tt+jets and tt+b-jets modelling for ttH(bb)

Stefan Guindon
State University of New York at Albany

On behalf of ATLAS ttH WG

ttH/tH LHC Subgroup meeting
May 31, 2016

UNIVERSITY AT ALBANY

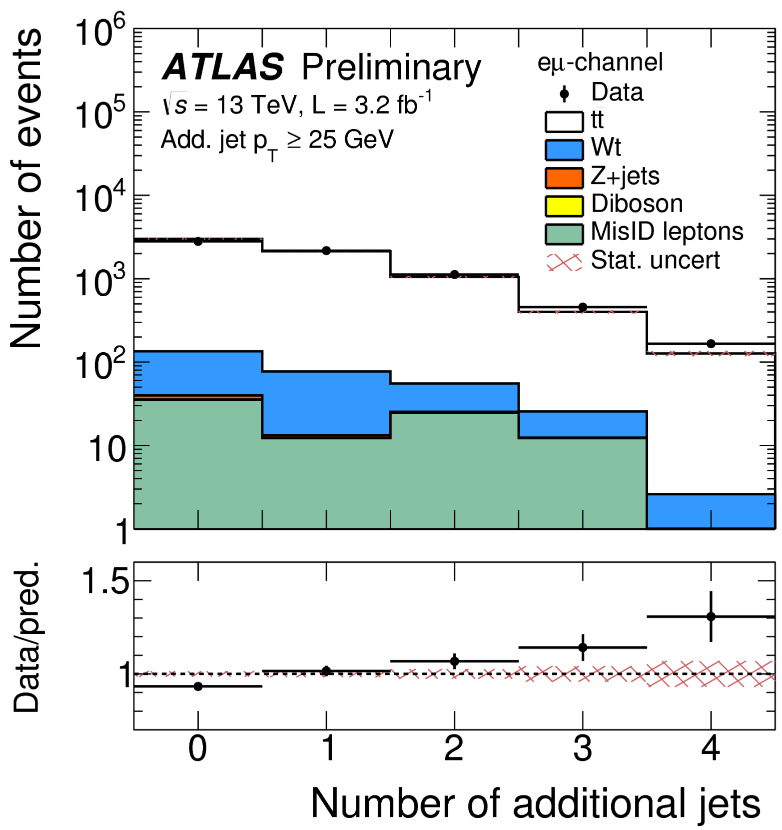
ATLAS EXPERIMENT

tt+jets modelling ATLAS

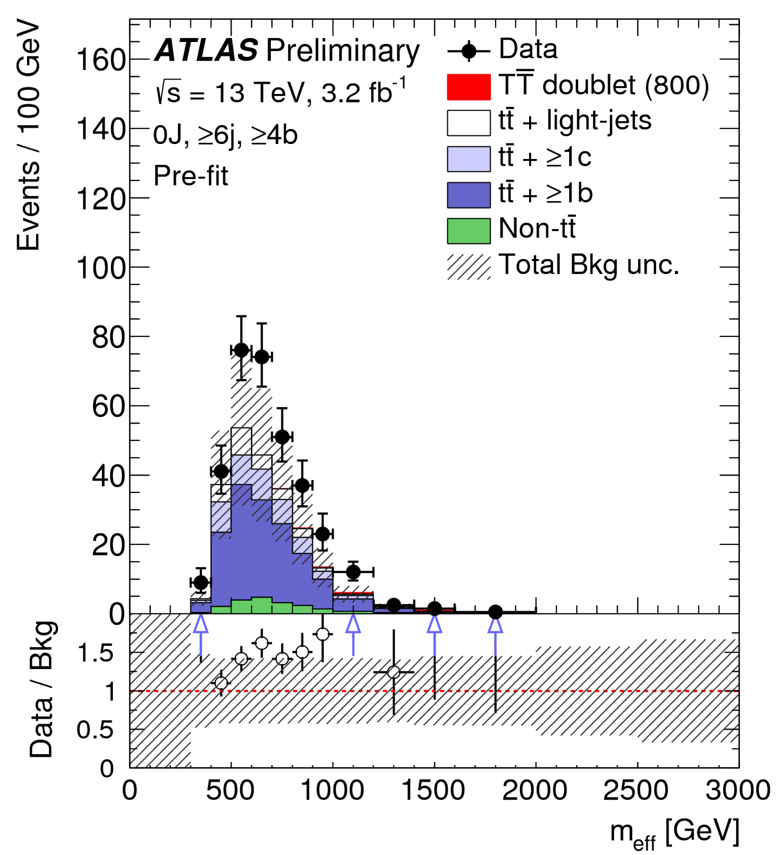
- Several preliminary results have been released by ATLAS for Moriond 2016
  - tt+jets dominated (tt+additional jets and tt dilepton cross section)
  - tt+b-jets background dominated (VLQ / 4 tops searches)

- **All show excesses in additional light and b-jet regions**
  - Similar trend using several different b-tagging working points and additional jet pT requirements

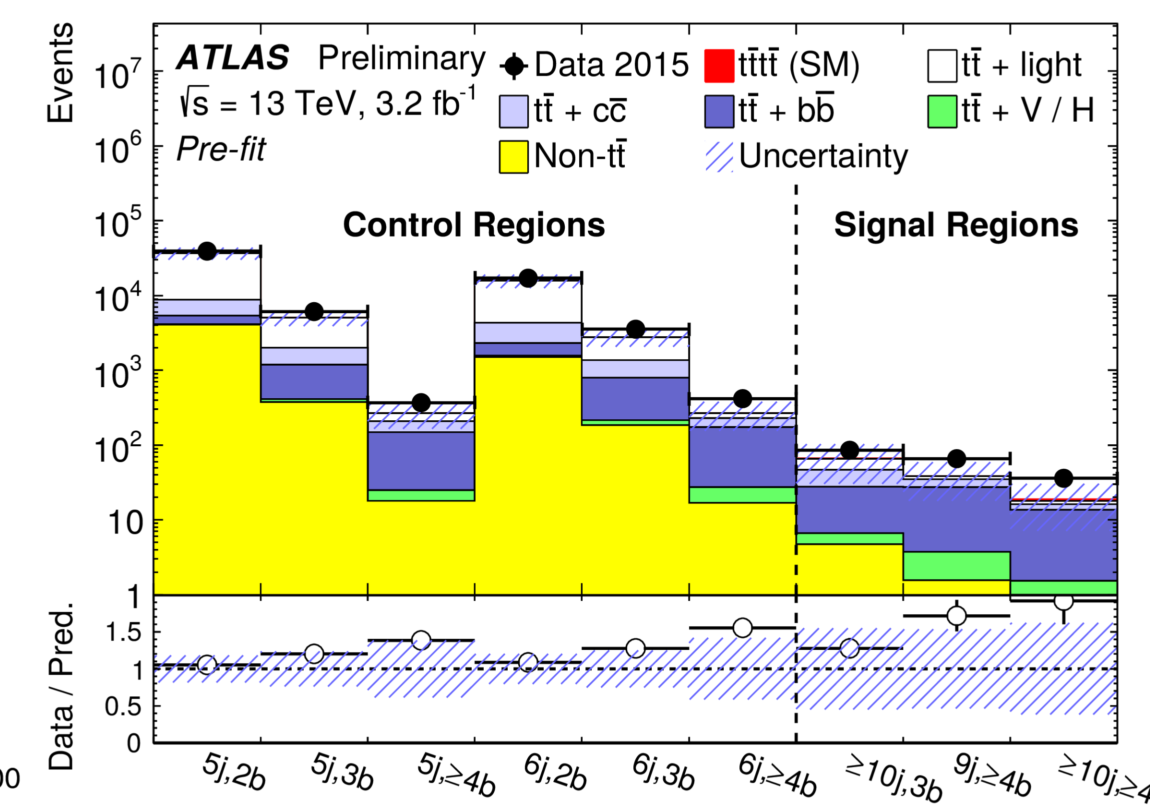
ATLAS-CONF-2015-065



ATLAS-CONF-2016-013



ATLAS-CONF-2016-020



- **ATLAS tt+cc PUB note for TopLHCWG with MG5_aMC@NLO 3FS:** ATL-PHYS-PUB-2016-011

