Search of ttH production mode in multi lepton signatures in ATLAS and CMS

Djamel BOUMEDIENE – LPC

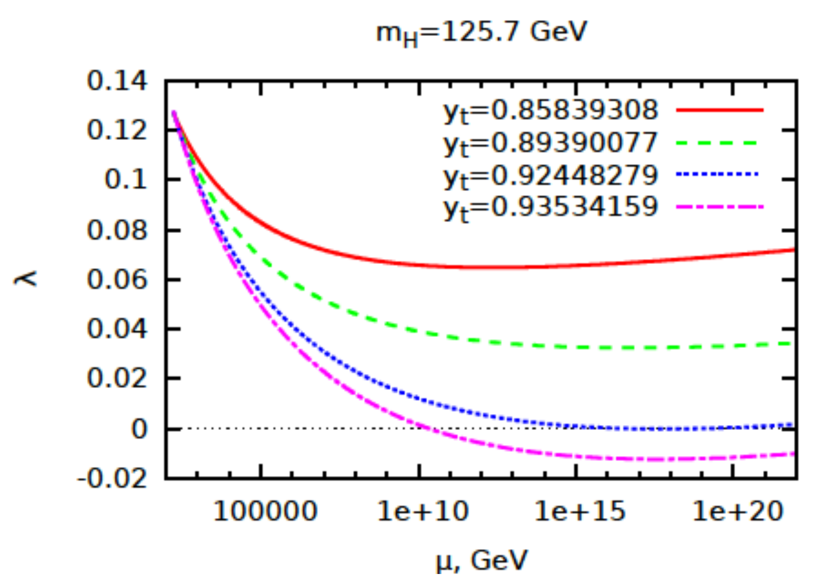
March 30, 2015

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 (λ) sensitive to Top yukawa coupling (y_t)

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



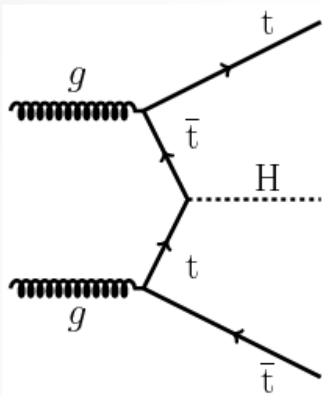
F. Bezrukov, M. Shaposhnikov ArXiv 1411.1923v2

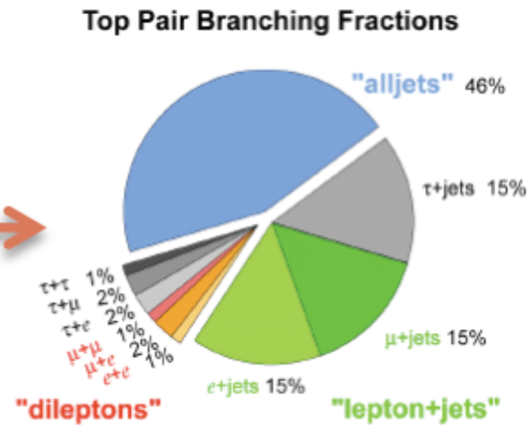
- Existence of a critical y_t above which vacuum is unstable:

  How does it compare to measured y_t ?

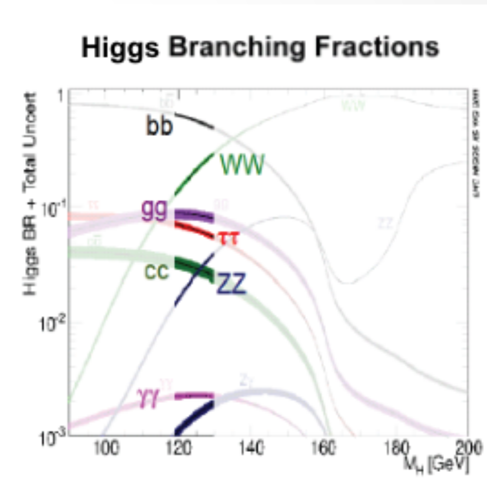
- y_t can be determined:
  - From Top mass measurement (requires interpretation of the MC Top mass)
  - From Higgs production and γγ decay (assuming SM decay)
  - From ttH – a test at the tree level. It will provide evidence of the coupling existence

ttH signature





⊗

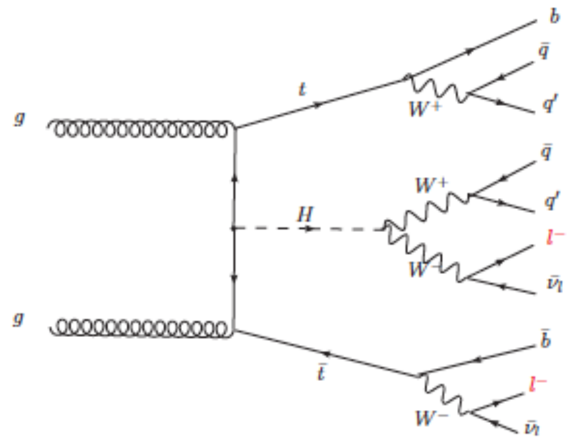


√s=8TeV: σ(ttH) = 130 fb^{-1}
~2700 events / experiment

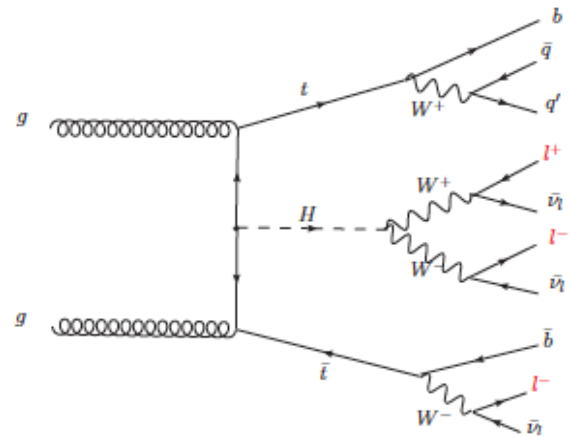
- ttH final state combines top pair decay signature and Higgs decay signature → large number of possible final states
- 3 families of signatures: 4b+X (H→bb), 2b+2γ (H→γγ), 2b+leptons (H→WW, ZZ,ττ)
- Leptonic signatures based on flavour (e,μ,τ) and charge can be used to select a ttH enriched sample

Leptonic channels

- 5 (ATLAS)/3(CMS) channels defined by the number of leptons (e,μ,τ)



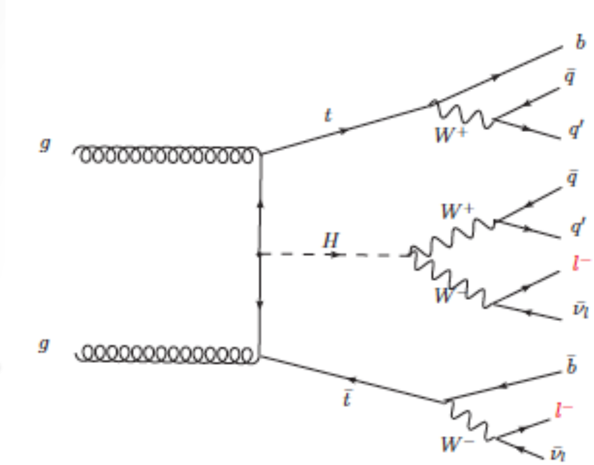
2l (same-sign ee/eμ/μμ)



