

# Search of ttH production with in multi lepton signatures at LHC Run1

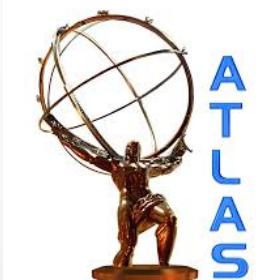
Top LHC France 2015 - Lyon

Djamel BOUMEDIENE – LPC

May 19, 2015



Laboratoire de Physique Corpusculaire



# Motivation

- Evidence of the Higgs coupling to fermions is a milestone in Higgs studies
- Top Yukawa coupling is the most important one – several motivations
- Running of Higgs self coupling ( $\lambda$ ) sensitive to Top yukawa coupling ( $y_t$ )

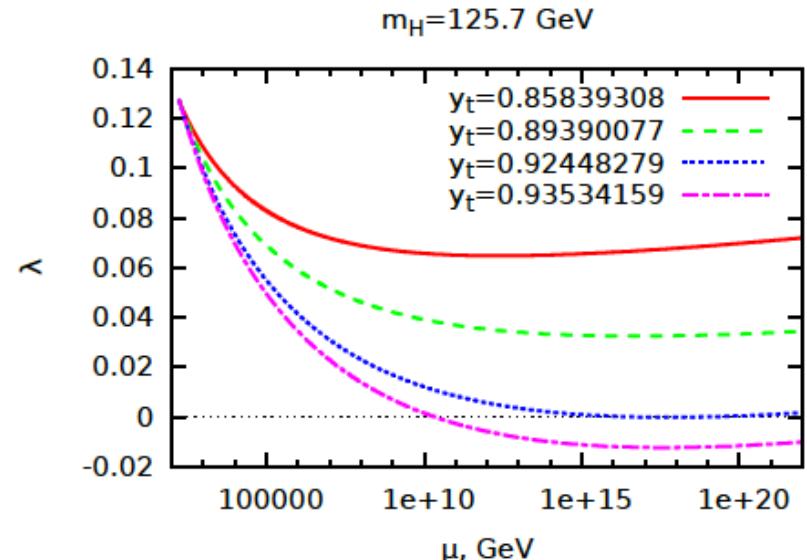
$$16\pi^2 \frac{d\lambda}{d\ln\mu} = 24\lambda^2 + 12\lambda y_t^2 - 9\lambda(g^2 + \frac{1}{3}g'^2) - 6y_t^4 + \frac{9}{8}g^4 + \frac{3}{8}g'^4 + \frac{3}{4}g^2g'^2$$

- Existence of a critical  $y_t$  above which vacuum is unstable:

How does it compare to measured  $y_t$  ?

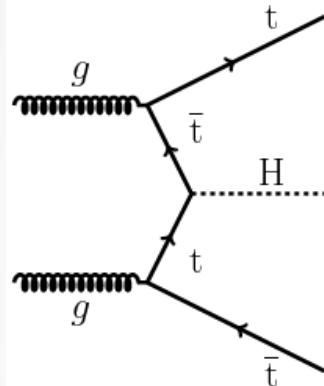
- $y_t$  can be determined:

- From Top mass measurement (requires a careful interpretation of the MC Top mass)
- From Higgs production and  $\gamma\gamma$  decay (assumes SM decay)
- From  $t\bar{t}H$  – a test at the tree level. It will provide evidence of the coupling existence and its value

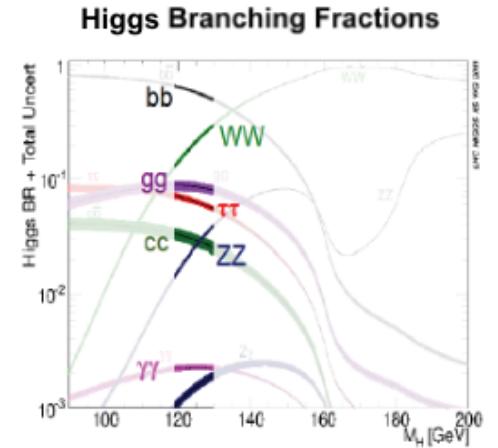
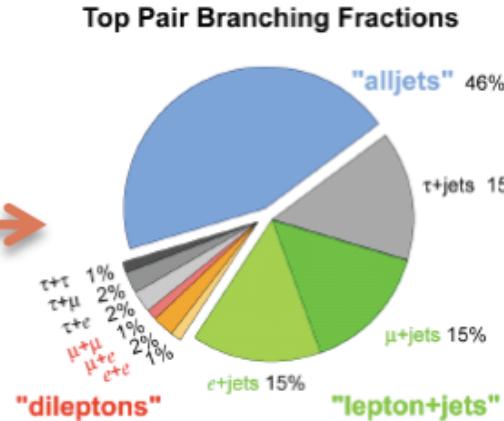


F. Bezrukov, M. Shaposhnikov

# ttH signature



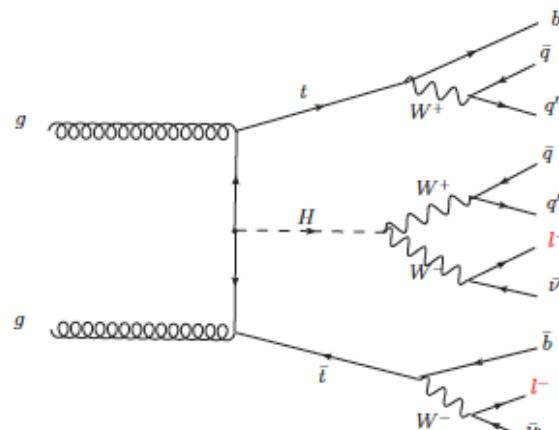
$\sqrt{s}=8\text{TeV}$ :  $\sigma(\text{ttH}) = 130 \text{ fb}^{-1}$   
 ~2700 events / experiment



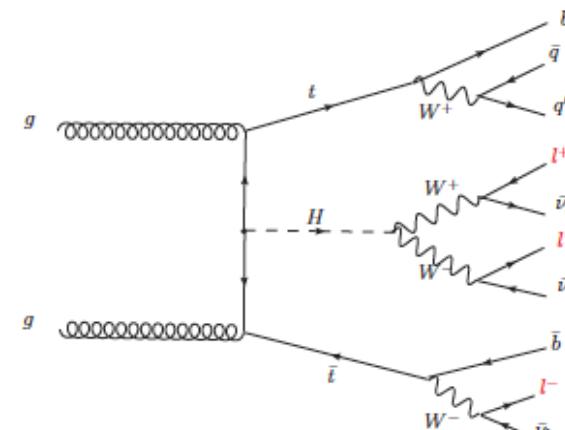
- ttH final state combines top pair decay signature and Higgs decay signature → large number of possible final states
- 3 families of signatures: 4b+X ( $H \rightarrow bb$ ), 2b+2 $\gamma$  ( $H \rightarrow \gamma\gamma$ ), **2b+leptons ( $H \rightarrow WW, ZZ, \tau\tau$ )**  
 ↴ *See Daniele Madaffari talk*
- Leptonic signatures based on flavour (e, $\mu$ , $\tau$ ) and charge can be used to select a ttH enriched sample

# Leptonic channels

- 5 channels defined by the number of leptons ( $e, \mu, \tau$ )