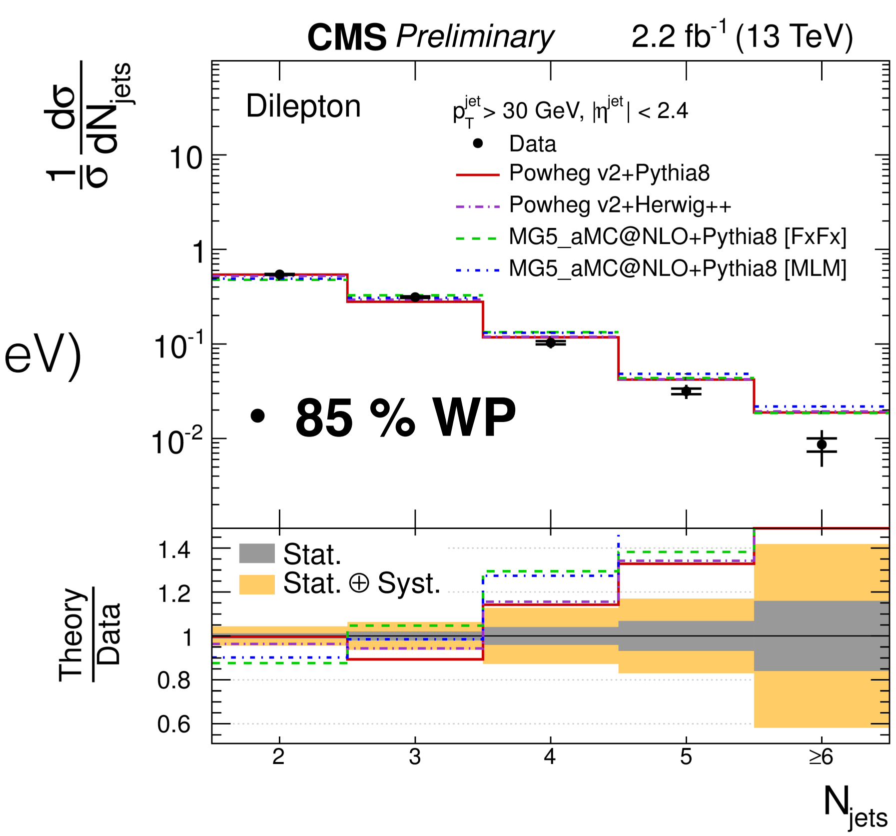
ATLAS and CMS: tt Modelling

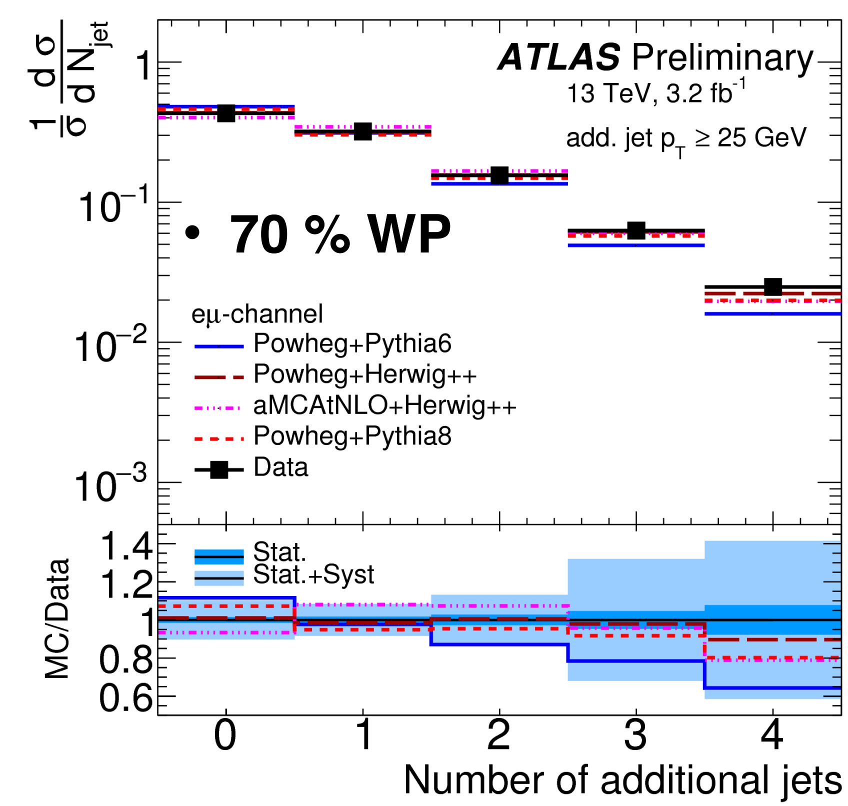
CMS TOP-16-011
ATLAS-CONF-2015-065

- MC generator comparisons with 2015 data in dilepton channel (ATLAS and CMS)

	Generators
CMS	• **Powheg+Pythia8 with CUETP8M1 tune (default tt generator for Moriond EW)** • MG5_aMC@NLO+Pythia8 with FxFx matching (up to 2 extra partons at NLO) • Powheg+Herwig++ with EE5C tune
ATLAS	• **Powheg+Pythia6 with P2012 tune (default tt generator for Moriond EW)** • MG5_aMC@NLO+Herwig++ with UE-EE-5 • Powheg+Herwig++ with UE-EE-5 • Powheg+Pythia8 with A14 tunes (*Main31*, pThard = 0 and hdamp = mtop)

- Jet pT CMS (ATLAS) 30 GeV (25 GeV)
- **ATLAS and CMS have opposite trends in data/MC**