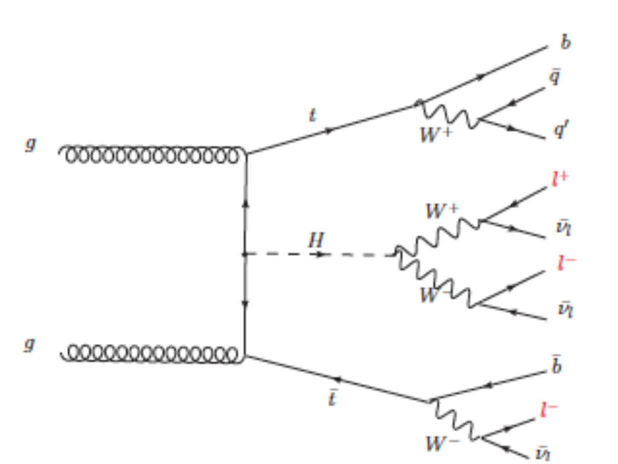
3l (3 e/μ)

Leptonic channels

- 5 (ATLAS)/3(CMS) channels defined by the number of leptons (e,μ,τ)

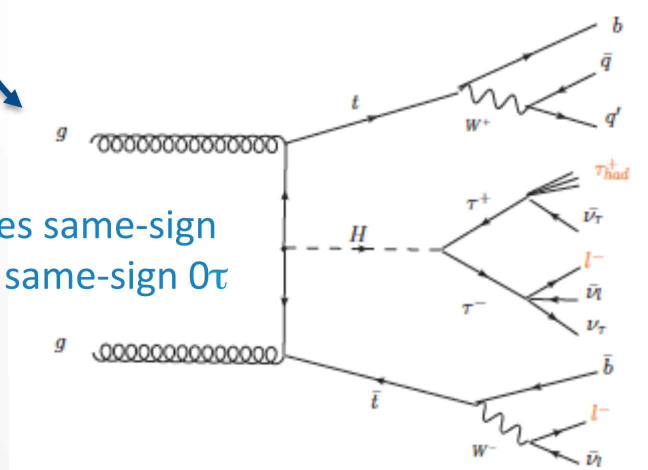


2l (same-sign ee/eμ/μμ)

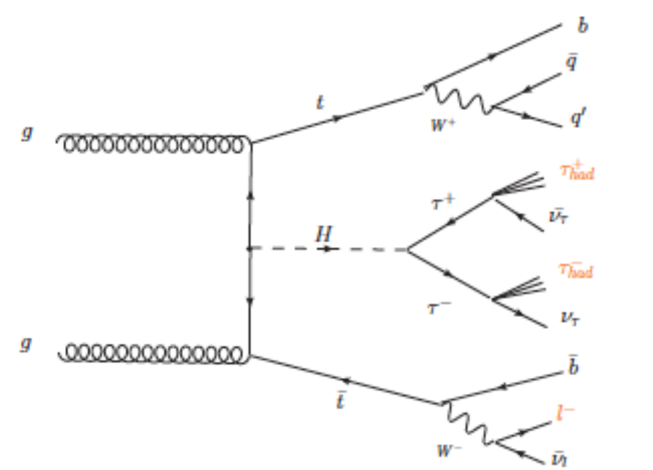


3l (3 e/μ)

ATLAS categorises same-sign
In same-sign+1t same-sign 0τ



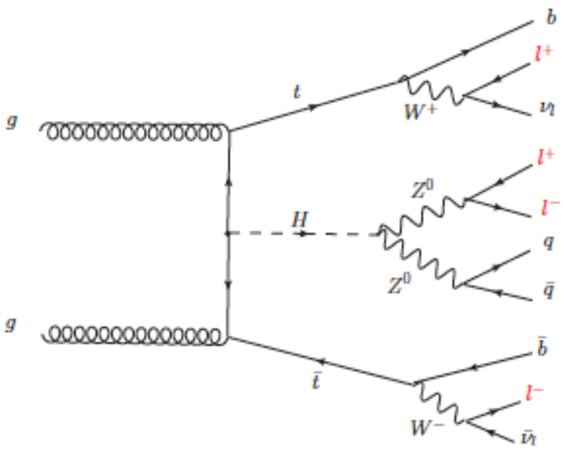
2l1τ (same-sign ee/eμ/μμ+1τ)



2τ1 (1 e/μ)

Leptonic channels

- 5 (ATLAS)/3(CMS) channels defined by the number of leptons (e,μ,τ)



4 leptons

Analysis strategy

ttH into lepton channels are sensitive to several Higgs decays:

Measurement of several Higgs coupling products

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

Fractions of Higgs decays/signature

ATLAS CONF-2015-006

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

B. Explicit coupling fit combining all Higgs inputs (including ttH)
- Cf. ATLAS-CONF-2015-007 / JHEP 1409 (2014) 087 (arXiv: 1412.8662)

Channel and Object definitions

- Selection of e, μ, τ, 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

- ATLAS: use of counting experiment
  - 2l: >=1 b-tagged jet, =4 jets and >= 5 jets x (ee,μμ,eμ)
  - 3l: = 2 b-tagged jets and 3 jets or >= 1 b-tagged jet >=4 jets
  - 2l1τ: >=1 b-tagged jet, >=4 jets
  - 2τ1l: >= 1-btagged jet, >=3 jets
  - 4l: >=1 b-tagged jet, >=2 jets

- CMS: Selection + BDT discriminant
  - All: >=2 b-tagged jet (1 medium + 1 loose), >=2 jets
  - 2l, 3l: 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:

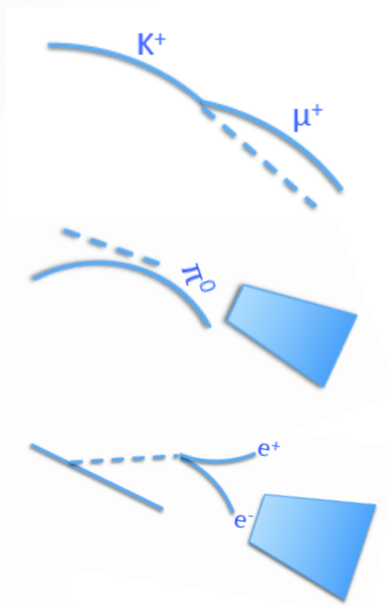
Signature	Top	ttV	VV	V+jets
2l	**30%**	**50%**	10%	10%
3l	15%	**65%**	10%	7%
2l1τ	**43%**	**45%**	10%	< 1%
2τ1l	**90%**	5%	3%	2%
4l	<1%	**78%**	10%	< 1%

approx numbers from MC
Full estimate in slides 14/15

- Main expected background processes:
  - Irreducible: ttV (leads to very similar signature with prompt leptons) → use of MC simulation
  - Reducible: top (non prompt leptons or charge flips) → use of data-driven techniques

