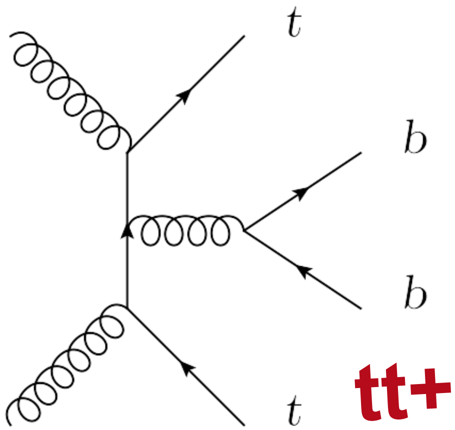
Top quark pair production with additional heavy flavour jets

LHCTopWG Workshop, November 2017

Theory: Stefano Pozzorini (Zurich)
Laura Reina (Florida State)
ATLAS: María Moreno Llácer (CERN)
CMS: Benjamin Stieger (Nebraska Lincoln)

on behalf of the whole LHC Higgs tt+H/t+H XS group



tt+bb

Introduction

Twiki: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWGTTH
Mailing list: lhc-higgs-xsbr-tth
Subconveners: Theory: Stefano Pozzorini (Zurich)
Laura Reina (Florida State)
ATLAS: María Moreno Llácer (CERN)
CMS: Benjamin Stieger (Nebraska Lincoln)

Activities:

- Signal modelling studies: tt+H and t+H
- More emphasis to precision theory predictions/simulations for tt+H/t+H (irreducible) backgrounds
  - for tt+H (H→WW,$\tau\tau$): tt+W/Z/γ^*
  - for tt+H (H→ $\gamma\gamma$): tt+$\gamma\gamma$
  - for tt+H (H→bb): tt+bb
- Address non-trivial aspects of theory uncertainties that play an important role in the experimental analyses (e.g. signal and background shape uncertainties)
- Try to involve more key experts (from both experimental and theory side)
- Foster communication between ATLAS and CMS

Meetings with topical discussions:

Indico: https://indico.cern.ch/category/5847/ (usually Monday's at 17h)
Next meeting: November 6th, topic: tt+b-jet backgrounds to ttH(bb)

Deliverables:

Handbook of LHC Higgs Cross Sections: 4. Deciphering the nature of the Higgs sector, arXiv: 1610.07922
No more CERN Higgs Yellow Report are planned in the near future, but another kind of deliverables (TBD).

Activities within the LHC Higgs *tt+H/tH* XS subgroup

tt+H **modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposaltTH

tt+Z/W **modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposaltTV

tt**+heavy flavour modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposalTtbb

Activities within the LHC Higgs *tt+H/tH* XS subgroup

***tt+H* modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposaltTH

***tt+Z/W* modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposaltTV

***tt*+heavy flavour modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposalTtbb

- Goal: study *tt+bb* production with several MC generators (4F and 5F) and compare with data
- Of great interest for both the Higgs and Top groups
  - Joined experimental discussion (https://indico.cern.ch/event/638184/, May 2017)

Available tt+bb cross-section measurements

ATLAS

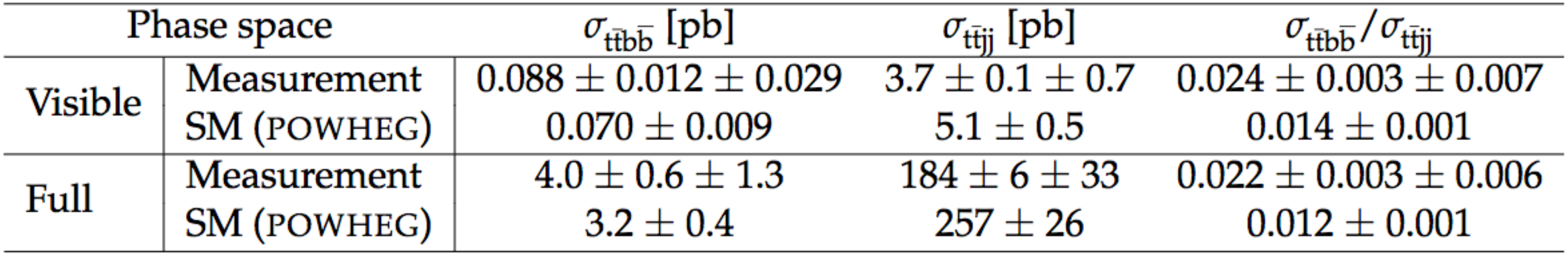
8 TeV: Eur. Phys. J. C76 (2016) 11

Precision of measurements ~30-36%
Precision NLO calculation ~20%

CMS

8 TeV: CMS PAS TOP-13-016, Eur. Phys. J. C76 (2016) 379
13 TeV: arXiv: 1705.10141

13 TeV, 2.3 fb^{-1}



Phase space		$\sigma_{t\bar{t}b\bar{b}}$ [pb]	$\sigma_{t\bar{t}jj}$ [pb]	$\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$
Visible	Measurement	$0.088 \pm 0.012 \pm 0.029$	$3.7 \pm 0.1 \pm 0.7$	$0.024 \pm 0.003 \pm 0.007$
	SM (POWHEG)	0.070 ± 0.009	5.1 ± 0.5	0.014 ± 0.001
Full	Measurement	$4.0 \pm 0.6 \pm 1.3$	$184 \pm 6 \pm 33$	$0.022 \pm 0.003 \pm 0.006$
	SM (POWHEG)	3.2 ± 0.4	257 ± 26	0.012 ± 0.001

tt+bb differential cross-section measurements (CMS)

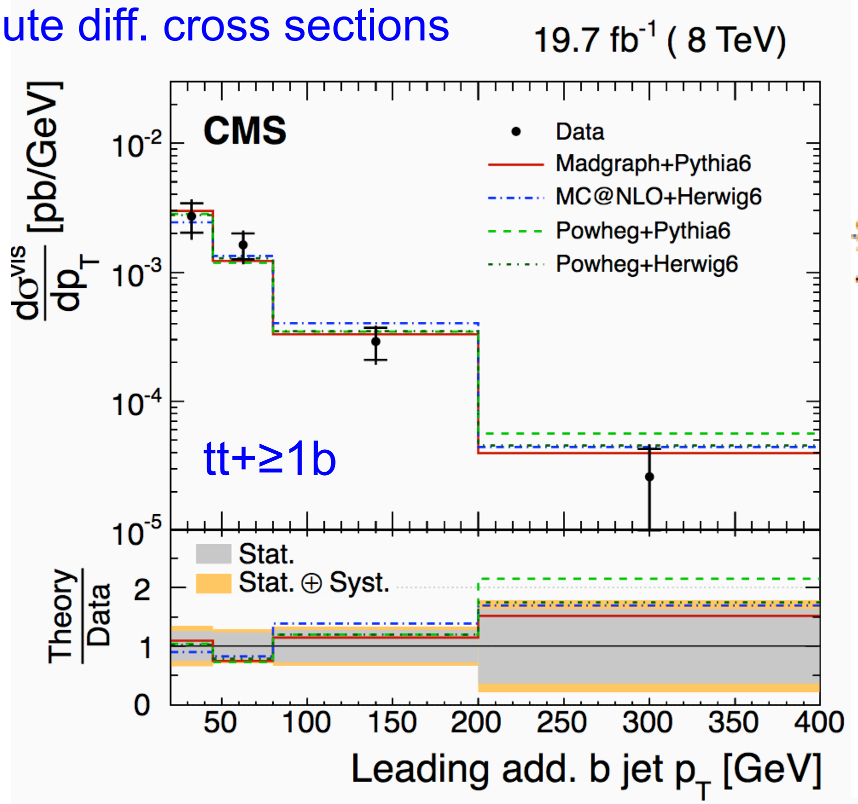
8 TeV: Eur. Phys. J. C76 (2016) 379

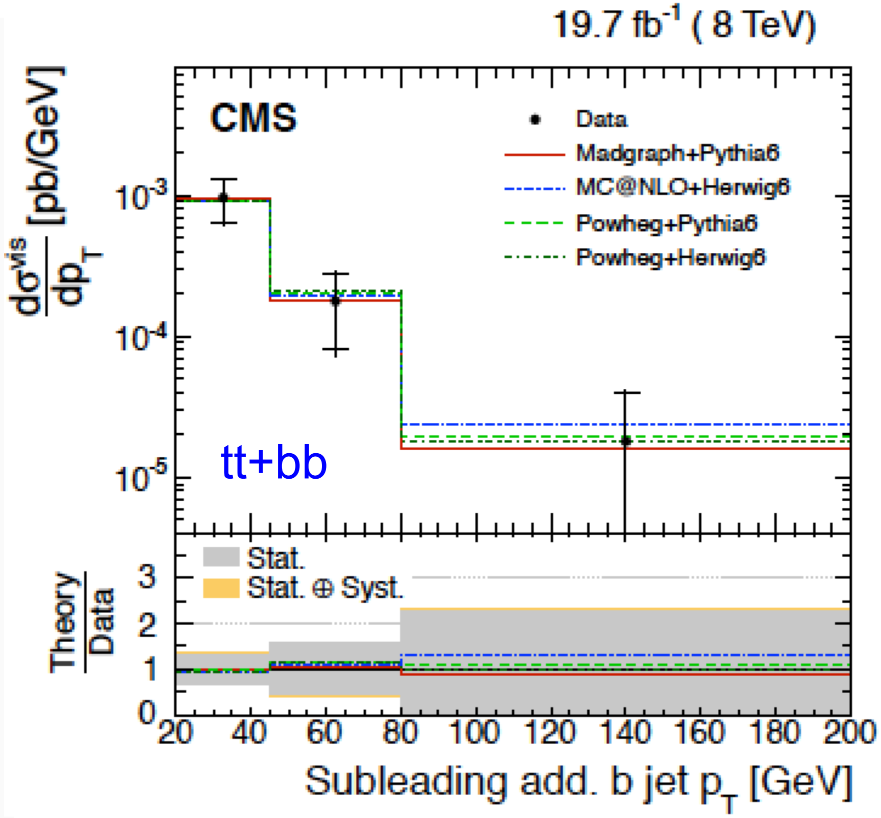
→ *tt+bb* absolute and normalized differential cross-sections measured as a function of the jet multiplicity for different jet transverse momentum thresholds and the kinematic properties of the leading additional jets.

→ First differential *tt+b* and *tt+bb* cross sections as a function of the kinematic properties of the leading additional *b*-jets.

* Data/MC(tt+jets 5F) for *tt+b* ~1.3
* Data/MC for *tt+bb* ~1.8
in agreement with other CMS and ATLAS results. unc. dominated by the stat. unc. (20-100%).

Absolute diff. cross sections





Do we have Rivet routines?

A critical piece in *tt+H* (*H→bb*) searches: *tt*+jets modelling

***ttH* (*H*→bb) signal produces 0-2 leptons and 4-8 jets, 4 of them *b*-jets** → very challenging

Strategy: categorize events according to # jets and *b*-jets → define control and signal regions with different background composition

tt+jets events classified into several categories (tt+light / c / b), and subcategories, based on the flavour of additional jets and number of hadrons in each of them.

