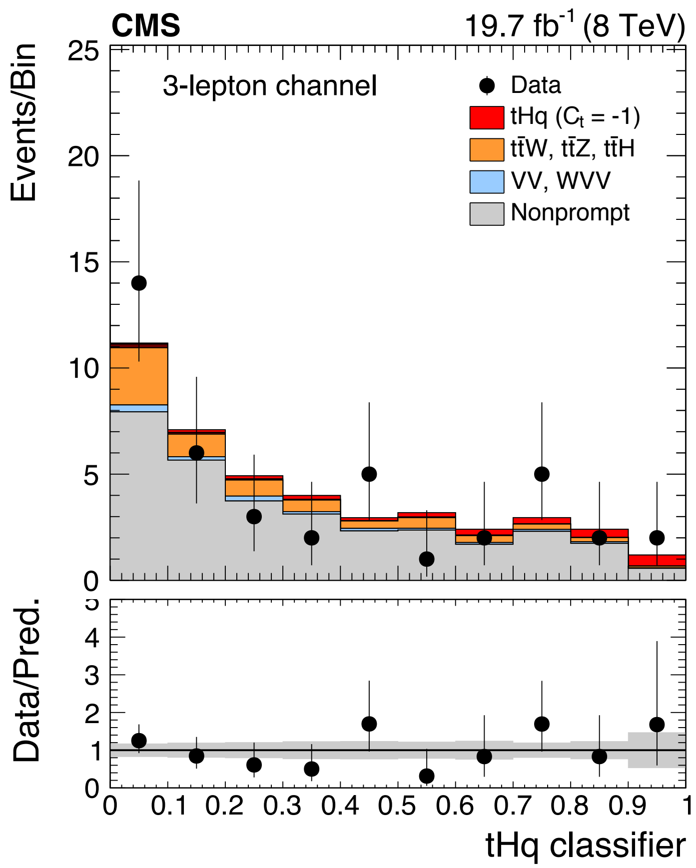
$$L(x) = \frac{L_S(x)}{L_S(x) + L_B(x)}$$