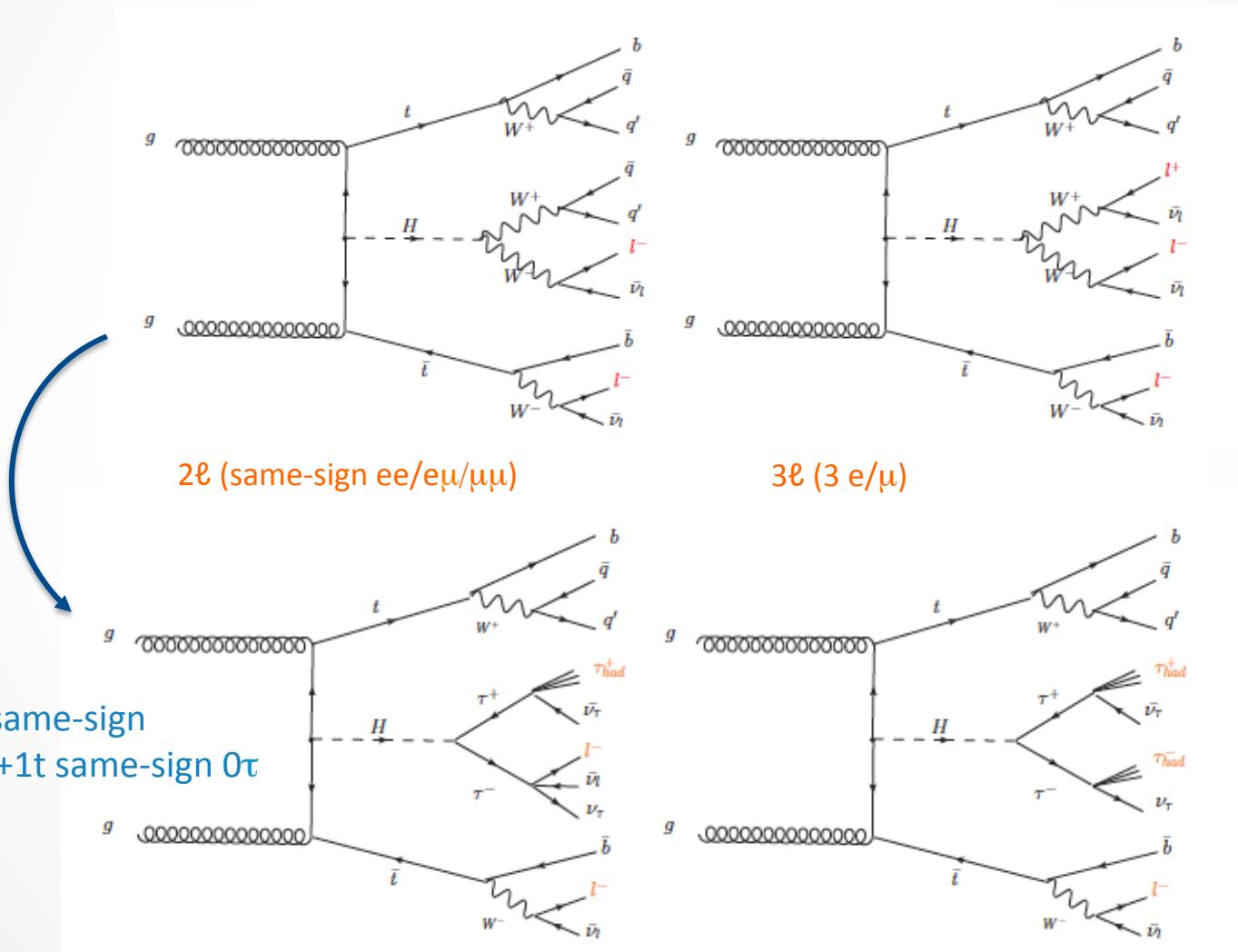
2 $\ell$  (same-sign ee/emu/mumu)



3 $\ell$  (3 e/  $\mu$ )

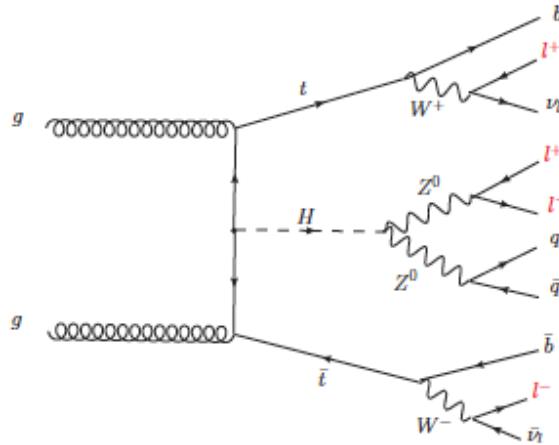
# Leptonic channels

- 5 channels defined by the number of leptons ( $e, \mu, \tau$ )



# Leptonic channels

- 5 channels defined by the number of leptons ( $e, \mu, \tau$ )



4 leptons

# Analysis strategy

- ttH into lepton channels are sensitive to several Higgs decays
- ttH production process combines several Higgs coupling products:

Category	Higgs boson decay mode			
	$WW^*$	$\tau\tau$	$ZZ^*$	other
$2\ell 0\tau_{\text{had}}$	80%	15%	3%	2%
$3\ell$	74%	15%	7%	4%
$2\ell 1\tau_{\text{had}}$	35%	62%	2%	1%
$4\ell$	69%	14%	14%	4%
$1\ell 2\tau_{\text{had}}$	4%	93%	0%	3%

Fractions of Higgs decays/signature

- Two approaches:
  - A. Assume that Higgs decay branching fractions are known
    - Search for inclusive ttH production
    - Assume that Higgs decay branching fractions are known (determined at NNLO and applied in MC simulation)
    - Consider other processes sensitive to Top Yukawa coupling constant ( $tH$ ) – small impact however  
[JHEP 1409 \(2014\) 087](#) [ATLAS CONF-2015-006](#)
  - B. Explicit coupling fit combining all Higgs inputs (including ttH)
    - Cf. [ATLAS-CONF-2015-007](#) / [JHEP 1409 \(2014\) 087 \(arXiv: 1412.8662\)](#)
- D. BOUMEDIENE (LPC)

# Analysis strategy

- ttH into lepton channels are sensitive to several Higgs decays
- ttH production process combines several Higgs coupling products:

Category	Higgs boson decay mode			
	$WW^*$	$\tau\tau$	$ZZ^*$	other
$2\ell 0\tau_{had}$	80%	15%	3%	2%
$3\ell$	74%	15%	7%	4%
$2\ell 1\tau_{had}$	35%	62%	2%	1%
$4\ell$	69%	14%	14%	4%
$1\ell 2\tau_{had}$	4%	93%	0%	3%

Fractions of Higgs decays/signature

- Two approaches:

## A. Assume that Higgs decay branching fractions are known

- Search for inclusive ttH production
- Assume that Higgs decay branching fractions are known (determined at NNLO and applied in MC simulation)
- Consider other processes sensitive to Top Yukawa coupling constant ( $tH$ ) – small impact however

Shown Today



JHEP 1409 (2014) 087 ATLAS CONF-2015-006

# Channel and Object definitions

- Selection of  $e, \mu, \tau, \text{jets}$  optimised on MC simulation against main background processes: ttV, ttbar, VV
- Main object properties (ID,  $pT$ , isolation,  $d0$ , ...) scanned in order to maximise sensitivity to the signal → optimisation of the combination (take into account event migration across channels) → Clermont
- For a first measurement/search: use of counting experiments (ATLAS)
  - $2\ell$ :  $\geq 1$  b-tagged jet,  $=4$  jets and  $\geq 5$  jets  $\times (ee, \mu\mu, e\mu)$  → 6 categories → Clermont
  - $3\ell$ :  $= 2$  b-tagged jets and 3 jets or  $\geq 1$  b-tagged jet  $\geq 4$  jets → 1 category → Strasbourg
  - $2\ell 1\tau$ :  $\geq 1$  b-tagged jet,  $\geq 4$  jets → 1 category
  - $2\tau 1\ell$ :  $\geq 1$ -btagged jet,  $\geq 3$  jets → 1 category
  - $4\ell$ :  $\geq 1$  b-tagged jet,  $\geq 2$  jets → 2 categories
- For comparison: CMS uses selection + BDT discriminant (CMS)
  - All:  $\geq 2$  b-tagged jet (1 medium + 1 loose),  $\geq 2$  jets
  - $2\ell, 3\ell$ : categorised in ++ and - - (exploits SM charge symmetry of ttV, single top, W+jets)
  - BDT applied to selected events and used as discriminant

# Main background processes

- Typical background composition per channel:

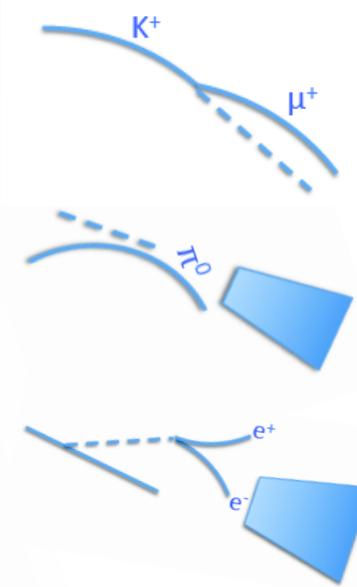
Signature	Top	ttV	VV	V+jets
2 $\ell$	30%	50%	10%	10%
3 $\ell$	15%	65%	10%	7%
2 $\ell$ 1t	43%	45%	10%	< 1%
2t1 $\ell$	90%	5%	3%	2%
4 $\ell$	<1%	78%	10%	< 1%

*approx numbers from MC  
Full estimate in slide 14*

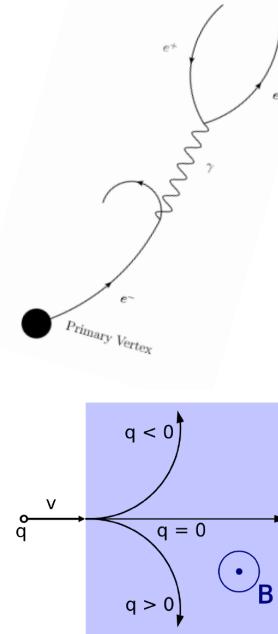
- Main expected background processes:
  - Irreducible**: ttV, VV (leads to very similar signature with prompt leptons) → use of MC simulation / fit to the data  
→ CMS: use of template method for WZ estimate. → **Performed by Strasbourg**
  - Reducible**: top (non prompt leptons or charge flips) → use of data-driven techniques