Two distributions crucial to model correctly:

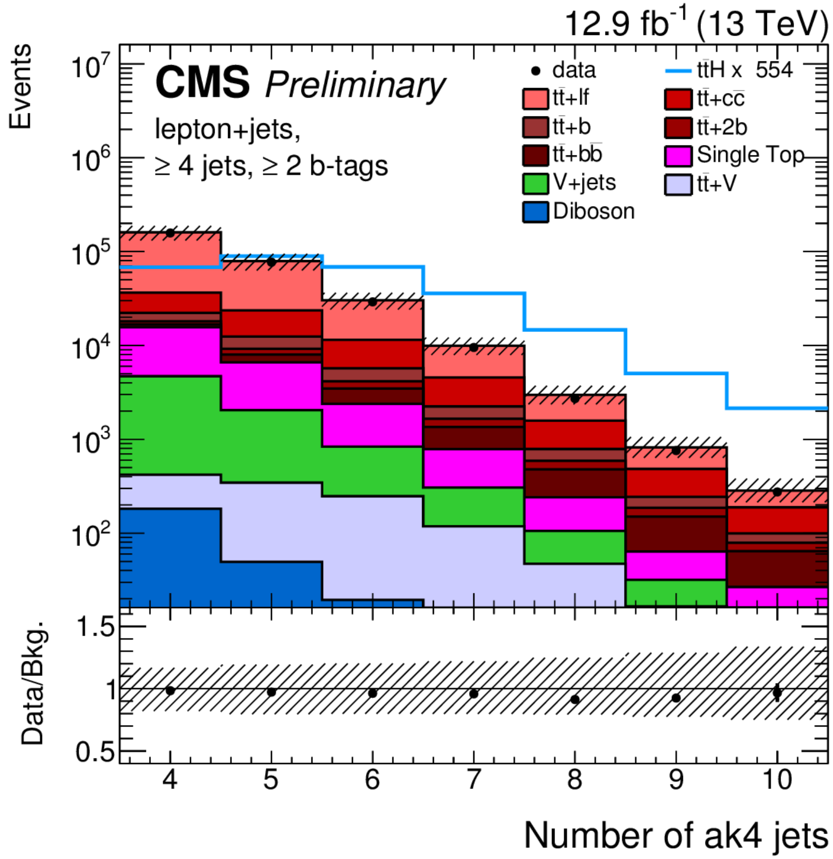
* **ttbar p_T** (mainly affects jet multiplicity): improved thanks to differential measurements with several observables sensitive to different effects (matrix element, radiation, hadronisation)

- → well described with tuned **Powheg+Pythia8** with h_{damp} ~1.5-1.58 m_{top} (nominal)

* **top p_T** (mainly affects jets p_T): largely improved by NNLO computations

→ More details in top modelling and tuning talks

CMS-PAS-HIG-16-038



tt+HF modelling: very challenging task

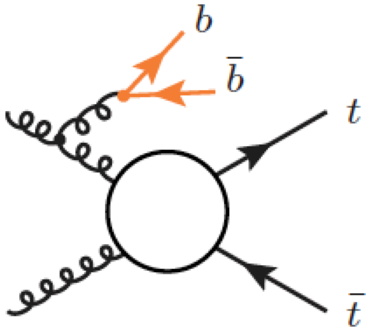
- pure QCD process, very complicated that involves several scales and massive quarks
    - challenging for the MC generator community
    - implementation of latest theoretical developments crucial
- studies ongoing in both experiments in close collaboration with theorists (documented in YR4)
    - 5F *tt*+jets (m_b=0) vs. 4F *tt*+*bb* predictions
    - 4F *tt*+*bb* NLO+PS predictions (with massive *b*-quarks in ME) with novel generators
        - recommended but need further studies (settings, associated shape uncertainties, etc.)
    - how to merge 4F and 5F samples?
    - heavy flavour classification

tt+HF modelling: very challenging task

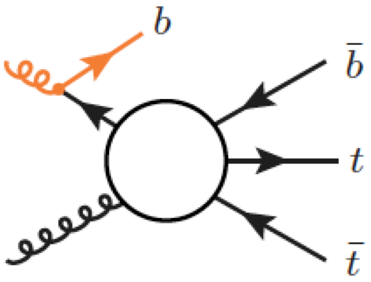
- pure QCD process, very complicated that involves several scales and massive quarks
  - challenging for the MC generator community
  - implementation of latest theoretical developments crucial
- studies ongoing in both experiments in close collaboration with theorists (documented in YR4)
  - 5F *tt*+jets (m_b=0) vs. 4F *tt*+*bb* predictions
  - 4F *tt*+*bb* NLO+PS predictions (with massive *b*-quarks in ME) with novel generators
    - recommended but need further studies (settings, associated shape uncertainties, etc.)
  - how to merge 4F and 5F samples?
  - heavy flavour classification

tt+jets 5F NLO+PS or NLOmultileg

$t\bar{t}b\bar{b}$ described through $t\bar{t}j$ tree MEs plus $g \to b\bar{b}$ shower splittings



$gg \to t\bar{t}g$ + FS splittings



$\bar{b}g \to t\bar{t}\bar{b}$ + IS splittings

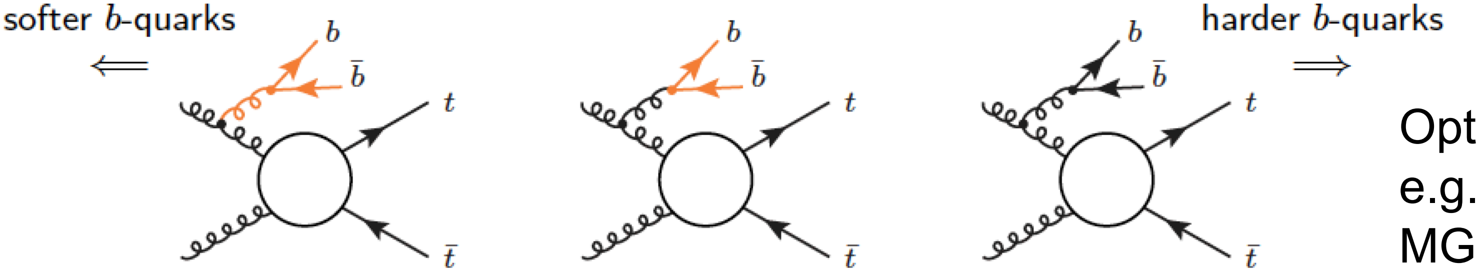
Option 1: NLO+PS 5F tt+jets (m_b=0)
e.g. Powheg+Pythia8 used as
nominal sample in experiments

tt+jets ME cannot describe collinear *g*→*bb* splittings

Precision vs accuracy

- precision lower than LO (parton shower allows for accurate tuning to data)

$t\bar{t}b\bar{b}$ described through $t\bar{t} + 0, 1, 2$ jet MEs and $g \to b\bar{b}$ shower splittings



$t\bar{t}$ + PS

$t\bar{t}g$ + PS

$t\bar{t}b\bar{b}$

Option 2: NLOmultileg
e.g. Sherpa+OpenLoops or
MG5_aMCNLO+Py8/HW7 FxFx

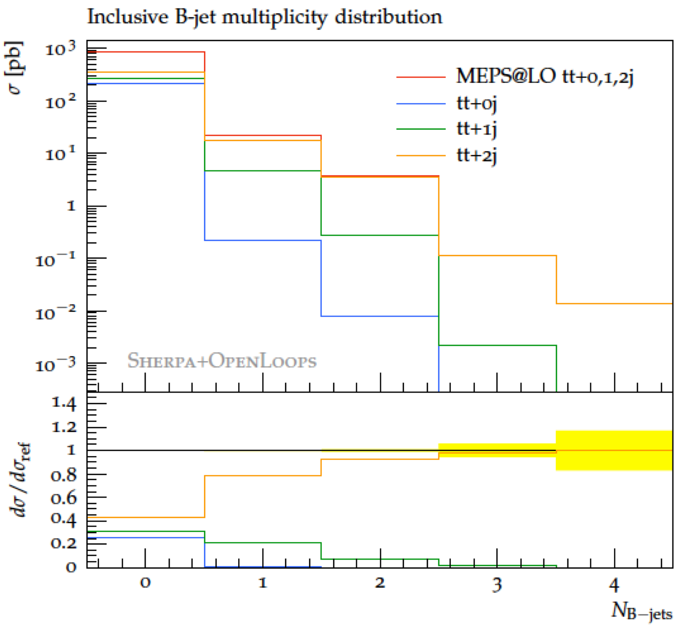
CPU consuming,
challenging for *tt*+HF,
still based on m_b=0 MEs
+ shower in collinear regions

Precision and CPU cost strongly depend on choice of merging cut Q_{cut}

- separates ME regions ($k_T > Q_{\text{cut}}$) from shower regions ($k_T < Q_{\text{cut}}$)

Does this describe $t\bar{t}$+b-jet production mostly through $t\bar{t}b\bar{b}$ MEs?

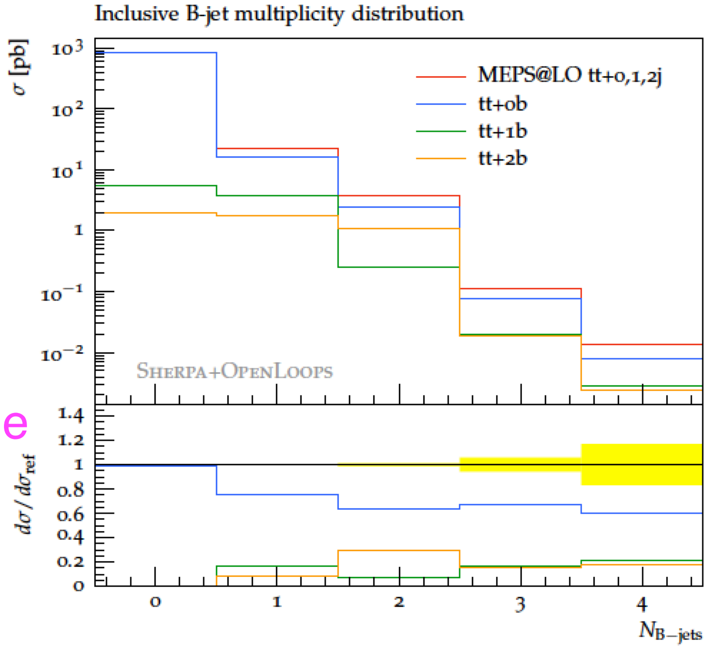
tt+jets 5F NLOmultileg



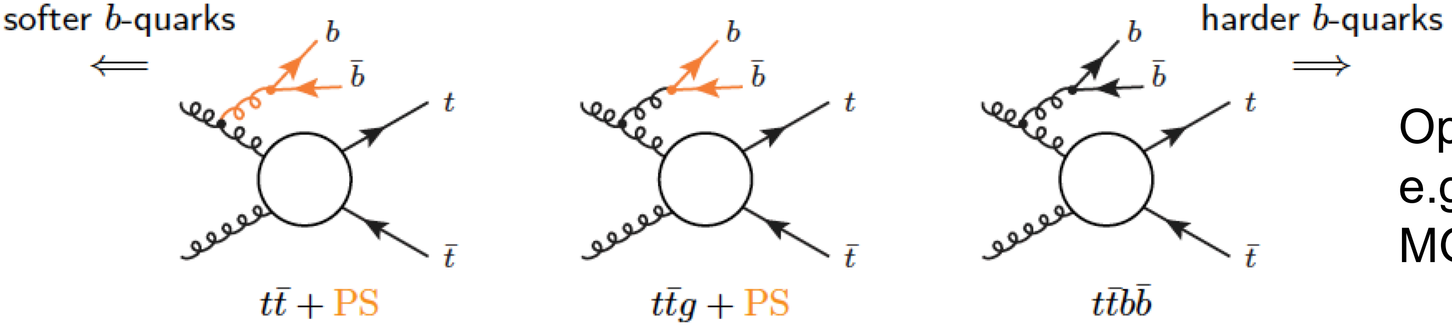
dominated by
MEs with 2 extra jets

but 2 extra light jets

→ direct description in terms of **ttbb** MEs seems preferable



$t\bar{t}b\bar{b}$ described through $t\bar{t}+0,1,2$ jet MEs and $g \to b\bar{b}$ shower splittings



Option 2: NLOmultileg
e.g. Sherpa+OpenLoops or
MG5_aMCNLO+Py8/HW7 FxFx

CPU consuming,
challenging for *tt*+HF,
still based on m_b=0 MEs
+ shower in collinear regions

Precision and CPU cost strongly depend on choice of merging cut $Q_{\rm cut}$

- separates ME regions ($k_T > Q_{\rm cut}$) from shower regions ($k_T < Q_{\rm cut}$)

Does this describe $t\bar{t}$+b-jet production mostly through $t\bar{t}b\bar{b}$ MEs?