$$L^i(x) = \prod_j p_j^i(x^j),$$

i = signal or bkg. process
$p_{j}(x_j)$ pdf of variable j evaluated at x





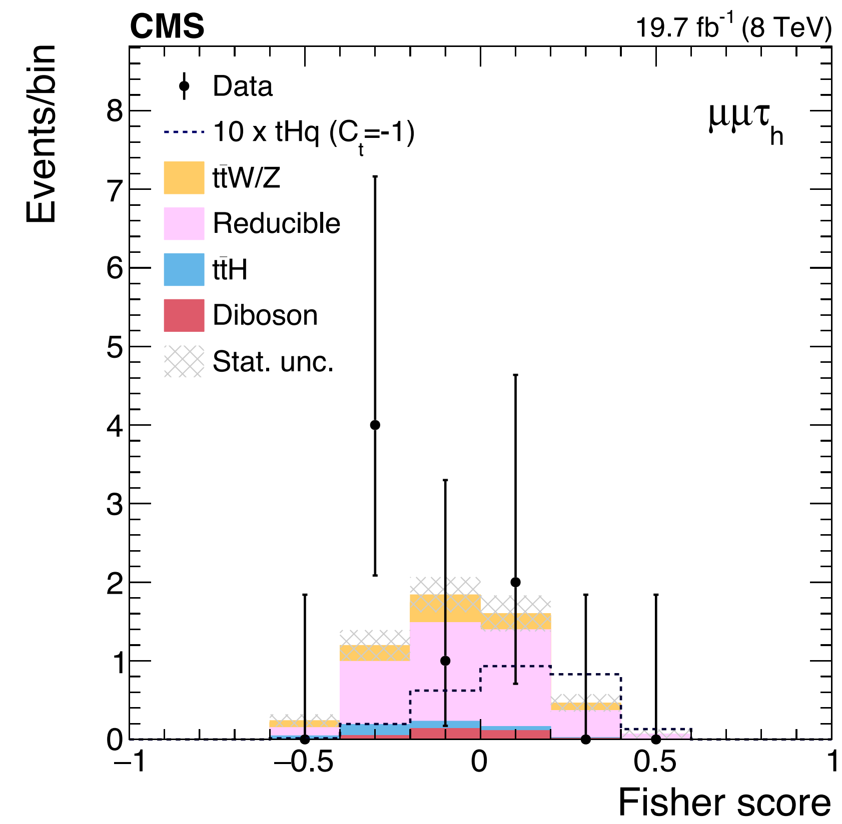


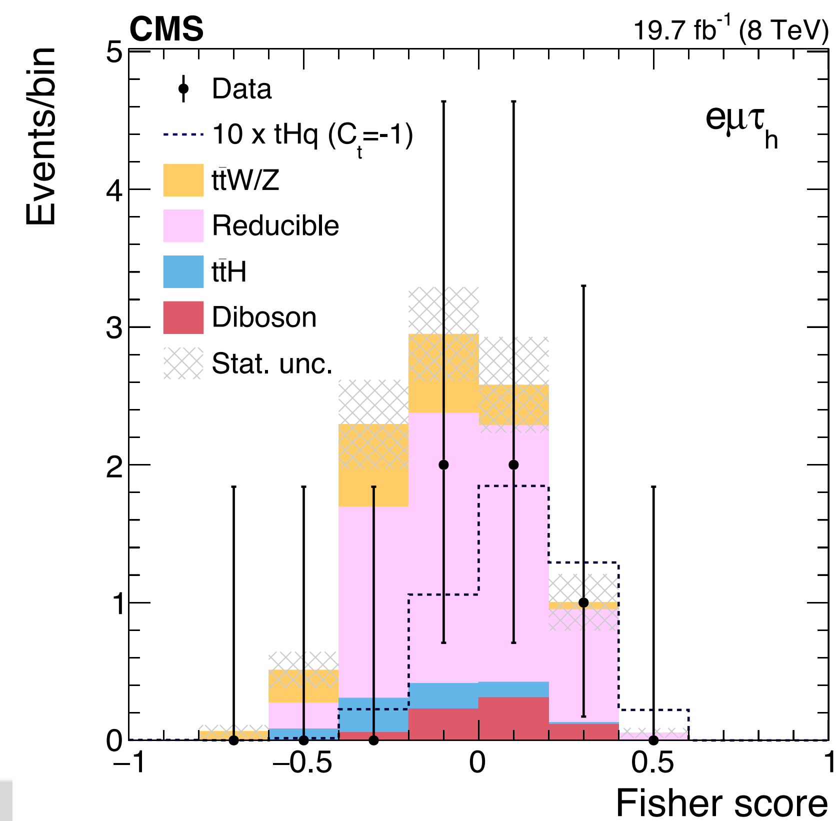
Single-top + Higgs Production

$H \to \tau_h \tau_l$

- Target $H \to \tau_\mu \tau_h$ and $H \to \tau_e \tau_h$ final states with additional same-sign lepton from top decay
- **Irreducible backgrounds** from WZ, ZZ, ttH, and $t\bar{t}$ + V production modelled with MC
- **Reducible backgrounds** from $t\bar{t}$, single-t, V+jets, Z+jets and QCD multijet, measured using a fake-rate method in data
- Signal extraction uses a linear Fisher discriminant exploiting forward jet properties, b-jet multiplicity and other kinematic variables