# Two types of reducible backgrounds

## Fake/non-prompt leptons



## Charge mis-Identification



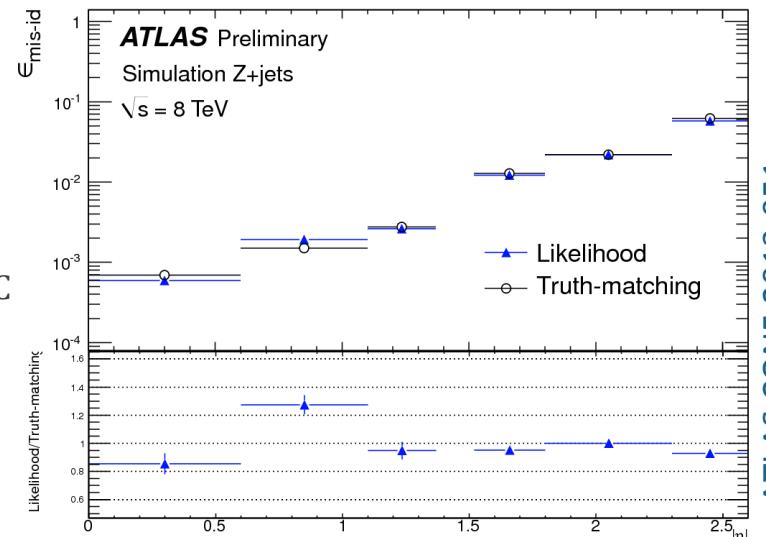
- Non prompt leptons: any lepton not produced by  $W/Z/\tau$  decay is source of background
- Strong contribution from top pair decays (mainly due to  $B$  decays)

- Misidentification of the lepton charge
- Concerns electrons mainly
- Affect same-sign channels

Cannot rely on Monte Carlo simulation for their estimate → use of data-driven methods

# Charge mis-identification

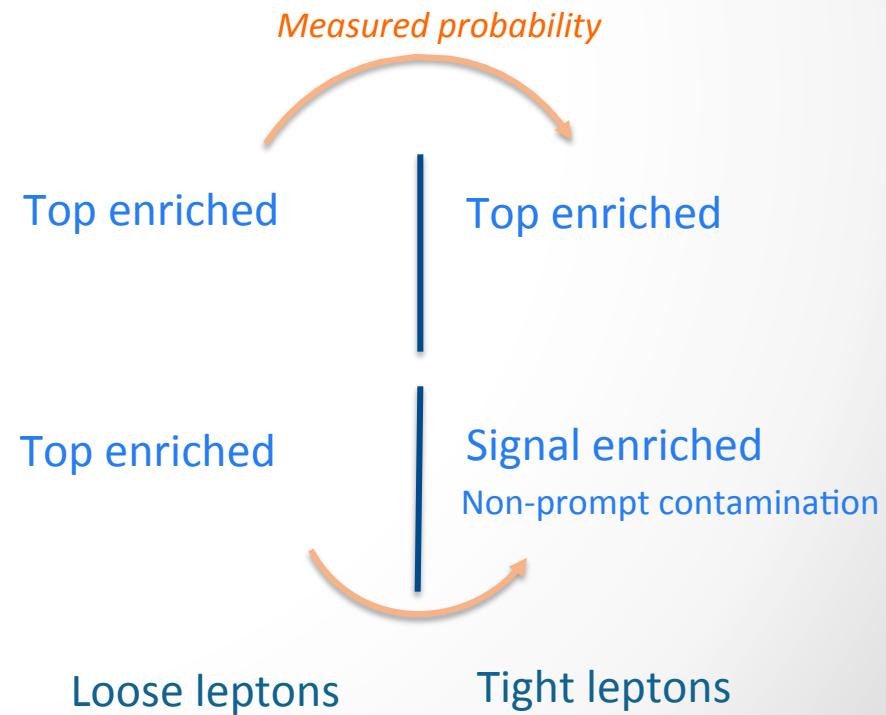
- Same approach is used in exotic same-sign studies
- Electron mis-identified charge:
  - Due to trident events or large curvature → depends on lepton properties and interaction with detector ( $p_T$ ,  $|\eta|$ )
  - Negligible for muons
- Affects  $2\ell$  ee/e $\mu$  channels
- Likelihood method:
  - Z peak OS and SS events used to compute charge flip probability
  - Probability binned in  $|\eta|$  and  $p_T$
  - ATLAS: Probability extrapolation to high  $p_T$  based on MC
- Systematic uncertainty:
  - Includes likelihood statistic, Z peak definition, extrapolation, closure test, ...
  - Total systematic  $\sim 40\%$  (mainly due to statistics)



→ Clermont, Daniela Paredes (LPC - PhD 2013)

# Non-prompt leptons

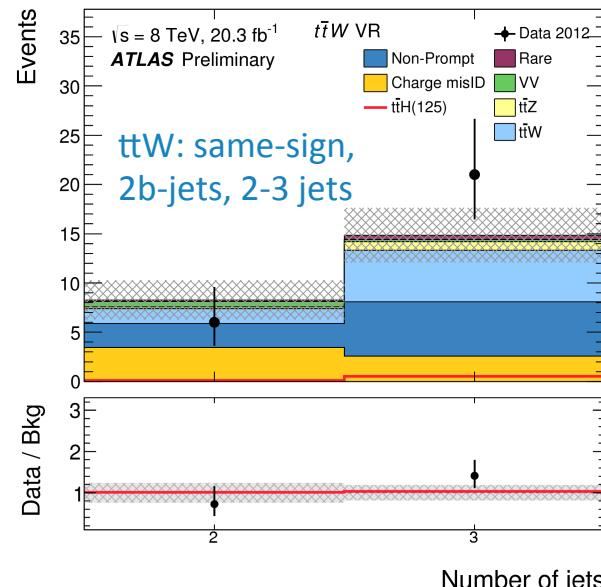
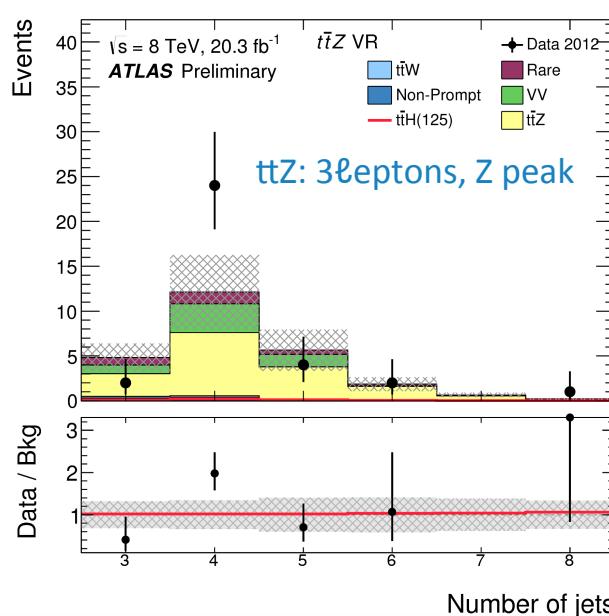
- In-situ Technique:
  - use of a control region enriched in non-prompt leptons (loose lepton region)
  - Reweighted using a probability to predict non-prompt lepton contamination in signal region
  - Probabilities measured in a dedicated fake region:  
Top enriched region, low jet multiplicity



→ Implemented by Clermont

# ttV estimate

- In ATLAS (and CMS): ttV estimate based on MC simulation:
  - $t\bar{t}Z$  ( $t\bar{t}Z/\gamma^*$ ),  $t\bar{t}W$ ,  $t\bar{t}WW$
  - Data does not allow precise constraint, degenerated with SR (especially  $t\bar{t}W$ )