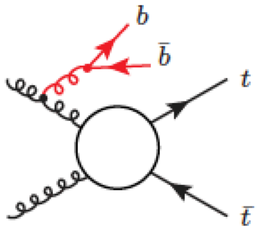
tt+bb 4F NLO+PS

NLO precision for *tt* + 2 *b*-jets and 1 *b*-jet! [Cascioli et al '13]
(80% LO uncertainty reduced to 20-30% at NLO)
can be applied to full phase space (no generation cuts)

Pre-YR4

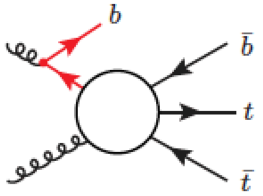
Dominant topologies in 4F $t\bar{t}b\bar{b}$ (FS vs IS $g \to b\bar{b}$)

$t\bar{t}b\bar{b}$ topologies with FS $g \to b\bar{b}$ splittings



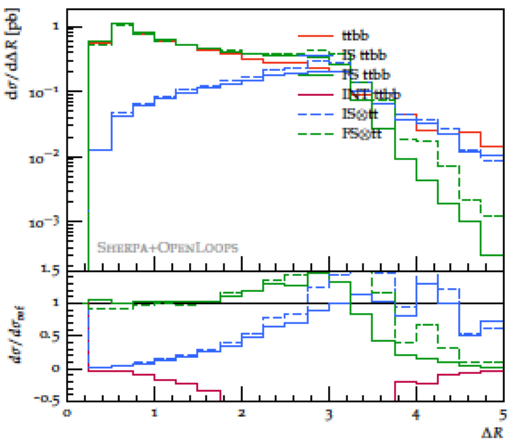
- dominant in full ttbb and ttb phase space
- notion of $g \to b\bar{b}$ splittings and IS/FS separation seems ill defined at large ΔR_{bb}, m_{bb}, $p_{T,b}$ due to sizable interferences

$t\bar{t}b\bar{b}$ topologies with IS $g \to b\bar{b}$ splittings

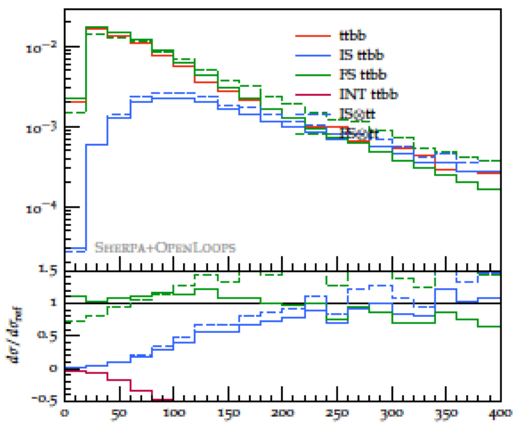


- mostly clearly subdominant (no need for 5F scheme resummation)

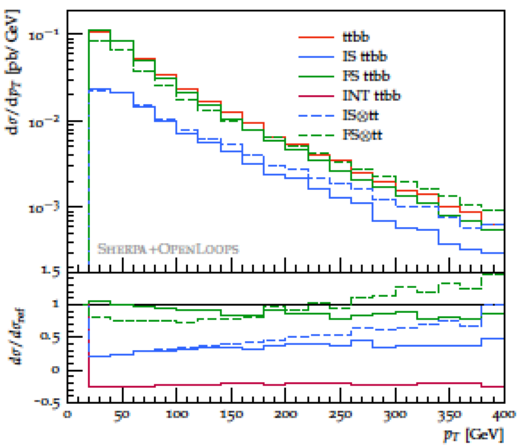
ΔR_{bb} with ttbb cut



m_{bb} with ttb cuts



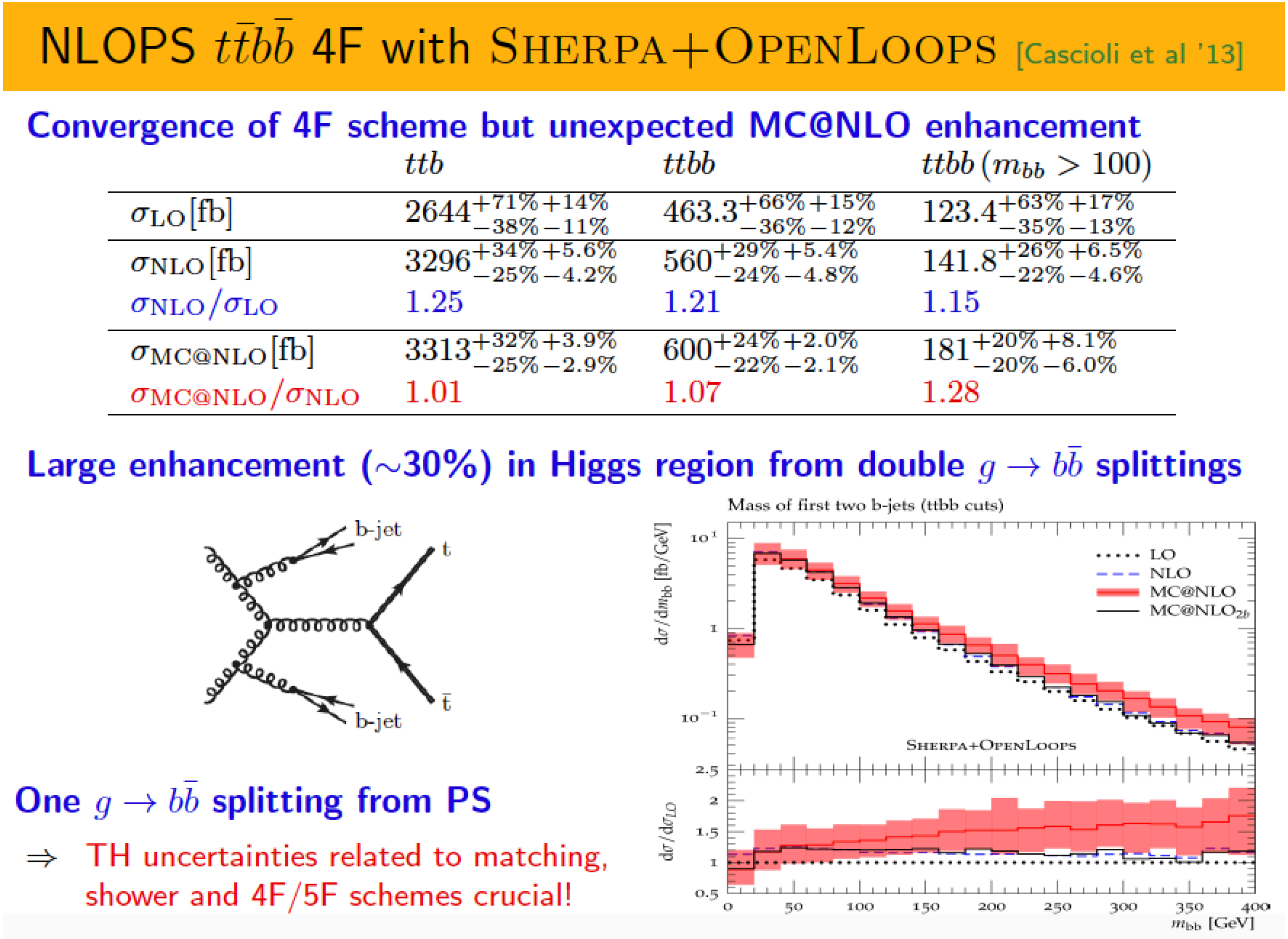
p_{T,b_1} with ttb cuts



supports choice of 4F scheme with $m_b > 0$ and no b-quark PDF

tt+bb 4F NLO+PS

Pre-YR4



NLOPS $t\bar{t}b\bar{b}$ 4F with SHERPA+OPENLOOPS [Cascioli et al '13]

Convergence of 4F scheme but unexpected MC@NLO enhancement

	ttb	$ttbb$	$ttbb\,(m_{bb} > 100)$
$\sigma_{\rm LO}$[fb]	$2644^{+71\%+14\%}_{-38\%-11\%}$	$463.3^{+66\%+15\%}_{-36\%-12\%}$	$123.4^{+63\%+17\%}_{-35\%-13\%}$
$\sigma_{\rm NLO}$[fb]	$3296^{+34\%+5.6\%}_{-25\%-4.2\%}$	$560^{+29\%+5.4\%}_{-24\%-4.8\%}$	$141.8^{+26\%+6.5\%}_{-22\%-4.6\%}$
$\sigma_{\rm NLO}/\sigma_{\rm LO}$	1.25	1.21	1.15
$\sigma_{\rm MC@NLO}$[fb]	$3313^{+32\%+3.9\%}_{-25\%-2.9\%}$	$600^{+24\%+2.0\%}_{-22\%-2.1\%}$	$181^{+20\%+8.1\%}_{-20\%-6.0\%}$
$\sigma_{\rm MC@NLO}/\sigma_{\rm NLO}$	1.01	1.07	1.28

Large enhancement (~30%) in Higgs region from double $g \to b\bar{b}$ splittings

One $g \to b\bar{b}$ splitting from PS

⇒ TH uncertainties related to matching, shower and 4F/5F schemes crucial!

PS is very important: large enhancement in Higgs-boson region due to *g→bb* from PS

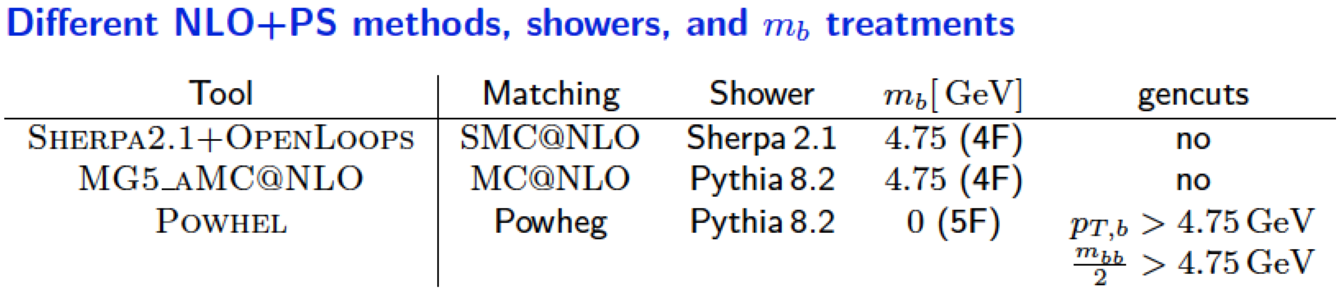
→ Theoretical uncertainties related to matching & shower crucial!

tt+HF modelling: very challenging task

- pure QCD process, very complicated that involves several scales and massive quarks
    - challenging for the MC generator community
    - implementation of latest theoretical developments crucial
- studies ongoing in both experiments in close collaboration with theorists (documented in YR4)
    - 5F *tt*+jets (m_b=0) vs. 4F *tt*+*bb* predictions
    - 4F *tt*+*bb* NLO+PS predictions (with massive *b*-quarks in ME) with novel generators
        - recommended but need further studies (settings, associated shape uncertainties, etc.)
    - how to merge 4F and 5F samples?
    - heavy flavour classification

Comparisons of different MC predictions (YR4 exercise)

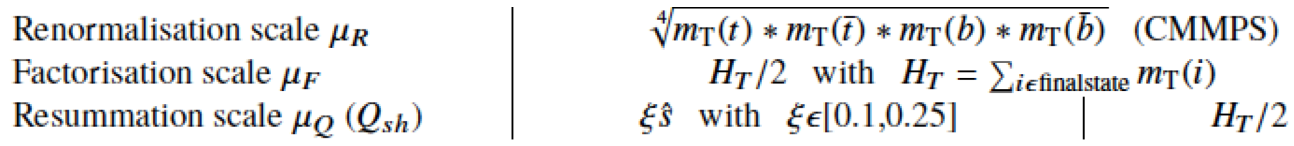
tt+bb 4F NLO+PS predictions (Sherpa+OpenLoops, MG5_aMC@NLO+Py8 and Powhel+Py8) with theory motivated shower settings for consistent comparisons:



Different NLO+PS methods, showers, and m_b treatments

Tool	Matching	Shower	m_b[GeV]	gencuts
SHERPA2.1+OPENLOOPS	SMC@NLO	Sherpa 2.1	4.75 (4F)	no
MG5_AMC@NLO	MC@NLO	Pythia 8.2	4.75 (4F)	no
POWHEL	Powheg	Pythia 8.2	0 (5F)	$p_{T,b} > 4.75\,\text{GeV}$ $\frac{m_{bb}}{2} > 4.75\,\text{GeV}$

using 4F scheme



Renormalisation scale μ_R	$\sqrt[4]{m_T(t)*m_T(\bar{t})*m_T(b)*m_T(\bar{b})}$ (CMMPS)	
Factorisation scale μ_F	$H_T/2$ with $H_T = \sum_{i\epsilon\text{finalstate}} m_T(i)$	
Resummation scale μ_Q (Q_{sh})	$\xi\hat{s}$ with $\xi\epsilon[0.1,0.25]$	$H_T/2$

for MG5_aMC@NLO Sherpa

PDF set: NNPDF3.0 4F
Top quarks are not decayed,
hadronisation and UE are swtiched off

More details in: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposalTtbb

Comparisons of different MC predictions (YR4 exercise)

tt+bb 4F NLO+PS predictions (Sherpa+OpenLoops, MG5_aMC@NLO+Py8 and Powhel+Py8) with theory motivated shower settings for consistent comparisons:

Different NLO+PS methods, showers, and m_b treatments

Tool	Matching	Shower	m_b[GeV]	gencuts
SHERPA2.1+OPENLOOPS	SMC@NLO	Sherpa 2.1	4.75 (4F)	no
MG5_AMC@NLO	MC@NLO	Pythia 8.2	4.75 (4F)	no
POWHEL	Powheg	Pythia 8.2	0 (5F)	$p_{T,b} > 4.75$ GeV $\frac{m_{bb}}{2} > 4.75$ GeV

using 4F scheme

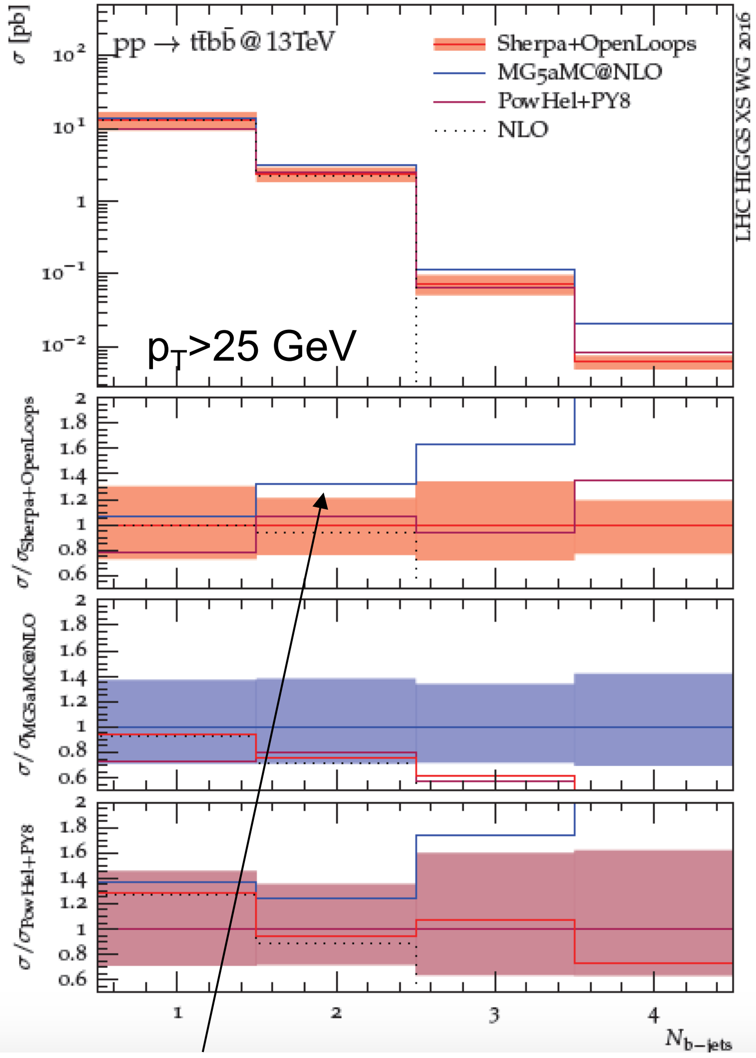
Renormalisation scale μ_R	$\sqrt[4]{m_T(t) * m_T(\bar{t}) * m_T(b) * m_T(\bar{b})}$ (CMMPS)	
Factorisation scale μ_F	$H_T/2$ with $H_T = \sum_{i \epsilon \mathrm{finalstate}} m_T(i)$	
Resummation scale μ_Q (Q_{sh})	$\xi\hat{s}$ with $\xi\epsilon[0.1,0.25]$	$H_T/2$
	for MG5_aMC@NLO	Sherpa

PDF set: NNPDF3.0 4F
Top quarks are not decayed,
hadronisation and UE are swtiched off

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

Selection	Tool	σ_{NLO} [fb]	$\sigma_{\mathrm{NLO+PS}}$ [fb]	$\sigma_{\mathrm{NLO+PS}}/\sigma_{\mathrm{NLO}}$
$n_b \geq 1$	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

additional b-jets (inclusive)



Differences of ≥40% for tt+≥2b

Comparisons of different MC predictions (YR4 exercise)

tt+bb 4F NLO+PS predictions (Sherpa+OpenLoops, MG5_aMC@NLO+Py8 and Powhel+Py8) with theory motivated shower settings for consistent comparisons:

Different NLO+PS methods, showers, and m_b treatments

Tool	Matching	Shower	m_b[GeV]	gencuts
SHERPA2.1+OPENLOOPS	SMC@NLO	Sherpa 2.1	4.75 (4F)	no
MG5_AMC@NLO	MC@NLO	Pythia 8.2	4.75 (4F)	no
POWHEL	Powheg	Pythia 8.2	0 (5F)	$p_{T,b} > 4.75\,\mathrm{GeV}$ $\frac{m_{bb}}{2} > 4.75\,\mathrm{GeV}$

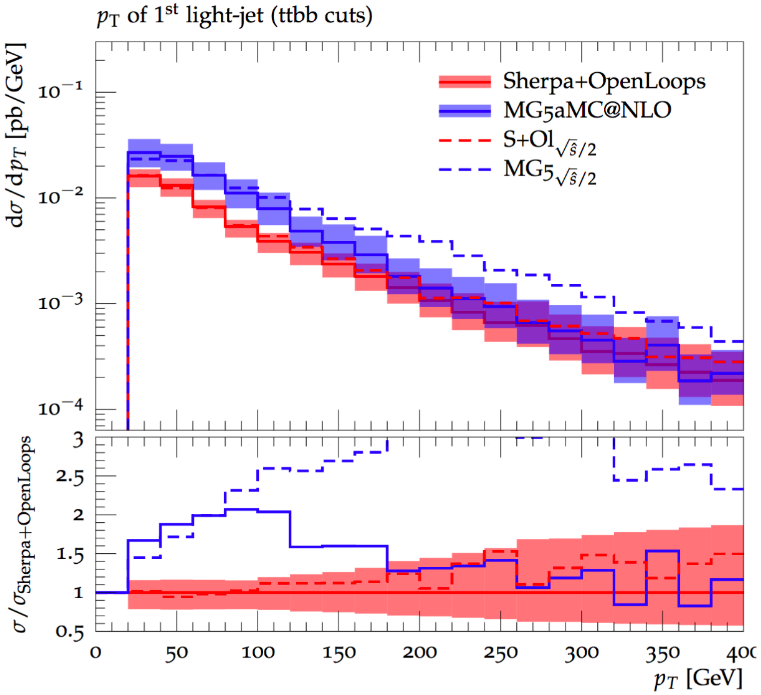
using 4F scheme

Renormalisation scale μ_R	$\sqrt[4]{m_T(t) * m_T(\bar{t}) * m_T(b) * m_T(\bar{b})}$ (CMMPS)	
Factorisation scale μ_F	$H_T/2$ with $H_T = \sum_{i \epsilon \mathrm{finalstate}} m_T(i)$	
Resummation scale μ_Q (Q_{sh})	$\xi \hat{s}$ with $\xi \epsilon [0.1,0.25]$	$H_T/2$
	for MG5_aMC@NLO	Sherpa

PDF set: NNPDF3.0 4F
Top quarks are not decayed,
hadronisation and UE are swtiched off

- differences of ≥40% for tt+≥2b cross section
- sizable differences in NLO radiation pattern
- mu$_Q$ (shower starting scale) dependence in MG5_aMC@NLO

p_T leading light-jet (radiation)



Comparisons of different MC predictions (YR4 exercise)

tt+bb 4F NLO+PS predictions (Sherpa+OpenLoops, MG5_aMC@NLO+Py8 and Powhel+Py8) with theory motivated shower settings for consistent comparisons:

Different NLO+PS methods, showers, and m_b treatments

Tool	Matching	Shower	m_b[GeV]	gencuts
SHERPA2.1+OPENLOOPS	SMC@NLO	Sherpa 2.1	4.75 (4F)	no
MG5_AMC@NLO	MC@NLO	Pythia 8.2	4.75 (4F)	no
POWHEL	Powheg	Pythia 8.2	0 (5F)	$p_{T,b} > 4.75$ GeV $\frac{m_{bb}}{2} > 4.75$ GeV

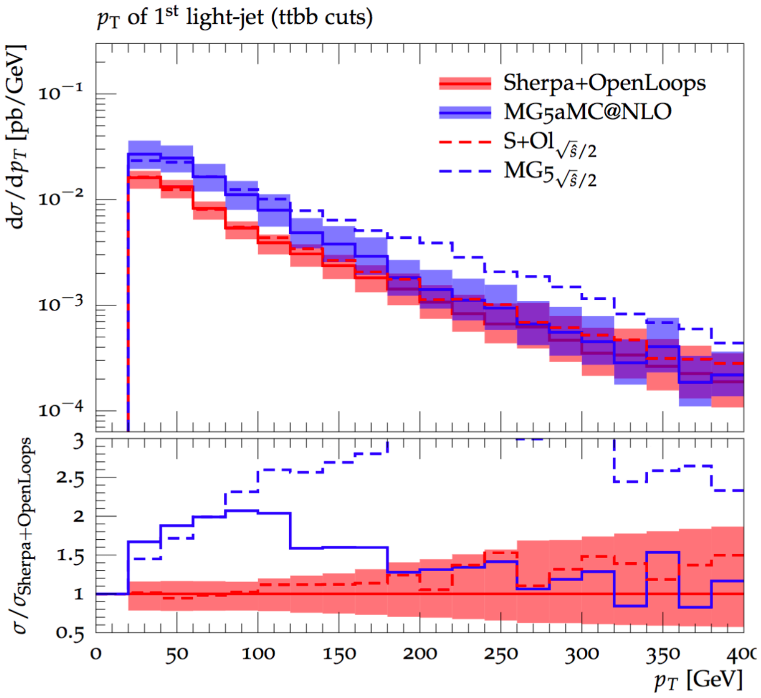
using 4F scheme

Renormalisation scale μ_R	$\sqrt[4]{m_T(t) * m_T(\bar{t}) * m_T(b) * m_T(\bar{b})}$ (CMMPS)	
Factorisation scale μ_F	$H_T/2$ with $H_T = \sum_{i \epsilon \mathrm{finalstate}} m_T(i)$	
Resummation scale μ_Q (Q_{sh})	$\xi \hat{s}$ with $\xi \epsilon [0.1, 0.25]$	$H_T/2$
	for MG5_aMC@NLO	Sherpa

PDF set: NNPDF3.0 4F
Top quarks are not decayed,
hadronisation and UE are swtiched off

- differences of ≥40% for tt+≥2b cross section
- sizable differences in NLO radiation pattern
- mu$_Q$ (shower starting scale) dependence in MG5_aMC@NLO
- much more studies and new setups (more in Nov.6th mtg.):
  - PowHel (arXiv: 1709.06915)
  - Powheg+OpenLoops+Pythia8/Herwig7
  - further studies regarding mu$_Q$ inMG5_aMC@NLO (+Py8/HW7)

p_T leading light-jet (radiation)



tt+bb 4F NLO+PS systematic uncertainties

Scale choices (YR4) and uncertainties (no proposal yet)

Factorisation (μ_Q) and resummation (μ_Q) scales

$$E_{T_i} = \sqrt{m_i^2 + p_{T,i}^2}$$

$$\mu_F = \mu_Q = \frac{H_T}{2} = \frac{1}{2} \sum_{i=t,\bar{t},b,\bar{b}} E_{T,i}$$

$\mu_Q \equiv$ shower starting scale is a free paramater in MC@NLO (not in Powheg)

CKKW-like (softer) renormalisation scale

$$\mu_R = \mu_{\mathrm{CKKW}} = \prod_{i=t,\bar{t},b,\bar{b}} E_{T,i}^{1/4}$$

Scale variations (leading uncertainty) ~20-30%

- factor-2 variations of μ_R and μ_F ⇔ normalisation
- "kinematic" variations of μ_R, μ_F, μ_Q ⇔ shape
- variations of μ_Q in MC@NLO and h_{damp} in Powheg ⇔ NLOPS matching

Other variations

- PDF variations (only few percent)
- shower variations: tune variations, shower recoil scheme, . . .

tt+HF modelling: very challenging task

- pure QCD process, very complicated that involves several scales and massive quarks
  - challenging for the MC generator community
  - implementation of latest theoretical developments crucial
- studies ongoing in both experiments in close collaboration with theorists (documented in YR4)
  - 5F *tt*+jets (m_b=0) vs. 4F *tt*+*bb* predictions
  - 4F *tt*+*bb* NLO+PS predictions (with massive *b*-quarks in ME) with novel generators
    - recommended but need further studies (settings, associated shape uncertainties, etc.)
  - how to merge 4F and 5F samples?
  - heavy flavour classification

Approach proposed in the LHCHiggs Yellow Report 4

* NLO+PS 4F *tt*+*bb* sample

- can be applied in full phase space (no generation cuts)
- inclusive description of *tt*+≥1*b*-quarks

* Inclusive 5F *tt*+jets sample

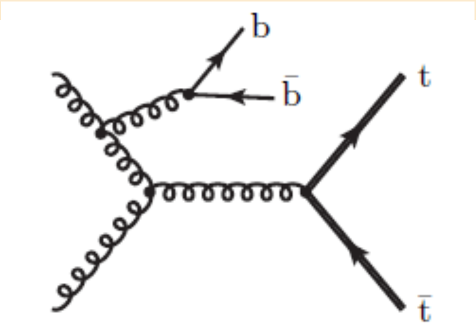
- needs to be restricted to *tt*+0 *b*-quarks to avoid double counting (veto events containing *b*-quarks not arising from showered top decays or MPI or UE)

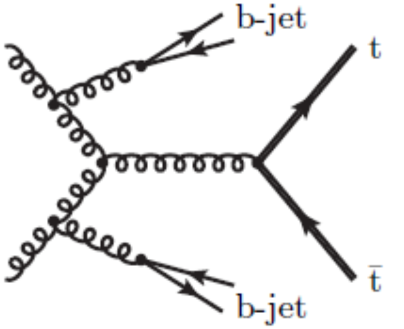
→ Ongoing discussions on possible implementations

tt+HF modelling: very challenging task

- pure QCD process, very complicated that involves several scales and massive quarks
  - challenging for the MC generator community
  - implementation of latest theoretical developments crucial
- studies ongoing in both experiments in close collaboration with theorists (documented in YR4)
  - 5F *tt*+jets (m_b=0) vs. 4F *tt+bb* predictions
  - 4F *tt+bb* NLO+PS predictions (with massive *b*-quarks in ME) with novel generators
    - recommended but need further studies (settings, associated shape uncertainties, etc.)
  - how to merge 4F and 5F samples?
  - heavy flavour classification

First, define *tt*+light, *tt*+≥1c and *tt*+≥1b.





