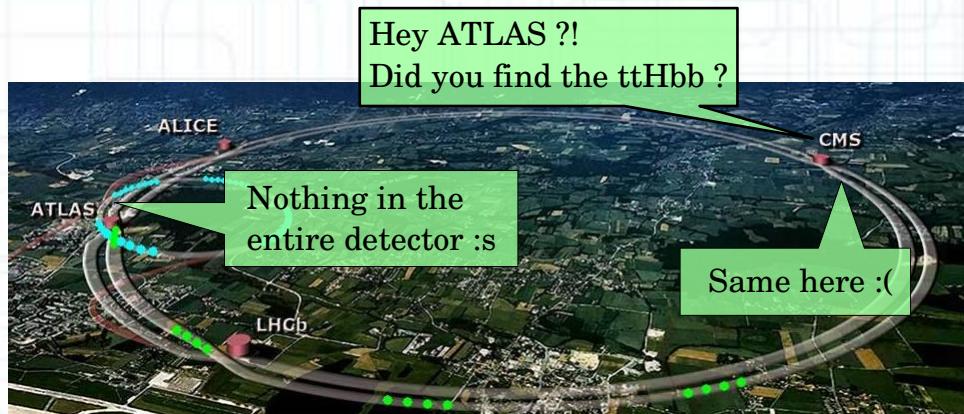


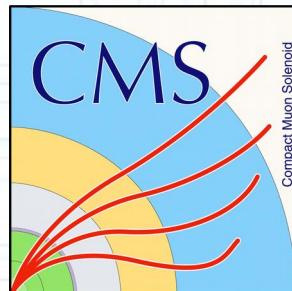
ttH($H \rightarrow bb$) analyses results in ATLAS and CMS.



Don't worry.
I'm coming !



Aix*Marseille
université



Thomas CALVET

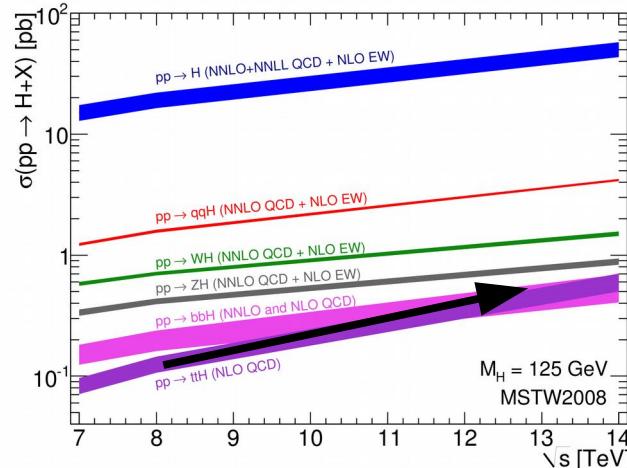
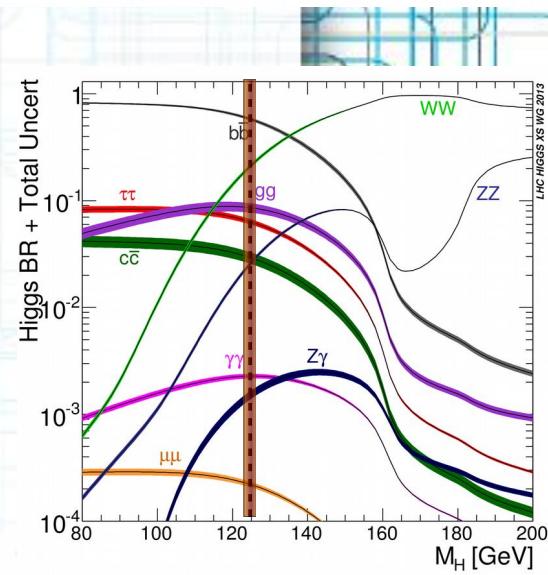
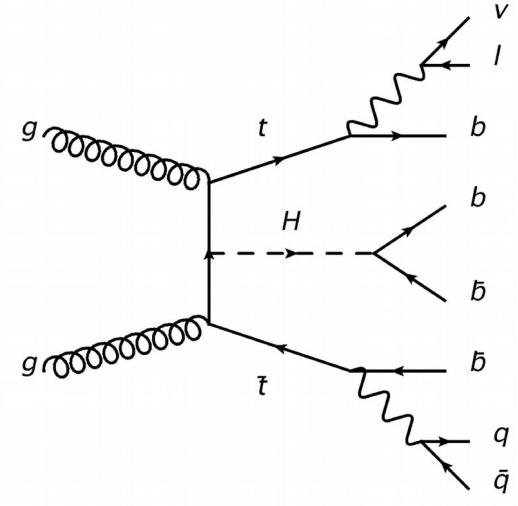
CPPM, Aix-Marseille Université, CNRS/IN2P3 (FR)

Top LHC France

May, 4th 2017

Introduction: Motivations

- **ttH** gives **direct measurement of H \leftrightarrow t coupling**.
- **H \rightarrow bb highest branching ratio**.
- Complex analysis:
 - 4b in the final state => **b-tagging** is one of the main ingredients.
 - Main background tt+jets (tt+light, tt+cc, tt+bb):
 - $\sigma(\text{tt+jets}) \gg \sigma(\text{ttH})$.
 - **tt+bb irreducible**.
- Increased sqrt(s) energy to **13 TeV for Run 2**:
 - $\sigma(\text{ttH})$ increased by **3.7**
 - $\sigma(\text{t}\bar{\text{t}})$ increased by **3.2**
 - In signal enriched regions $\sigma(\text{t}\bar{\text{t}})$ increases more than $\sigma(\text{ttH})$.
- Focus:
 - **2016 results includes 13.2 (11.4-12.9) fb $^{-1}$ data from ATLAS(CMS).**