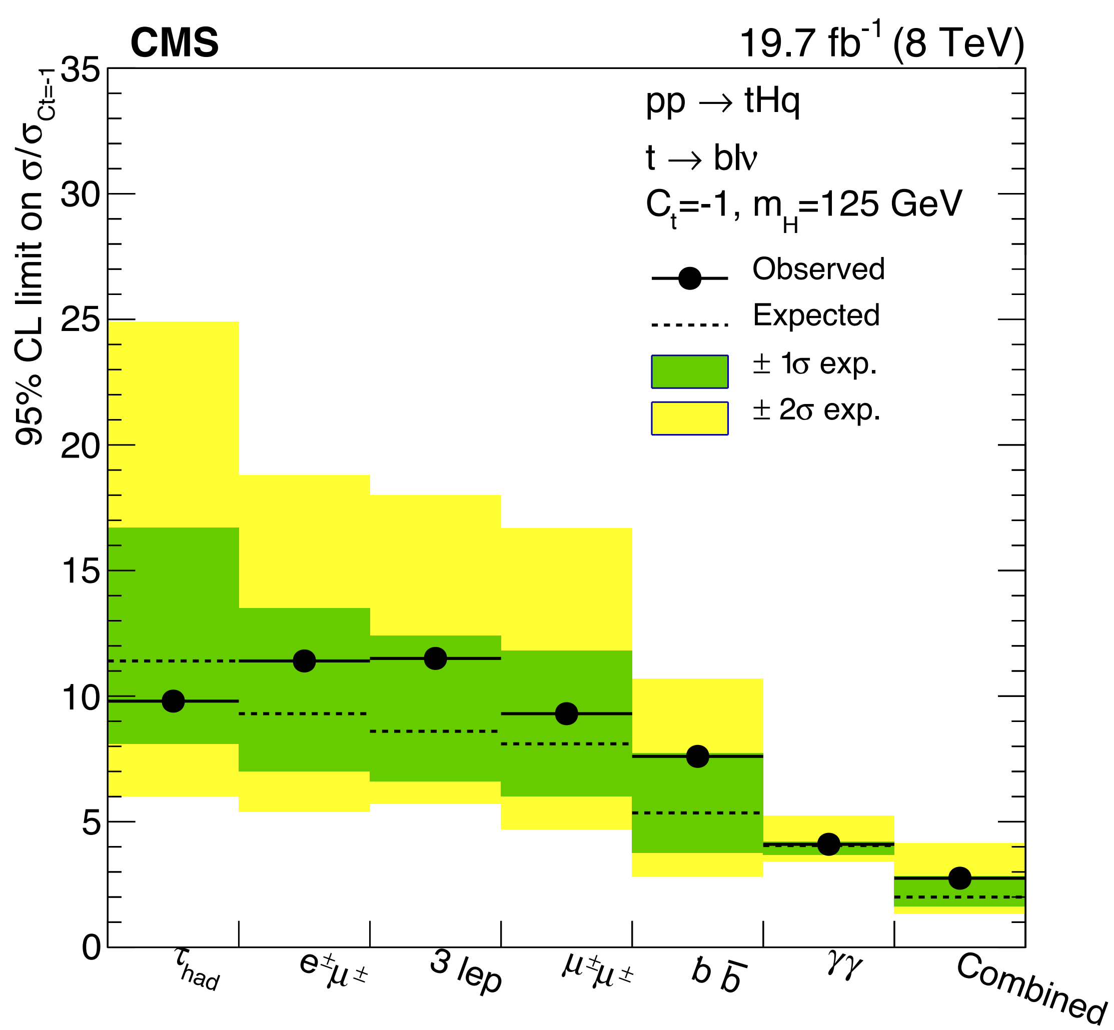
Process	$e\mu\tau_h$	$\mu\mu\tau_h$
tHq, $C_t = -1$	0.42 ± 0.05	0.26 ± 0.03
tHW, $C_t = -1$	0.06 ± 0.01	0.04 ± 0.01
$t\bar{t}H$	0.6 ± 0.1	0.3 ± 0.1
$t\bar{t}V$	1.8 ± 0.4	0.9 ± 0.2
VV	0.7 ± 0.1	0.3 ± 0.1
Reducible	6.3 ± 3.1	4.5 ± 1.9
Total background	9.5 ± 3.7	5.4 ± 2.4
Data	5	7