Then, tt+≥1b subcategories:

- *tt+bb*
- *tt+b*: mainly tt+bb with one b-jet out of acceptance
- *tt+B/2b* (ATLAS/CMS): 2 B-hadrons merged in same jet, collinear g→bb
- *tt+3b*: the rest
- *tt+bb* from MPI and FSR are treated separately (at least in ATLAS, small fraction only available in *tt*+jets inclusive 5F calculations)

HF definition and treatment of uncertainties

Reconstructed *tt*+jets events are classified into several categories and subcategories, based on the flavour of additional jets (at particle level) and number of hadrons in each of them.

* Only additional particle level jets above a p_T threshold are considered in the classification

* Jets flavour (b, c or light) is determined via a ghost or dR matching to hadrons.

- For b and c jets, kinematics cuts on leading hadron to which they are matched being studied.
- No p_T ratio p_T^{hadron}/p_T^{jet} cut is considered (so far) in the HF classification.

Cuts	ATLAS *	CMS
Reco-level jets	(all events are classified)	≥ two jets with p_T > 30 GeV
Particle level jets	15 GeV	20 GeV
Hadrons	5 GeV, no p_T^{hadron}/p_T^{jet} cut	No cuts
Particle-hadron matching	dR<0.3	Ghost matching

* From ongoing studies, the relative differences among generators in tt+jets fractions seem stable against these cuts

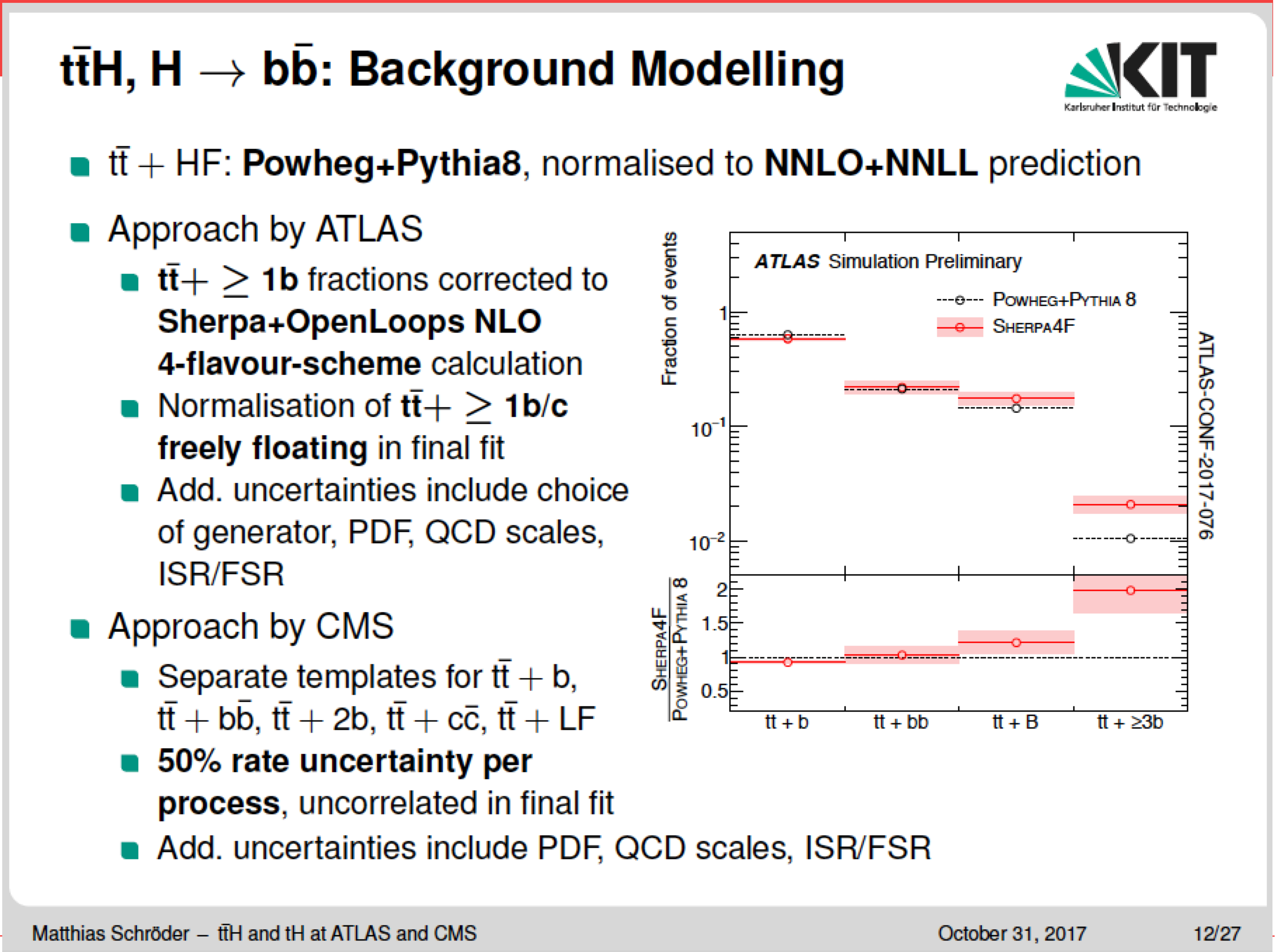
Subcategories

Treatment of uncertainties

ATLAS: reweighting fractions (not kinematics) for each subcategory in 5F sample to 4F SherpaOL

→ treating uncertainties as fully correlated among subcategories

CMS: shapes from 5F predictions → treating uncertainties as fully uncorrelated.



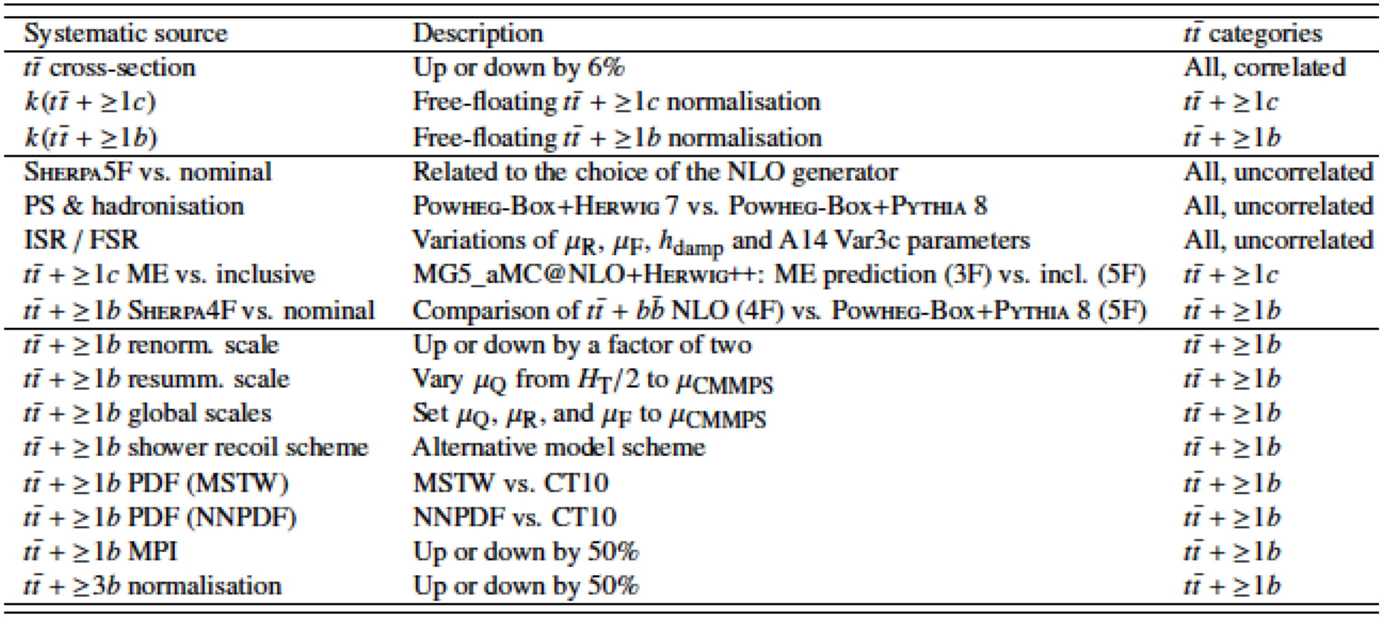
$t\bar{t}H$, $H \to b\bar{b}$: Background Modelling

- $t\bar{t}$ + HF: **Powheg+Pythia8**, normalised to **NNLO+NNLL** prediction
- Approach by ATLAS
  - **$t\bar{t}+ \geq 1b$** fractions corrected to **Sherpa+OpenLoops NLO 4-flavour-scheme** calculation
  - Normalisation of **$t\bar{t}+ \geq 1b/c$ freely floating** in final fit
  - Add. uncertainties include choice of generator, PDF, QCD scales, ISR/FSR
- Approach by CMS
  - Separate templates for $t\bar{t}+b$, $t\bar{t}+b\bar{b}$, $t\bar{t}+2b$, $t\bar{t}+c\bar{c}$, $t\bar{t}+$LF
  - **50% rate uncertainty per process**, uncorrelated in final fit
  - Add. uncertainties include PDF, QCD scales, ISR/FSR

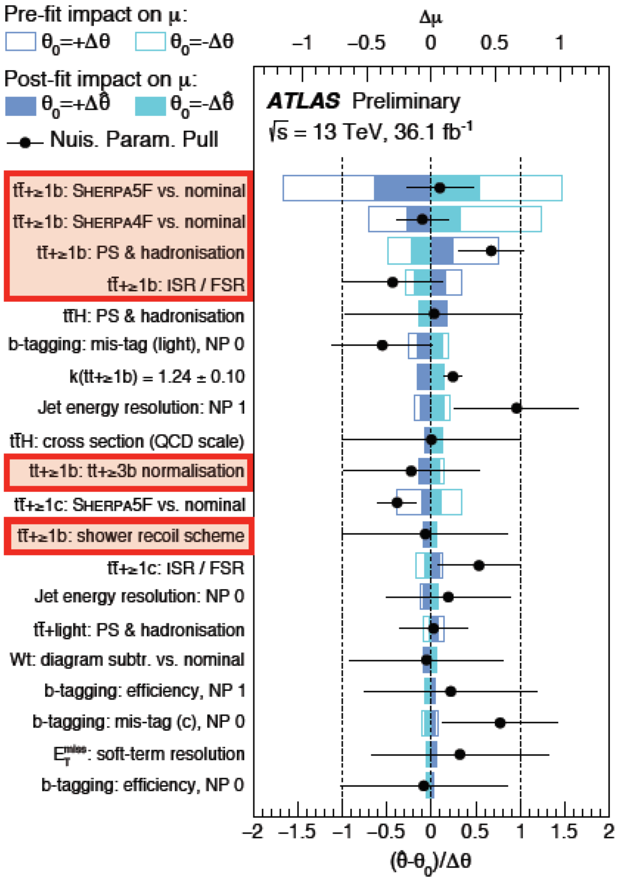
More on the treatment of systematics

ATLAS-CONF-2017-076

$t\bar{t}$ modelling uncertainties



Systematic source	Description	$t\bar{t}$ categories
$t\bar{t}$ cross-section	Up or down by 6%	All, correlated
$k(t\bar{t}+\geq 1c)$	Free-floating $t\bar{t}+\geq 1c$ normalisation	$t\bar{t}+\geq 1c$
$k(t\bar{t}+\geq 1b)$	Free-floating $t\bar{t}+\geq 1b$ normalisation	$t\bar{t}+\geq 1b$
SHERPA5F vs. nominal	Related to the choice of the NLO generator	All, uncorrelated
PS & hadronisation	POWHEG-BOX+HERWIG 7 vs. POWHEG-BOX+PYTHIA 8	All, uncorrelated
ISR / FSR	Variations of μ_R, μ_F, h_{damp} and A14 Var3c parameters	All, uncorrelated
$t\bar{t}+\geq 1c$ ME vs. inclusive	MG5_aMC@NLO+HERWIG++: ME prediction (3F) vs. incl. (5F)	$t\bar{t}+\geq 1c$
$t\bar{t}+\geq 1b$ SHERPA4F vs. nominal	Comparison of $t\bar{t}+b\bar{b}$ NLO (4F) vs. POWHEG-BOX+PYTHIA 8 (5F)	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ renorm. scale	Up or down by a factor of two	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ resumm. scale	Vary μ_Q from $H_T/2$ to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ global scales	Set μ_Q, μ_R, and μ_F to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ shower recoil scheme	Alternative model scheme	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF (MSTW)	MSTW vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF (NNPDF)	NNPDF vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ MPI	Up or down by 50%	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 3b$ normalisation	Up or down by 50%	$t\bar{t}+\geq 1b$