Two types of reducible backgrounds

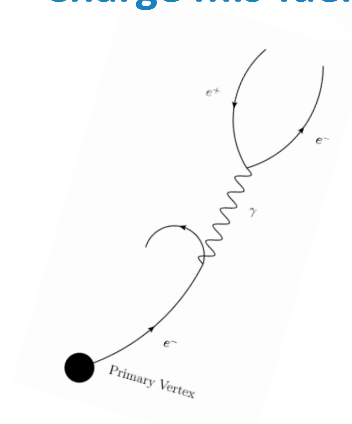
Fake/non-prompt leptons

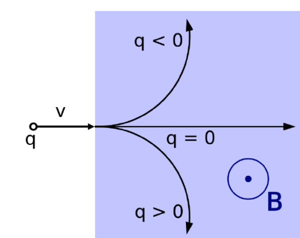


- Non prompt leptons: any lepton not produced by W/Z/τ decay is source of background

- Strong contribution from top pair decays (mainly due to B decays)

Charge mis-Identification





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

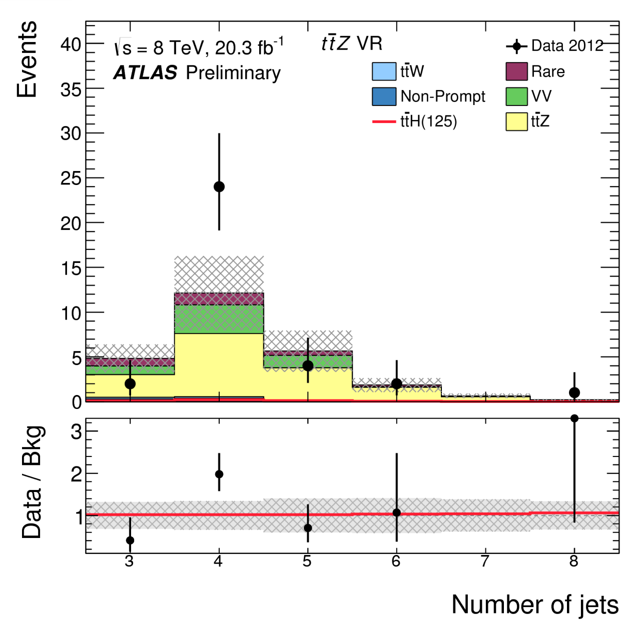
Cannot rely on Monte Carlo simulation for their estimate → use of data-driven methods

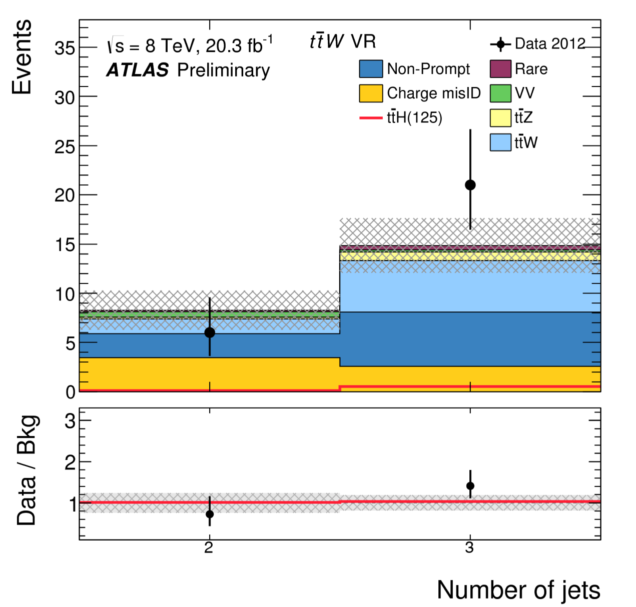
ttV estimate

- In ATLAS & CMS: ttV estimate based on MC simulation:
  - Data does not allow precise constraint, degenerated with SR (especially ttW)
  - Use of NLO cross-sections
  - LO generator Madgraph+Pythia6
  - Final result is also expressed versus ttV cross-section (ATLAS) *See slide 22*

see associated systematic errors slide 29

- Data validation regions:
  - ttZ: 3leptons, Z peak
  - ttW: same-sign, 2b-jets, 2-3 jets





ATLAS CONF-2015-006

Results: Yields (ATLAS)

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

ATLAS CONF-2015-006

- Observed data excess w.r.t. background in 7 regions over 11 regions
- 2l0τ:
  - Overall Data excess. Largest excess in μμ 4j
  - 2 leading backgrounds: ttV, Fakes

Results: Yields (CMS)

- Nominal background predictions (pre-fit) compared to observed number of events

	ee	eμ	μμ	3ℓ	4ℓ
$t\bar{t}H$, H → WW	1.0 ± 0.1	3.2 ± 0.4	2.4 ± 0.3	3.4 ± 0.5	0.29 ± 0.04
$t\bar{t}H$, H → ZZ	—	0.1 ± 0.0	0.1 ± 0.0	0.2 ± 0.0	0.09 ± 0.02
$t\bar{t}H$, H → ττ	0.3 ± 0.0	1.0 ± 0.1	0.7 ± 0.1	1.1 ± 0.2	0.15 ± 0.02
$t\bar{t}W$	4.3 ± 0.6	16.5 ± 2.3	10.4 ± 1.5	10.3 ± 1.9	—
$t\bar{t}Z/\gamma^*$	1.8 ± 0.4	4.9 ± 0.9	2.9 ± 0.5	8.4 ± 1.7	1.12 ± 0.62
$t\bar{t}WW$	0.1 ± 0.0	0.4 ± 0.1	0.3 ± 0.0	0.4 ± 0.1	0.04 ± 0.02
$t\bar{t}\gamma$	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