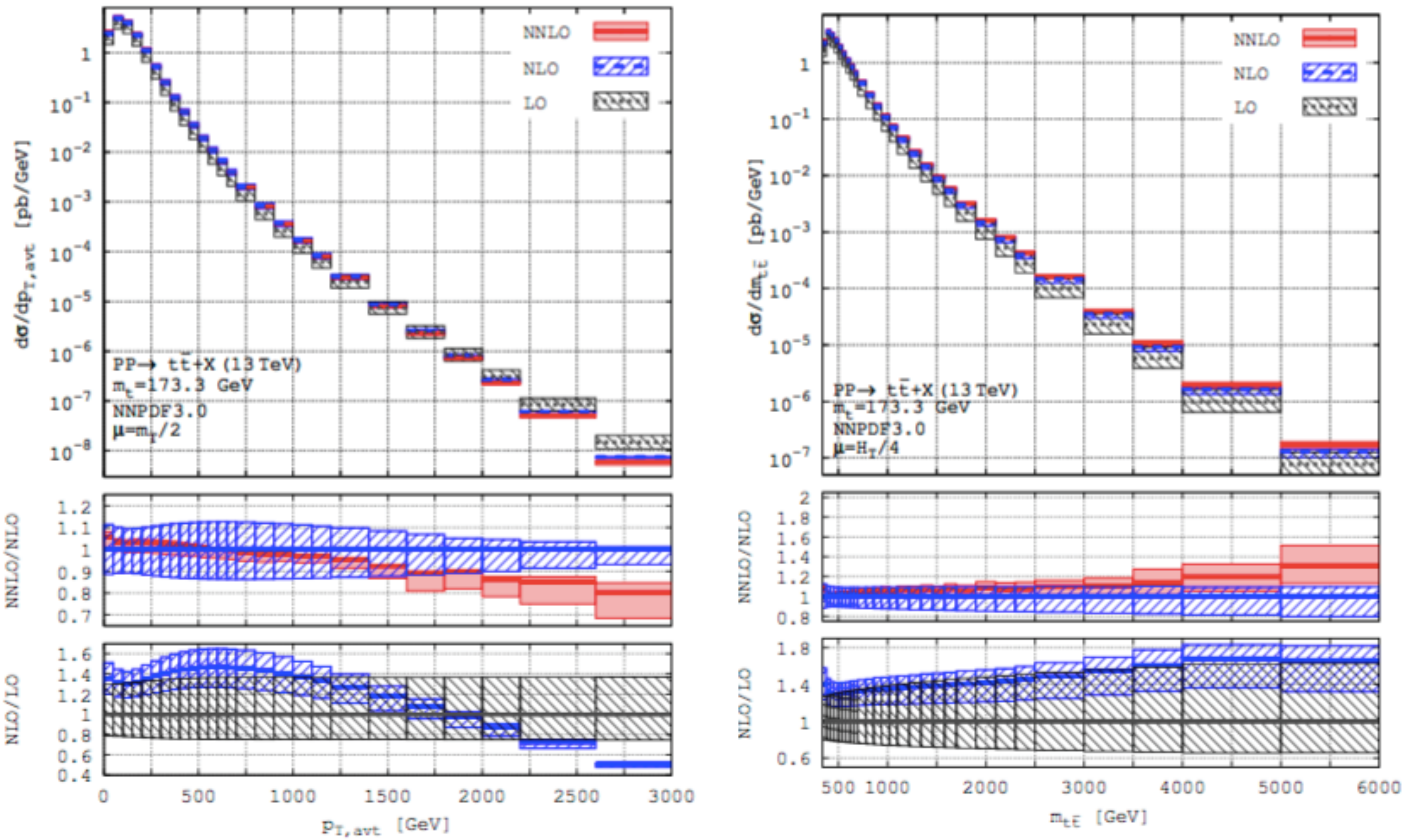




tt+jets modelling NNLO diff.

- Update of NNLO tt differential distributions
  - pT top and M_tt with NNPDF3.0
- **NNLO pT top softer than NLO/LO**

TopLHCWG

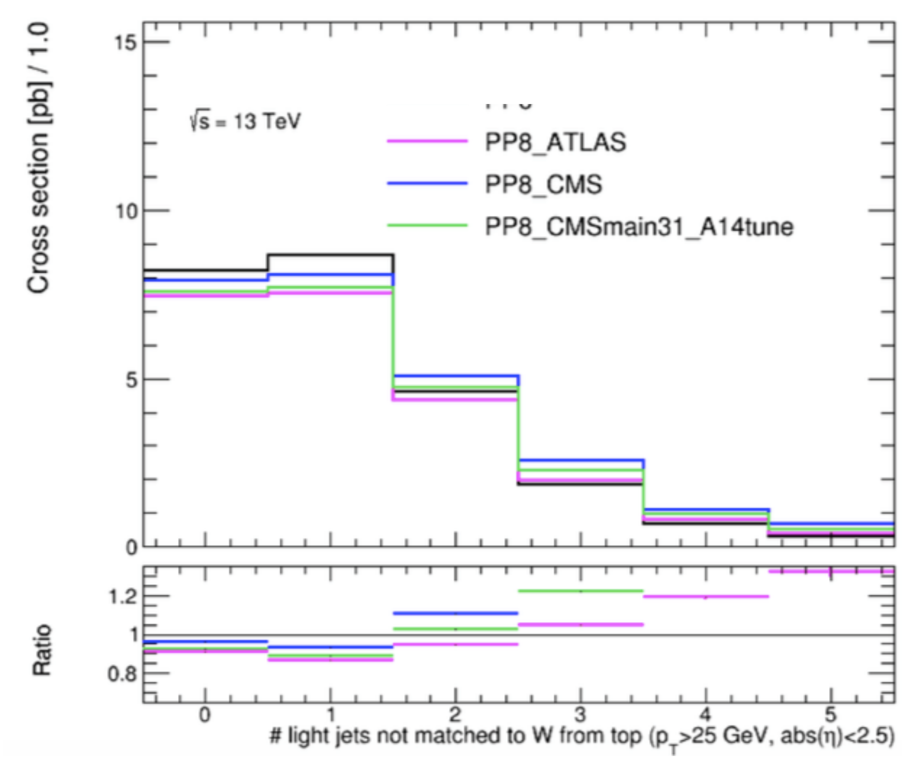


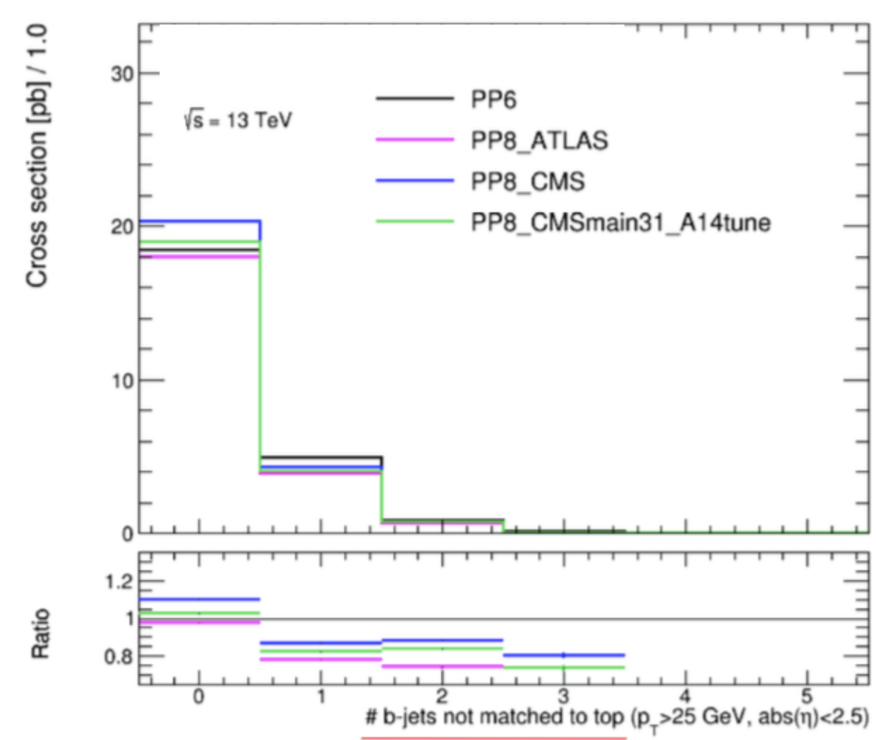
- **NNLO Corrections quite small**
- Small corrections from NLO -> NNLO (larger difference compared to generators)

tt+jets modelling

- **Checks performed with CMS settings**
  - Predicted cross section larger for PP8 than PP6 (even larger for CMS setup)
  - **More extra light jets with CMS PP8 setup than with PP6**
- Checks with alternative tt generator (Sherpa and MG5_aMC@NLO compared to Powheg+Pythia6):
  - Harder jets
  - Larger number of additional light jets (~ 20 %) in SR

