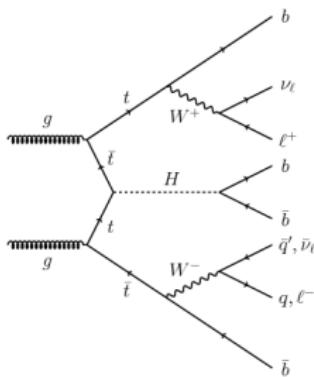


Run 1 results $t\bar{t}H(H \rightarrow b\bar{b})$

Daniele Madaffari

May 19, 2015

top LHC-France 2015



Discussing latest results from ATLAS (arXiv:1503.05066) and CMS (arXiv:1502.02485)



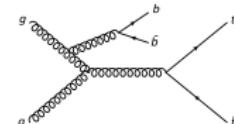
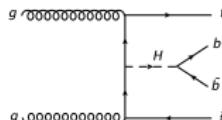
$t\bar{t}H$ at LHC, an introduction

► $t\bar{t}H (H \rightarrow b\bar{b})$

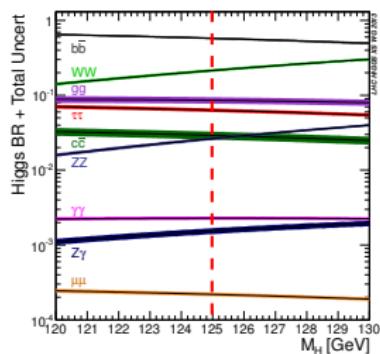
- $\sigma(t\bar{t}H) = 130 \text{ fb}$
 - ◊ NLO in QCD, $m_H = 125 \text{ GeV}$, $\sqrt{s} = 8 \text{ TeV}$
 - ◊ Uncertainties: $+3.8 / -9.3 \text{ (scale)} \pm 8.1 \text{ (PDF)}$
- Simulated with:
 - ◊ ATLAS: Helac-Oneloop at NLO + Powheg-Box
 - ◊ CMS: Pythia at LO

► Main physics background: $t\bar{t} + \text{jets}$

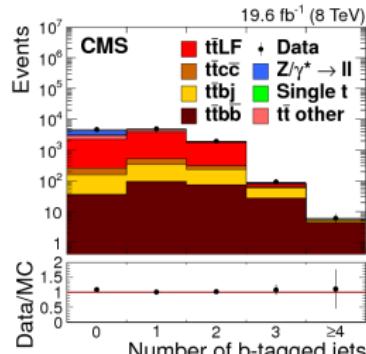
- Theoretical predictions at $\sqrt{s} = 8 \text{ TeV}$:
 - ◊ $\sigma(t\bar{t} + \text{jets}) = 245.8^{+6.2}_{-8.4} \text{ (scales)}^{+6.2}_{-6.4} \text{ (PDF) pb}$
- Simulated with:
 - ◊ ATLAS: Powheg-box NLO + Pythia
 - ◊ CMS: MadGraph + Pythia
- Measurement at $\sqrt{s} = 8 \text{ TeV}$ by CMS
 - ◊ $\frac{\sigma(t\bar{t}+b\bar{b})}{\sigma(t\bar{t}+\text{jets})} = 0.022 \pm 0.003 \text{ (stat)} \pm 0.005 \text{ (syst)}$



$H \rightarrow b\bar{b}$: branching ratio $\sim 56\%$



arXiv:1411.5621



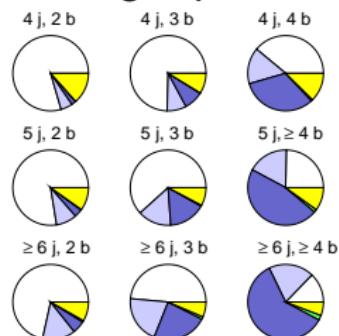
$t\bar{t}H$ ($H \rightarrow b\bar{b}$) event selection

- ▶ Trigger:
 - ATLAS : e isolated $p_T > 24$ GeV, non isolated $p_T > 60$ GeV;
 μ isolated $p_T > 24$ GeV, non isolated $p_T > 36$ GeV
 - CMS : e isolated $p_T > 27$ GeV; μ isolated $p_T > 24$ GeV;
Two e isolated $p_T > 17$ GeV and $p_T > 8$ GeV
- ▶ Muons requirements:
 - ATLAS: $p_T > 25$ GeV [SL], $p_T > 25$ GeV and $p_T > 15$ GeV [DL], $|\eta| < 2.5$
 - CMS: $p_T > 30$ GeV [SL], $p_T > 30$ GeV and $p_T > 20$ GeV [DL], $|\eta| < 2.1$
- ▶ Electron requirements:
 - ATLAS: $p_T > 25$ GeV [SL], $p_T > 25$ GeV and $p_T > 15$ GeV [DL], $|\eta| < 2.5$
 - CMS: $p_T > 30$ GeV [SL], $p_T > 30$ GeV and $p_T > 20$ GeV [DL], $|\eta| < 2.5$
- ▶ Jets:
 - ATLAS: anti- k_t 0.4, $|\eta| < 2.5$, $p_T > 25$ GeV
 - ◊ b -tagging 70% b -jets, 20% c -jets, 1% light-jets
 - CMS: anti- k_t 0.5, $|\eta| < 2.5$, $p_T > 30$ GeV
 - ◊ b -tagging 70% b -jets, 20% c -jets, 2% light-jets
- ▶ Additional selections:
 - ATLAS and CMS: $m_{ll} > 15$ GeV, $|m_{ll} - m_Z| > 8$ GeV