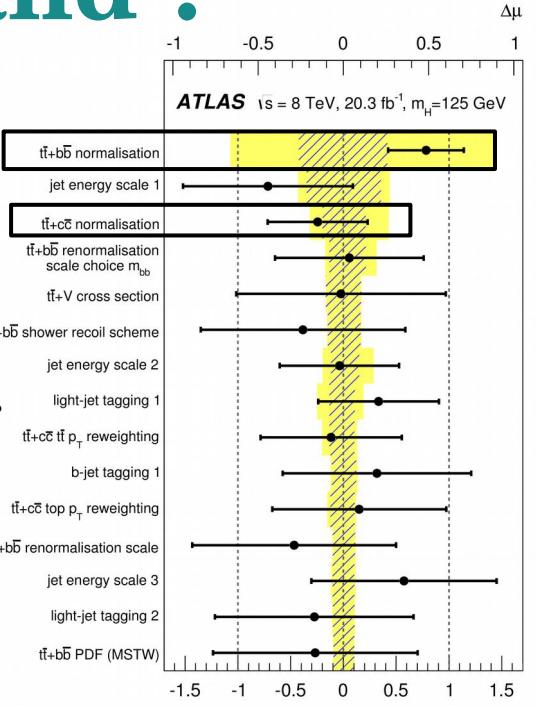
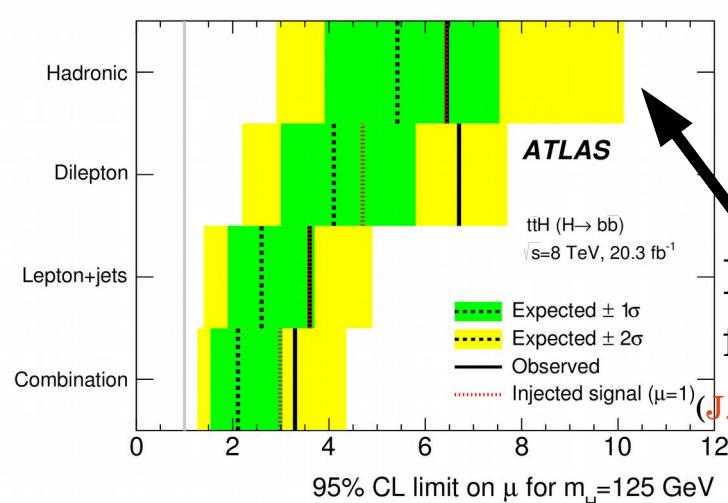
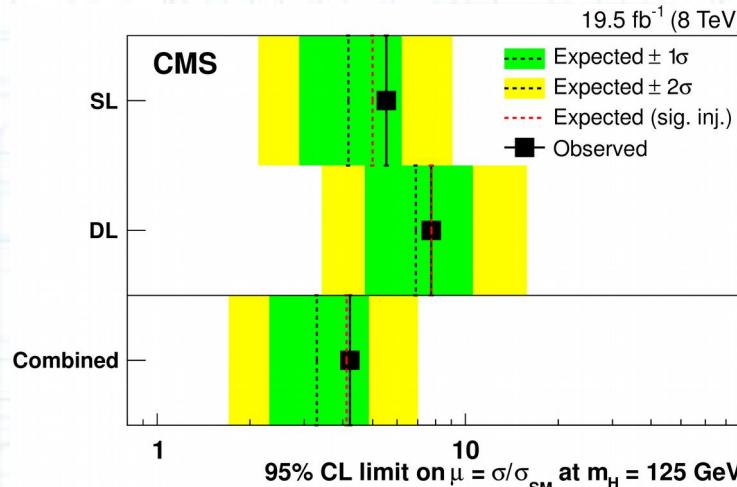
Introduction: Where do we stand ?

→ Run 1 analysis:

- ~20 fb^{-1} of 8 TeV data.
- No evidence of $t\bar{t}H$ or $H \rightarrow bb$.
- Best fit $\mu = 1.4 \pm 1.0$ (ATLAS: Eur. Phys. J. C (2015) 75: 349.)
and $\mu = 1.2 +1.6 -1.5$ (CMS: J. High Energ. Phys. (2014) 2014: 87.).

→ Dominated by systematic uncertainties:

- Large uncertainties from $t\bar{t}+bb$ modelling.
- Leading components of error on mu.



Recently (2016) added fully hadronic channel

(J. High Energ. Phys. (2016) 2016: 160.).

→ Preliminary 13 TeV result from CMS with 2.7 fb^{-1} data (CMS PAS HIG-16-004):

- Including a boosted channel.
- 2016 results includes 13.2 (11.4-12.9) fb^{-1} data from ATLAS(CMS).

Introduction: Towards Run 2

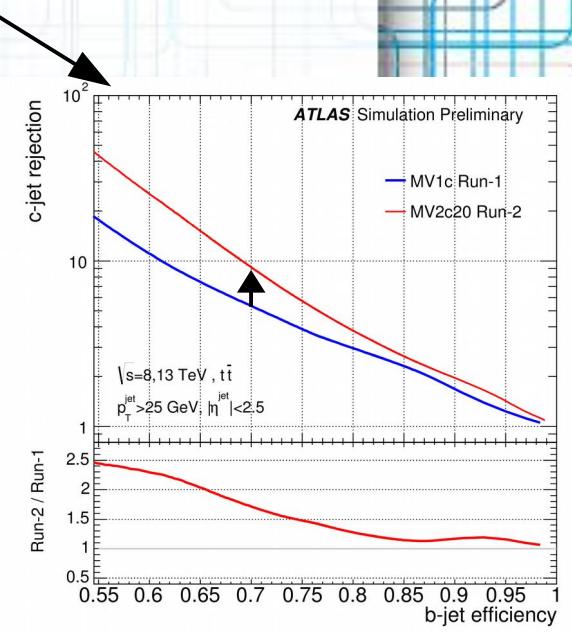
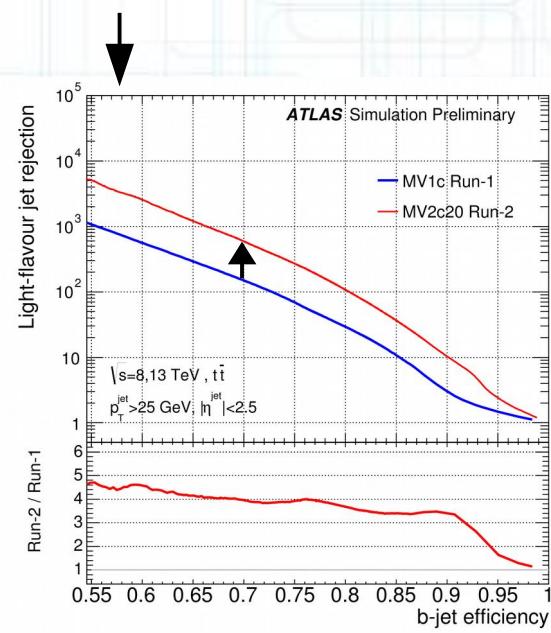
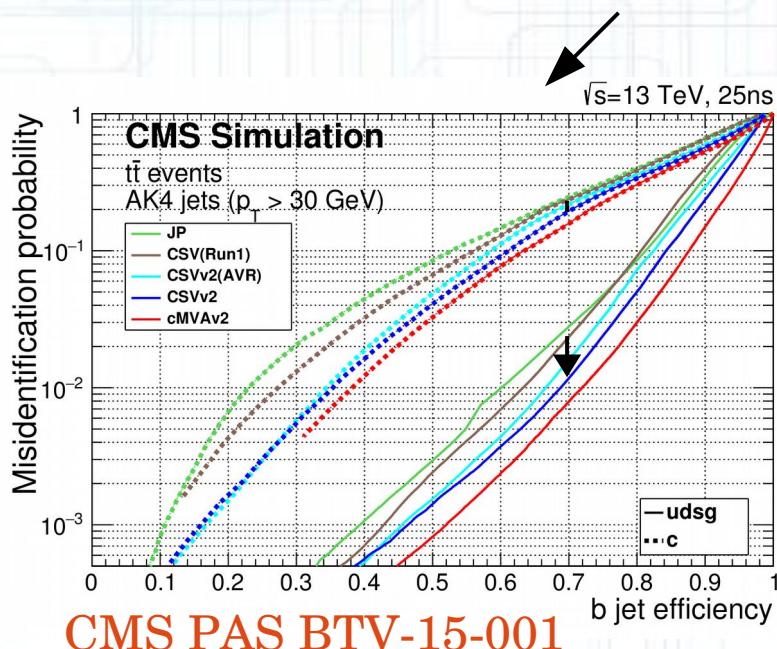
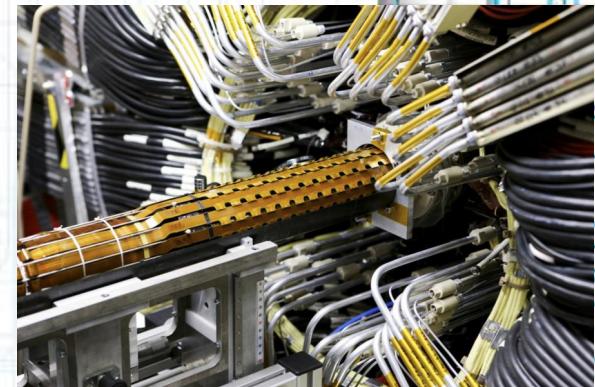
→ Identification of b-jets crucial for ttH(H->bb).