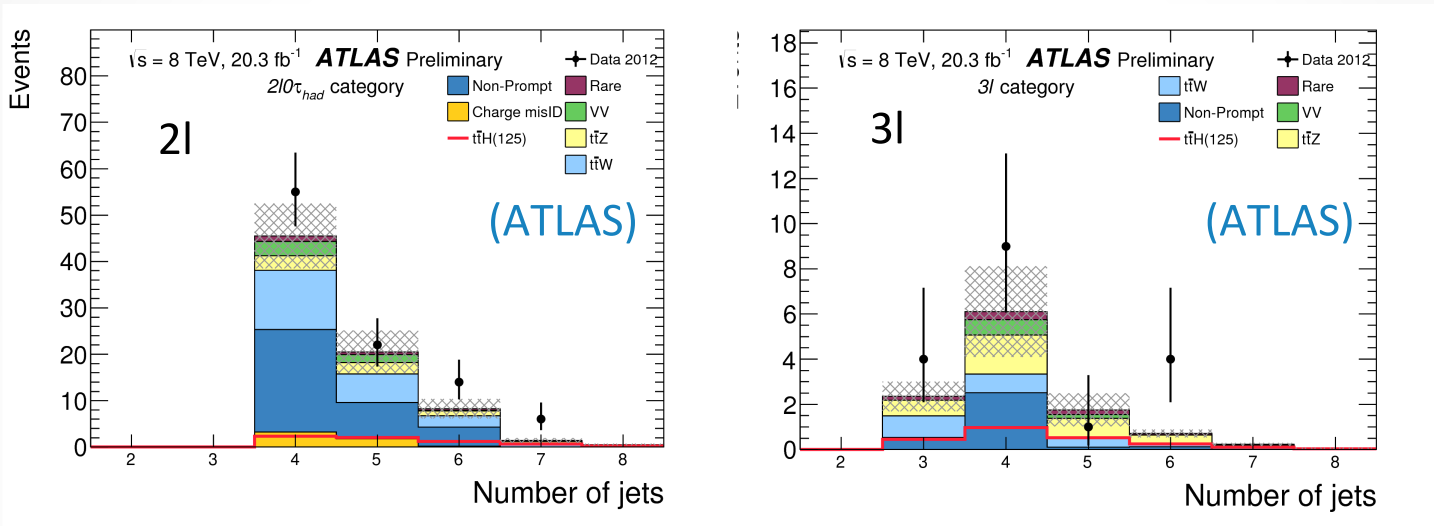
JHEP 1409 (2014) 087

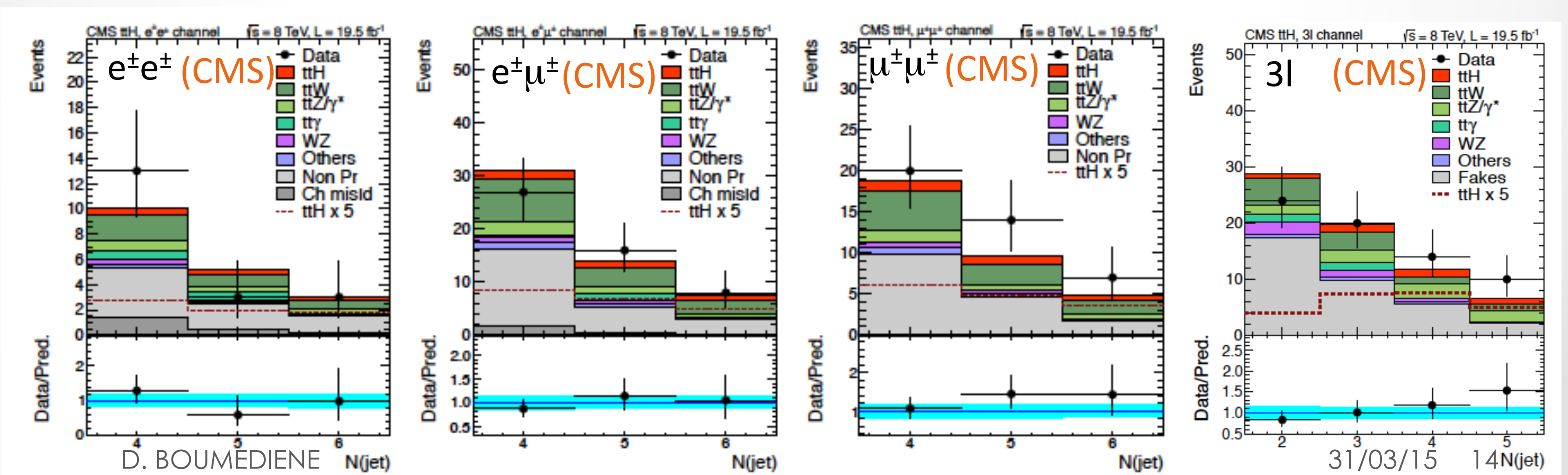
- Observed data excess w.r.t. background in 2 regions over 5

Results: distributions

- Jet multiplicity distributions in 2l/3l

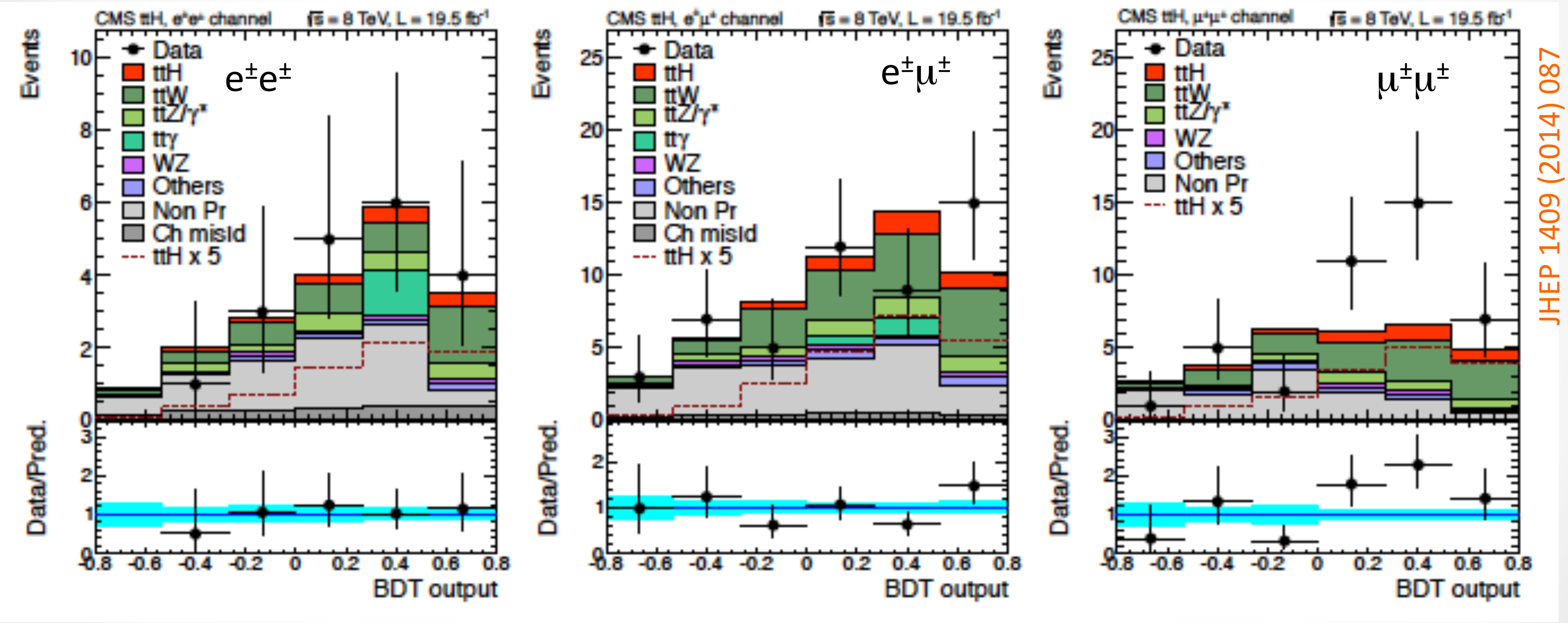
JHEP 1409 (2014) 087
ATLAS CONF-2015-006





BDT discriminant (CMS)

- BDT is applied on selected events in CMS in each channel
- BDT output is used as discriminant to measure signal strength (and put a limit)