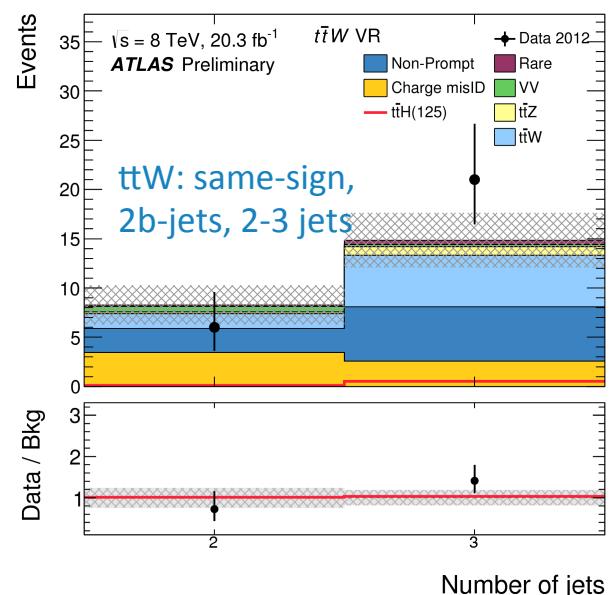
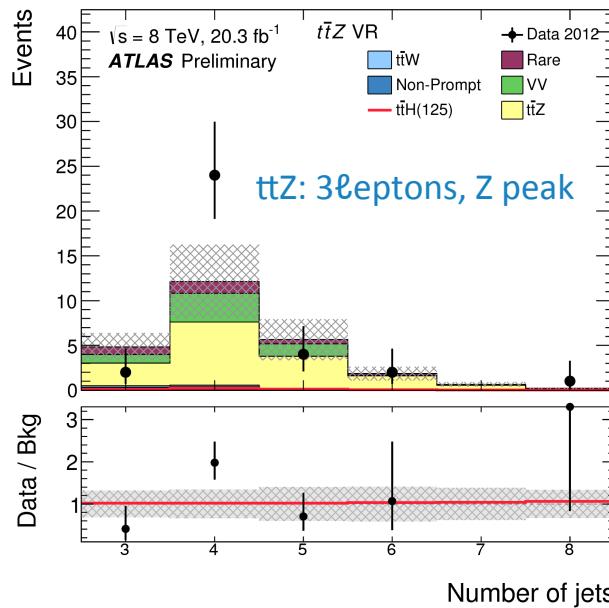


ATLAS CONF-2015-006

- ttV questions:**
  - Does it make a sense to use data-driven estimate when theoretical calculations do better ?
  - Is it true if the ttV process was not yet « discovered » ?
  - Does it make a sense to doubt about these calculation while the simulation is used to compute the signal acceptance ?
  - Different calculations for cross-section are available, which one to use ?

# ttV estimate

- In ATLAS (and CMS): ttV estimate based on MC simulation:
  - ttZ ( $t\bar{t}Z/\gamma^*$ ), ttW, ttWW
  - Data does not allow precise constraint, degenerated with SR (especially ttW)



- Use of NLO cross-sections
- LO generator Madgraph+Pythia6
- Two calculations are available:
  - J. M. Campbell & al. (ttW):  $\sigma=232 \pm 28$  (scale) fb → used for ttW (ATLAS)
  - M. Garzelli & al. (ttW, ttZ):  $\sigma=206 \pm 23$  (scale) fb → used for ttZ (ATLAS, CMS) for ttW (CMS)
- Final result is also expressed versus ttV cross-section (+ 2D fits for cross-check)
  - Clermont, Arthur Chomont - LPC (ongoing PhD)

# Results: Yields

- Nominal background predictions compared to observed number of events

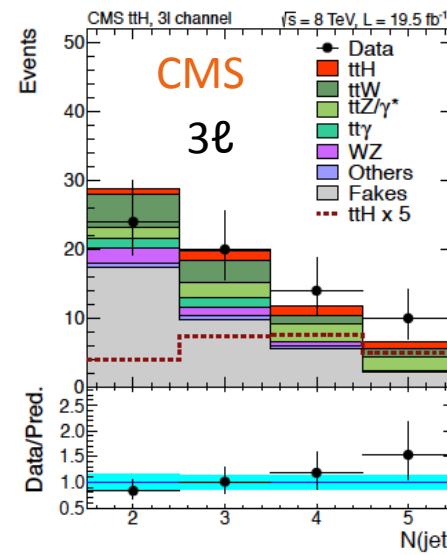
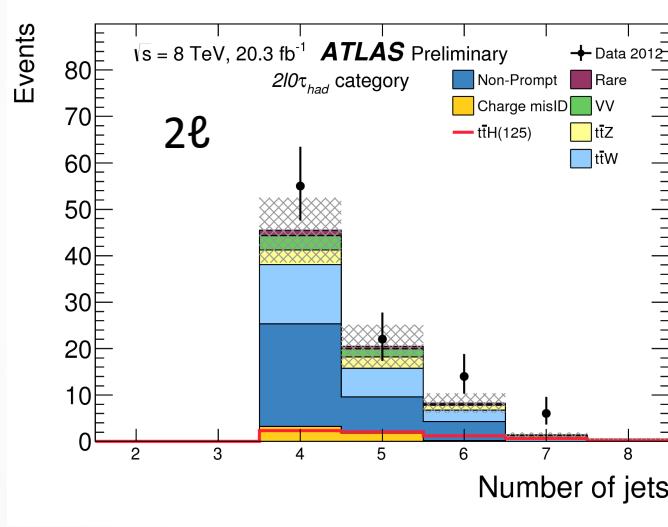
Category	$q$ mis-id	Non-prompt	$t\bar{t}W$	$t\bar{t}Z$	Diboson	Expected Bkg.	$t\bar{t}H (\mu = 1)$	Observed
$ee + \geq 5j$	$1.1 \pm 0.5$	$2.3 \pm 1.2$	$1.4 \pm 0.4$	$0.98 \pm 0.32$	$0.47 \pm 0.42$	$6.5 \pm 2.0$	$0.73 \pm 0.11$	10
$e\mu + \geq 5j$	$0.85 \pm 0.35$	$6.7 \pm 2.4$	$4.8 \pm 1.4$	$2.1 \pm 0.7$	$0.38 \pm 0.32$	$15 \pm 4$	$2.13 \pm 0.31$	22
$\mu\mu + \geq 5j$	–	$2.9 \pm 1.4$	$3.8 \pm 1.1$	$0.95 \pm 0.31$	$0.69 \pm 0.63$	$8.6 \pm 2.5$	$1.41 \pm 0.21$	11
$ee + 4j$	$1.8 \pm 0.7$	$3.4 \pm 1.7$	$2.0 \pm 0.4$	$0.75 \pm 0.25$	$0.74 \pm 0.58$	$9.1 \pm 2.3$	$0.44 \pm 0.06$	9
$e\mu + 4j$	$1.4 \pm 0.6$	$12 \pm 4$	$6.2 \pm 0.9$	$1.5 \pm 0.2$	$1.9 \pm 1.2$	$24.0 \pm 4.5$	$1.16 \pm 0.14$	26
$\mu\mu + 4j$	–	$6.3 \pm 2.6$	$4.7 \pm 0.9$	$0.80 \pm 0.26$	$0.53 \pm 0.30$	$12.7 \pm 3.0$	$0.74 \pm 0.10$	20
$3\ell$	–	$2.6 \pm 0.6$	$2.3 \pm 0.9$	$3.9 \pm 0.9$	$0.86 \pm 0.59$	$11.4 \pm 3.1$	$2.34 \pm 0.32$	18
$2\ell 1\tau_{\text{had}}$	–	$0.4^{+0.6}_{-0.4}$	$0.38 \pm 0.15$	$0.37 \pm 0.09$	$0.12 \pm 0.15$	$1.4 \pm 0.6$	$0.47 \pm 0.02$	1
$1\ell 2\tau_{\text{had}}$	–	$15 \pm 5$	$0.17 \pm 0.07$	$0.37 \pm 0.10$	$0.41 \pm 0.42$	$16 \pm 6$	$0.68 \pm 0.07$	10
$4\ell Z\text{-enr.}$	–	$\lesssim 10^{-3}$	$\lesssim 3 \times 10^{-3}$	$0.43 \pm 0.13$	$0.05 \pm 0.02$	$0.55 \pm 0.17$	$0.17 \pm 0.01$	1
$4\ell Z\text{-dep.}$	–	$\lesssim 10^{-4}$	$\lesssim 10^{-3}$	$0.002 \pm 0.002$	$\lesssim 2 \times 10^{-5}$	$0.007 \pm 0.005$	$0.03 \pm 0.00$	0

ATLAS CONF-2015-006

- Observed data excess w.r.t. background in 7 regions over 11 regions
- $2\ell 0\tau$ :
  - Overall Data excess. Largest excess in  $\mu\mu 4j$
  - 2 leading backgrounds:  $t\bar{t}V$ , Fakes