→ Improved b-tagging from Run 1 to Run 2:

→ ATLAS 70% b-tag eff: total ~3(2) times higher light-(c-)jet rejection.

→ Mainly the IBL introduction.

→ Comparing Run 1 and Run 2 setup on 13 TeV simulations.



Selection and Categorization

Selections

Targets ttbar like events.

Object selections:

- Leptons:

- Baseline:** Isolated, with $p_T > 25$ GeV for leading lepton, 15 GeV for second.
- CMS tighter in single lepton channel.**
- ATLAS: looser second lepton** in $e\mu$ and $\mu\mu$ channels.

- Jets:

- CMS: 30 GeV threshold and $|\eta| < 2.4$. 20 GeV cuts after 2nd jets in di-lepton.**
- ATLAS: 25 GeV threshold with $|\eta| < 2.4$.**
- b-tagging: 70% b-jet tagging efficiency** for both.

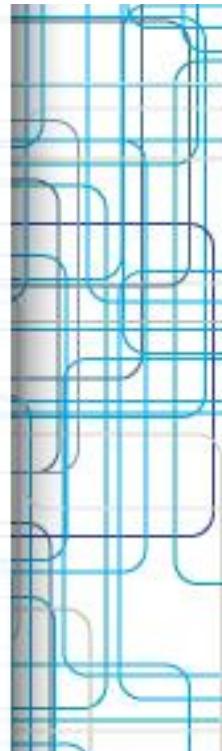
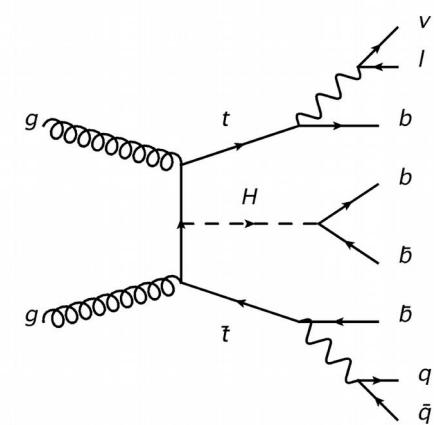
Event selections:

- Di-lepton:

- ATLAS: (two opposite sign leptons) + (≥ 3 -jets ≥ 2 b-tags).
- CMS : (two opposite sign leptons) + (≥ 3 -jets ≥ 3 b-tags).

- Single lepton:

- ATLAS & CMS: (exactly one lepton) + (≥ 4 -jets ≥ 2 b-tags).
- CMS: Exclude low N(jets)xN(b-tags) regions.



Experiment	ATLAS		CMS	
Channel	Single lepton	Di-lepton	Single lepton	Di-lepton
First event selection: triggers				
Electrons: 2015(2016) data	Single e with ID+ISO $p_T > 24$ GeV or ID $p_T > 60$ GeV or loose-ID $p_T > 120(140)$ GeV		Single e with $p_T > 27$ GeV $ \eta < 2.1$	Two isolated e $p_T > X$, $X \in [8, 23]$ GeV
Muons: 2015(2016) data	Single μ with ID at $p_T > 20(24)$ GeV or $p_T > 40(50)$ GeV		Single μ with $p_T > 22$ GeV $ \eta < 2.1$	Two isolated μ $p_T > X$, $X \in [8, 23]$ GeV
Objects:				
Leptons:				
Electrons:	ISO+tight-ID $p_T > 25$ GeV, not in $1.37 < \eta < 1.52$	ISO+tight-ID not in $1.37 < \eta < 1.52$, lead(second in ee [$e\mu, \mu\mu$]) at $p_T > 25(15[10])$ GeV	ISO $p_T > 30$ GeV $ \eta < 2.4$	lead(second) $p_T > 25(15)$ GeV and $ \eta < 2.4$
Muons:	ISO+tight-ID $p_T > 25$ GeV	ISO+tight-ID, lead at $p_T > 25$ GeV, sublead at $p_T > 15(10)$ GeV in ee ($e\mu, \mu\mu$)	ISO $p_T > 25$ GeV $ \eta < 2.4$	lead(second) $p_T > 25(15)$ GeV and $ \eta < 2.4$
Jets:				
Kinematic cuts:		$p_T > 25$ GeV, $ \eta < 2.5$, pile-up suppression	$p_T > 30$ GeV, $ \eta < 2.4$	two leads at $p_T > 30$ GeV and third at $p_T > 20$ GeV, $ \eta < 2.4$
b-tagging cuts:	Require 70% b-jet efficiency leading to 8, 0.3% c-, light- jet efficiencies	Require 70% b-jet efficiency leading to 20, 1% c-, light- jet efficiencies	Require 70% b-jet efficiency leading to 20, 1% c-, light- jet efficiencies	
Second event selection:				
Leptons:	Not di-lepton with one lepton	Exactly two leptons with τ veto	Single lepton	Opposite sign lepton pair with $m_{ll} > 20$ GeV and $76 < m_{ee \text{ or } \mu\mu} < 106$ GeV
N(jets) \times N(b-tags)	$\geq 4j \geq 2b$	$\geq 3j \geq 2b$	$\geq 4j \geq 2b$	$\geq 3j \geq 3b$

Table 1: Object and event selections for the $t\bar{t}H(H \rightarrow bb)$ analyses in ATLAS and CMS.

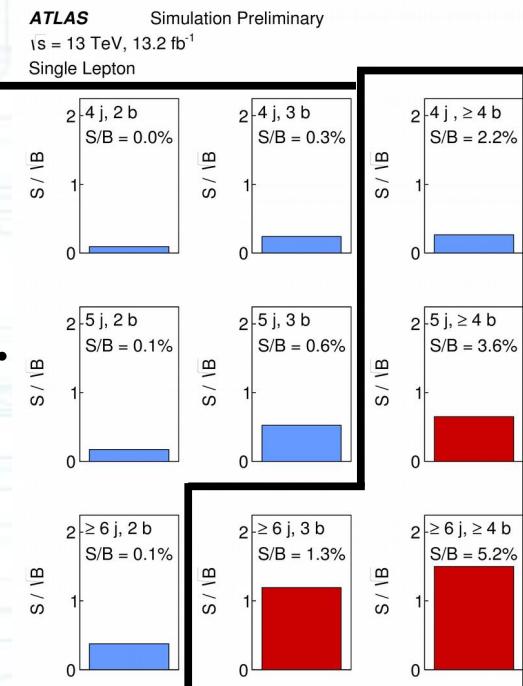
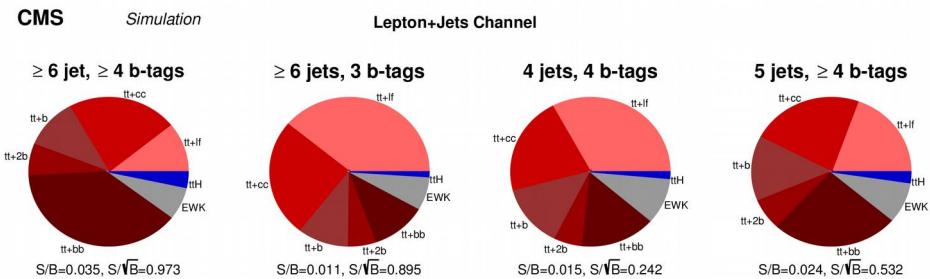
Strategy: Single lepton channel

ATLAS and CMS: divide and conquer strategy.