Fit procedure

- Maximum likelihood fit of signal strengh
- Floating systematics uncertainties

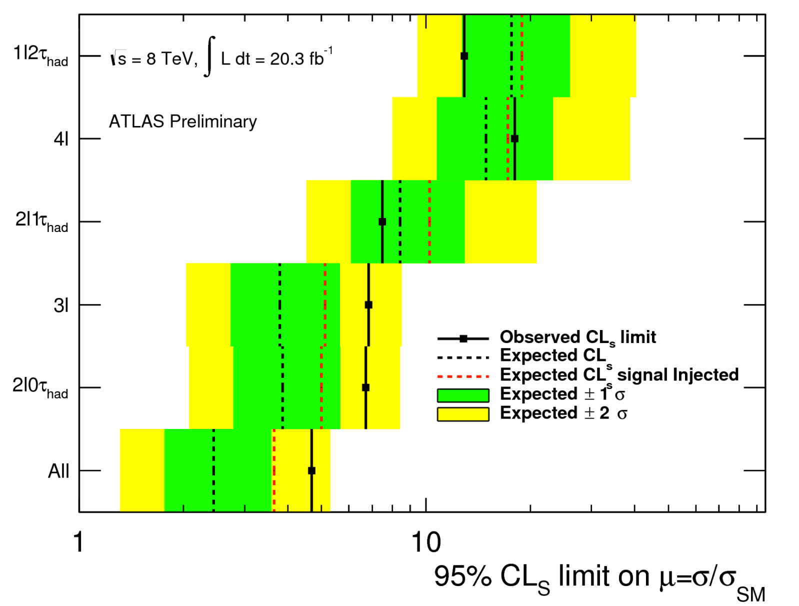
- 5 Nuisance Parameters with largest impact on signal strengh (ATLAS):
  - Uncertainty on Fake μ in 2l
  - ttW acceptance uncertainty
  - ttH cross section uncertainty
  - Jet energy scale uncertainty
  - Uncertainty on Fake e in 2l

Source	$\Delta\mu$	
$2\ell 0\tau_{\rm had}$ 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
$2\ell 0\tau_{\rm had}$ 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 (ATLAS)

95% CL limits on signal strengh (in SM unit) from CLs method: **4.7**

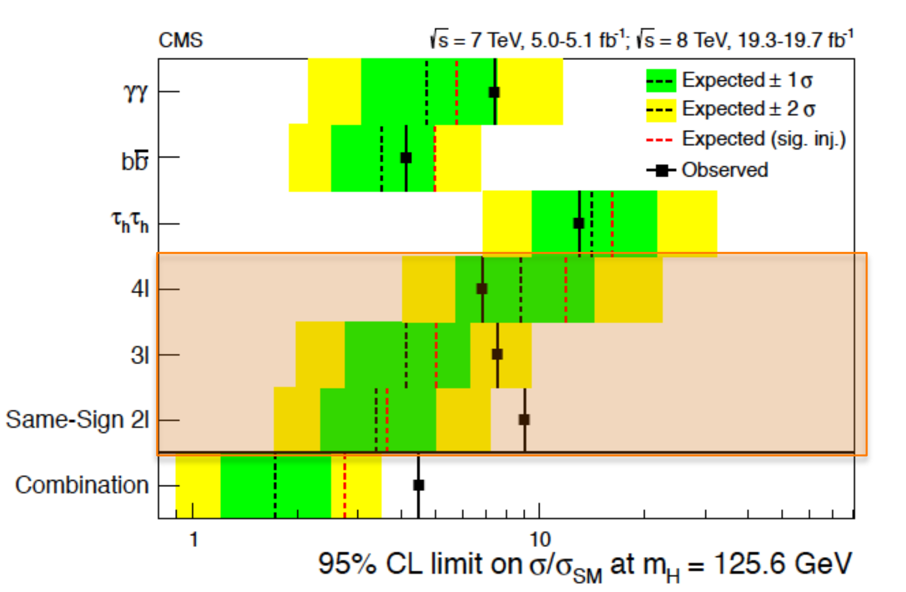


ATLAS CONF-2015-006

Channel	Observed Limit	Expected Limit					
		-2σ	-1σ	Median	$+1\sigma$	$+2\sigma$	Median ($\mu = 1$)
$2\ell 0\tau_{\text{had}}$	6.7	2.1	2.8	3.9	5.7	8.4	5.0
3ℓ	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
4ℓ	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, exclusion limit (CMS)



$t\bar{t}$H channel	Best-fit μ	95% CL upper limits on $\mu = \sigma/\sigma_{SM}$ ($m_H = 125.6$ GeV)				
			Expected			
	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]

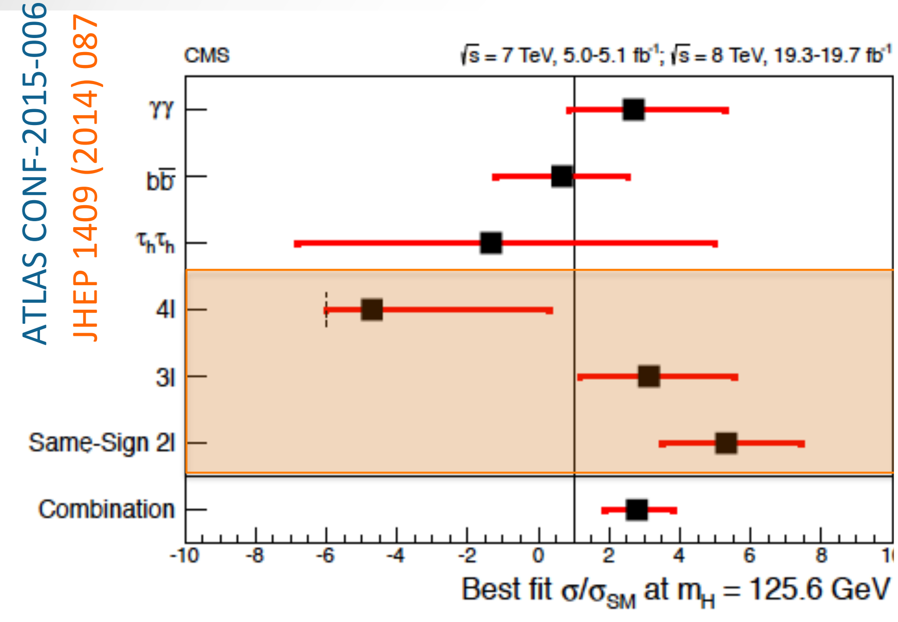
Expected Median in presence of SM signal