# Results: distributions

- Jet multiplicity distributions in  $2\ell/3\ell$



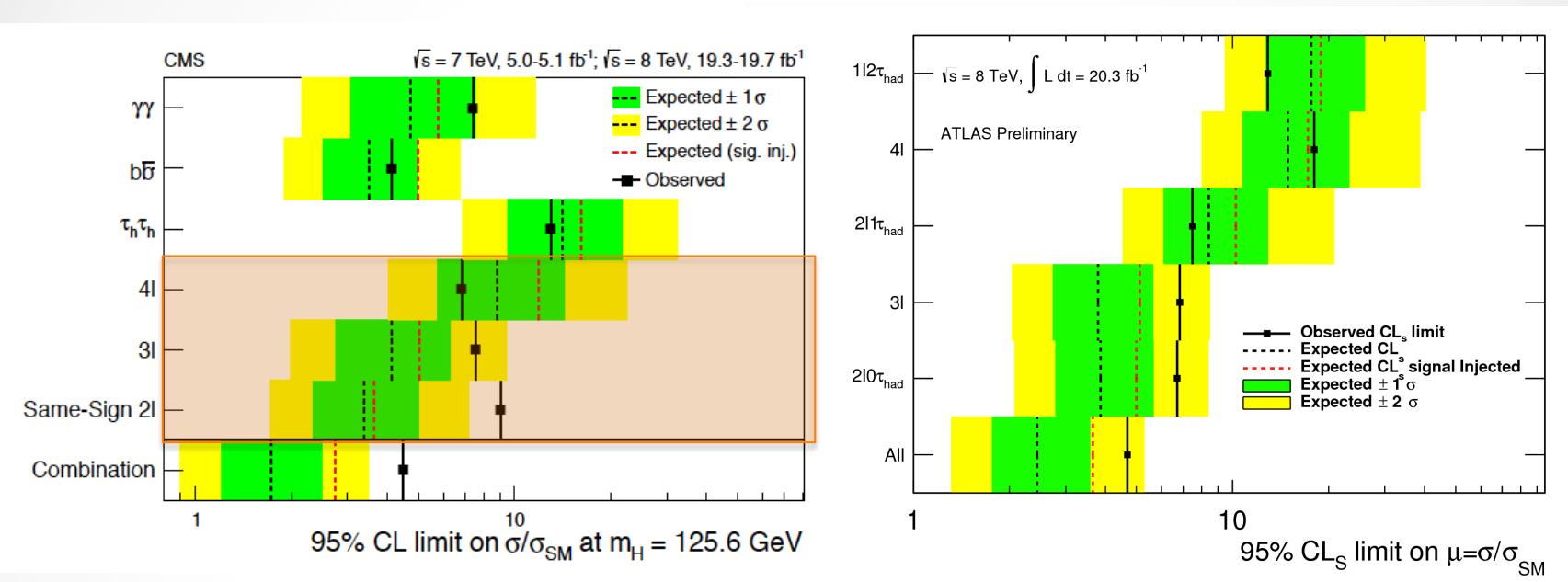
# Fit procedure

- Maximum likelihood fit of signal strength
- Floating systematics uncertainties
- 5 Nuisance Parameters with largest impact on signal strength:
  - Uncertainty on Fake  $\mu$  in  $2\ell$
  - $t\bar{t}W$  acceptance uncertainty
  - $t\bar{t}H$  cross section uncertainty
  - Jet energy scale uncertainty
  - Uncertainty on Fake e in  $2\ell$

Source	$\Delta\mu$	
$2\ell 0\tau_{\text{had}}$ non-prompt muon transfer factor	+0.38	-0.35
$t\bar{t}W$ acceptance	+0.26	-0.21
$t\bar{t}H$ inclusive cross section	+0.28	-0.15
Jet energy scale	+0.24	-0.18
$2\ell 0\tau_{\text{had}}$ non-prompt electron transfer factor	+0.26	-0.16
$t\bar{t}H$ acceptance	+0.22	-0.15
$t\bar{t}Z$ inclusive cross section	+0.19	-0.17
$t\bar{t}W$ inclusive cross section	+0.18	-0.15
Muon isolation efficiency	+0.19	-0.14
Luminosity	+0.18	-0.14

ATLAS CONF-2015-006

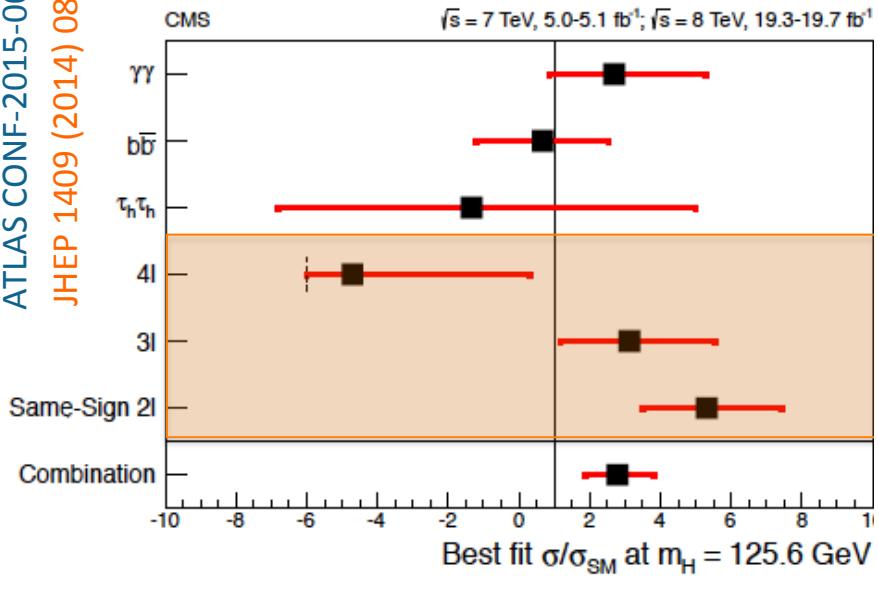
# Fit result, exclusion limit



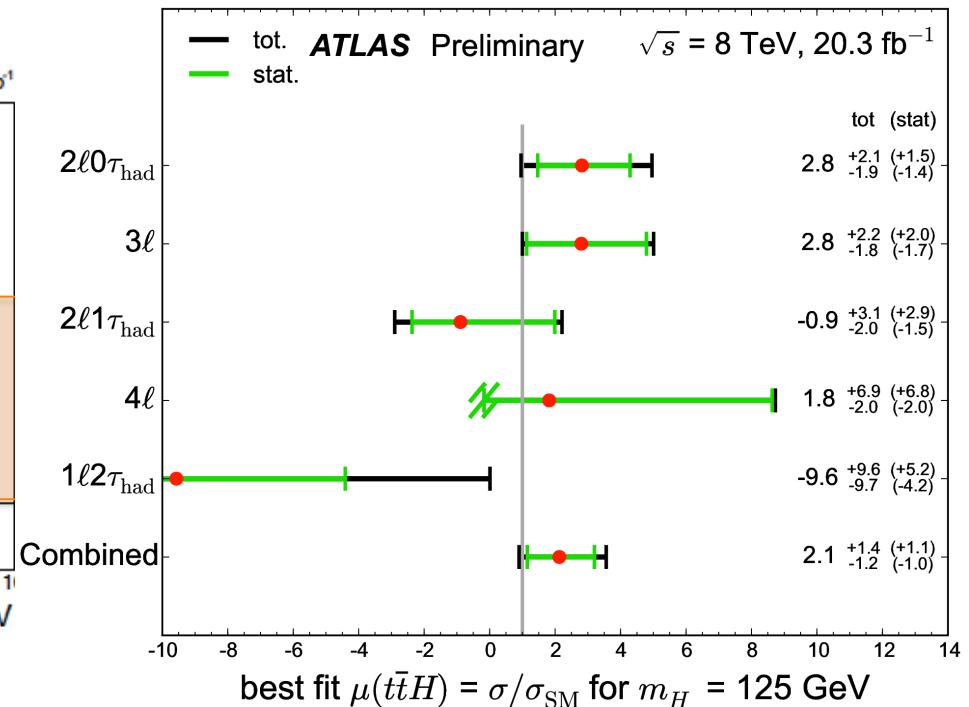
Channel	Observed Limit	Expected Limit						Median ( $\mu = 1$ )
		$-2\sigma$	$-1\sigma$	Median	$+1\sigma$	$+2\sigma$		
$2\ell 0\tau_{\text{had}}$	6.7	2.1	2.8	3.9	5.7	8.4	5.0	
$3\ell$	6.8	2.0	2.7	3.8	5.7	8.5	5.1	
$2\ell 1\tau_{\text{had}}$	7.5	4.5	6.1	8.4	13.0	20.8	10.3	
$4\ell$	18.1	8.0	10.8	14.9	23.3	38.8	17.2	
$1\ell 2\tau_{\text{had}}$	12.9	9.5	12.7	17.6	26.1	40.4	18.9	
Combined	4.7	1.3	1.8	2.4	3.6	5.3	3.7	