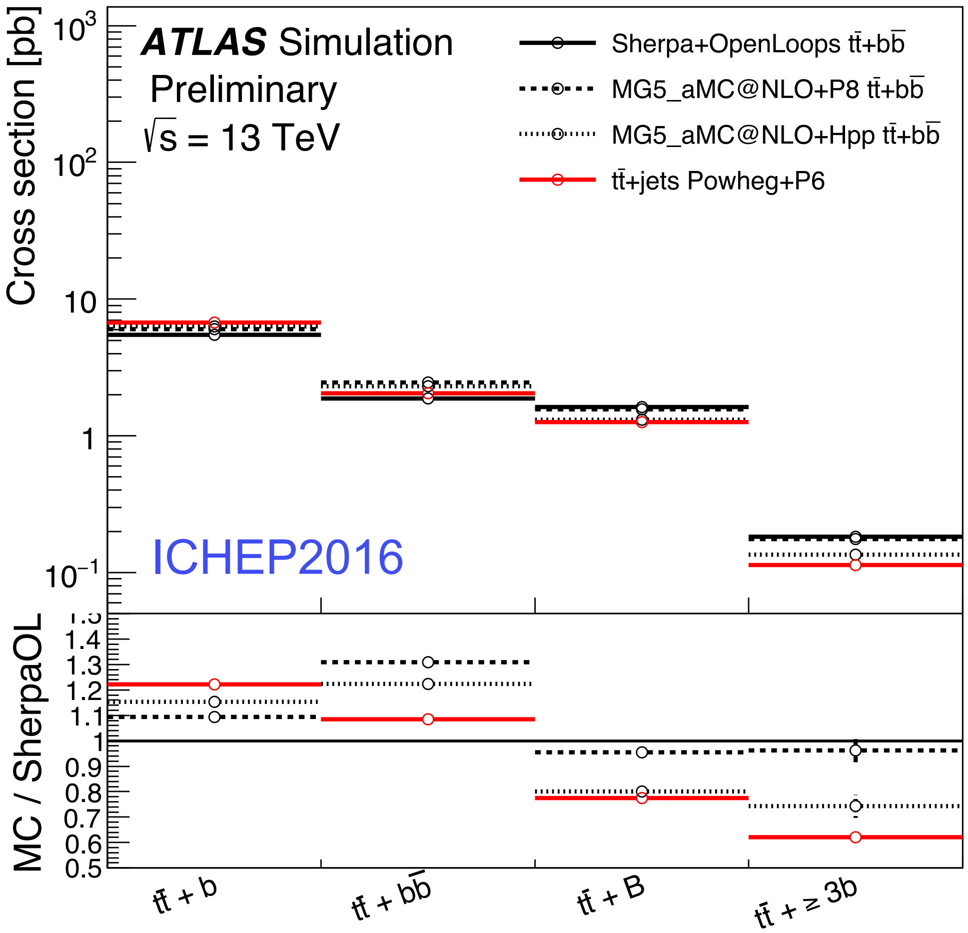
- Many sources of modelling uncertainty considered:
  - Generator: Powheg+Pythia8 vs. Sherpa (5F)
  - Parton shower: Powheg+Pythia8 vs. Powheg+Herwig7
  - 5F vs. 4F in Sherpa+OpenLoops
  - Scale variations in Sherpa+OpenLoops
- All $t\bar{t}$+jets modelling uncertainties uncorrelated between $t\bar{t}+\geq 1b/\geq 1c$/light
- Scale variation uncertainties correlated across each $t\bar{t}+\geq 1b$ sub-component

→ tt + HF modelling dominant unc.
→ others: *b*-tagging, JER
→ also limited MC sample size in background modelling

ATLAS reweighting to 4F Sherpa+OL for tt+≥1b

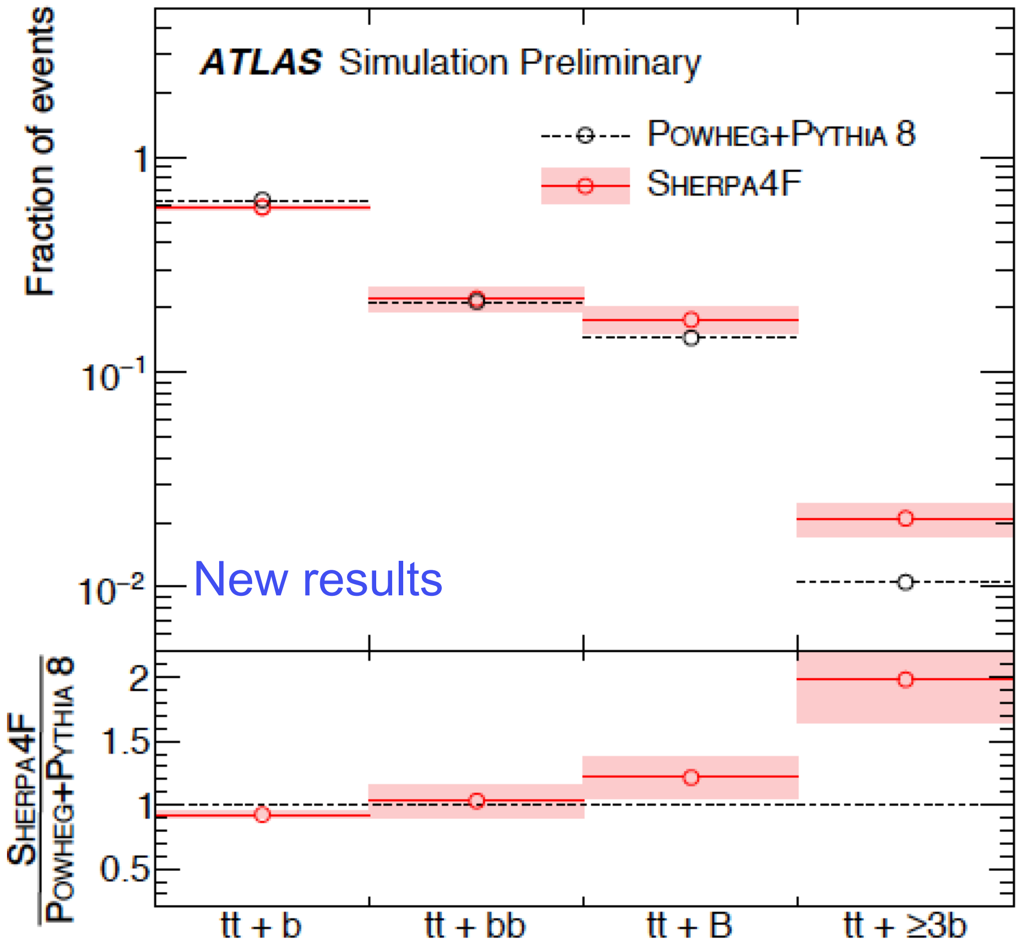
- Correct normalisation of the different subcategories
- ~~Small kinematic corrections in each category~~

ATL-PHYS-PUB-2016-016



ICHEP2016

ATLAS-CONF-2017-076



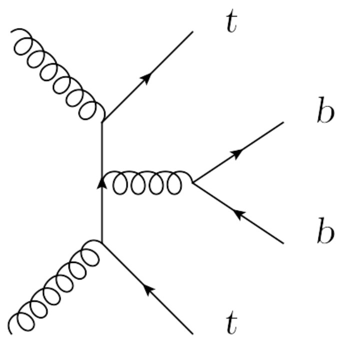
New results

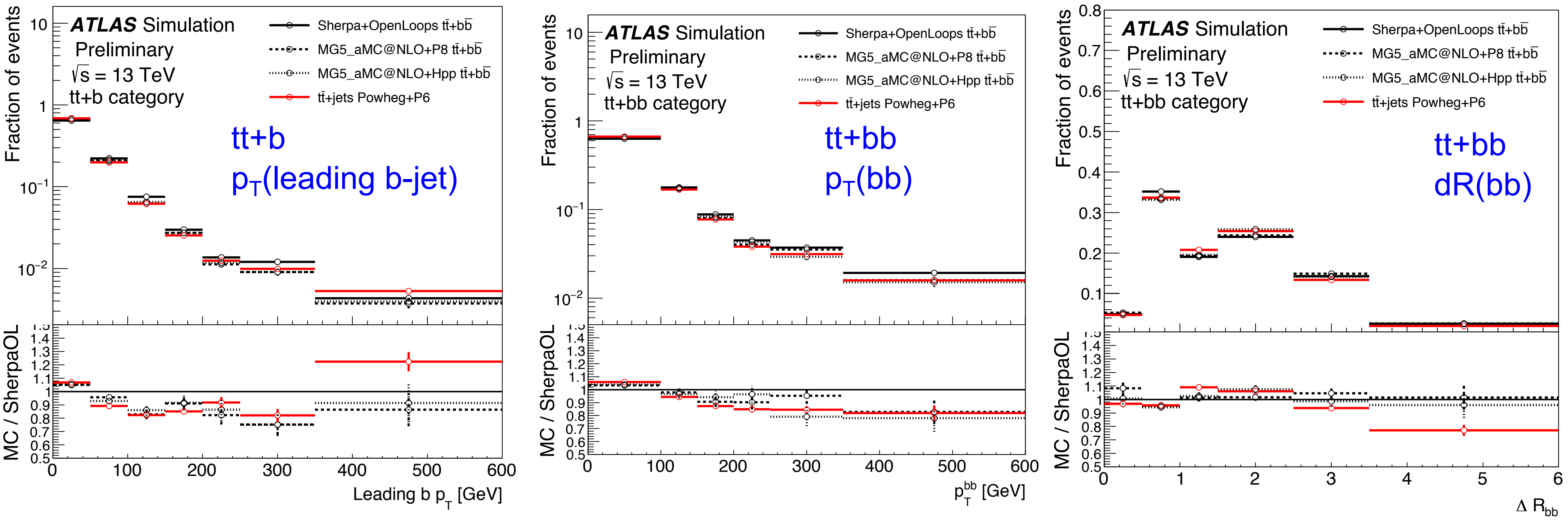
All samples compared predict more events in tt+b/bb categories than SherpaOL 4F.

ATLAS reweighting to 4F Sherpa+OL for tt+≥1b

$t\bar{t}b\bar{b}$, $\sqrt{s}$=13 TeV, m_t= 172.5 GeV, m_b= 4.75 GeV			
ME gen.	aMC@NLO + MadSpin	aMC@NLO + MadSpin	SHERPA+OPENLOOPS -
PS/UE gen.	HERWIG++	PYTHIA8	SHERPA
Renormalisation scale μ_R	$\sqrt[4]{m_T(t) * m_T(\bar{t}) * m_T(b) * m_T(\bar{b})}$ (CMMPS)		
Factorisation scale μ_F	$H_T/2$ with $H_T = \sum_{i \epsilon \mathrm{finalstate}} m_T(i)$		
Resummation scale μ_Q (Q_{sh})	$\xi\hat{s}$ with $\xi\epsilon$[0.1,0.25]		$H_T/2$
ME PDF	NNPDF3.0nlo 4F	NNPDF3.0nlo 4F	CT10nlo 4F
PS/UE PDF	CTEQ6L1	NNPDF2.3	
Tune	UE-EE-5	A14	author's tune
Cross-section × BR($t\bar{t} \rightarrow \mu\mu$)[pb]	0.322 ± 0.020	0.320 ± 0.020	0.315 ± 0.020

ATL-PHYS-PUB-2016-016





* Differences (up to 20%) in p_T(leading b-jet) for tt+b category: SherpaOL 4F is harder.