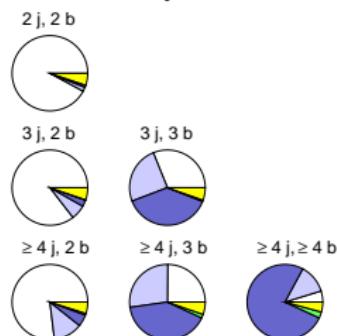
Analysis strategy - ATLAS

- ▶ Signal and control regions are defined according to jet and b -tagged jet multiplicity
 - Analysed separately and combined statistically to maximise the overall sensitivity
- ▶ Different background composition in each region allows to constrain systematics
 - 2 b -jet regions provide constraints on $t\bar{t}$ + light-jets modelling systematics
 - The region with exactly 4 jets and exactly 3 b -tagged jet in SL channel gives constraints on c -tagging due to $W \rightarrow cs$ decays from $t\bar{t}$ + light-jets

Single-lepton



Dilepton



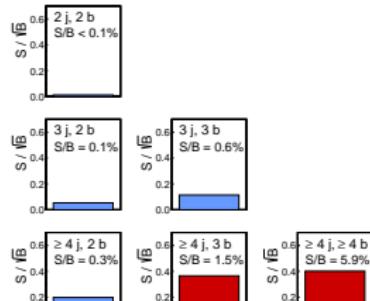
ATLAS Simulation

$m_H = 125 \text{ GeV}$
 $\sqrt{s} = 8 \text{ TeV}$

- $t\bar{t}+\text{light}$
- $t\bar{t}+c\bar{c}$
- $t\bar{t}+b\bar{b}$
- $t\bar{t}+V$
- non- $t\bar{t}$

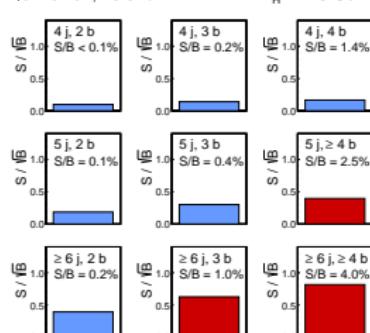
ATLAS Simulation

$\sqrt{s} = 8 \text{ TeV}, 20.3 \text{ fb}^{-1}$



ATLAS Simulation

$\sqrt{s} = 8 \text{ TeV}, 20.3 \text{ fb}^{-1}$



Analysis strategy - CMS

- ▶ Regions are defined according to number of jets, leptons, **Likelihood ratio \mathcal{F}**

Single-lepton :

- Category 1 : ≥ 6 jets, mass of untagged jets required to be compatible with W mass
 - ◊ $W \rightarrow q\bar{q}$ decay is fully reconstructed
- Category 2 : ≥ 6 jets, **Inversion** of mass constraint wrt Category 1
 - ◊ One of the quark form W decay has failed reconstruction
- Category 3 : == 5 jets
 - ◊ One of the quark form W decay has failed reconstruction

Dilepton:

- One category: at least four jets

Likelihood ratio \mathcal{F} , definition:

- $\vec{\xi}$ = Values of flavour discriminator calculated for all the jets in the event
 - ◊ In **single lepton** and **dilepton** channel only the **six** and **four** jets with the highest CSV are taken into account
- Likelihood f to observe $\vec{\xi}$ is evaluated for two hypotheses
 - ◊ $f(\vec{\xi} | t\bar{t} + hf)$: $t\bar{t}$ + heavy-flavours
 - ◊ $f(\vec{\xi} | t\bar{t} + lf)$: $t\bar{t}$ + light-flavours

- Likelihood ratio is defined as:

$$\mathcal{F}(\vec{\xi}) = \frac{f(\vec{\xi} | t\bar{t} + hf)}{f(\vec{\xi} | t\bar{t} + hf) + f(\vec{\xi} | t\bar{t} + lf)}$$

- ▶ Events with high \mathcal{F} are classified as high-purity (H), otherwise low-purity (L)