Expected Median in presence of SM signal

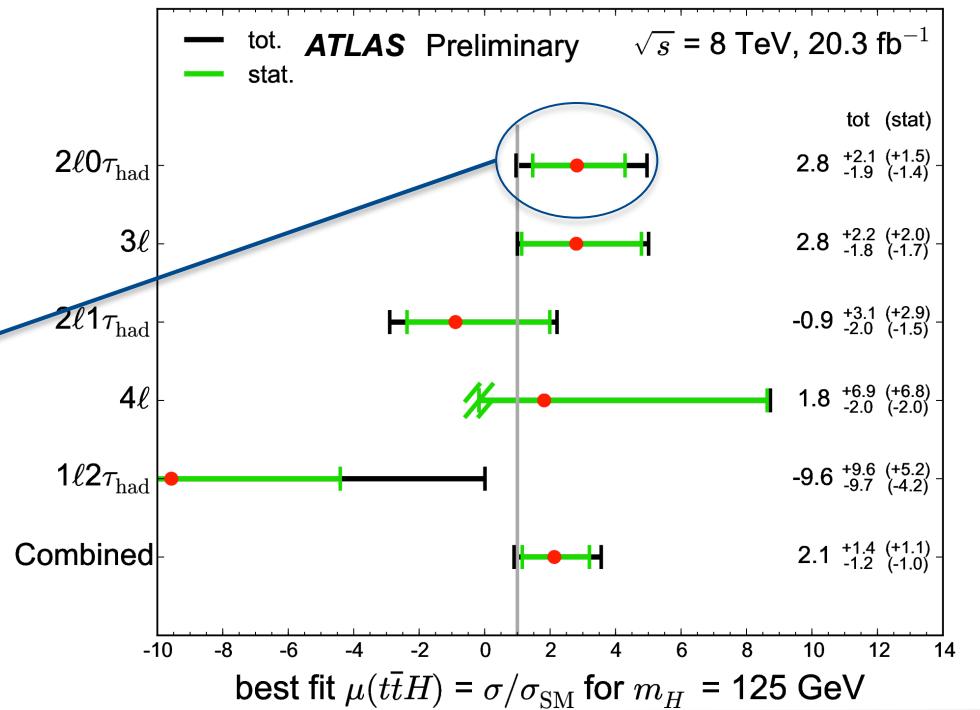
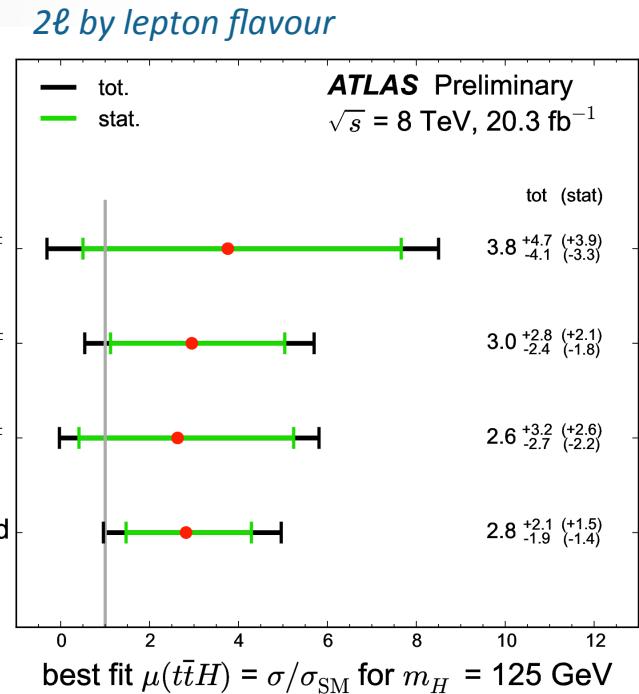
# Fit result: signal strength



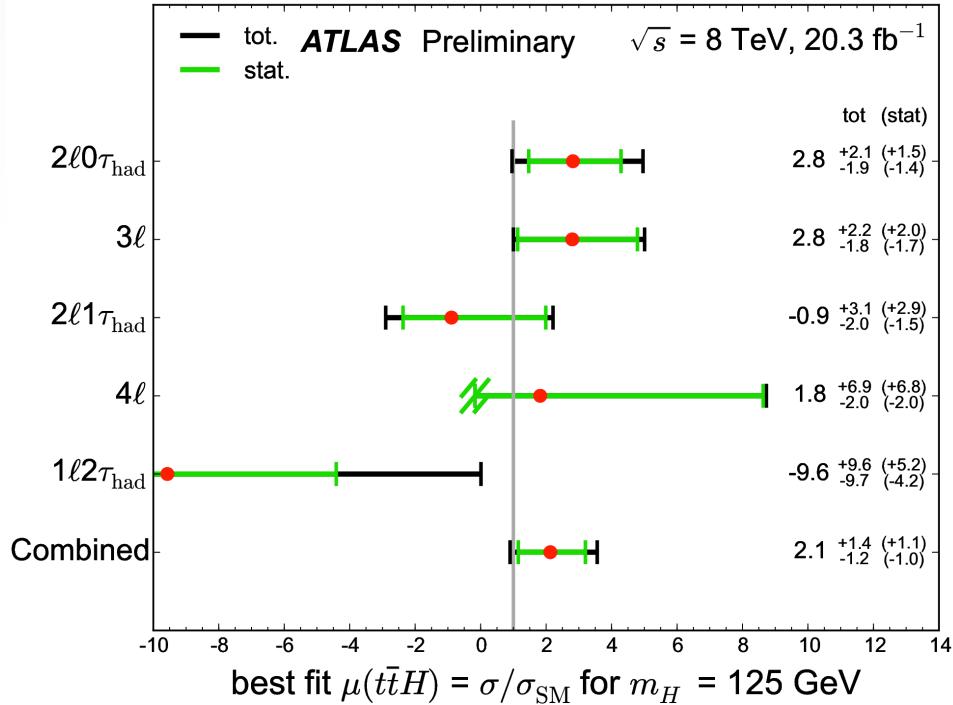
$$\mu(t\bar{t}H) = 3.7 + 1.6/-1.4 \text{ (CMS)}$$



# Fit result: signal strength



# Fit result: signal strength



- Best fit  $\mu$ :  $\mu(t\bar{t}H) = 2.1 + 1.4/-1.2$  (ATLAS) /  $\mu(t\bar{t}H) = 3.7 + 1.6/-1.4$  (CMS)
- Excess w.r.t. background only hypothesis:  $1.8 \sigma$  /  $1.8 \sigma$
- Excess w.r.t. background and SM signal:  $0.9 \sigma$  /  $1.8 \sigma$
- Dependence of best  $\mu$  to ttV cross-section:  
(ATLAS), Arthur Chomont (Clermont)

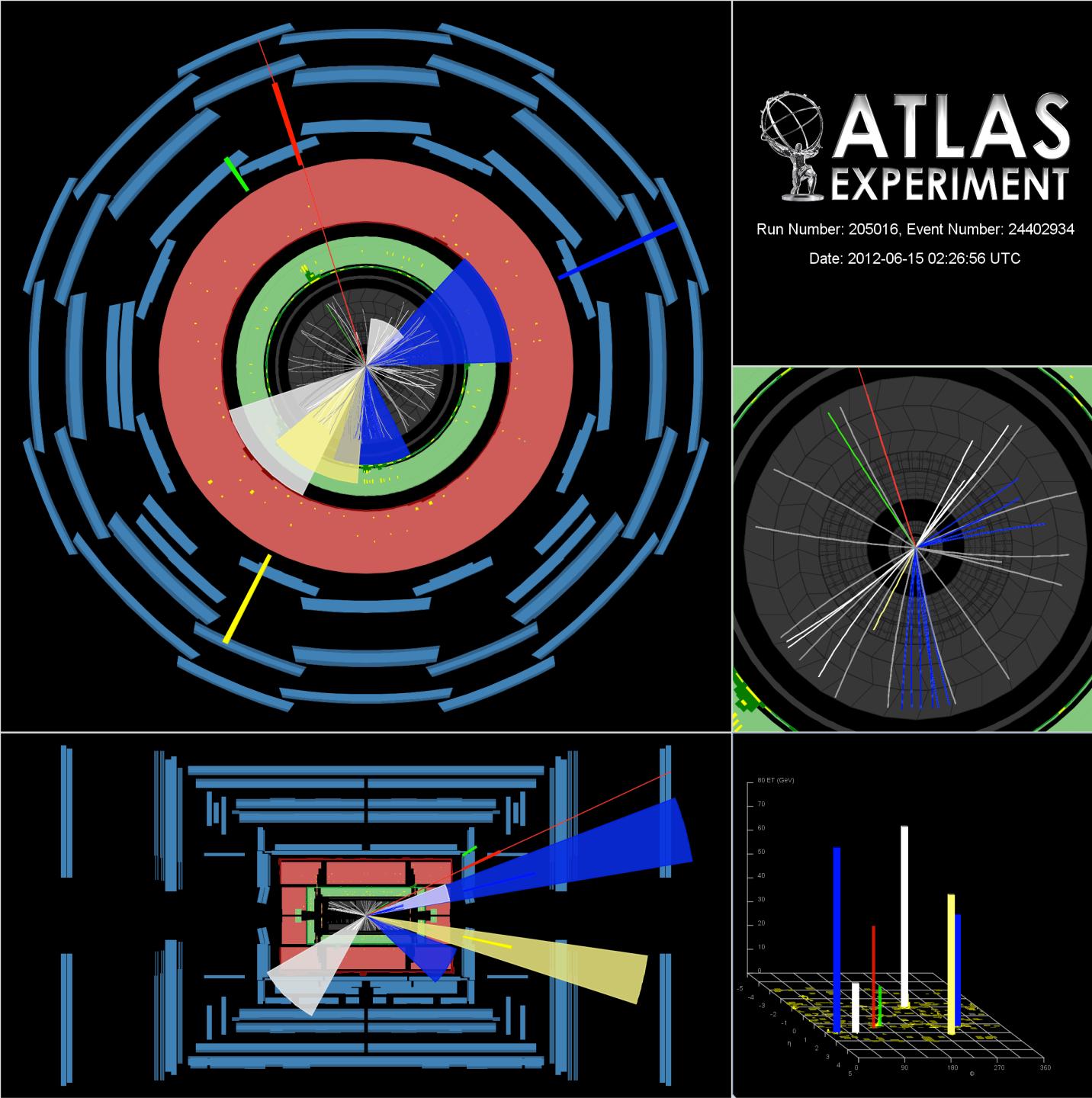
$$\mu(t\bar{t}H) = 2.1 - 1.4 \left( \frac{\sigma(t\bar{t}W)}{232 \text{ fb}} - 1 \right) - 1.3 \left( \frac{\sigma(t\bar{t}Z)}{206 \text{ fb}} - 1 \right)$$