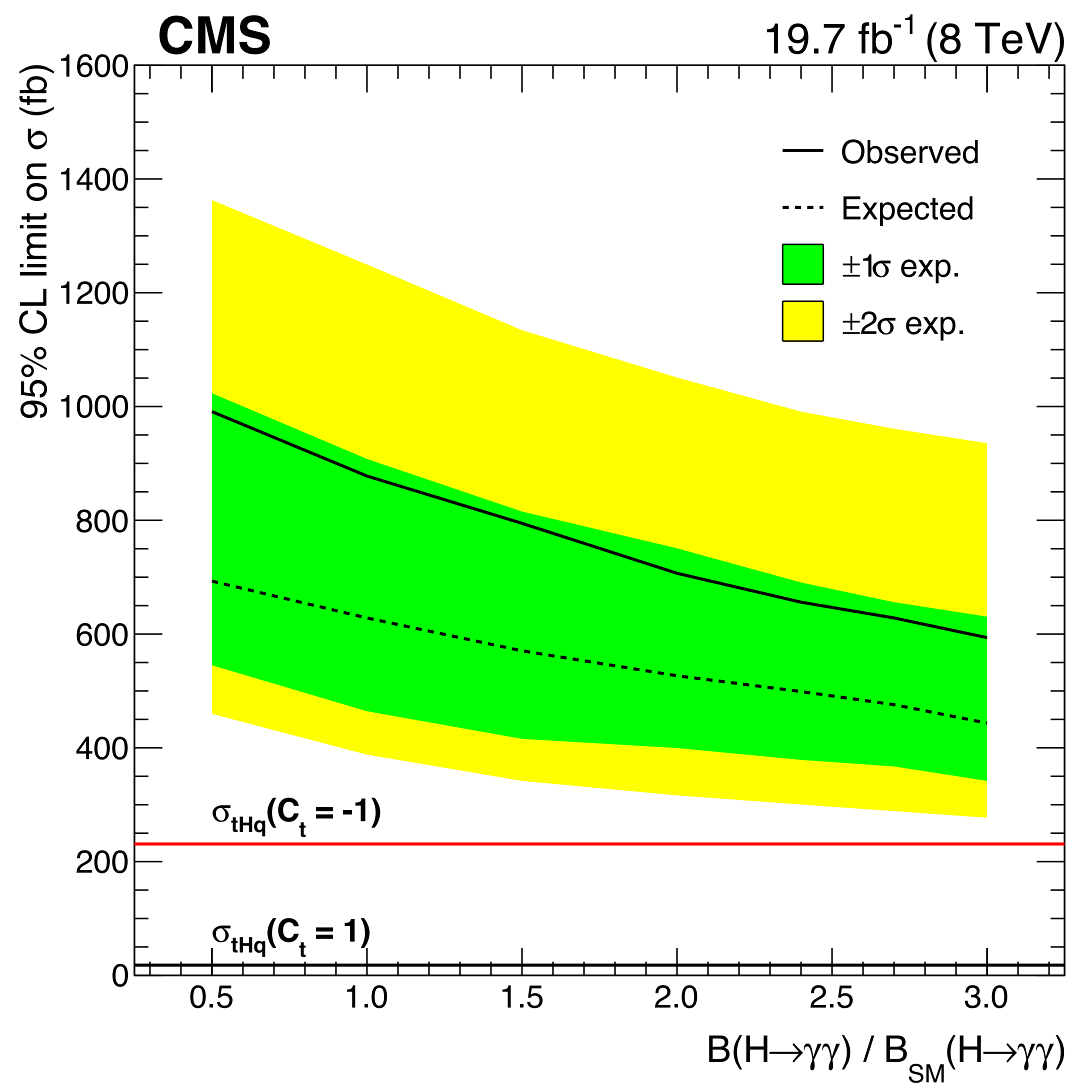




Results

- No significant excess of events observed over background-only expectation
- Set limits on **tHq production cross section relative to C_t = -1** expectation and assuming SM branching fractions: observed 95% CL limit of 2.8 (2.0 expected) on $\sigma/\sigma_{Ct=-1}$