Yield comparison

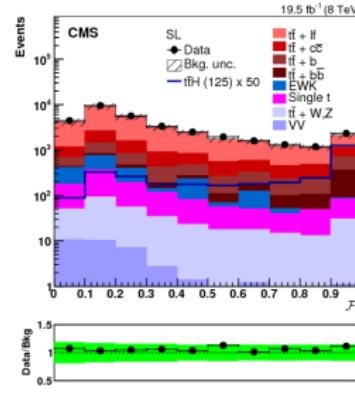
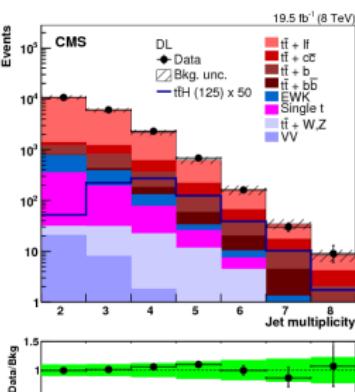
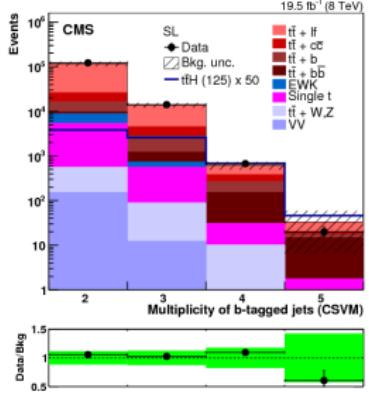
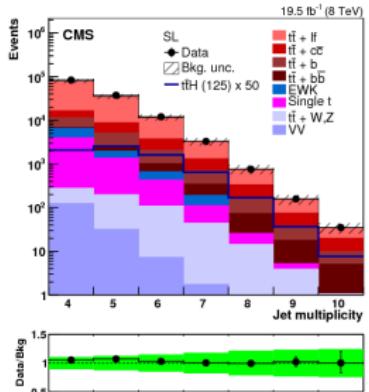
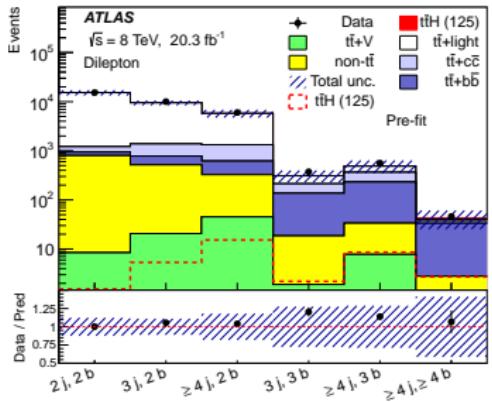
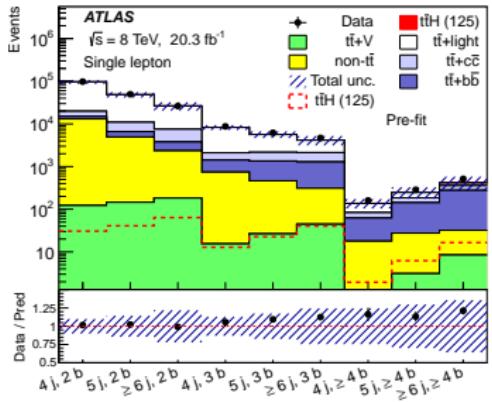
CMS

Post-fit	SL			
	Cat-1 (H)	Cat-2 (H)	Cat-3 (H)	DL (H)
$t\bar{t}H$	3.5 ± 0.7	5.9 ± 1.1	7.5 ± 1.4	4.5 ± 0.7
$t\bar{t} + l\bar{l}$	17 ± 3	70 ± 13	152 ± 21	84 ± 11
$t\bar{t} + c\bar{c}$	22 ± 8	66 ± 20	81 ± 24	85 ± 24
$t\bar{t} + b\bar{b}$	16 ± 8	44 ± 23	70 ± 32	47 ± 23
$t\bar{t} + b\bar{b}$	43 ± 11	75 ± 17	69 ± 18	50 ± 13
Other bkgns	3.3 ± 1.9	10 ± 3	18 ± 4	12 ± 5
Tot bkg	103 ± 11	265 ± 24	390 ± 28	283 ± 24
Data	107	272	401	279
S/B	3.4%	2.2%	1.9%	1.6%
S/\sqrt{B}	0.34	0.36	0.38	0.27