- tt+jets studies with Powheg+Pythia8:
  - Variations in ISR or g->bb settings, the yields in SR can increase up to ~20 %





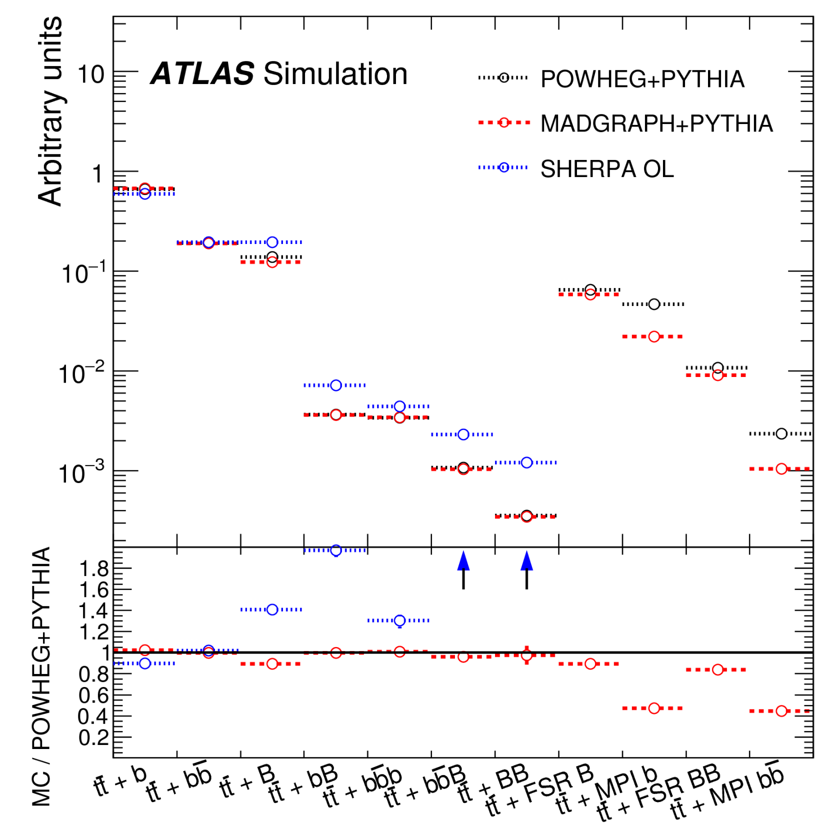
  - CMS main 31 settings: more extra light jets -> up to 20% increase in 6j4b SR
  - With CMS tune: slightly softer jets
  - EvtGen: No differences

tt+>=1b definition in ATLAS (4FS vs 5FS)

- Run-1 definition of tt+bb events in ttH(bb) analysis
  - Truth-level definition based on matching b-hadrons to particle jets

	ATLAS
tt+bb	b-hadron (pT > 5 GeV) not from top or W within DR 0.4 of particle jet (pT >15 GeV) Re-weighted to Sherpa OL calculation
tt+cc	c-hadron not from top/W within DR 0.4 of particle jet (pT >15 GeV)

- **Re-weighting of Powheg+Pythia6 (5FS) to Sherpa+OpenLoops (4FS)**
  - **Small normalization correction to most HF categories**
    - Largest in unresolved gluon spitting to bb
  - 2D re-weighting:
    - **top/ttbar pT (SherpaOL)**
  - 2D re-weighting:
    - **ΔR(bb)/pT(bb)**
    - **OR: q1 pT/η (for tt+b/B)**