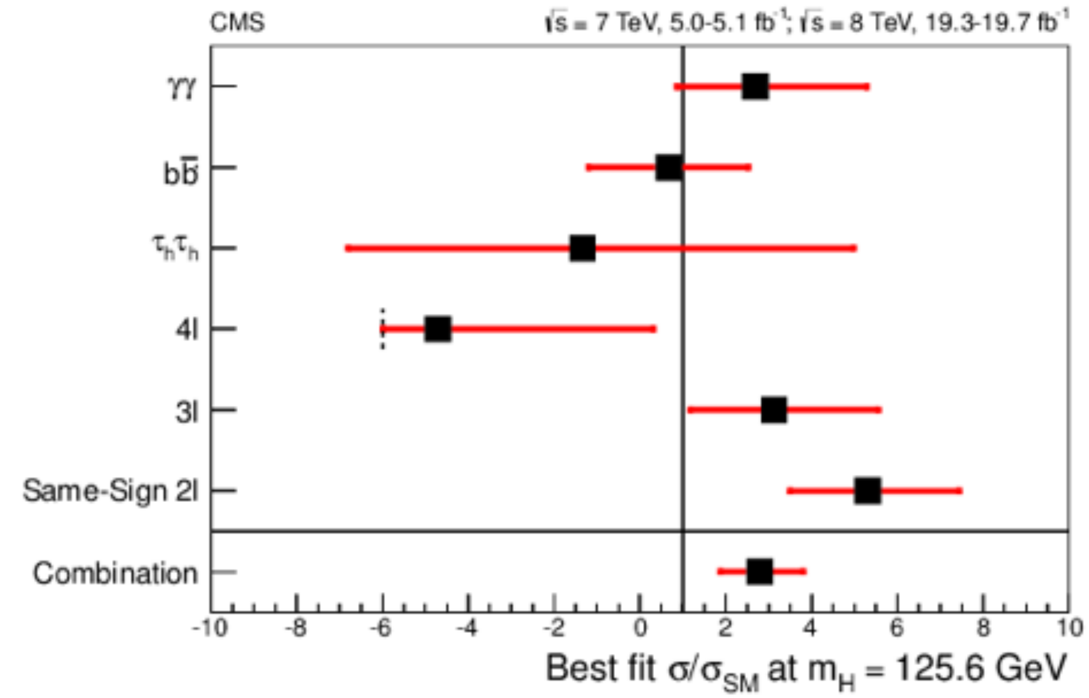


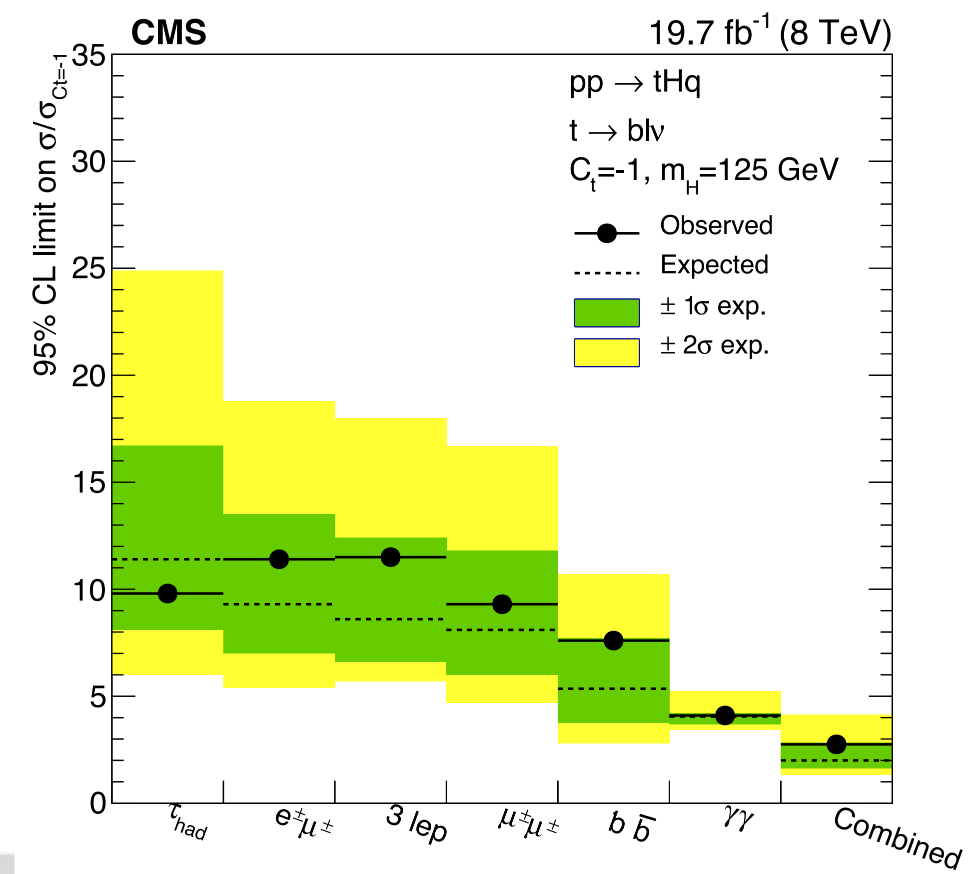


- As C_t = -1 also enhances H→γγ for all production modes, also set limits on **tHq cross section as a function of B(H→γγ)/B_{SM}(H→γγ)**. Observed limits in range 600-1000 fb (450-700 fb expected)

Summary

- CMS has performed a search for the ttH process utilising leptonic and hadronic tt final states and several Higgs decay modes
- Combined signal strength:
  - $\hat{\mu} = 2.8^{+1.0}_{-0.9}$ @ 125.6 GeV
- Good prospects for LHC Run 2 due to factor ~4 enhancement in signal cross section (though backgrounds also enhanced)
- Dedicated searches for the enhancement of single-top + Higgs production
  - Observed limit of 2.8 (2.0 expected) on $\sigma/\sigma_{Ct=-1}$





Backup

Systematic Uncertainties