(exactly one lepton) + (≥ 4 -jets ≥ 2 b-tags) \Rightarrow ttbar dominated



Split in N(jets) and N(b-jets) to increase ttH sensitivity



Define Signal enriched Regions (SR) at high N(jets)/N(b-tags).

CMS:

→ Keep only Signal Regions.

ATLAS:

→ All sub-categories are kept:

- More sensitivity to smaller background (e.g. fakes).
- High constraining power on ttbar uncertainties.

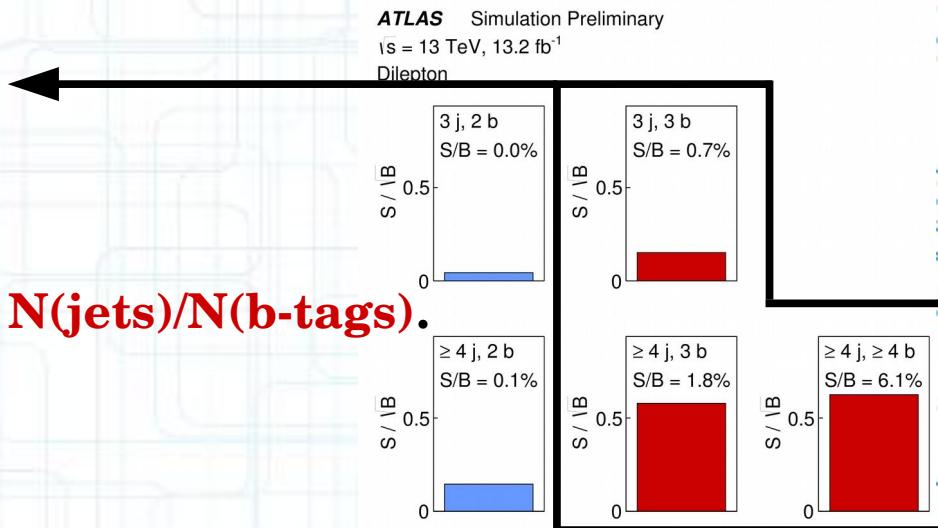
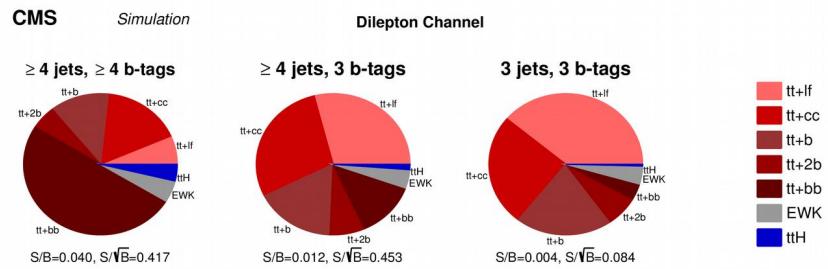
Low purity: reach at most S/B~5%

=> Use MVAs in relevant regions (ATLAS-SR, CMS-All).

Strategy: Di-lepton channel

ATLAS and CMS: divide and conquer strategy.

(exactly two lepton) + (≥ 3 -jets $\geq 2(3)$ b-tags) \Rightarrow ttbar dominated
Split in N(jets) and N(b-jets) to increase ttH sensitivity



Define Signal enriched Regions (SR) at high N(jets)/N(b-tags).

CMS:

→ Keep only Signal Regions.

ATLAS:

→ All sub-categories are kept:

- More sensitivity to smaller background (e.g. fakes).
- High constraining power on ttbar uncertainties.

Low purity: reach at most S/B~6%

=> Use MVAs in relevant regions (ATLAS-SR, CMS-All).

Multi-variate Analysis

MVAs in CMS

Low ttH purity after selection/categorization => Use MVAs.

Two step MVAs in all categories:

→ Step 1: **BDT for ttbar VS ttH separation:**

→ Gives high inclusive ttbar VS ttH separation.

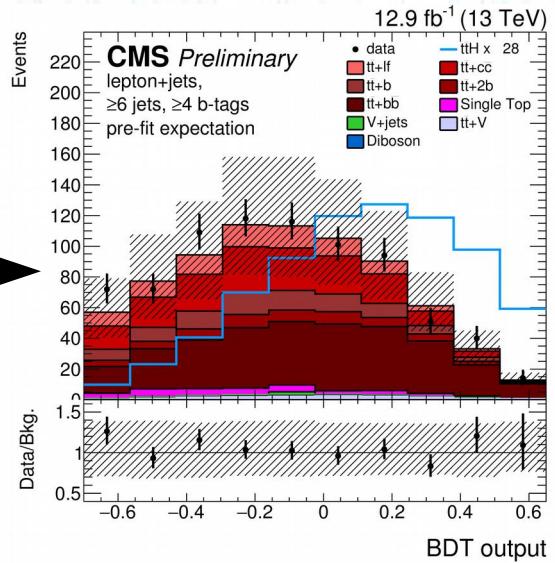
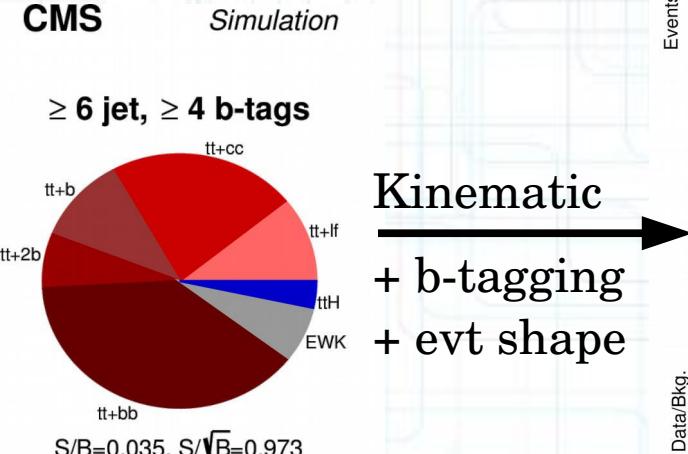
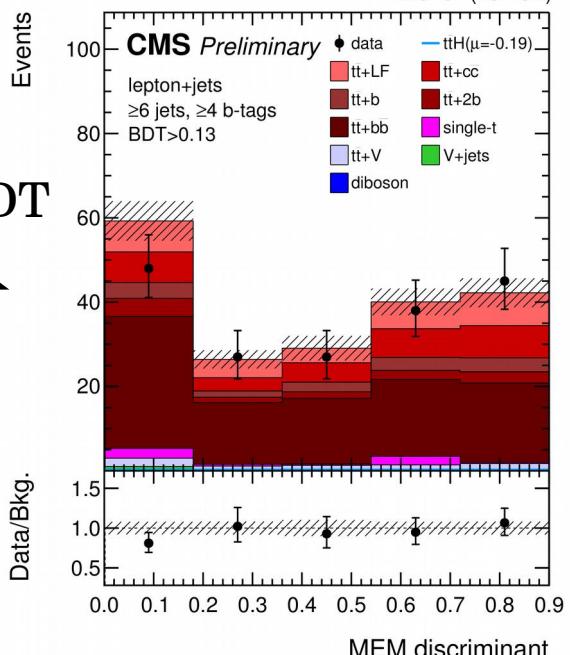
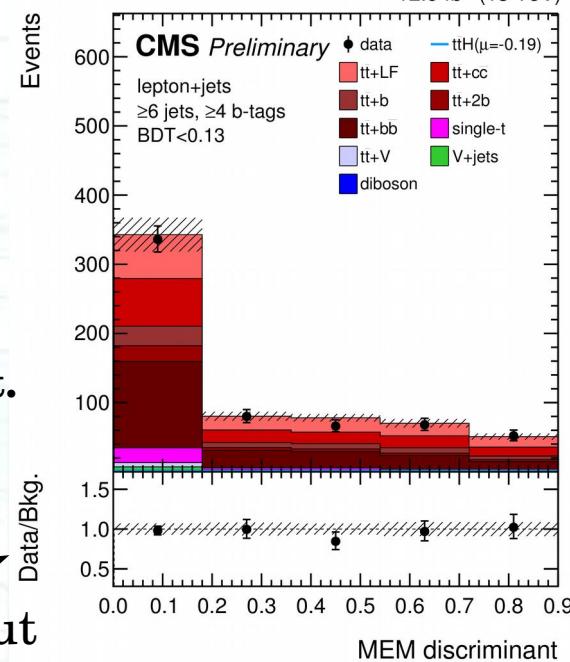
→ Split categories in high and low BDT regions:

→ No split of di-lep 3j3b, use BDT as final discriminant.

→ Step 2: **MEM discriminant:**

→ Gives high tt+ ≥ 1 b VS ttH separation.

→ Used as a **final discriminant**.



Low BDT
output

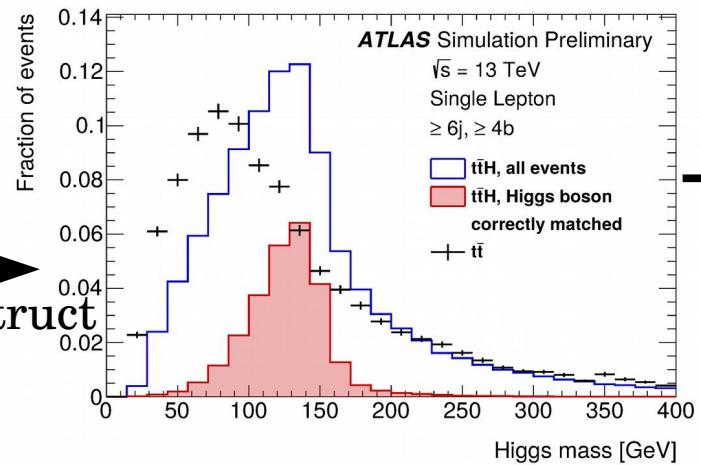
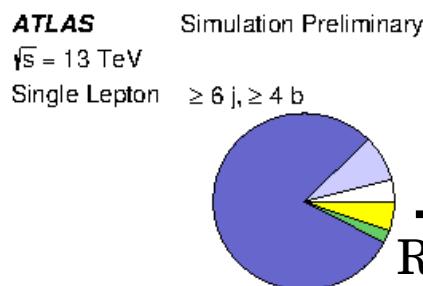
High BDT
output

MVAs in ATLAS

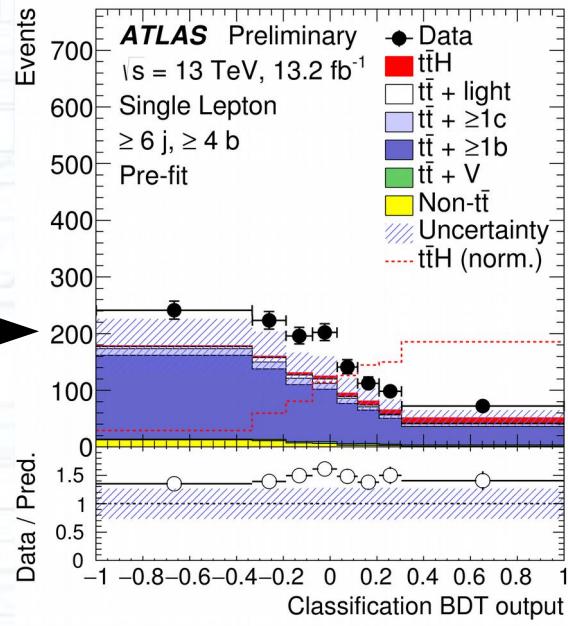
Low ttH purity after selection/categorization => Use MVAs.

Two step MVAs in signal regions:

- Step 1: **BDT for ttH system reconstruction:**
 - Find the right matching of jets/partons.
 - Define new variables for ttbar VS ttH separation:
 - Reco BDT output distribution gives the best bkg/sig separation before step 2.
- Step 2: **BDT for ttbar VS ttH separation:**
 - Gives ttbar VS ttH separation.
 - Used as a **final discriminant**.



Kinematic
+ reco
+ evt shape



Background modelling and systematic model

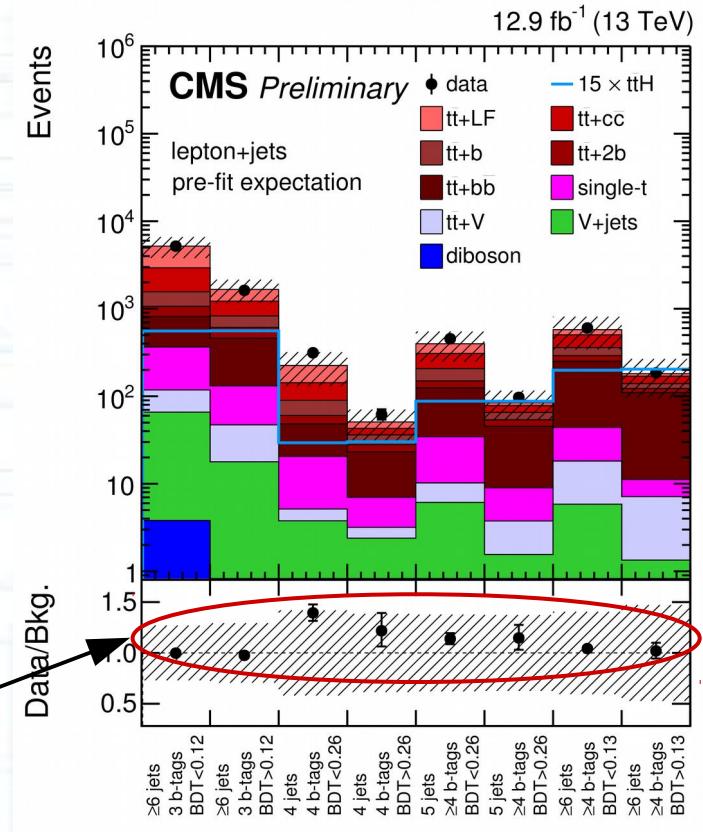
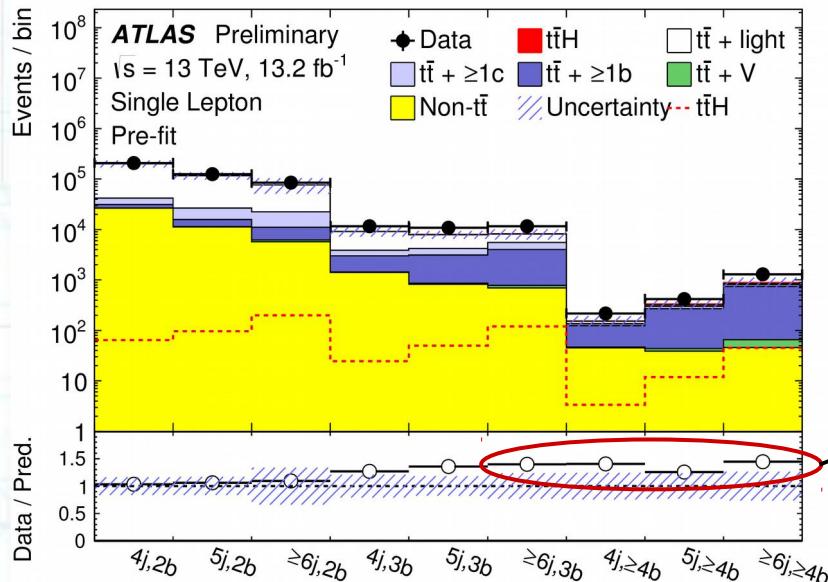
ttbar modelling