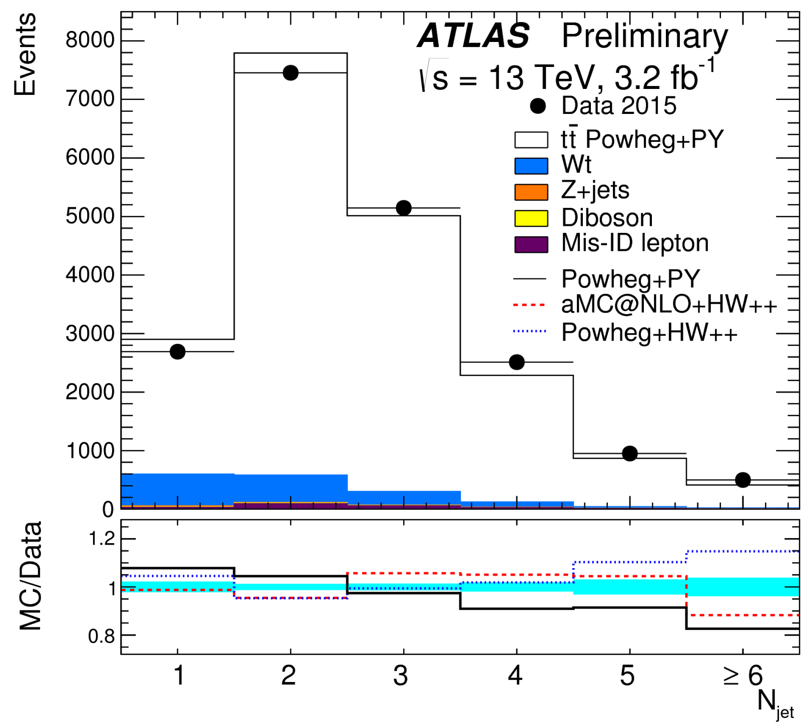


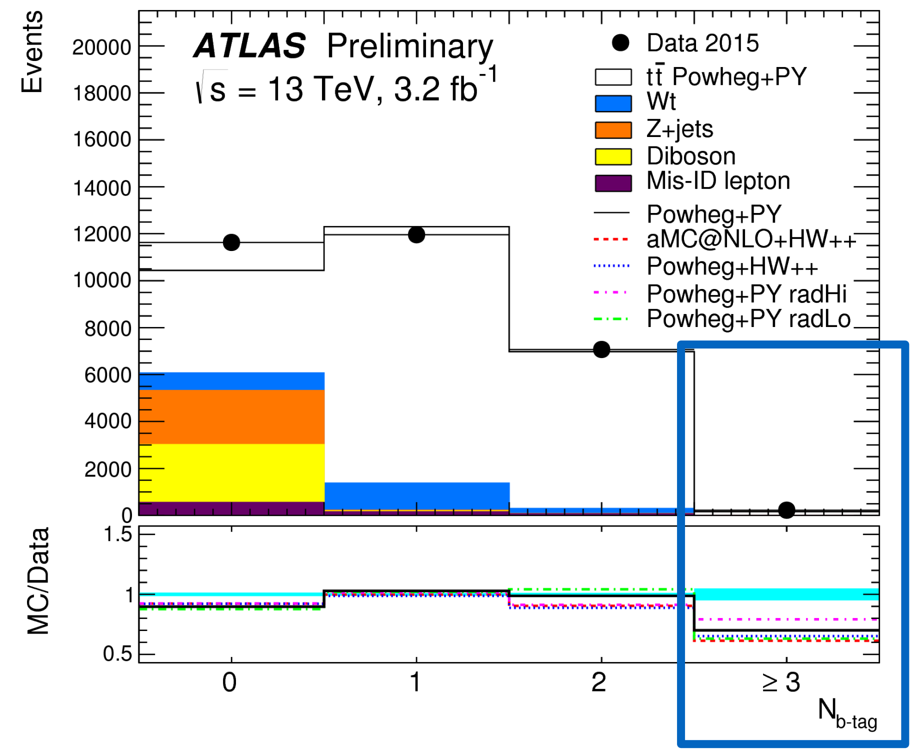
ATLAS: tt+bb in 2015

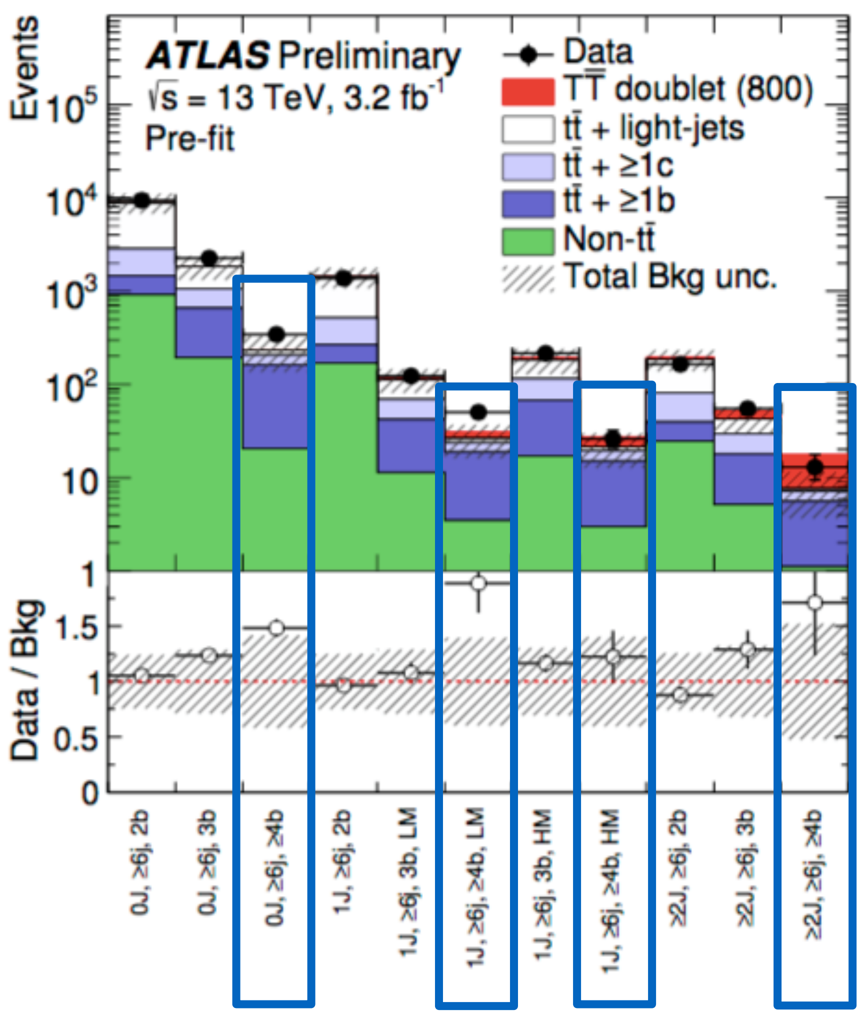
- Excess in tt+bb dominated regions seen in analyses presented at Moriond EW

ATLAS-CONF-2016-005

- **(1) Dilepton tt cross section**
  - lepton pTs > 25 GeV
  - jet pT > 25 GeV, b-tag 70 % WP







- **(2) VLQ Search from ATLAS**
  - Single lepton channel with b-tag 77 % WP
  - Similar final states at ttH(H->bb) separated into three categories
    - Large-R jet tagging and mbb min DR at 100 GeV
  - **See excess of ~ 1.7 times tt+bb**
    - Allowed by decrease in tt+cc from fit (~ 30 %)

ATLAS-CONF-2016-013

tt+b-jets modelling

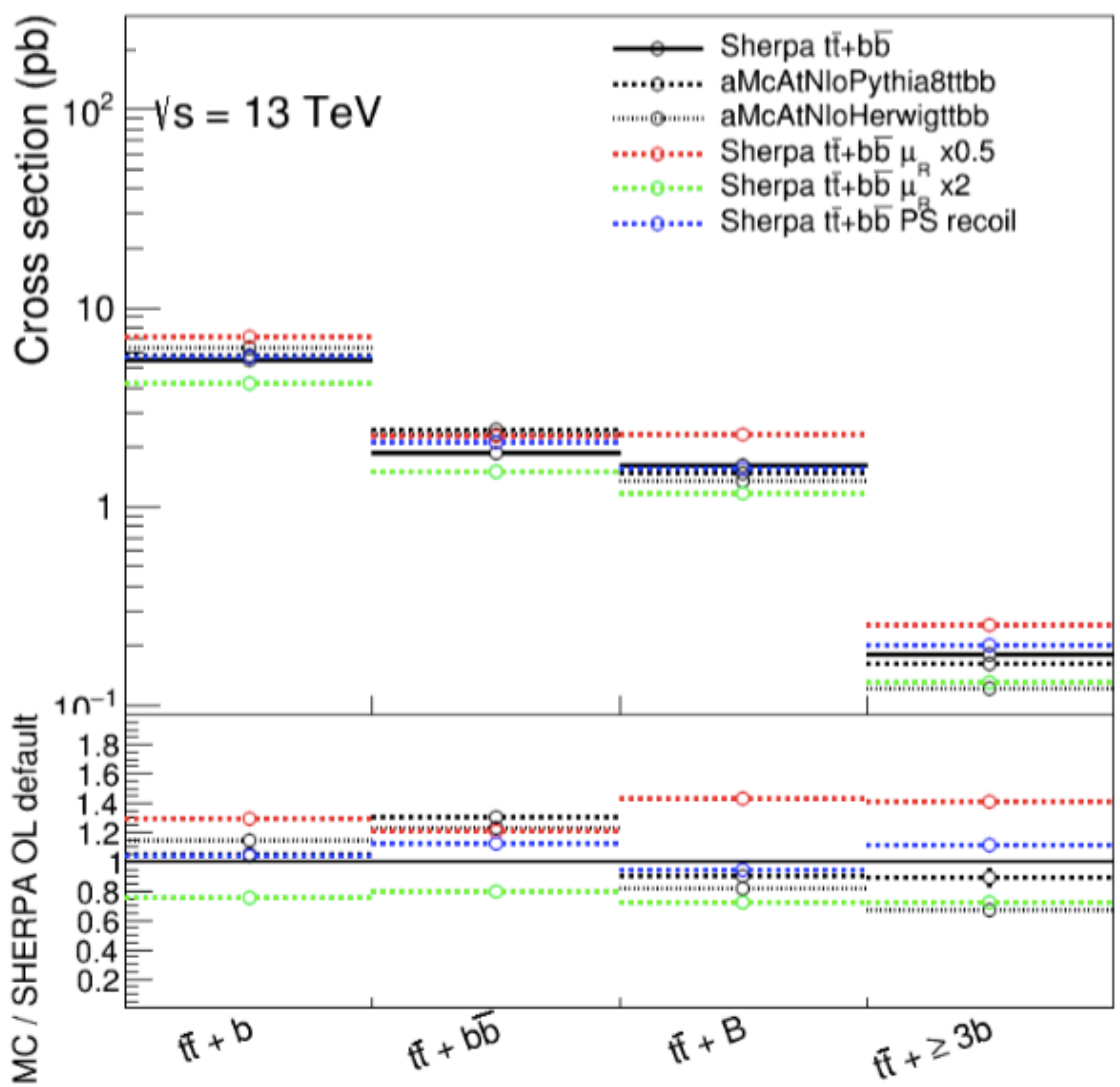
- Inclusive XS comparison for leptonic ttbar decay