# Summary

- ATLAS/CMS has used multileptonic signatures to search for ttH in run 1:
  - This signature has **good performances** (i.e. competitive w.r.t. 4b signature)
  - Estimate of ttH amplitude was estimated:
    - $1.8\sigma/2.6\sigma$  excess w.r.t background only hypothesis in ATLAS/CMS
    - $1\sigma/2\sigma$  excess w.r.t. Standard Model signal in ATLAS/CMS
    - A limit was set to 4.7 by ATLAS
  - ttH into leptons included in coupling fits
- Contribution from
  - **Strasbourg:** WZ control from the data in  $3\ell$
  - **Clermont:** Charge flip estimate, Fake estimate in  $2\ell$ , 2D fit with ttV in the combination
- Still room for improvements

# Outlook for Run2

- Promising measurement for run 2:
  - Higher cross-sections, higher s/b
  - More luminosity
  - Many systematics with large statistical components will be reduced
  - It will be possible to constrain ttV with data, ttV constraints to be smartly included in the ttH fits (find the right regions, good discriminants, good categorisations)
  - 2 Additional channels being considered in ATLAS for ttH into leptons
  - Intense activity on framework, background estimates, statistical framework, ...
- Many efforts already ongoing or foreseen:
  - From Clermont (Arthur Chomont & al.) on 2l: background estimate, fit framework, ...
  - From Marseille (Kun Liu & al.) on 2l/3l: framework, optimisation, ...
  - From Strasbourg (Kirill Skovpen, Xavier Coubez et al.) on 2l/3l: Matrix Element methods, ...
- See Georges Aad talk for run 2 prospectives



Event display:  $2\ell 1\tau$  candidate

ATLAS CONF-2015-006

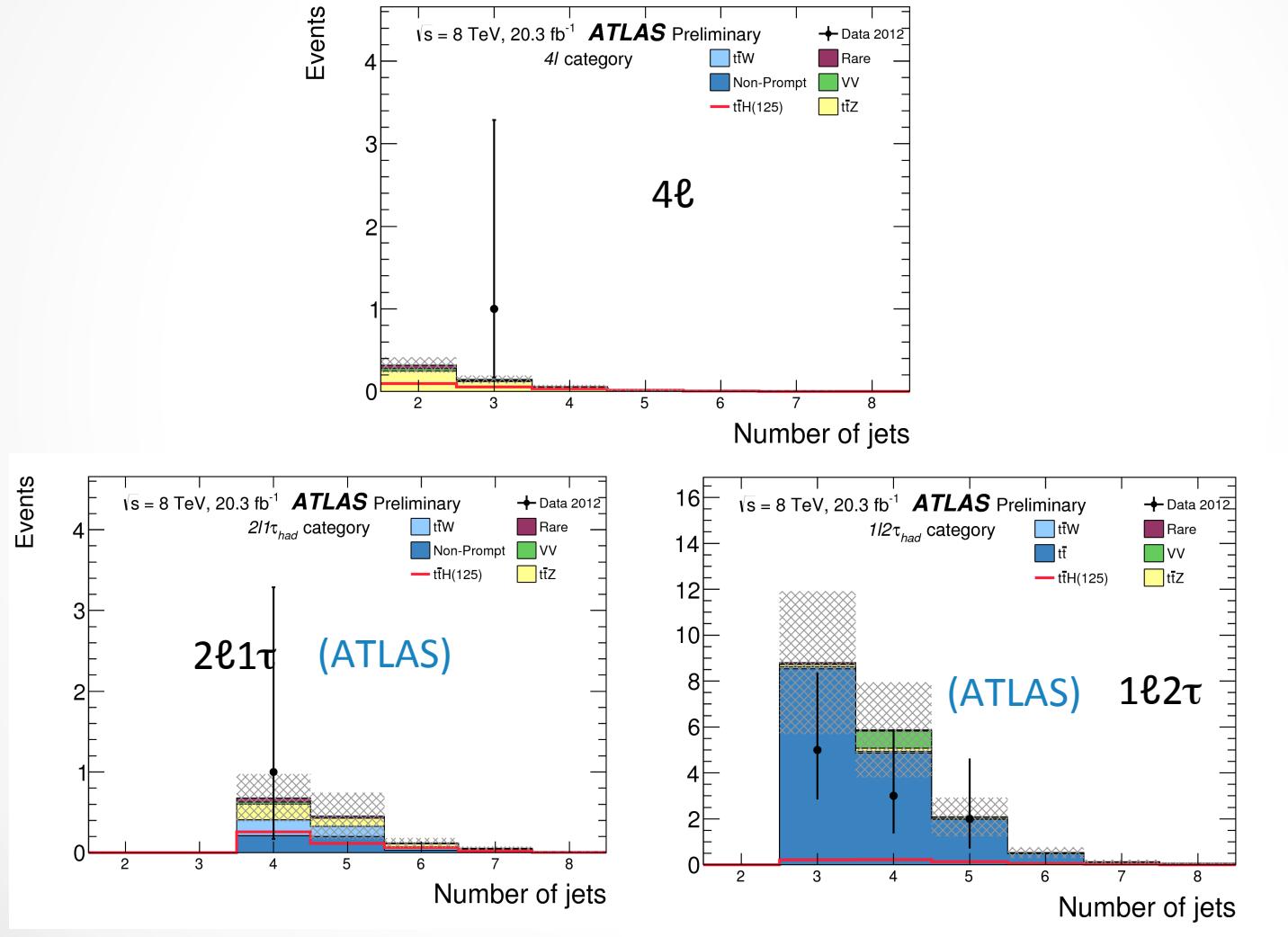
• • •

# Backup material

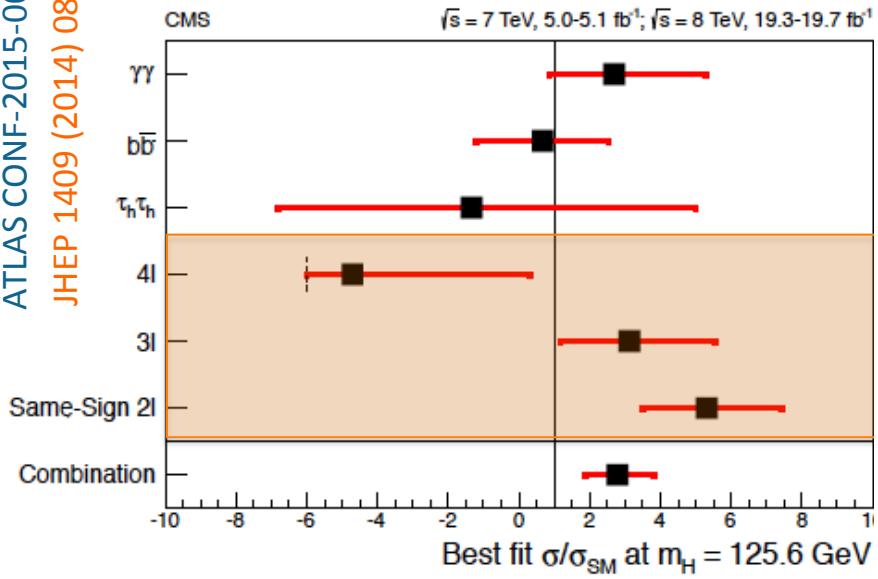
# Uncertainties

- Cross-section: QCD Scale uncertainty:
  - Based on NLO generator
  - ttH: +4-9%, ttW: 12%, ttZ: 11%
- Cross-section: PDF uncertainty:
  - Varying input parton distribution
  - ttH: 8%, ttW: 8%, ttZ: 9%
- Acceptance uncertainty (channel dependant):
  - PS algorithm: Comparison of different generators: ttV 5-23% ttH 1.5-13%
  - PDF impact on acceptance: ttV 1.3-6.7%, 0.3-1.4% ttH
  - QCD scale impact on acceptance: ttV 0.9-4.8%, ttH 0.1-2.7%