ATLAS:

- Powheg+Pythia6 nominal prediction:
 - Re-weight to latest theoretical predictions for ttbar, $tt+\geq 1c$ and $tt+\geq 1b$.
- Large differences between Data and MC:
 - Covered with large systematics => gives freedom to the fit.

CMS:

- Powheg+Pythia8 nominal prediction:
 - Tuned to match 8 TeV data.
- Good Data/MC agreement
- Limit amount of systematics.



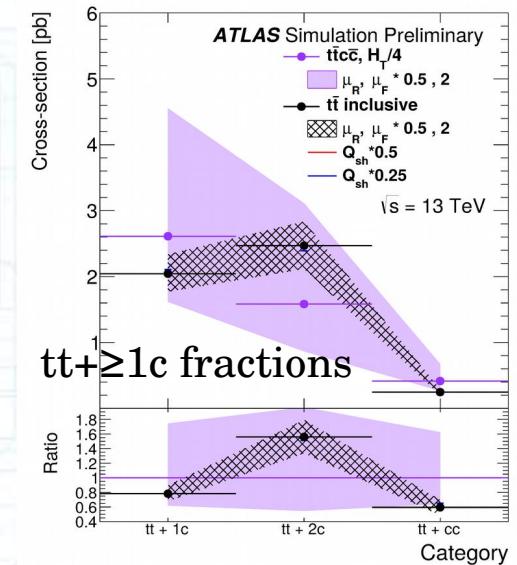
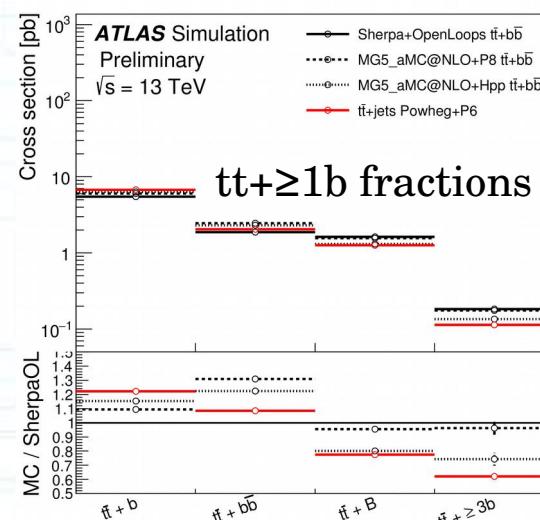
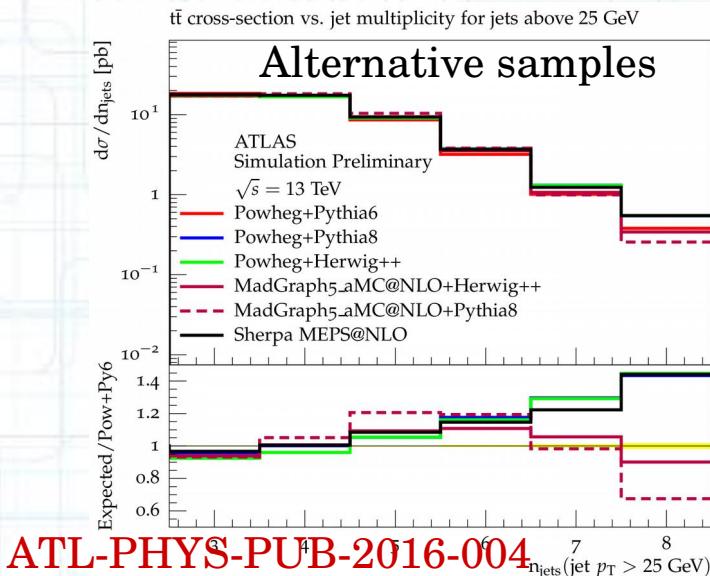
ATLAS ttbar modelling – Alt slide

Nominal ttbar
Powheg+Pythia 6
 \rightarrow tt+light/ $\geq 1c/\geq 1b$
fractions.
 \rightarrow Basic kinematics.

Cross section: NNLO QCD + NNLL soft gluons.

$tt+\geq 1b$: Re-weight additional b-jets fractions and kinematics to dedicated Sherpa+OL 4FS.

$tt+\geq 1c/light$: Re-weight top/ttbar p_T to NNLO.



Systematics:

- 6% on incl ttbar norm + free to float $tt+\geq 1c/\geq 1b$.
- $tt+\geq 1b$ RW: 6variations of sherpa + 2 alternative RW to aMC@NLO 4FS.
- $tt+\geq 1c$: diff in add c-jets kin and fractions w.r.t aMC 3FS.
- Residuals: alternative samples comparisons with re-weightings applied

Generator: aMC+H7 vs PH7

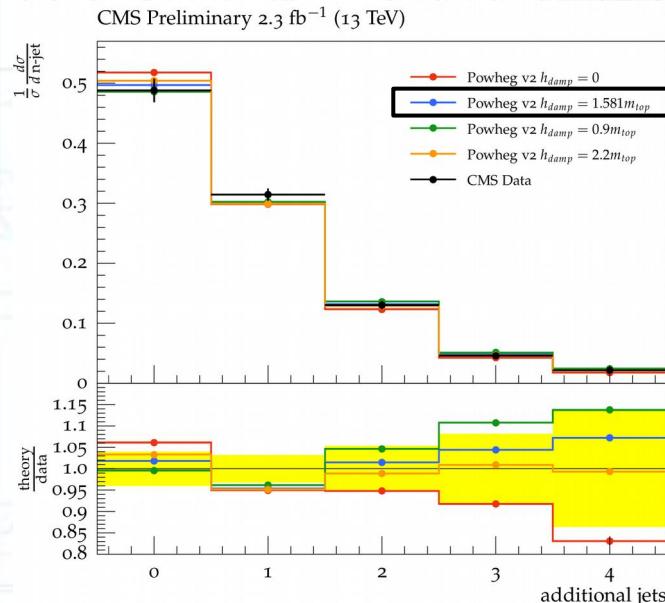
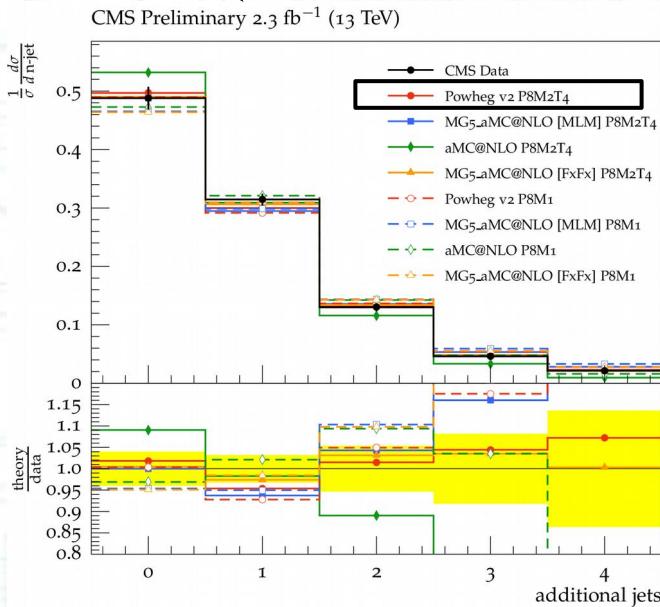
Parton Shower: PH7 vs PH++