ATLAS

Pre-fit	SL				DL	
	$\geq 6 j, 3 b$	$\geq 6 j, \geq 4b$	$\geq 4 j, 3 b$	$\geq 4 j, \geq 4b$		
$t\bar{t}H$	40 ± 3	16 ± 2	8.6 ± 0.6	2.7 ± 0.3		
$t\bar{t} + l\bar{l}$	2000 ± 460	52 ± 17	120 ± 31	1.9 ± 0.8		
$t\bar{t} + c\bar{c}$	850 ± 480	79 ± 46	130 ± 74	5.0 ± 3		
$t\bar{t} + b\bar{b}$	970 ± 530	250 ± 130	200 ± 100	31 ± 17		
Other bkgns	308 ± 77	32 ± 8	34 ± 8	3 ± 0.5		
Tot bkg	4200 ± 1000	430 ± 150	490 ± 140	43 ± 18		
Data	4701	516	561	46		
S/B	1%	4%	1.5%	5.9%		
S/\sqrt{B}	0.63	0.815	0.365	0.401		

Pre-fit yields



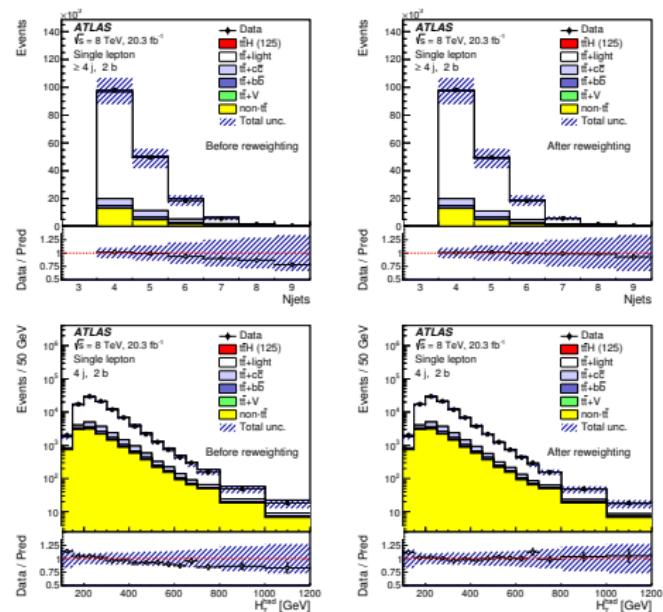
$t\bar{t} + \text{jets}$ & $t\bar{t} + c\bar{c}$ modelling - ATLAS

► Reweighting applied to Powheg-Pythia simulations

- Extracted from differential $t\bar{t}$ cross section measurement at $\sqrt{s} = 7 \text{ TeV}$
Phys.Rev. D90 (2014) 072004

► $t\bar{t}$ system p_T

- Affects mainly the number of jets

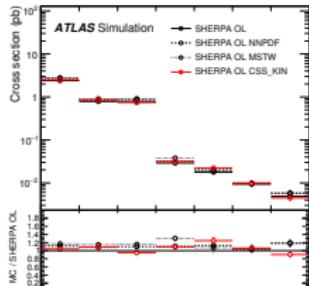
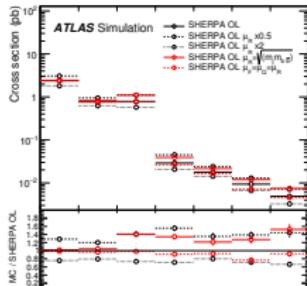
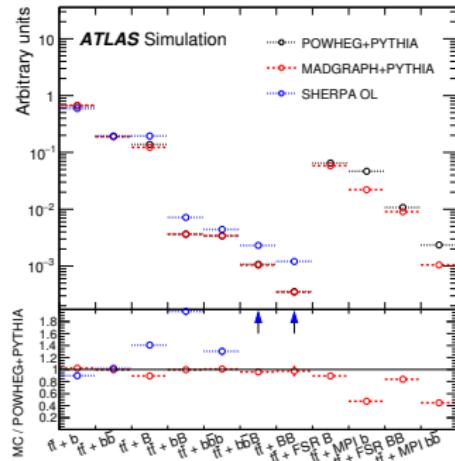


► top quark p_T

- Affects mainly jet p_T variables, like H_T^{had}

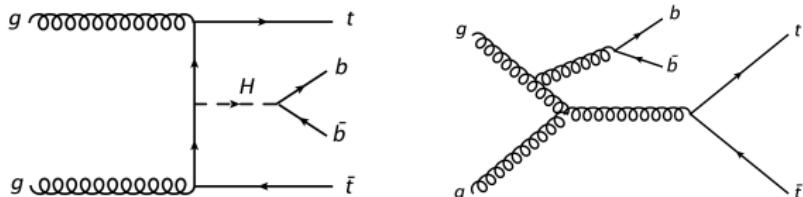
$t\bar{t} + b\bar{b}$ modelling - ATLAS

- After a finer categorisation, different Monte Carlo generators are compared to understand the impact on $t\bar{t}$ + HF simulation
- Powheg+Pythia** generator has been compared to **Madgraph LO + Pythia**
 - Overall good agreement
- Comparing with **Sherpa + Openloops** NLO with massive b -quarks:
 - SherpaOL** predicts higher contribution when production of a second $b\bar{b}$ pair is required
 - Differences are observed in low region of mass and p_T of $b\bar{b}$ pair and in the p_T of top quark and $t\bar{t}$ system
- Systematics take into account:
 - Renormalization scale
 - Resummation scale
 - Factorization scale
 - PDF and shower recoil scheme