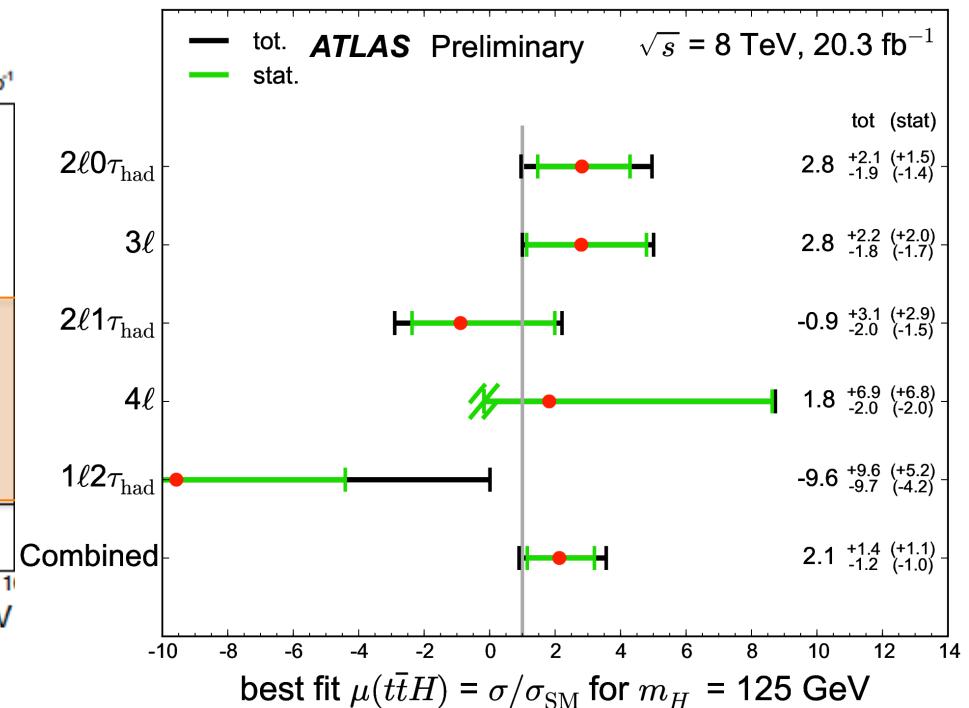
# Results: distributions



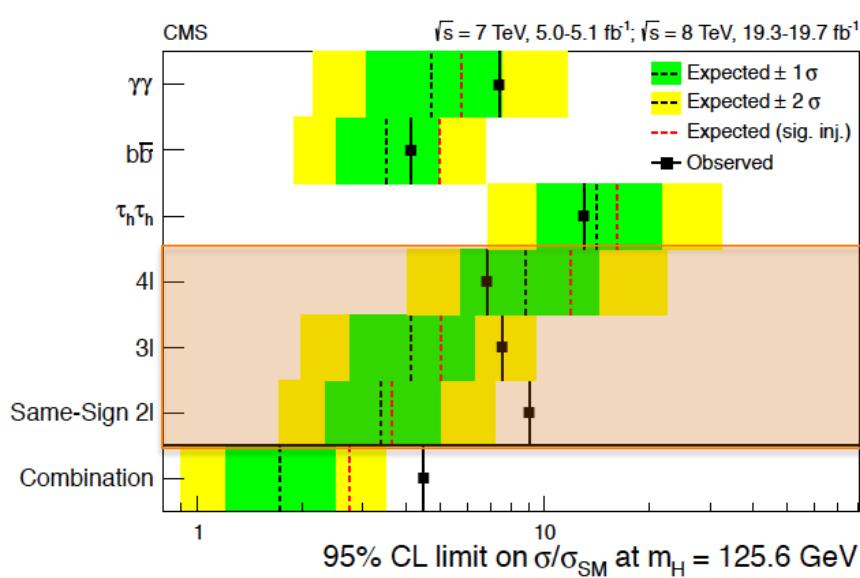
# Fit result: signal strength



$$\mu(t\bar{t}H) = 3.7 + 1.6/-1.4 \text{ (CMS)}$$



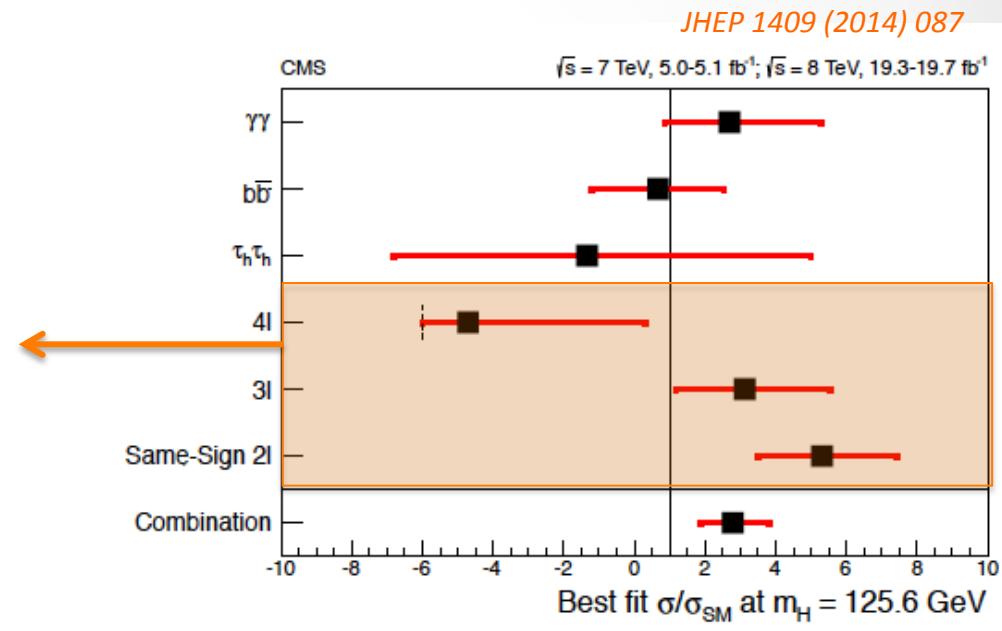
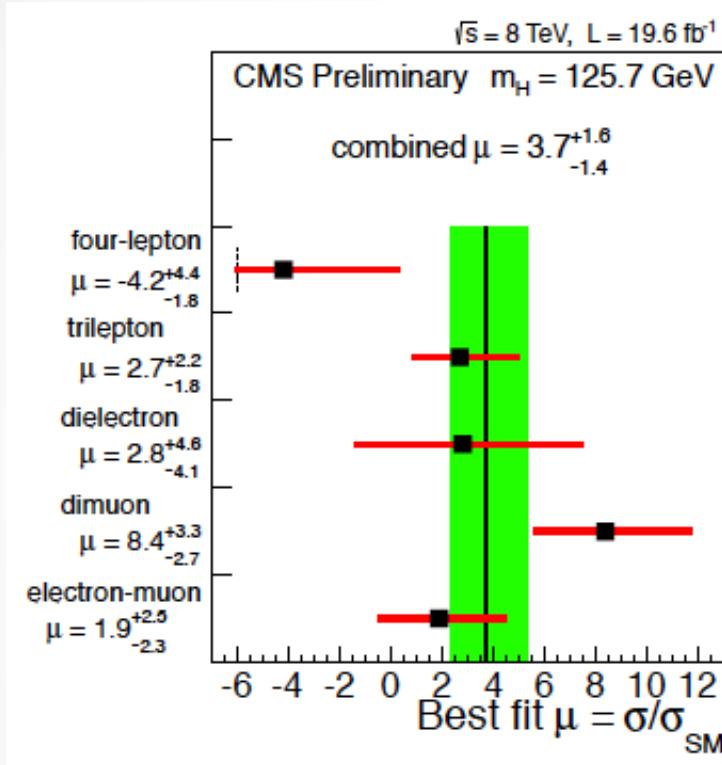
# Fit result, exclusion limit in CMS



$t\bar{t}H$ channel	Best-fit $\mu$		95% CL upper limits on $\mu = \sigma/\sigma_{\text{SM}} (m_H = 125.6 \text{ GeV})$				
	Observed	Observed	Median signal-injected	Median	68% CL range	95% CL range	
4l	$-4.7^{+5.0}_{-1.3}$	6.8	11.9	8.8	[5.7, 14.3]	[4.0, 22.5]	
3l	$+3.1^{+2.4}_{-2.0}$	7.5	5.0	4.1	[2.8, 6.3]	[2.0, 9.5]	
Same-sign 2l	$+5.3^{+2.1}_{-1.8}$	9.0	3.6	3.4	[2.3, 5.0]	[1.7, 7.2]	

Expected Median in presence of SM signal

# Fit result: signal strength



CMS PAS HIG-13-020

JHEP 1409 (2014) 087 is more up to date  
however breakdown in flavor not available.