Activities within the LHC Higgs *tt+H/tH* XS subgroup

tt+H **modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposaltTH

tt+Z/W **modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposaltTV

tt**+heavy flavour modelling:** https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ProposalTtbb

tt+Z and *tt+W* cross-sections predictions

Recent developments reported in YR4:

* NLO QCD+EW corrections to *tt+H/Z/W*

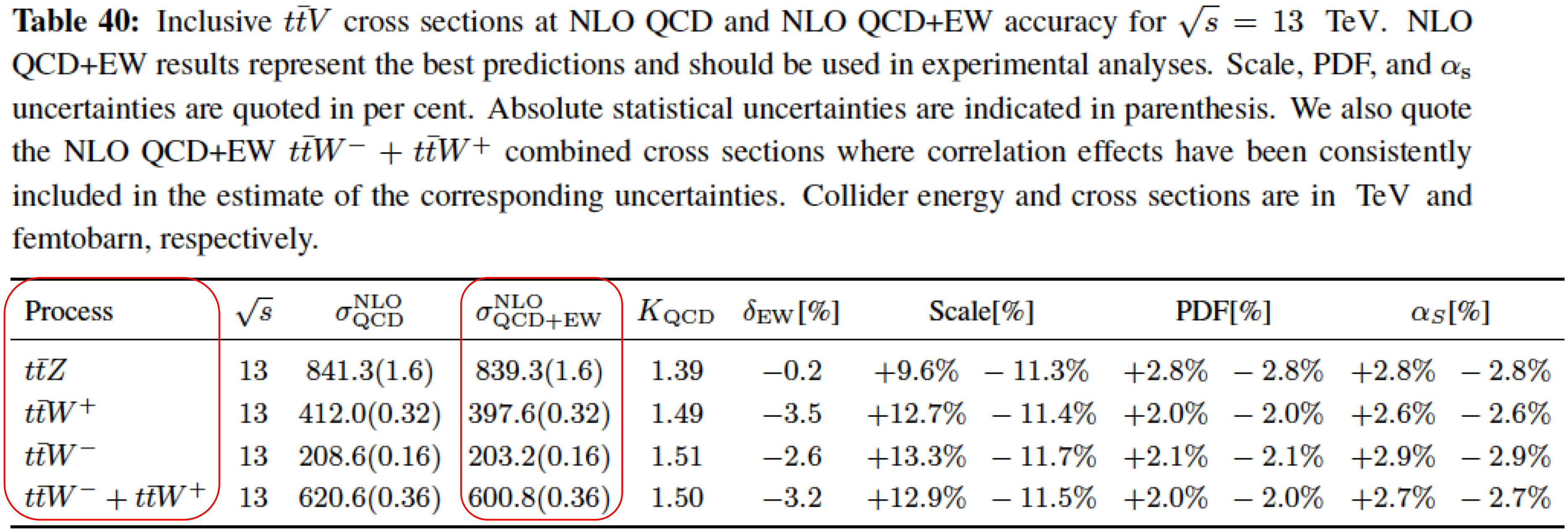


Table 40: Inclusive $t\bar{t}V$ cross sections at NLO QCD and NLO QCD+EW accuracy for $\sqrt{s} = 13$ TeV. NLO QCD+EW results represent the best predictions and should be used in experimental analyses. Scale, PDF, and α_s uncertainties are quoted in per cent. Absolute statistical uncertainties are indicated in parenthesis. We also quote the NLO QCD+EW $t\bar{t}W^- + t\bar{t}W^+$ combined cross sections where correlation effects have been consistently included in the estimate of the corresponding uncertainties. Collider energy and cross sections are in TeV and femtobarn, respectively.

Process	$\sqrt{s}$	$\sigma^{\text{NLO}}_{\text{QCD}}$	$\sigma^{\text{NLO}}_{\text{QCD+EW}}$	K_{QCD}	$\delta_{\text{EW}}[\%]$	Scale[%]		PDF[%]		$\alpha_S[\%]$	
$t\bar{t}Z$	13	841.3(1.6)	839.3(1.6)	1.39	−0.2	+9.6%	− 11.3%	+2.8%	− 2.8%	+2.8%	− 2.8%
$t\bar{t}W^+$	13	412.0(0.32)	397.6(0.32)	1.49	−3.5	+12.7%	− 11.4%	+2.0%	− 2.0%	+2.6%	− 2.6%
$t\bar{t}W^-$	13	208.6(0.16)	203.2(0.16)	1.51	−2.6	+13.3%	− 11.7%	+2.1%	− 2.1%	+2.9%	− 2.9%
$t\bar{t}W^- + t\bar{t}W^+$	13	620.6(0.36)	600.8(0.36)	1.50	−3.2	+12.9%	− 11.5%	+2.0%	− 2.0%	+2.7%	− 2.7%

- Values for fixed scale $\mu = m_t + m_V/2$ (replacing by a dynamic scale $\mu = H_T/2$ shifts cross-sections by -7%, within unc. quoted)
- For ttW production, QCD+EW corrections as well as the NLO scale uncertainties are slightly more pronounced than for ttZ.
- Scale variations range from 10 to 13% and represent the dominant source of uncertainty.

* Experiments are using these cross-section values to normalise their samples,
but currently available MC simulated do not include EW corrections.

* *tt+Z* values include on-shell contribution only,
but experiments include off-shell *tt*$\gamma^* \rightarrow$*ll* and thus
some approximations are made to derive a *tt+ll* K-factor.

* Precision of exp. measurements close to theory unc.

$\Delta\sigma/\sigma$ (%), Obs.Sign.	tt+Z	tt+W
CMS 2015+2016	14%, 9.9σ (stat~syst. unc.)	15%, 5.5σ (stat~syst. unc.)

tt+Z and *tt+W* cross-sections predictions

Recent developments reported in YR4:

* NLO QCD+EW corrections to *tt+H/Z/W*

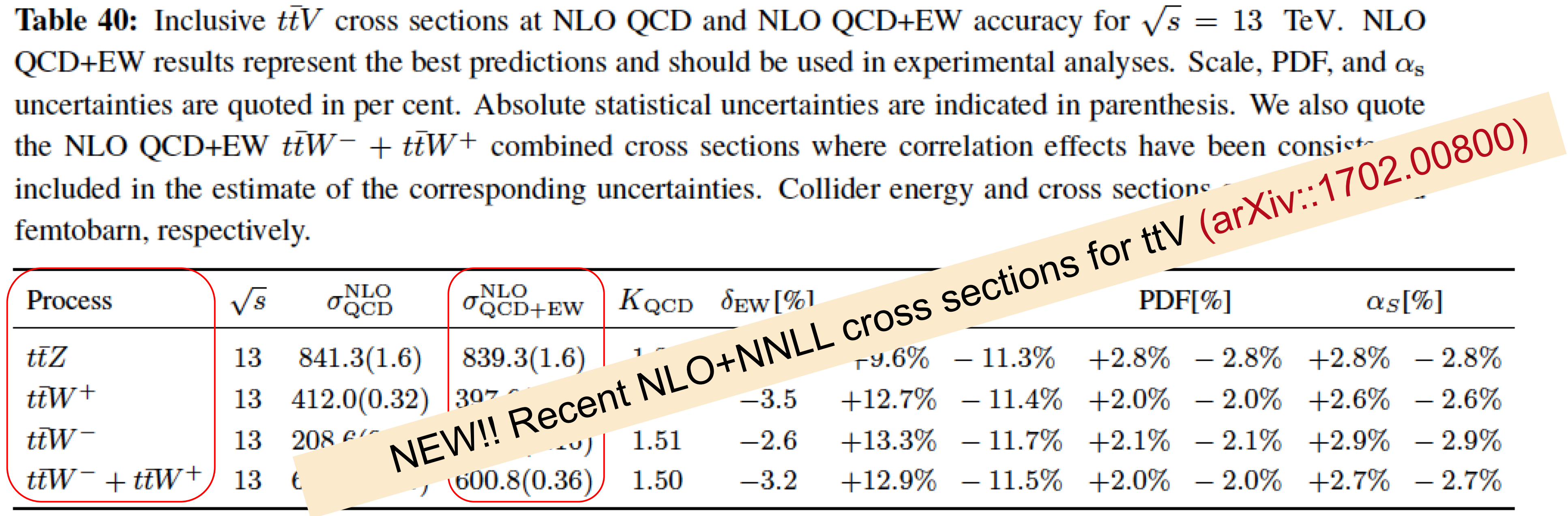


Table 40: Inclusive $t\bar{t}V$ cross sections at NLO QCD and NLO QCD+EW accuracy for $\sqrt{s}$ = 13 TeV. NLO QCD+EW results represent the best predictions and should be used in experimental analyses. Scale, PDF, and α_s uncertainties are quoted in per cent. Absolute statistical uncertainties are indicated in parenthesis. We also quote the NLO QCD+EW $t\bar{t}W^-$ + $t\bar{t}W^+$ combined cross sections where correlation effects have been consist[illegible] included in the estimate of the corresponding uncertainties. Collider energy and cross sections [illegible] femtobarn, respectively.

Process	$\sqrt{s}$	$\sigma^{\mathrm{NLO}}_{\mathrm{QCD}}$	$\sigma^{\mathrm{NLO}}_{\mathrm{QCD+EW}}$	K_{QCD}	$\delta_{\mathrm{EW}}\,[\%]$	[illegible]		PDF[%]		$\alpha_S[\%]$	
$t\bar{t}Z$	13	841.3(1.6)	839.3(1.6)	[illegible]	[illegible]	[illegible]	− 11.3%	+2.8%	− 2.8%	+2.8%	− 2.8%
$t\bar{t}W^+$	13	412.0(0.32)	[illegible]	[illegible]	−3.5	+12.7%	− 11.4%	+2.0%	− 2.0%	+2.6%	− 2.6%
$t\bar{t}W^-$	13	[illegible]	[illegible]	1.51	−2.6	+13.3%	− 11.7%	+2.1%	− 2.1%	+2.9%	− 2.9%
$t\bar{t}W^- + t\bar{t}W^+$	13	[illegible]	600.8(0.36)	1.50	−3.2	+12.9%	− 11.5%	+2.0%	− 2.0%	+2.7%	− 2.7%

NEW!! Recent NLO+NNLL cross sections for ttV (arXiv::1702.00800)

- Values for fixed scale $\mu = m_t + m_V/2$ (replacing by a dynamic scale $\mu = H_T/2$ shifts cross-sections by -7%, within unc. quoted)
- For ttW production, QCD+EW corrections as well as the NLO scale uncertainties are slightly more pronounced than for ttZ.
- Scale variations range from 10 to 13% and represent the dominant source of uncertainty.

* Experiments are using these cross-section values to normalise their samples, but currently available MC simulated do not include EW corrections.

* tt+Z values include on-shell contribution only, but experiments include off-shell *tt*$\gamma^*\rightarrow$*ll* and thus some approximations are made to derive a *tt+ll* K-factor.

* Precision of exp. measurements close to theory unc.

$\Delta\sigma/\sigma$ (%), Obs.Sign.	**tt+Z**	**tt+W**
CMS 2015+2016	14%, 9.9σ (stat~syst. unc.)	15%, 5.5σ (stat~syst. unc.)

Outlook

The program of *tt+X* production at the LHC is well underway

BUT entering regime of results being systematically limited (bkg. and signal modelling)

→ one of the main focus of the LHCHiggs *tt+H/t+H* XS subgroup

- implementation of the latest theoretical developments is crucial to reduce unc.
- will continue comparing with data to further tune and improve the MC generators

tt+bb process is of interest for both Higgs and Top WGs

→ try to exploit possible synergies between *tt+H* and *tt+bb* measurements

Many developments from theory side:

4F scheme preferable for NLO+PS simulations of tt + b-jet production

- Many studies summarized in the YR4
- Studies ongoing to understand significant matching/shower scale dependence and to provide reliable theoretical uncertainties
- Three 4F generators being compared: Sherpa, MG5aMC@NLO, Powheg

From the experimental side:

- provide unfolded measurements (and Rivet routines)
- start generating full MC samples (including top quark decays, detector simulation, etc.)
  - can be quite CPU consuming (possibilities to share LHE files?)
- start testing "matching of 5F *tt*+jets and *tt+bb* 4F" samples