Matrix Element Method

- Using a likelihood technique based on theoretical matrix elements for $t\bar{t}H$ process and $t\bar{t} + b\bar{b}$ background

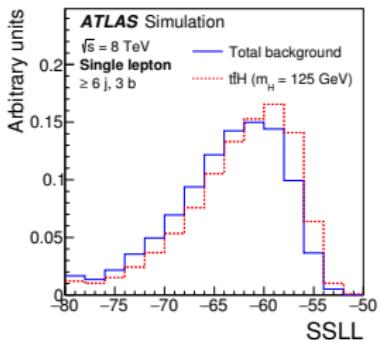
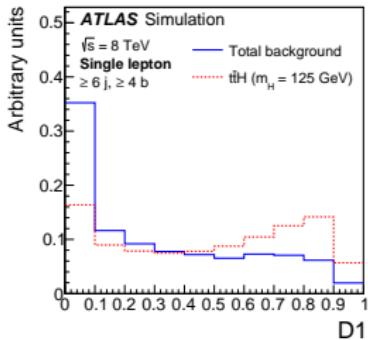


- Method links theoretical calculation to observed quantities
 - Assign to each event a probability of being consistent with background and signal hypothesis respectively

- Variable used:
 - ATLAS and CMS: Likelihood ratio

$$D1, P_{s/b} = \frac{\mathcal{L}_{t\bar{t}H}}{\mathcal{L}_{t\bar{t}H} + \alpha \mathcal{L}_{t\bar{t}+b\bar{b}}}$$

- ATLAS: Logarithm of summed signal likelihood:
 $SSLL = \mathcal{L}_{t\bar{t}H}$



Discriminating variables - CMS

- ▶ Two dimensional maximum likelihood fit, looking at joint distribution of:
 - $P_{h/I}$: build using b -tagging information
 - ◊ Discriminate heavy- and light-flavour components of $t\bar{t} + \text{jets}$

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

- $P_{s/b}$: MEM probability density function ratio
 - ◊ Discriminate $t\bar{t}H$ ($H \rightarrow b\bar{b}$) against $t\bar{t} + b\bar{b}$

$$P_{s/b} = \frac{\mathcal{L}_{t\bar{t}H}}{\mathcal{L}_{t\bar{t}H} + k_{s/b} \mathcal{L}_{t\bar{t}+b\bar{b}}}$$

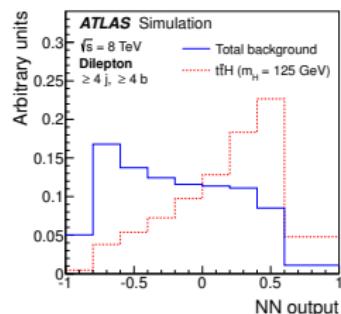
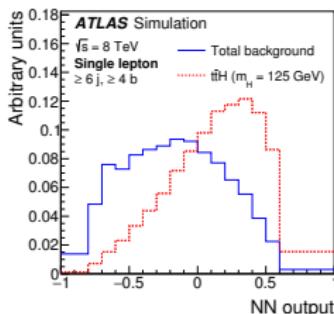
- ▶ Coefficients $k_{s/b}$ and $k_{h/I}$ can differ among the categories

Discriminating variables - ATLAS

- ▶ Neural network (NN) is trained in each of the most signal rich regions to discriminate $t\bar{t}H$ against the other backgrounds
- ▶ In less sensitive regions:
 - H_T^{had} is used in the single-lepton channel
 - H_T is used in the dilepton channel
- ▶ In region (5j,3b) of single lepton channel the NN is trained to discriminate $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$ against $t\bar{t} + \text{light-jets}$
- ▶ Variable categories entering NNs:
 - Object kinematics
 - Global event variables
 - Event shape variables
 - Object pair properties
 - **Matrix element method variables**, only in *