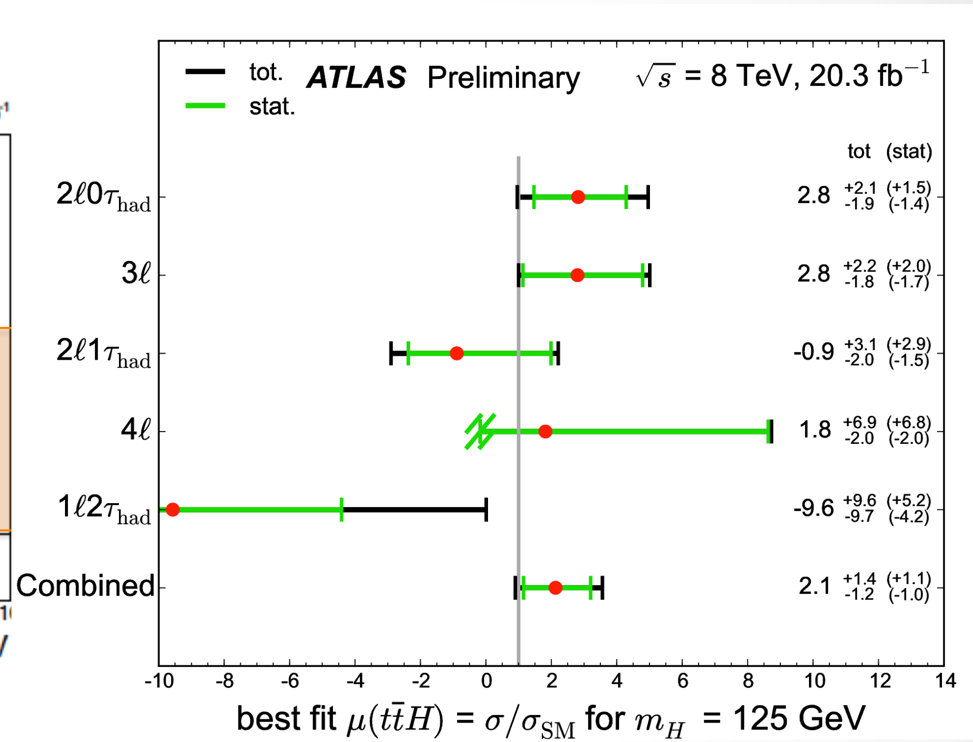
JHEP 1409 (2014) 087

Fit result: signal strength

ATLAS CONF-2015-006

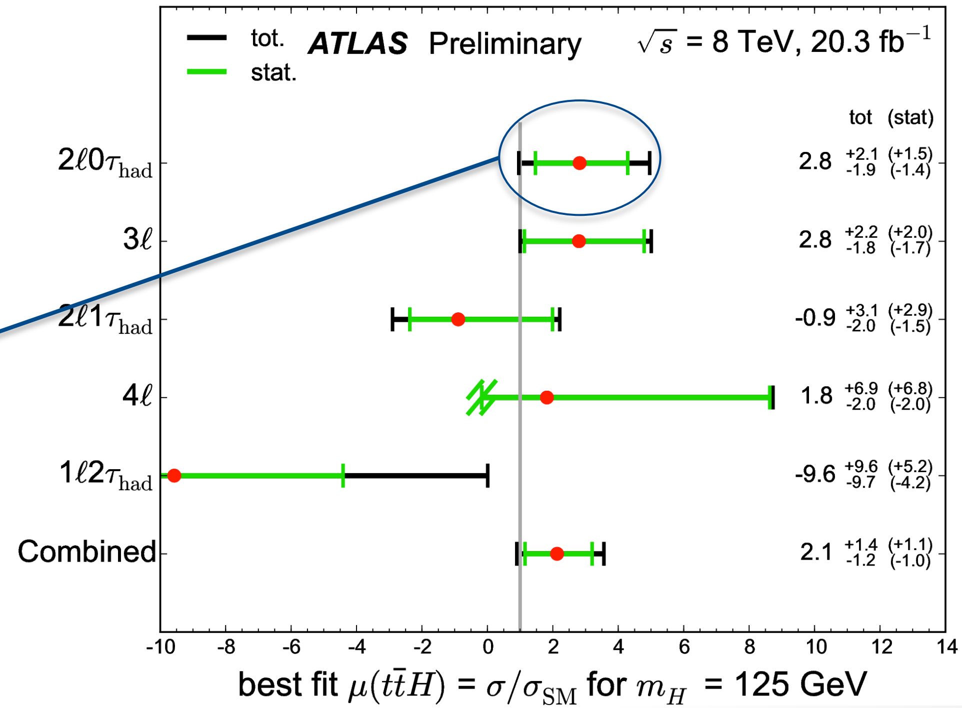
JHEP 1409 (2014) 087



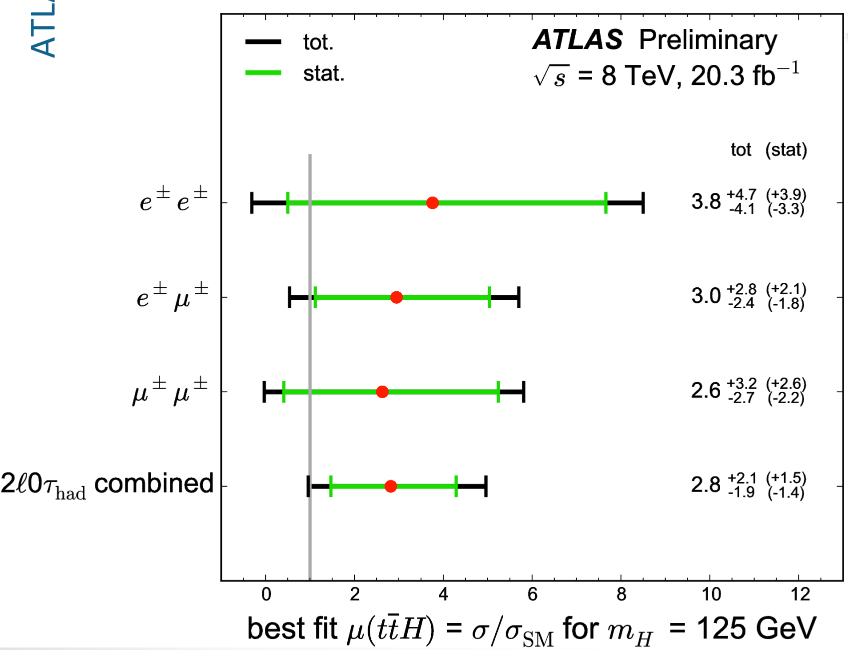


Fit result: signal strength

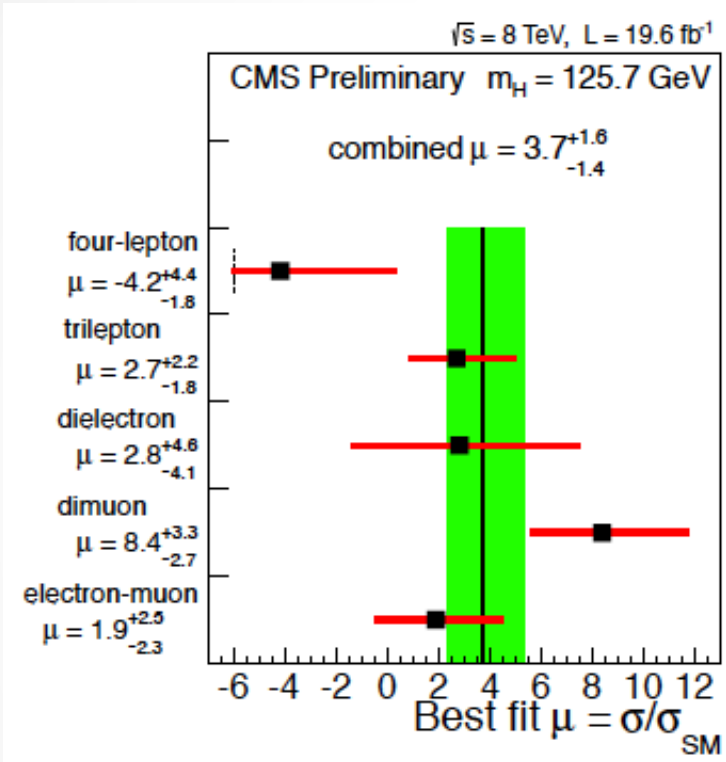
ATLAS CONF-2015-006



2l by lepton flavour

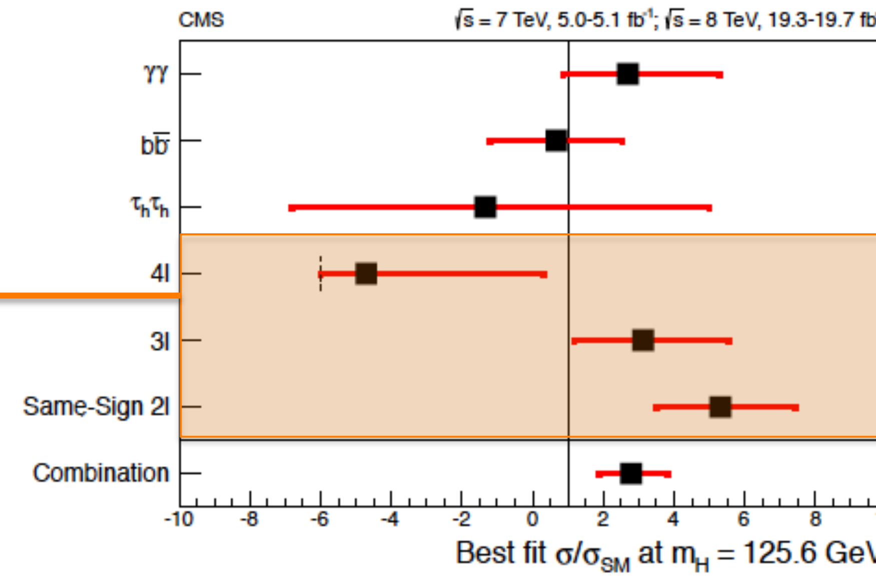


Fit result: signal strength



CMS PAS HIG-13-020

JHEP 1409 (2014) 087

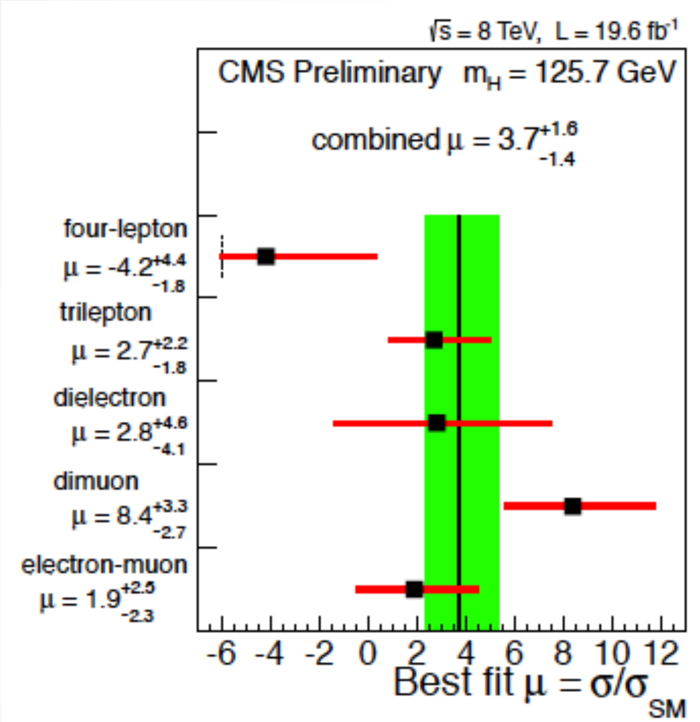


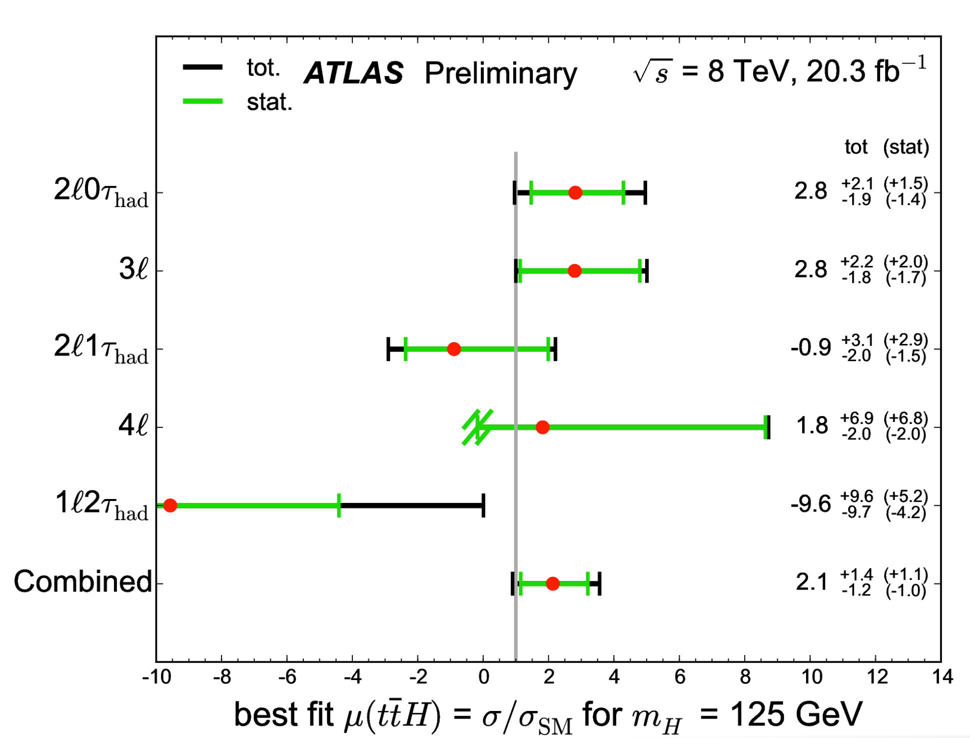
JHEP 1409 (2014) 087 is more up to-date however beakdown in flavor not available.

Fit result: signal strength

ATLAS CONF-2015-006

CMS PAS HIG-13-020





- Best fit μ: ***μ (ttH)*** **= 3.7 + 1.6/-1.4 (CMS)** ***μ (ttH)*** **= 2.1 + 1.4/-1.2 (ATLAS)**
- Excess w.r.t. background only hypothesis: **2.6 σ / 1.8 σ**
- Excess w.r.t. background and SM signal: **1.9 σ / 0.9 σ**
- Dependence of best μ to ttV cross-section: $\mu(t\bar{t}H) = 2.1 - 1.4\left(\frac{\sigma(t\bar{t}W)}{232\text{ fb}} - 1\right) - 1.3\left(\frac{\sigma(t\bar{t}Z)}{206\text{ fb}} - 1\right)$