Radiation: QCD scale + hdamp vars

CMS ttbar modelling

- Based Powheg+Pythia 8 nominal sample:

→ α_s^{ISR} and h_{damp} tuned from 8 TeV ttbar analysis (CMS PAS TOP-16-021).



- Systematic model:

- **ttbar XS: Inclusive QCD scale + 50% on tt+HF subset.**
- **Q^2 scale: ME unc. from renormalization/factorization scale variations:**
 - Add theoretical shape unc. on ttbar.
- **PS scale: ISR/FSR scales variations → Norm Only.**
- **PDF unc. on gg initiated processes.**

Additional processes and systematics

- ATLAS and CMS common processes:

Experiment	ATLAS		CMS	
Channel	ME generator	Parton Shower	ME generator	Parton Shower
$t\bar{t}H$	MG5_aMC@NLO	Pythia 8	Powheg	Pythia 8
$W/Z+jets$	Sherpa 2.1.1		MG5_aMC@NLO	Pythia 8 (FxPx ME+PS matching)
$t\bar{t} + V$	Madgraph	Pythia 6	MG5_aMC@NLO	Pythia 8
Di-boson	Sherpa 2.1.1			Pythia
Single top (tW only in CMS)	Powheg	Pythia 6	Powheg	Pythia 8

Table 1: Non-ttbar background and signal MC samples for the $t\bar{t}H(H \rightarrow bb)$ analyses in ATLAS and CMS.

- ATLAS includes low N(jets)xN(b-tags) regions:

=> Need to include QCD events with fake/non-prompt lepton events.

→ Data driven estimate in fake enriched regions.

Additional processes and systematics

Background systematics

ATLAS

CMS

ttH: PDF + BR norm unc.

→ PS from comp to aMC+H++

→ QCD scale: with norm and shape decorr.

Single top: overall **+5/-4% XS unc.**

→ Wt: add PS unc. and unc. for **overlap with ttbar**.

V+jets:

→ **30% XS decorr across N(jets).**

→ **30% V+HF XS decorr across N(b-tags).**

Di-boson: **50% XS unc.**

Fakes: **50% XS unc.**

ttV: **15% XS unc. decorr PDF and QCD scale.**

QCD scales:

→ **ttH:** rate unc. for NLO prediction

→ **single top:** rate unc. for NLO prediction

→ **V+jets:** rate unc. for NNLO prediction

→ **VV:** rate unc. for NLO prediction

PDF uncertainties:

→ **gg(ttH):** rate unc. on **ttH**.

→ **gg initiated:** rate unc.

correlated across **ttbar and ttZ**.

→ **qq initiated:** rate unc.

correlated across **ttW, V+jets, VV**.

→ **qg initiated:** rate unc. on **single top**.

Detector Systematics:

→ Luminosity and pile-up.

→ Jet energy scale and resolution: 2(19) unc. in CMS(ATLAS).

→ b-tagging: decorrelated for b-, c-, light- tagging.

→ Leptons ID/iso and trigger efficiencies.

Results

Results: signal strength

Combined fit of all categories:

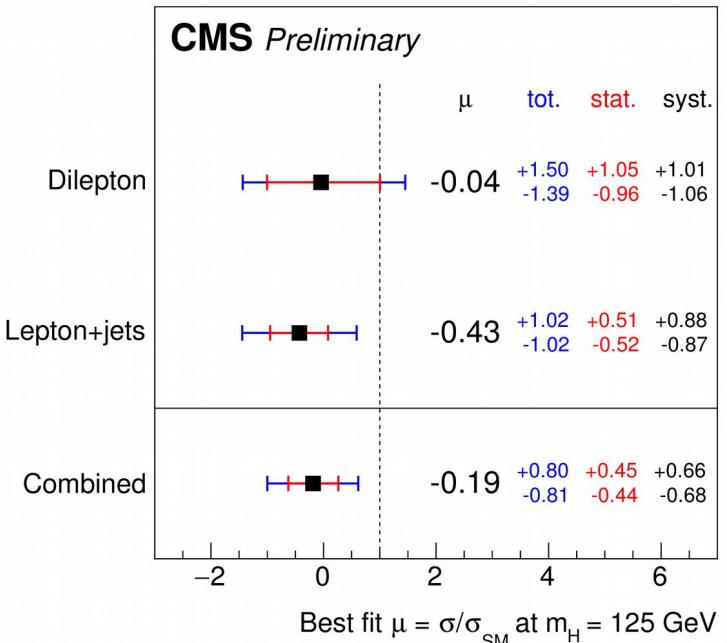
→ Binned Maximum Likelihood Fit.

→ Parametrized in **signal strength: $\mu = \sigma/\sigma_{\text{SM}}$** .

Sensitivity lead by lepton+jet channel:

→ **Higher statistic** => more constraining power on ttbar.

→ **~10/20% gain combining with di-lepton channel.**



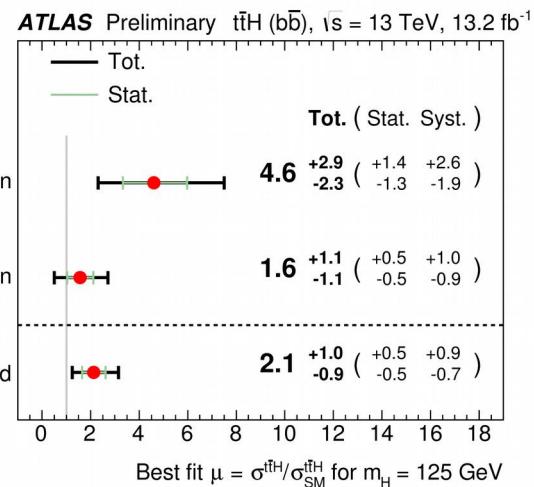
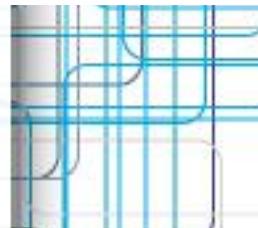
ATLAS Preliminary

Uncertainty source	$\Delta\mu$
$t\bar{t} + \geq 1b$ modelling	+0.53 -0.53
Jet flavour tagging	+0.26 -0.26
$t\bar{t}H$ modelling	+0.32 -0.20
Background model statistics	+0.25 -0.25
$t\bar{t} + \geq 1c$ modelling	+0.24 -0.23
Jet energy scale and resolution	+0.19 -0.19
$t\bar{t}$ +light modelling	+0.19 -0.18
Other background modelling	+0.18 -0.18
Jet-vertex association, pileup modelling	+0.12 -0.12
Luminosity	+0.12 -0.12
$t\bar{t}Z$ modelling	+0.06 -0.06
Light lepton (e, μ) ID, isolation, trigger	+0.05 -0.05
Total systematic uncertainty	+0.90 -0.75
$t\bar{t} + \geq 1b$ normalisation	+0.34 -0.34
$t\bar{t} + \geq 1c$ normalisation	+0.14 -0.14
Statistical uncertainty	+0.49 -0.49
Total uncertainty	+1.02 -0.89

tt+HF main source of unc.