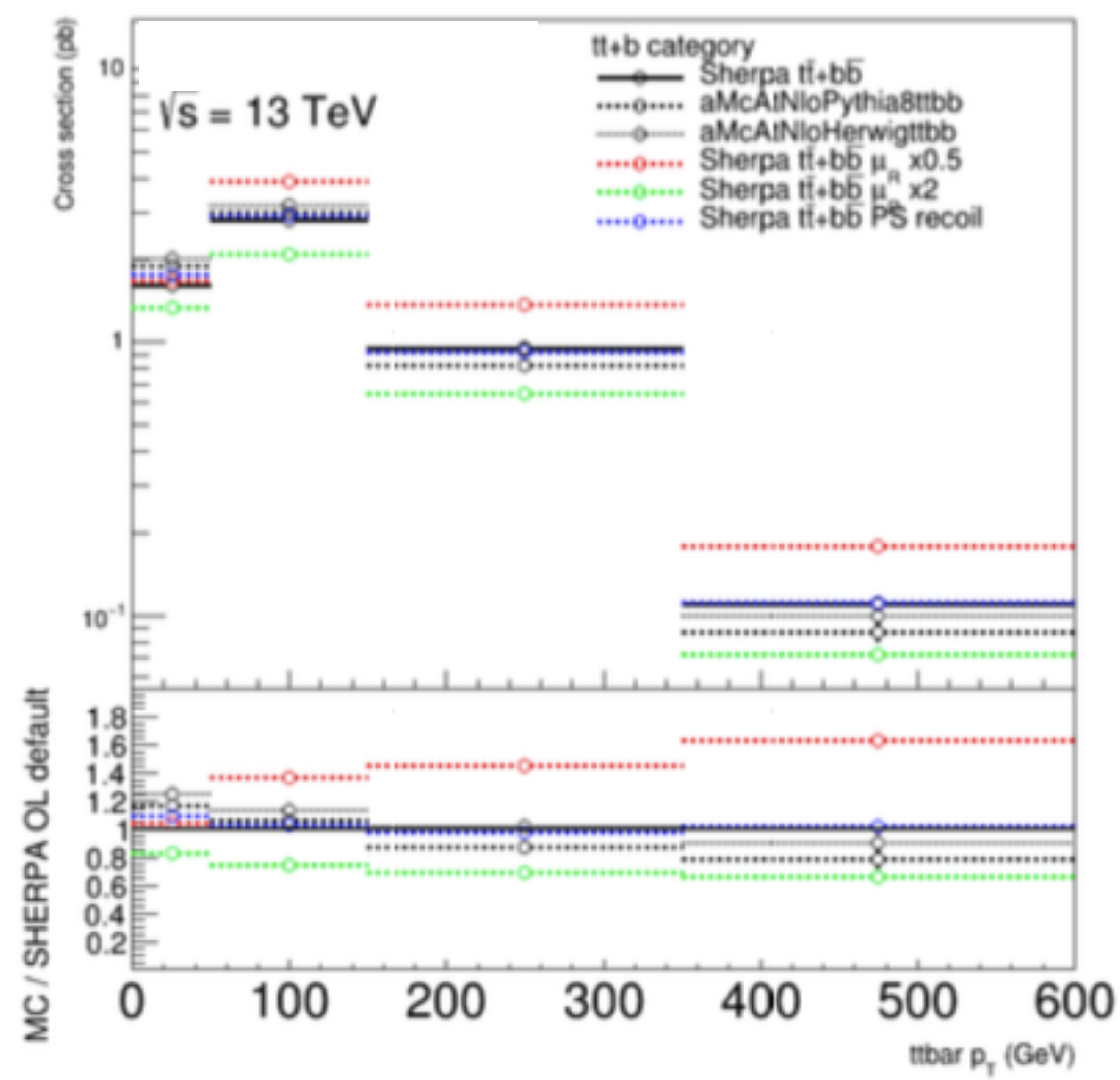
Sample	tt+light [pb]	tt+bb [pb]	tt+cc [pb]
Powheg+Pythia6	407,3	11,6	32,7
Powheg+Pythia8 ATLAS	402,7	11,15	36,7
Powheg+Pythia8 CMS	392,6	13,2	45,8
Sherpa ttbb		9,2	
aMcAtNlo+Pythia8 ttbb		10	

- MG5_aMC@NLO has higher tt+≥1b cross section than Sherpa
- **Still lower than 5F scheme ttbar inclusive samples.**

tt+b-jets modelling (kinematics)

- **Comparisons of tt+bb kinematics between Sherpa+OL and aMC@NLO_MG5 (following YR4)**
  - MG5_aMC@NLO: tt+bb XS higher
  - Generator differences within the scale variations on Sherpa+OL