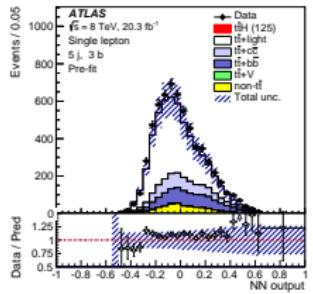
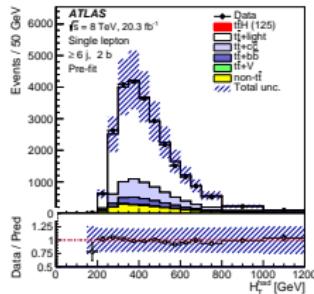
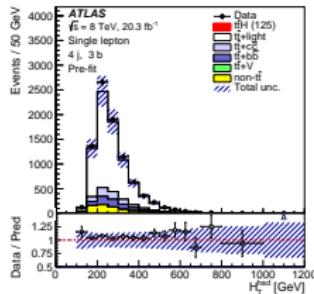
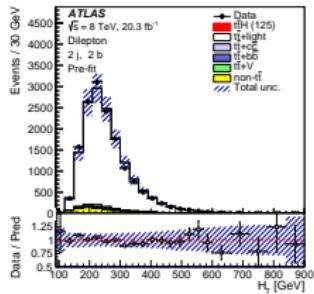
		Single-lepton		
		2 b	3 b	≥ 4 b
4 j	H_T^{had}	H_T^{had}	H_T^{had}	
	H_T^{had}	NN	NN	
	H_T^{had}	NN *	NN *	
5 j	H_T^{had}	NN	NN	
$\geq 6j$	H_T^{had}	NN *	NN *	

		Dilepton		
		2 b	3 b	≥ 4 b
2 j	H_T	—	—	—
	H_T	NN	—	—
	H_T	NN	NN	NN
3 j	H_T	NN	—	—
$\geq 4 j$	H_T	NN	NN	NN

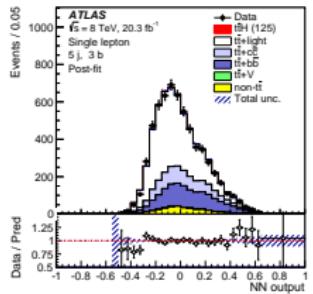
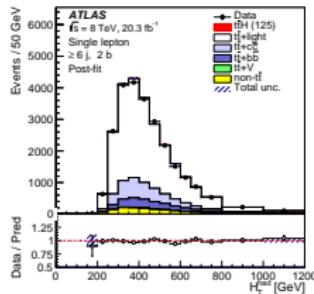
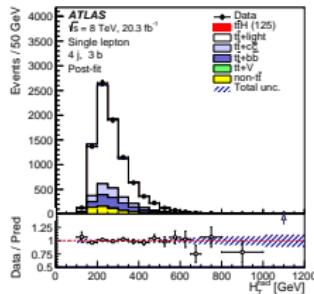
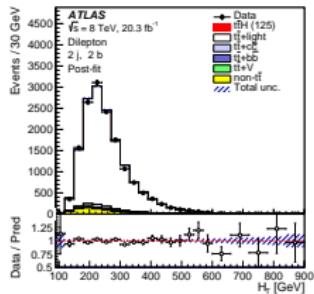


Pre-fit/post-fit comparisons

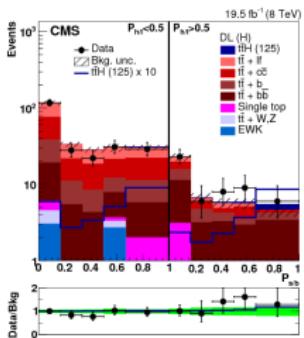
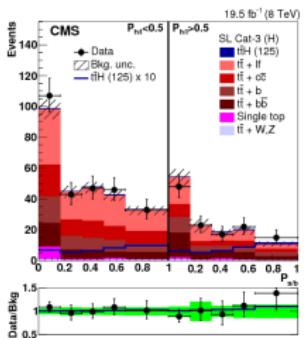
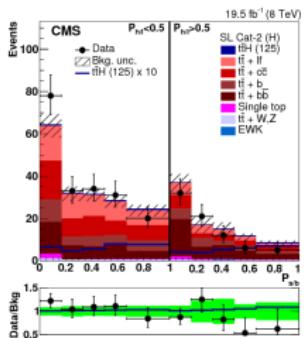
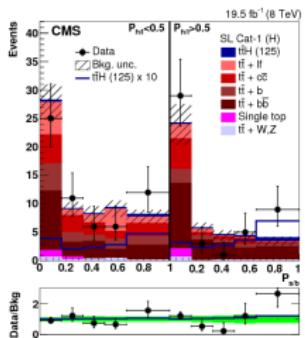
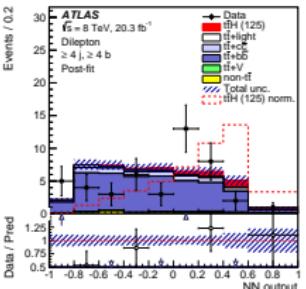
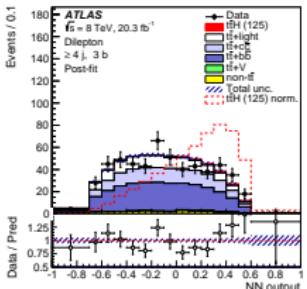
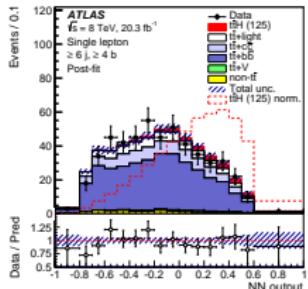
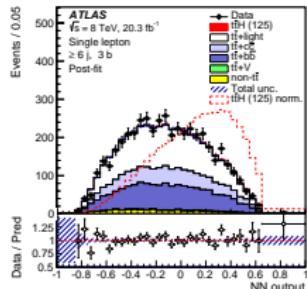
Pre-fit