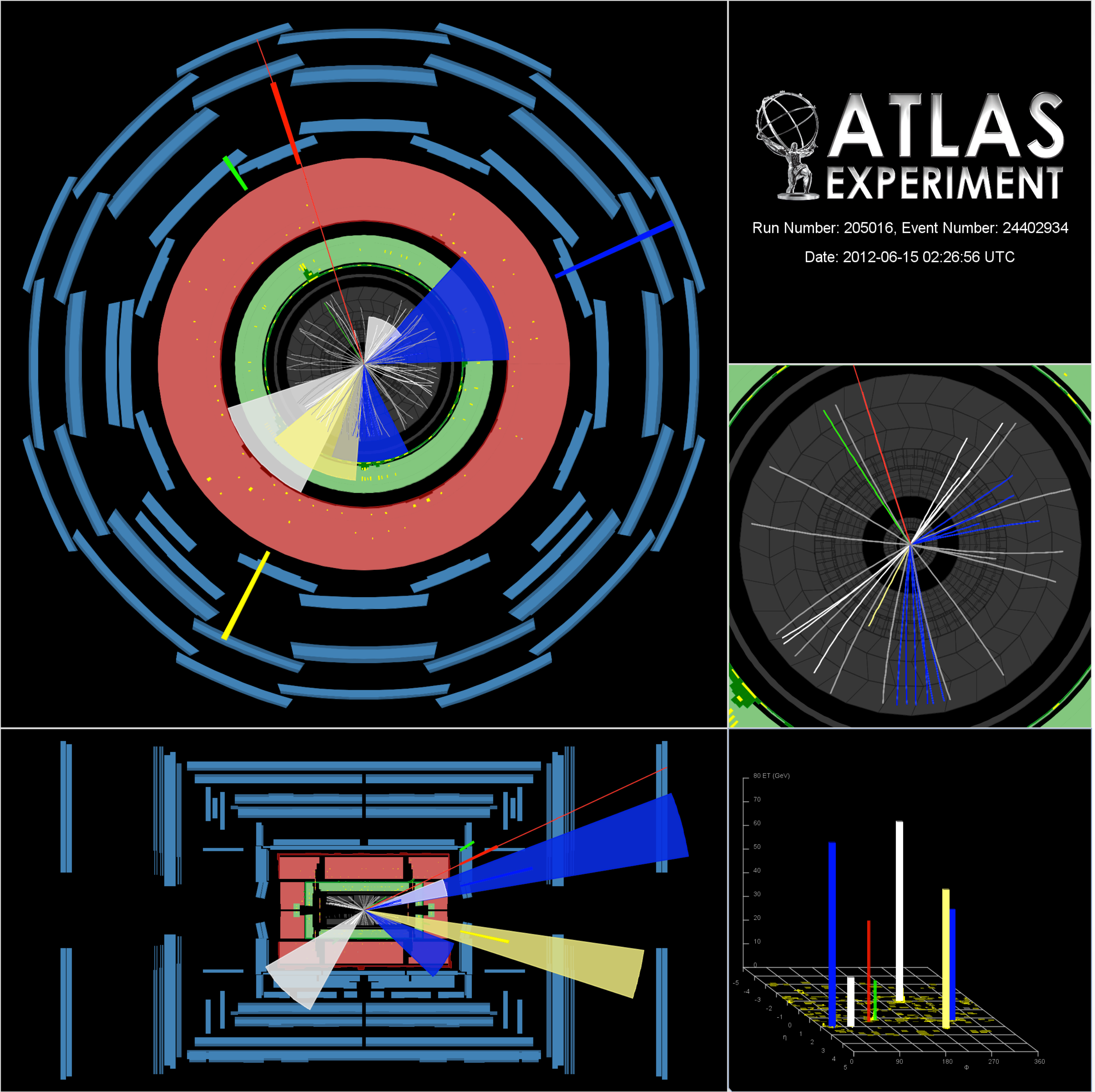
Summary

- ATLAS and 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:
    - 1.8σ/2.6σ excess w.r.t background only hypothesis in ATLAS/CMS
    - 1 to 2σ excess w.r.t. Standard Model signal in ATLAS/CMS
    - A limit was set to 4.7/6.6 in ATLAS/CMS
  - ttH into leptons included in coupling fit

- Promising measurement for run 2:
  - Higher cross-sections
  - More luminosity
  - Many systematics with large statistical components will be reduced
  - It will be possible to constrain ttV with data

Event display: 2l1τ candidate

ATLAS CONF-2015-006



…

Backup material

Charge mis-identification

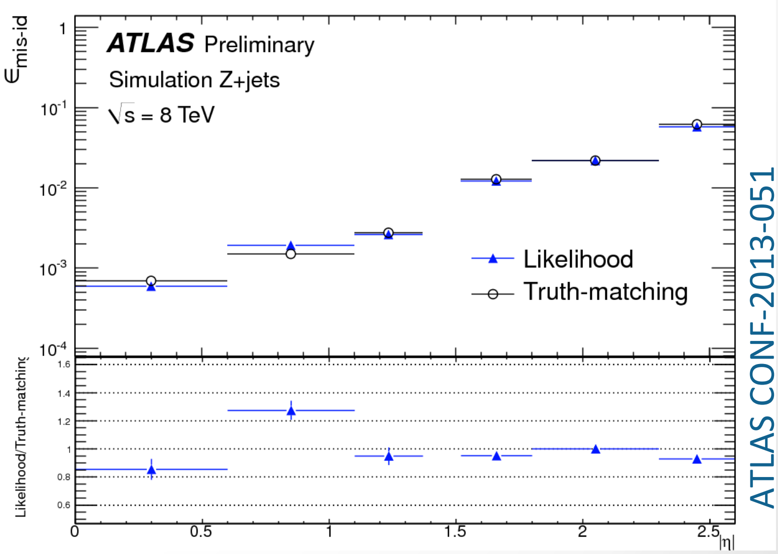
- Same approach is used in CMS and ATLAS

- Electron mis-identified charge:
  - Due to trident events or large curvature → depends on lepton properties and interaction with detector (pT, $|\eta|$)
  - Negligible for muons

- Affects 2l ee/eμ 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

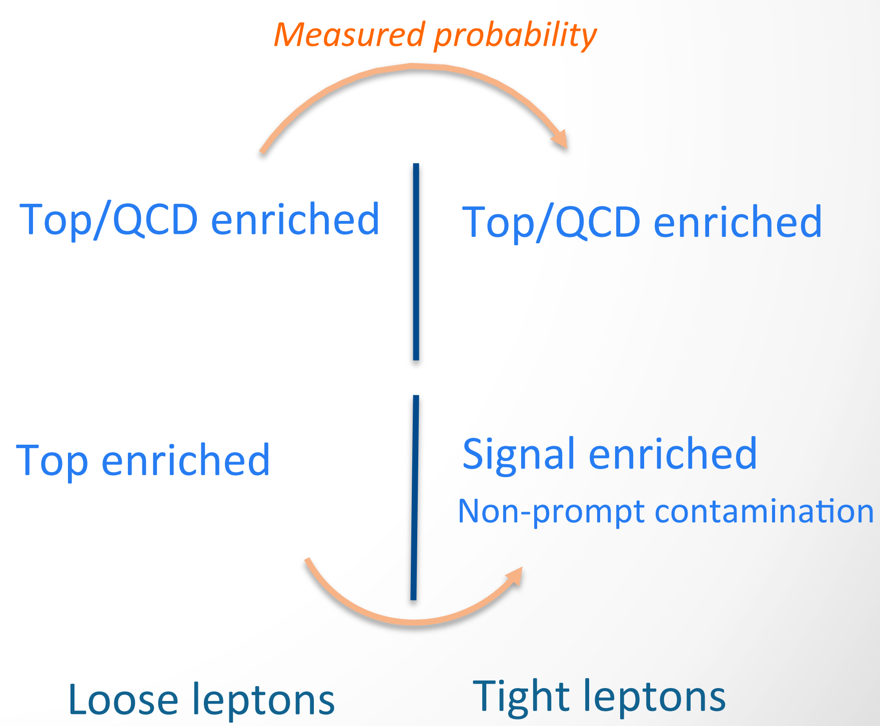
- Systematic uncertainty:
  - Includes likelihood statistic, Z peak definition, extrapolation, closure test, …
  - Total systematic ~40% (mainly due to statistics)



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:
    - ATLAS: Top enriched region, low jet multiplicity
    - CMS: QCD enriched region

- Main uncertainties:
  - Size of data control region
  - Closure tests on MC simulation



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

JHEP 1409 (2014) 087

ATLAS CONF-2015-006