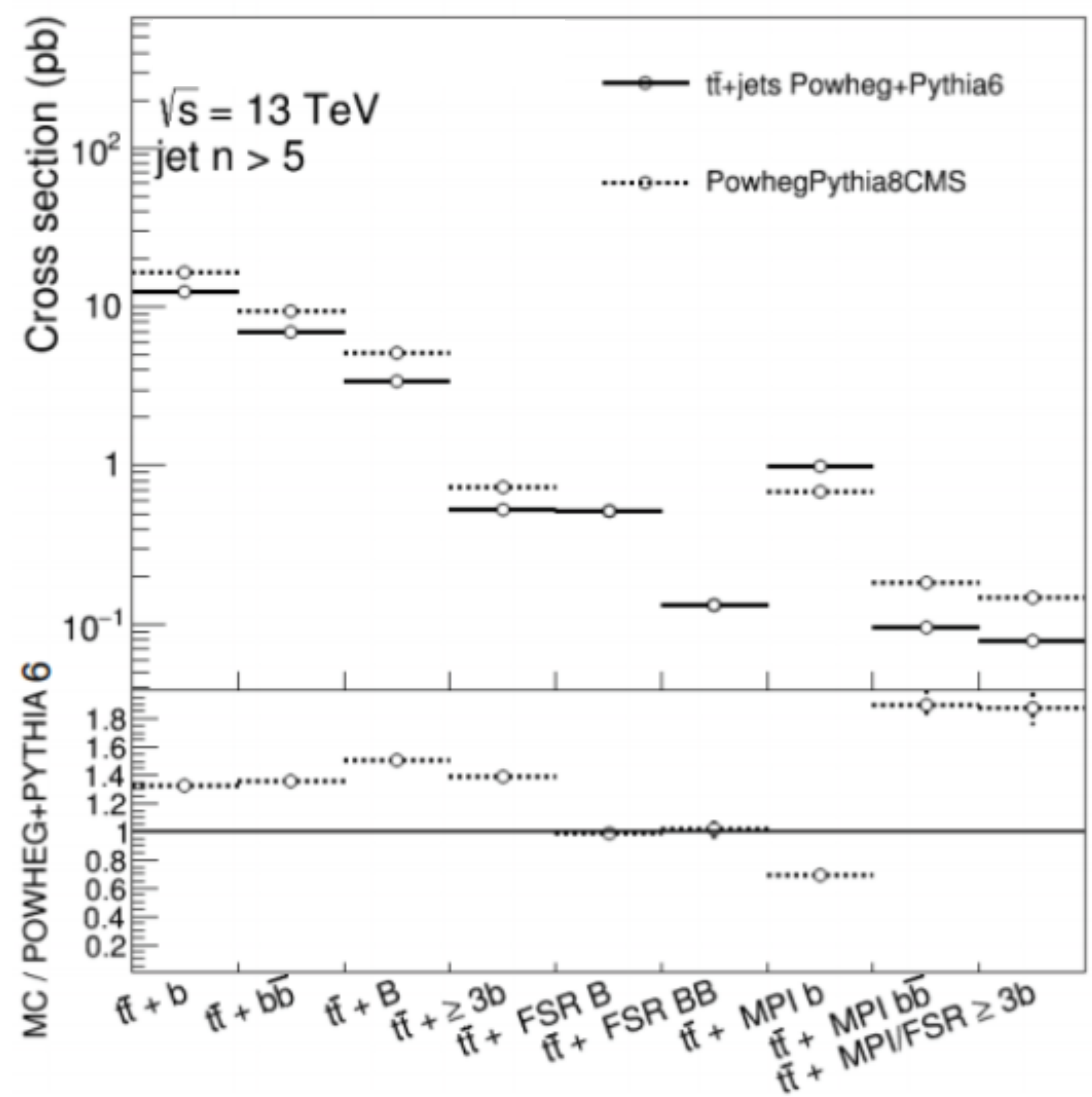


Comparison of ATLAS / CMS tt+bb

- Comparison of CMS settings (received from our CMS colleagues)
- Truth-level samples produced with CMS settings:
  - Main31+tune, Main31-only, tune-only

Sample	tt+light [pb]	tt+bb [pb]	tt+cc [pb]
Powheg+Pythia6	44.8	3.7	8.9
Powheg+Pythia8 CMS	50.0	4.8	12.6
PP8 CMS / PP6	1.1	1.3	1.4

also higher tt̄+≥1c!

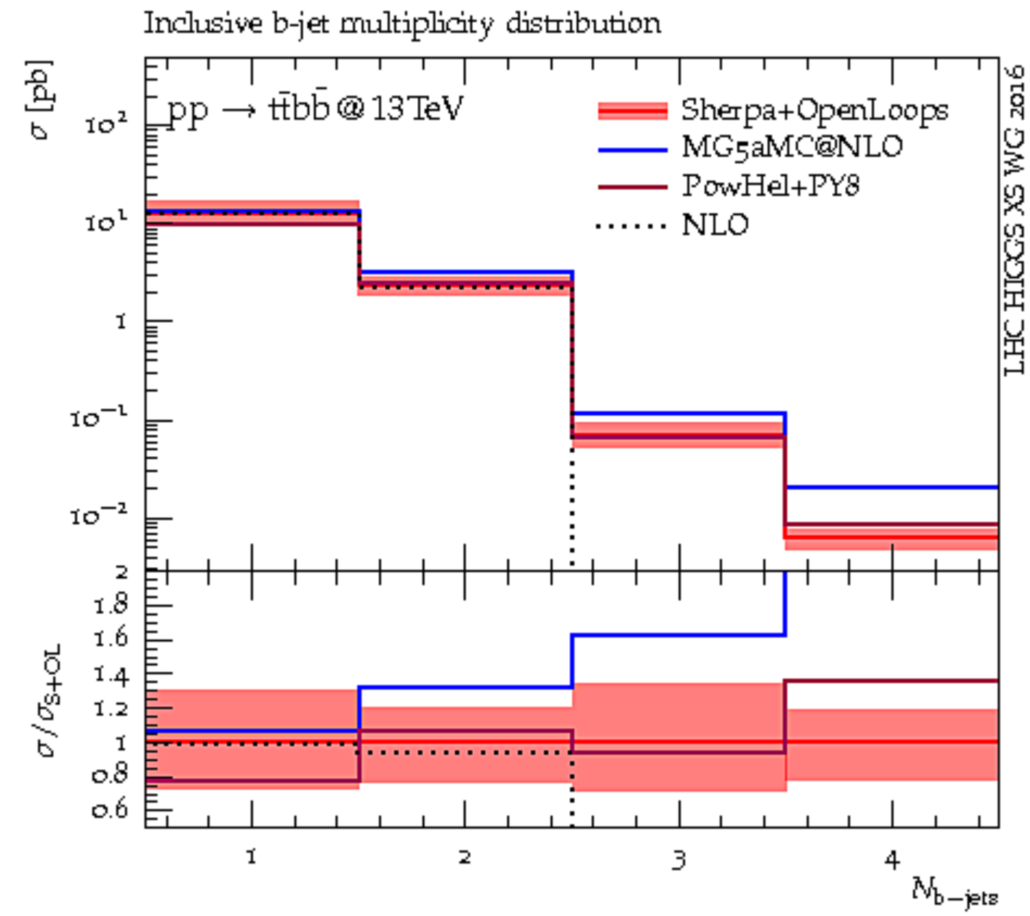


- **Large increase in tt+bb and tt+cc XS in CMS MC settings compared to ATLAS**
  - Still not enough to account for tt+bb increase seen in VLQ analysis

tt+b-jets YR4 studies

- **Alternative tt+bb NLO generator**
  - MG5_aMC@NLO tt+bb events generated, with same settings as in Yellow Report 4 (4F scheme with massive b, as SherpaOL)

- Truth level comparison shows:
  - ~ same inclusive tt+bb cross-section as SherpaOL
  - Larger XS in nb >= 2 regions
  - Systematics from scale variations (mainly renormalisation, on aMC@NLO+Py8) large: +50% / -30%



Inclusive b-jets not from top quarks
parton level pT > 25 GeV, Iηl<2.5

Selection	Tool	σ_{NLO} [fb]	σ_{NLO+PS} [fb]	$\sigma_{NLO+PS}/\sigma_{NLO}$
$n_b \geq 1$	SHERPA+OPENLOOPS	$12820^{+35\%}_{-28\%}$	$12939^{+30\%}_{-27\%}$	1.01
	MADGRAPH5_AMC@NLO		$13833^{+37\%}_{-29\%}$	1.08
	POWHEL		$10073^{+45\%}_{-29\%}$	0.79
$n_b \geq 2$	SHERPA+OPENLOOPS	$2268^{+30\%}_{-27\%}$	$2413^{+21\%}_{-24\%}$	1.06
	MADGRAPH5_AMC@NLO		$3192^{+38\%}_{-29\%}$	1.41
	POWHEL		$2570^{+35\%}_{-28\%}$	1.13

parton shower	on
hadronisation	off
UE	off
top decays	off

- **Need to be careful of fiducial phase-space in experimental searches**
  - Requirement of 4b-tags

tt+bb Generator Comparison in 4j4b

* b-jets are defined at particle level with p_T>25 GeV and h<2.5

* "add. b-jets" are defined as jets that are not matched (dR<0.3) to b-quarks from top quark decays

Selection	Tool	$\sigma_{\mathrm{NLO+PS}}$ [fb]
$n_b \geq 1$	SHERPA+OPENLOOPS	$12939^{+30\%}_{-27\%}$
	MADGRAPH5_AMC@NLO	$13833^{+37\%}_{-29\%}$
	POWHEL	$10073^{+45\%}_{-29\%}$
$n_b \geq 2$	SHERPA+OPENLOOPS	$2413^{+21\%}_{-24\%}$
	MADGRAPH5_AMC@NLO	$3192^{+38\%}_{-29\%}$
	POWHEL	$2570^{+35\%}_{-28\%}$

tt+bb, 4F, NLO+PS	4JI4BI
>=2 add b-jets	**Sherpa+OL** σxBR(tt→μμ)=0.0087 pb Inclusive σ(tt+bb)~0.73 pb **MG5_aMC@NLO+Pythia8** σxBR(tt→μμ)=0.0152 pb Inclusive σ(tt+bb)~1.27 pb **MG5_aMC@NLO+Herwig++** σxBR(tt→μμ)=0.0160 pb Inclusive σ(tt+bb)~1.33 pb

In 4JI4BI, tt+bb cross-section is ~1.7 larger in MG5_aMC@NLO samples.