Post-fit

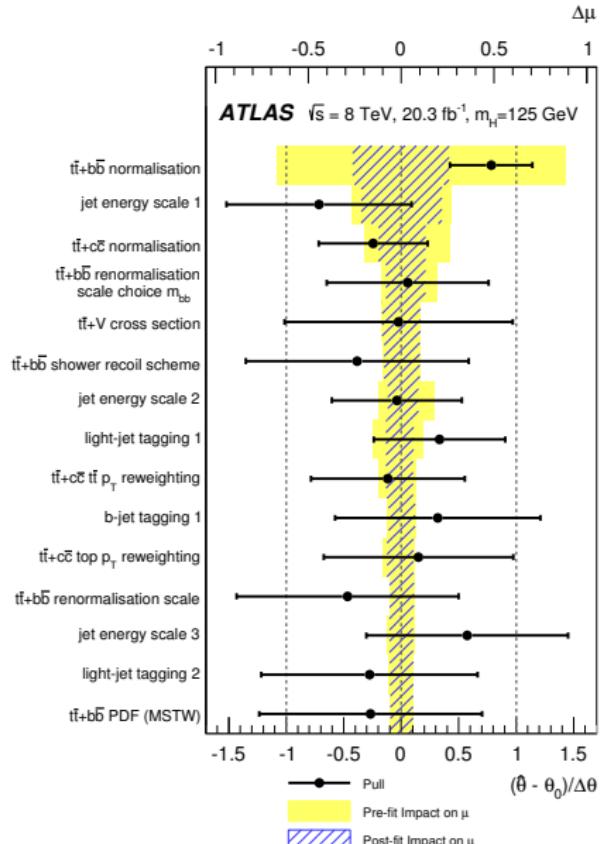


Post-fit variables



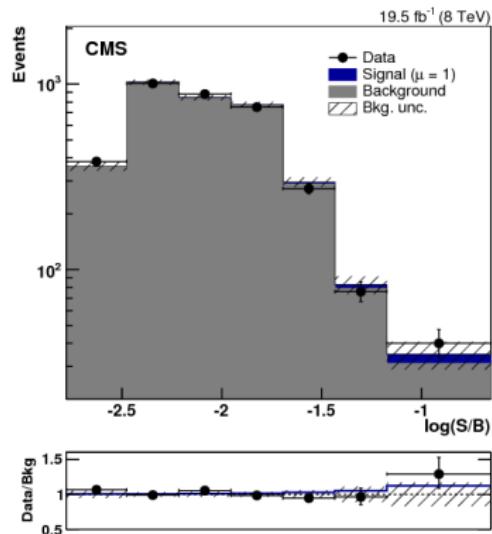
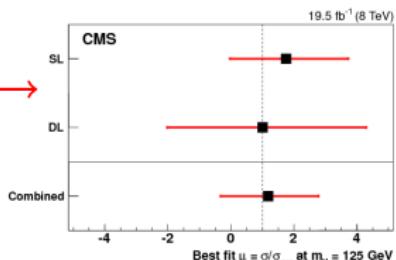
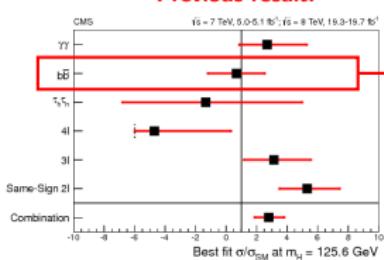
Systematic uncertainties

- ▶ Jet Energy scale:
 - ATLAS : Split in 22 uncorrelated components
 - CMS : 1% to 8% depending on p_T and η
- ▶ b -tagging:
 - ATLAS: b/c /light-tagging split into 6/6/12 uncorrelated components
 - CMS : from 2% to 17%
- ▶ $t\bar{t} + \text{jets}$ modelling
 - Shape and normalization uncertainties derived from variation of renormalization scale and PDF
- ▶ $t\bar{t} + \text{HF}$ normalizations
 - ATLAS : 50% on $t\bar{t} + b\bar{b}$, 50% on $t\bar{t} + c\bar{c}$
 - CMS : 50% on $t\bar{t} + b\bar{b}$, 50% on $t\bar{t} + b$, 50% on $t\bar{t} + c\bar{c}$
- ▶ top p_T reweighting
 - CMS : Doubling or removing the correction
 - ATLAS :
 - ◊ Scale variation, shower model and PDF for $t\bar{t} + b\bar{b}$ reweighting
 - ◊ 9 leading systematic uncertainties from differential $t\bar{t}$ cross-section measurement