$H\to\tau_h\tau_h$ | $H\to bb$ | $H\to WW/\tau\tau/ZZ$ | $H\to\gamma\gamma$

Source	Rate uncertainty: Signal	Rate uncertainty: Backgrounds	Shape
Experimental			
Integrated luminosity	2.2–2.6%	2.2–2.6%	No
Jet energy scale	0.0–8.4%	0.1–11.5%	Yes
CSV b-tagging	0.9–21.7%	3.0–29.0%	Yes
Lepton reco. and ID	0.3–14.0%	1.4–14.0%	No
Lepton misidentification rate ($H\to$ leptons)	—	35.1–45.7%	Yes
Tau reco. and ID ($H\to$ hadrons)	11.3–14.3%	24.1–28.8%	Yes
Photon reco. and ID ($H\to$ photons)	1.6–3.2%	—	Yes
MC statistics	—	0.2–7.0%	Yes
Theoretical			
NLO scales and PDF	9.7–14.8%	3.4–14.7%	No
MC modeling	2.3–5.1%	0.9–16.8%	Yes
Top quark p_T	—	1.4–6.9%	Yes
Additional hf uncertainty ($H\to$ hadrons)	—	50%	No
H contamination ($H\to$ photons)	36.7–41.2%		No
WZ (ZZ) uncertainty ($H\to$ leptons)	—	22% (19%)	No