→ **ATLAS**: especially $t\bar{t}+\geq 1b$ syst/norm.

→ **CMS**: **tt+HF 50% norm unc.**



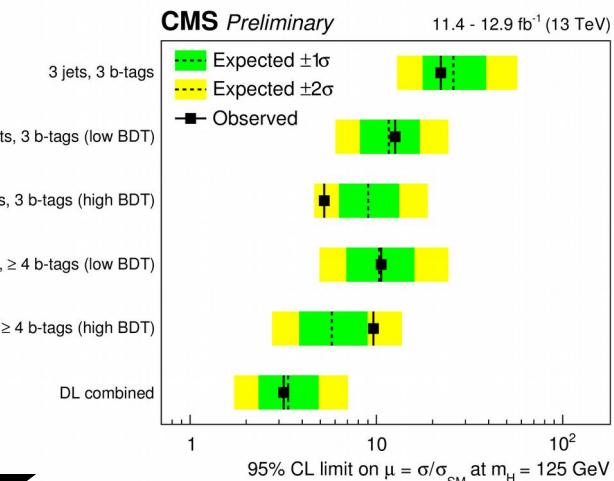
Results: limit on μ

Similar expected limits:

- 11% better combined expected limit from CMS.
- ATLAS: limit computed using post-fit background estimate:
 - Include lost of performance from increased tt+HF content.

	Observed		Expected $\mu = 0$			
	Median	$+/-1\sigma$	Median	$+/-1\sigma$		
Experiment	ATLAS	CMS	ATLAS - 13.2 fb^{-1}	$[3.8, 7.9]$	3.4	$[2.4, 4.9]$
Di-lepton	10.1	3.2	5.3	$[3.8, 7.9]$	3.4	$[2.4, 4.9]$
Single lepton	3.6	1.8	2.2	$[1.6, 3.2]$	2.1	$[1.5, 3.1]$
Combined	4.0	1.5	1.9	$[1.4, 2.8]$	1.7	$[1.2, 2.4]$

Table 1: Observed and expected 95% CL upper limits on signal strength in ATLAS and CMS analysis.



CMS: additional category split

- Upper limit per category in 1+jets and di-lep.
- Sensitivity driven by high purity bins.
- Higher differences Exp/Obs in high purity reg.

Conclusion

Challenging analysis:

- Lots of jets/b-jets in the final state.
- Overwhelming tt+bb background:
 - Badly constrained by data and difficult to model theoretically.
- ATLAS:
 - Use latest theoretical predictions.
 - Add large systematic uncertainties to cover differences with data.
- CMS:
 - Tune prediction to match data.
 - Add unc. from variation of best simulation.

Run 2 analysis with 13.2 fb^{-1} for ATLAS and $11.4\text{-}12.9 \text{ fb}^{-1}$ for CMS:

- Reach Run 1 sensitivity with 2/3 the amount data.
- Released in **ATLAS conf note (ATLAS-CONF-2016-080)**
and CMS-PAS (CMS PAS HIG-16-038).

	Run 1			Run 2		
	Observed limit	Expected Limit	Best fit μ	Observed limit	Expected Limit	Best fit μ
ATLAS combined fit	3.3	$2.1^{+0.9}_{-0.5}$	$1.4^{+1.0}_{-1.0}$	4.0	$1.9^{+0.9}_{-0.5}$	$2.1^{+1.0}_{-0.9}$
CMS combined fit	4.2	$3.3^{+1.6}_{-1.0}$	$1.2^{+1.6}_{-1.5}$	1.5	$1.7^{+0.7}_{-0.5}$	$-0.19^{+0.80}_{-0.81}$

backup

ATLAS ttbar modelling

- Strategy:

- Match ttbar components to latest predictions.
- Decorrelate all components in the fit.

- Inclusive ttbar based on Powheg+Pythia6:

- XS normalized to NNLO QCD with NNLL soft gluon terms.
- top and ttbar p_T from NNLO predictions:

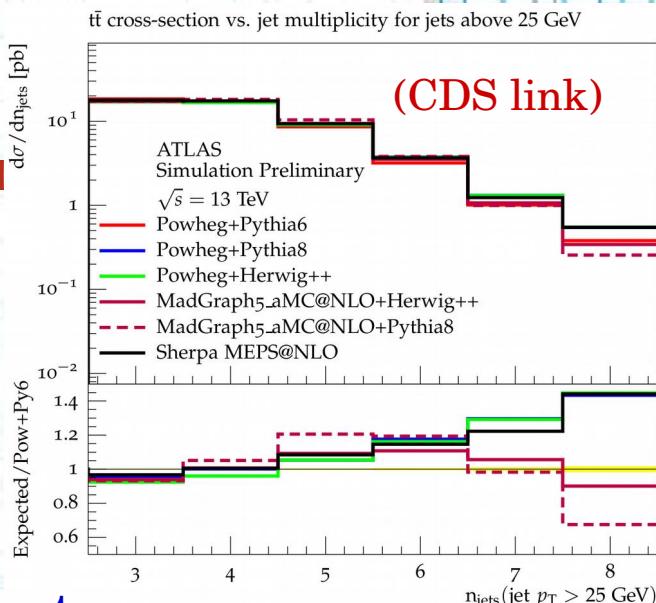
- Improves N(jets) and kinematic distributions.
- Not applied on tt+≥1b: dedicated RW procedure.

→ Assign corresponding systematics:

- 6% norm on ttbar XS.
- top/ttbar p_T : largest difference between NNLO and available samples.

- Residual uncertainties: choice of generator, PS, and radiation

- Use alternative sample all re-weighted to theoretical predictions:
 - NNLO top/ttbar p_T for tt+light/≥1c or dedicated tt+≥1b RW.
- Generator: Powheg+Herwig7 vs aMC@NLO+Herwig7
- Parton Shower: Powheg+Pythia6 vs Powheg+Herwig++
- Radiation: Nominal sample with increased/reduced QCD scale and hdamp.



ATLAS tt+HF modelling

- Nominal tt+light/tt+ ≥ 1 c/tt+ ≥ 1 b fractions from inclusive sample:
 - tt+ ≥ 1 c and tt+ ≥ 1 b norms are left free to float.
 - Large MC undershoot in tt+HF enriched regions.
- tt+ ≥ 1 c:
 - Dedicated MC study with 3 FS tt+cc sample (CDS link).
 - Differences with inclusive sample taken as a systematic:
 - Account for additional c-jet fractions (tt+1c-jet, tt+2c-jets, tt+1cc-jet) and their kinematics.
- tt+ ≥ 1 b:
 - Inclusive sample re-weighted to Sherpa+OL 4FS:
 - Account for additional b-jet fractions (tt+1b-jet, tt+2b-jets, tt+1bb-jet) and their kinematics.
 - Syst from Sherpa+OL parameter variations (7 systs).
 - Alternative aMC@NLO + Py8/H++ 4 FS sample:
 - Differences between 4FS samples in additional b-jet fractions/kinematics used as systematics.