Results - CMS

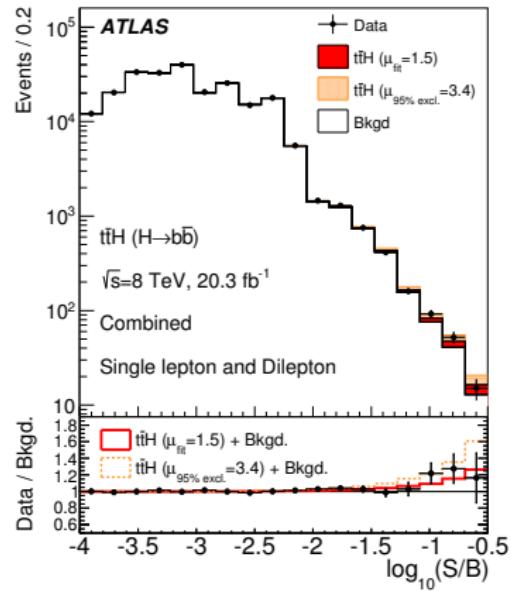
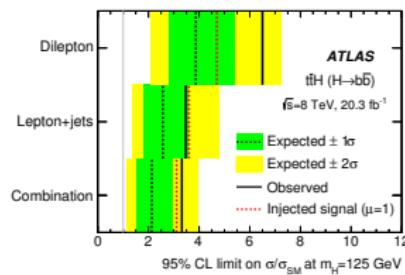
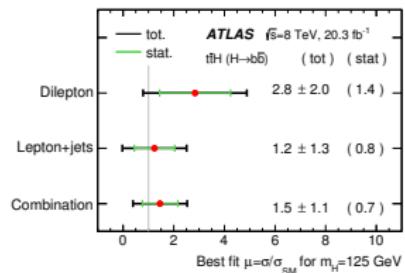
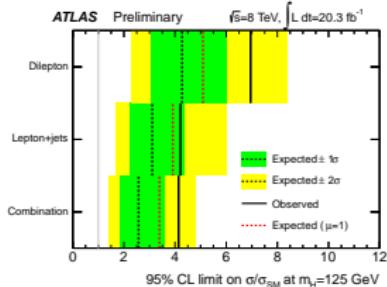
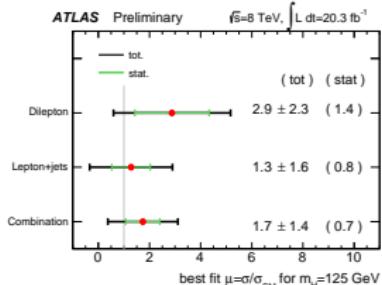
Previous result:



- ▶ 95% observed (expected) Upper Limit on $t\bar{t}H$ cross section: $4.2(3.3) \times \text{SM}$, best fit value $\hat{\mu} = 1.2^{+1.6}_{-1.5}$
 - Improvement by using Matrix Element Method and *b*-tagging information via a continuous discriminant:
 - ◊ 95% CL U.L.: 6% ($3.5 \rightarrow 3.3$) $\times \text{SM}$

Results - ATLAS

Previous result:



- ▶ 95% observed (expected) Upper Limit on $t\bar{t}H$ cross section: $3.4(2.2) \times \text{SM}$, best fit value $\hat{\mu} = 1.5 \pm 1.1$
 - Improvement by using Matrix Element Method:
 - ◊ 95% CL U.L.: 15% ($2.6 \rightarrow 2.2$) $\times \text{SM}$

Conclusions and outlook

- ▶ ATLAS and CMS searched for $t\bar{t}H$ ($H \rightarrow b\bar{b}$) in their 2012 datasets at $\sqrt{s} = 8$ TeV
 - Both collaboration use lepton+jets and dilepton $t\bar{t}$ decays
 - ◊ Run 1 Data continue being looked at by $t\bar{t}H$ ($H \rightarrow b\bar{b}$) in lepton+ τ and all-hadronic final states analyses
 - MEM successfully been used to improve discrimination between $t\bar{t}H$ and $t\bar{t} + b\bar{b}$ by both collaborations
 - CMS : 95% C.L. observed (expected) U.L. : $4.2(3.3) \times \text{SM}$
 - ATLAS : 95% C.L. observed (expected) U.L. : $3.4(2.2) \times \text{SM}$
- ▶ Main sources of systematics
 - Modelling and normalization of $t\bar{t}$ + heavy-flavour backgrounds, JES
- ▶ Run 2... next talk