Backup

Category	Signature	Trigger	Signature
H → Hadrons $H \to b\bar{b}$ $H \to \tau_h\tau_h$ $H \to WW$	Lepton + Jets ($t\bar{t}H \to \ell\nu jjbbbb$)	Single Lepton	1 e/μ, $p_T > 30$ GeV $\geq$4 jets + $\geq$2 b-tags, $p_T > 30$ GeV
	Dilepton ($t\bar{t}H \to \ell\nu\ell\nu bbbb$)	Dilepton	1 e/μ, $p_T > 20$ GeV 1 e/μ, $p_T > 10$ GeV $\geq$3 jets + $\geq$2 b-tags, $p_T > 30$ GeV
	Hadronic τ ($t\bar{t}H \to \ell\nu\tau_h[\nu]\tau_h[\nu]jjbb$)	Single Lepton	1 e/μ, $p_T > 30$ GeV 2 τ_h, $p_T > 20$ GeV $\geq$2 jets + 1-2 b-tags, $p_T > 30$ GeV
H → Photons $H \to \gamma\gamma$	Leptonic ($t\bar{t}H \to \ell\nu jjbb\gamma\gamma$, $t\bar{t}H \to \ell\nu\ell\nu bb\gamma\gamma$)	Diphoton	2 γ, $p_T > m_{\gamma\gamma}/2\,(25)$ GeV for 1^{st} (2^{nd}) $\geq$1 e/μ, $p_T > 20$ GeV $\geq$2 jets + $\geq$1 b-tags, $p_T > 25$ GeV
	Hadronic ($t\bar{t}H \to jjjjbb\gamma\gamma$)	Diphoton	2 γ, $p_T > m_{\gamma\gamma}/2\,(25)$ GeV for 1^{st} (2^{nd}) 0 e/μ, $p_T > 20$ GeV $\geq$4 jets + $\geq$1 b-tags, $p_T > 25$ GeV
H → Leptons $H \to WW$ $H \to \tau\tau$ $H \to ZZ$	Same-Sign Dilepton ($t\bar{t}H \to \ell^{\pm}\nu\ell^{\pm}[\nu]jjj[j]bb$)	Dilepton	2 e/μ, $p_T > 20$ GeV $\geq$4 jets + $\geq$1 b-tags, $p_T > 25$ GeV
	3 Lepton ($t\bar{t}H \to \ell\nu\ell[\nu]\ell[\nu]j[j]bb$)	Dilepton, Trielectron	1 e/μ, $p_T > 20$ GeV 1 e/μ, $p_T > 10$ GeV 1 $e(\mu)$, $p_T > 7(5)$ GeV $\geq$2 jets + $\geq$1 b-tags, $p_T > 25$ GeV
	4 Lepton ($t\bar{t}H \to \ell\nu\ell\nu\ell[\nu]\ell[\nu]bb$)	Dilepton, Trielectron	1 e/μ, $p_T > 20$ GeV 1 e/μ, $p_T > 10$ GeV 2 $e(\mu)$, $p_T > 7(5)$ GeV $\geq$2 jets + $\geq$1 b-tags, $p_T > 25$ GeV

Event yields in the H→leptons search channels

	ee	$e\mu$	$\mu\mu$	3ℓ	4ℓ
$\mathrm{t\bar{t}H}$, H → WW	1.0 ± 0.1	3.2 ± 0.4	2.4 ± 0.3	3.4 ± 0.5	0.29 ± 0.04
$\mathrm{t\bar{t}H}$, H → ZZ	—	0.1 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	0.09 ± 0.02
$\mathrm{t\bar{t}H}$, H → $\tau\tau$	0.3 ± 0.0	1.0 ± 0.1	0.7 ± 0.1	1.1 ± 0.2	0.15 ± 0.02
$\mathrm{t\bar{t}}\,W$	4.3 ± 0.6	16.5 ± 2.3	10.4 ± 1.5	10.3 ± 1.9	—
$\mathrm{t\bar{t}\,Z}/\gamma^*$	1.8 ± 0.4	4.9 ± 0.9	2.9 ± 0.5	8.4 ± 1.7	1.12 ± 0.62
$\mathrm{t\bar{t}}\,WW$	0.1 ± 0.0	0.4 ± 0.1	0.3 ± 0.0	0.4 ± 0.1	0.04 ± 0.02
$\mathrm{t\bar{t}}\,\gamma$	1.3 ± 0.3	1.9 ± 0.5	—	2.6 ± 0.6	—
WZ	0.6 ± 0.6	1.5 ± 1.7	1.0 ± 1.1	3.9 ± 0.7	—
ZZ	—	0.1 ± 0.1	0.1 ± 0.0	0.3 ± 0.1	0.47 ± 0.10
Rare SM bkg.	0.4 ± 0.1	1.6 ± 0.4	1.1 ± 0.3	0.8 ± 0.3	0.01 ± 0.00
Non-prompt	7.6 ± 2.5	20.0 ± 4.4	11.9 ± 4.2	33.3 ± 7.5	0.43 ± 0.22
Charge misidentified	1.8 ± 0.5	2.3 ± 0.7	—	—	—
All signals	1.4 ± 0.2	4.3 ± 0.6	3.1 ± 0.4	4.7 ± 0.7	0.54 ± 0.08
All backgrounds	18.0 ± 2.7	49.3 ± 5.4	27.7 ± 4.7	59.8 ± 8.0	2.07 ± 0.67
Data	19	51	41	68	1