- **Difference in MG5_aMC@NLO and Sherpa+OL is increased when requiring 4j4b in the dilepton channel**

Questions / Comments

- Summary of ATLAS Preliminary results from Moriond EW
  - Excess in data seen in higher jet/b-jet multiplicities where ttbar modelled with Powheg+Pythia6

- tt+jets modelling
  - Powheg+Pythia8 predicts more additional jets than Powheg+Pythia6
  - Using CMS settings (main31), even higher additional jets are predicted (~ 20 %)
  - Any recommendations in understanding ATLAS + CMS differences in ttbar modelling?
  - Will also help with modelling of future LHC combination

- tt+bb differences between aMC@NLO_MG5 and Sherpa+OpenLoops
  - **Large difference in predicted tt+bb XS (~ 1.7) in analysis signal region phase-space - *very important to understand***
  - In higher additional b-jet regions: MG5_aMC@NLO predicts more jets (softer)
  - Early Run-2 analyses see larger discrepancies in tt+bb enriched regions than Run-1
    - **With such differences seen in normalization, how can we be certain of shapes?**
    - A fitting bias can be induced with a 50 % prior on tt+HF given discrepancies
  - Run-1 procedure: Re-weight tt+bb to Sherpa+OL kinematics
  - Run-2: now have an alternative generator with kinematic differences to compare to Sherpa+OL

Back-up

ATLAS +CMS: ttH(H->bb) labelling

- **tt background broken down into 5 categories**
  - tt+lf, tt+cc, tt+b, tt+2b(B) and tt+bb
  - Truth-level matching to determine event categorization

	ATLAS	CMS
tt+bb	b-hadron (pT > 5 GeV) not from top or W within DR 0.4 of particle jet (pT >15 GeV) Re-weighted to Sherpa OL calculation	b-hadron not from top or W ghost match particle jet (pT > 20 GeV) No re-weighting applied
tt+cc	c-hadron not from top/W within DR 0.4 of particle jet (pT >15 GeV)	c-hadron not from top/W ghost match particle jet (pT > 20 GeV)

- **tt+b:** 1 jet with 1 b-hadron
- **tt+bb:** 2 jets with each 1 b-hadron
- **tt+B/tt+2b:** 1 jet with 2 b-hadrons
  - ***Theoretically b vs B categories are quite different, and deserve separate treatment!***

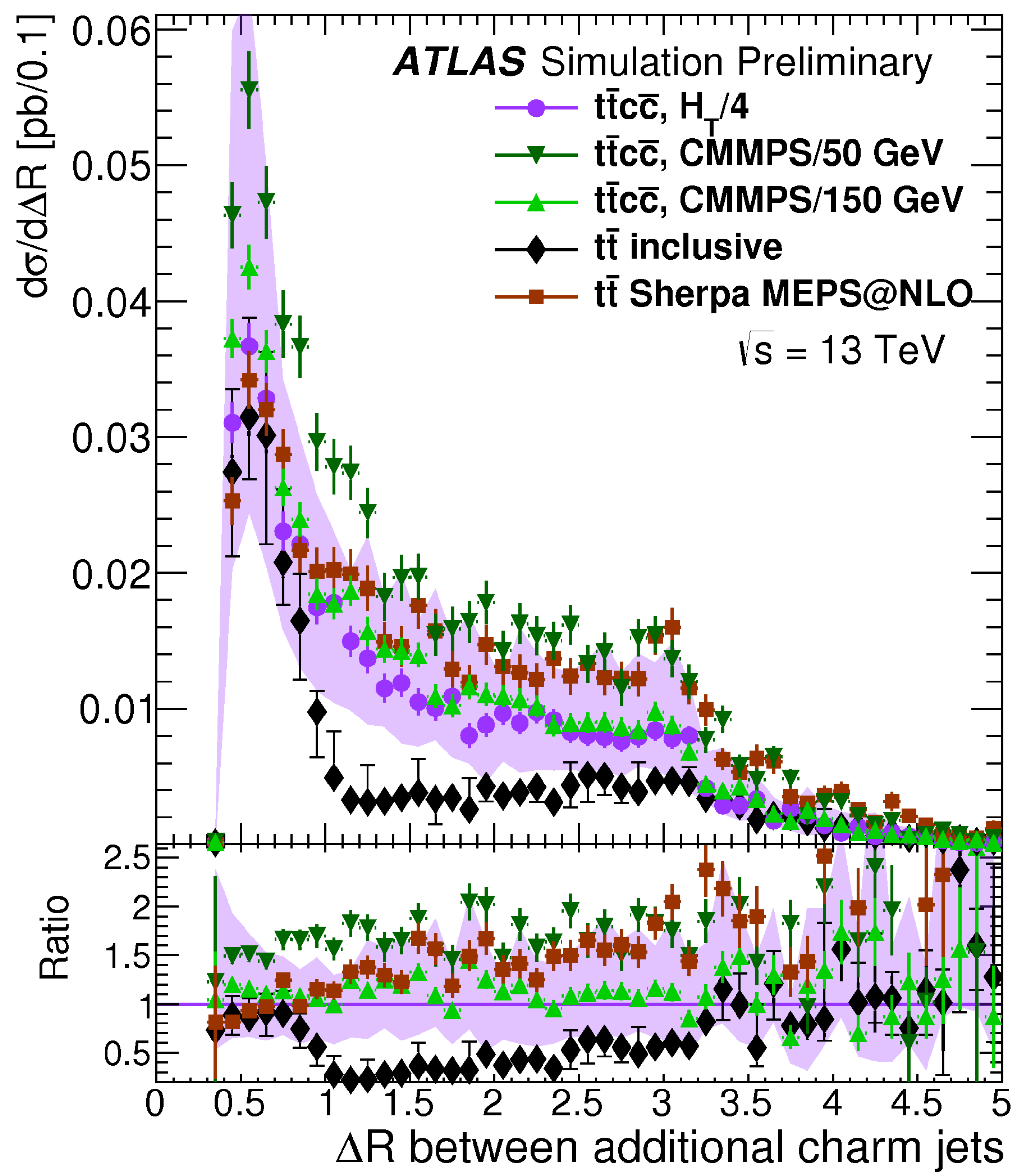
tt+cc modelling

- tt+cc is becoming more important in the signal region

ATL-PHYS-PUB-2016-011

- aMC@NLO tt+cc private samples with different scale choices studied
  - 3FS, massive c-quarks
  - Scales: HT/4, BDDP, and CMMPS
  - Instabilities in event generation

Scale name	Definition	Behaviour for $p_{T,i} \to 0$
$H_T/4$	$\mu = \frac{1}{4}\sum_{i\in FS} E_{T,i}$	$\mu \to \frac{1}{4}\sum_{i\in FS} m_i \approx 87.03$ GeV
BDDP	$\mu = \sqrt{m_t \cdot \sqrt{p_{T,c} \cdot p_{T,\bar{c}}}}$	$\mu \to 0$ GeV
CMMPS	$\mu = \prod_{i=t,\bar{t},c,\bar{c}} E_{T,i}^{1/4}$	$\mu \to \prod_{i=t,\bar{t},c,\bar{c}} m_i^{1/4} \approx 16.35$ GeV



- Studies documented in PUB note for TopLHC WG