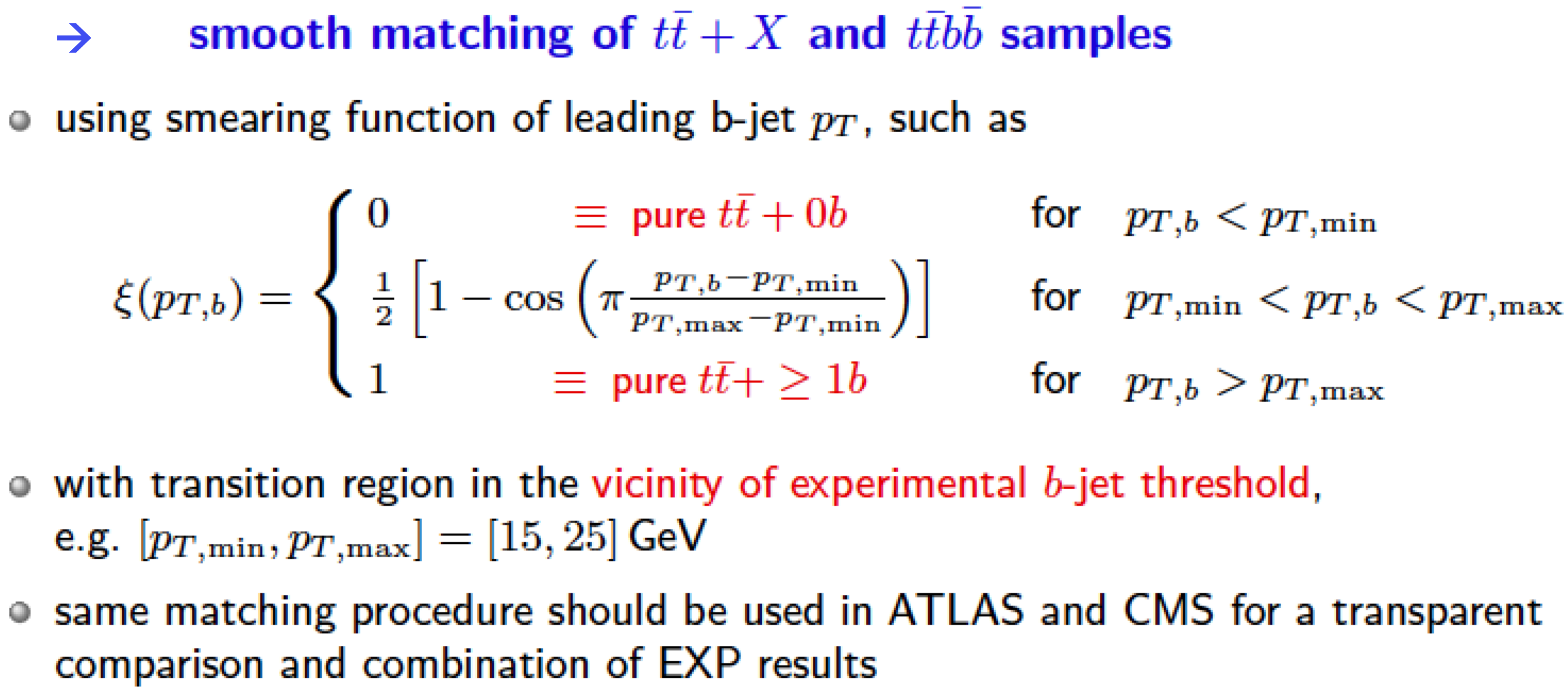
THANKS FOR YOUR ATTENTION

BACK-UP

How to merge 4F and 5F samples?

PROPOSAL:

tt+bb 4F for tt + b-jet categories
tt+X(jets) 5F for tt+light and c-jets



→ smooth matching of $t\bar{t}+X$ and $t\bar{t}b\bar{b}$ samples

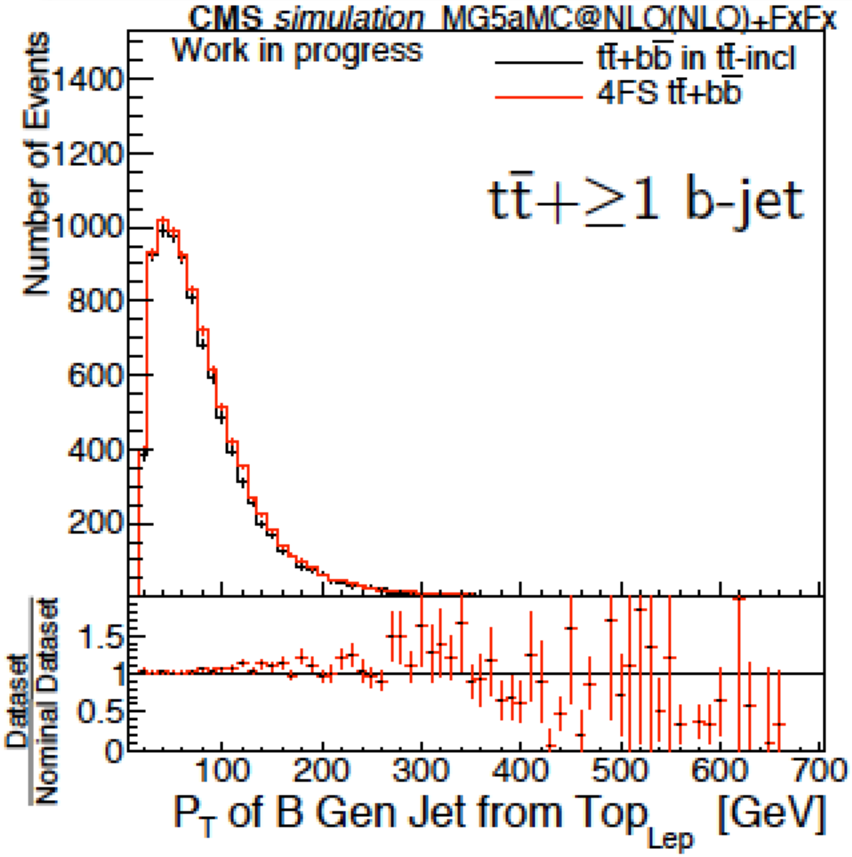
- using smearing function of leading b-jet p_T, such as

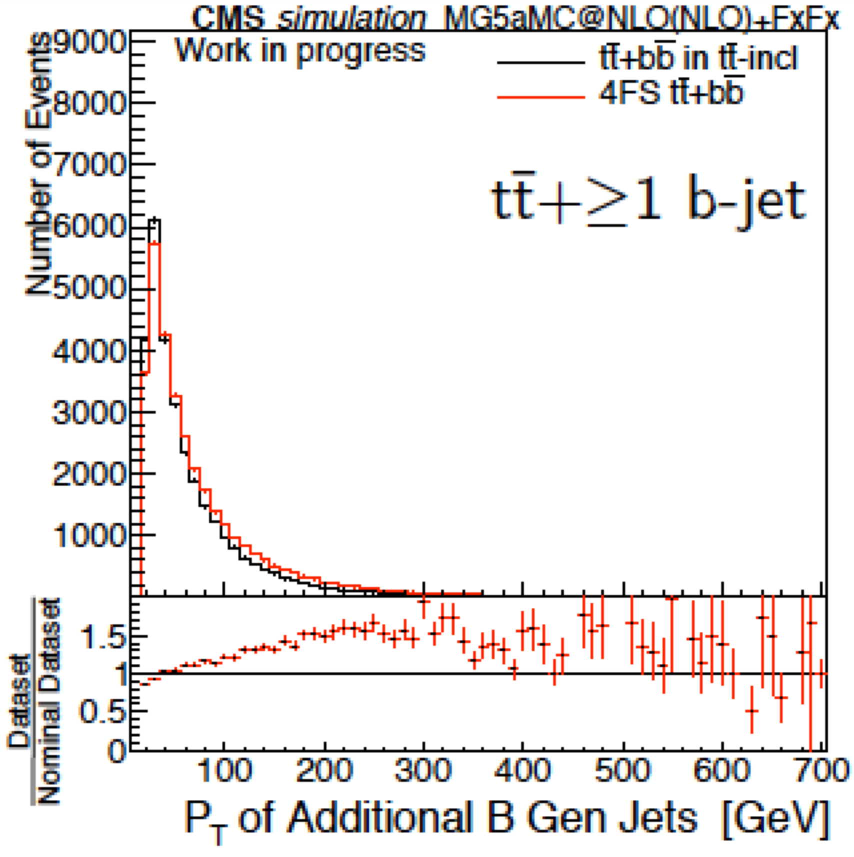
$$\xi(p_{T,b}) = \begin{cases} 0 & \equiv \text{ pure } t\bar{t}+0b & \text{for} \quad p_{T,b} < p_{T,\min} \\ \frac{1}{2}\left[1-\cos\left(\pi\frac{p_{T,b}-p_{T,\min}}{p_{T,\max}-p_{T,\min}}\right)\right] & & \text{for} \quad p_{T,\min} < p_{T,b} < p_{T,\max} \\ 1 & \equiv \text{ pure } t\bar{t}+\geq 1b & \text{for} \quad p_{T,b} > p_{T,\max} \end{cases}$$

- with transition region in the vicinity of experimental b-jet threshold, e.g. $[p_{T,\min}, p_{T,\max}] = [15, 25]$ GeV
- same matching procedure should be used in ATLAS and CMS for a transparent comparison and combination of EXP results

Comparisons *tt+bb* 4F and tt+jets NLOmultileg 5F (CMS)

- Comparison between MG5aMC@NLO+Pythia8 (FxFx merged) 5FNS $t\bar{t}$+0/1/2 jets and 4FNS $t\bar{t}+b\bar{b}$
- Require 1 b-gen-jet not from Top decay in $p_T \geq 20.0$GeV, $|\eta| < 2.4$
- b-gen-jets defined through jet-flavour-clustering (ghost hadrons)





tt+Z/W modelling

ATLAS: *tt+ll, tt+Z(→qq), tt+Z(→νν), tt+W*

tt+ll (includes off-shell *ttγ*→ll* production with m_{ll}>1 GeV for OSSF matrix element leptons)

Nominal: MadGraph5_aMC@NLO($\mu_R=\mu_F=H_T/2$, $\mu_Q=\xi\sqrt{\hat{s}}$, NNPDF3.0)+MadSpin+Py8 (A14 tune)

- Alternative MC generator: vs. Sherpa LOmultileg or MG5_aMC@NLO LOmultileg (N_p<=2)
- Tune variations: A14 eigentunes for Pythia8
- Scale choice & PDF set: using multiple event weights

CMS: *tt+ll, tt+Z(→qq), tt+Z(→νν), tt+W*

tt+ll (includes off-shell *ttγ*→ll* production with m_{ll}>10 GeV for OSSF matrix element leptons)

Nominal: MadGraph5_aMC@NLO (LOmultileg_MLMmatching, NNPDF3.0)+MadSpin+Pythia8

- Alternative MC generator:
  ttZ: MadGraph5_aMC@NLO NLOmode vs. LOmultileg_MLMmatching
  ttZ: MadGraph5_aMC@NLO NLOmode vs. LOmultileg_MLMmatching
- Scale choice & PDF set: using multiple event weights

Other samples: tt+photon, tZq, tWZ, tHq, tWH

- tt+photon:

  MG5_aMC@NLO+Py8 (LO mode),
  including photons radiated from the top quarks as well as from their decay products
  (Note: MadSpin was NOT used since it does NOT include photon radiation in top decay products)

  ATLAS cuts at generation level: $p_T(\gamma)>15$ GeV, $dR(lep, \gamma)>0.2$ and $dR(jet, \gamma)>0.2$
  CMS cuts at generation level: $p_T(\gamma)>13$ GeV, $dR(lep, \gamma)>0.3$ and $dR(jet, \gamma)>0.3$

- tZq

  ATLAS: MG5_aMC@NLO (LO mode)+Py6, Perugia2012, 4 FS → moving to NLO & Py8
  CMS: MG5_aMC@NLO (NLO mode)+Py8, 4 FS

- tWZ

  ATLAS: MG5_aMC@NLO (NLO mode)+Py8, 5FS, interference w. ttH removed with DR1 & DR2
  CMS: MG5_aMC@NLO (LO mode)+Py8, 5FS

[arXiv: 1607.05862, ATL-PHYS-PUB-2016-020]

$$|\mathcal{M}_{\mathrm{tot}}|^2 = \underbrace{\overbrace{|\mathcal{M}_{\mathrm{sr}}|^2}^{\mathrm{DR1}} + 2Re(\mathcal{M}_{\mathrm{sr}} \cdot \mathcal{M}_{\mathrm{dr}})}_{\mathrm{DR2}} + |\mathcal{M}_{\mathrm{dr}}|^2$$

- tHq

  ATLAS: MG5_aMC@NLO+Py8/HW++ (LO mode), 4 FS
  CMS: MG5_aMC@NLO+Py8(LO mode), 4 FS

- tWH

  ATLAS: MG5_aMC@NLO (NLO mode)+HWpp, A14, 5 FS, interference with ttH removed with DR1
  CMS: MG5_aMC@NLO+Py8(LO mode)+Py8, 5FS

tt+H modelling

NLO QCD+PS matched setups used in both experiments.

ATLAS

Nominal: MadGraph5_aMC@NLO* (μ_R=μ_F=H_T/2, μ_Q=$\xi\sqrt{\hat{s}}$, NNPDF3.0)+MadSpin+Py8 (A14 tune)

- Showering & hadronization: compared to MG5_aMC@NLO+MadSpin+HWpp (UE-EE5 tune)
- Tune variations: A14 eigentunes for Pythia8
- Scale choice & PDF set: using multiple event weights

→ Currently also studying (no official samples available yet):

- Powheg+Pythia8
- Sherpa(NLO)+OpenLoops

CMS

Nominal is different for ttH(bb) and ttH(multilepton, $\gamma\gamma$) to be consistent with main background in each of the channels:

ttH(bb): Powheg+Pythia8 (h_{damp} ~1.58*$m_{t,}$ CUETP8M2 tune) [as used for tt+jets]

ttH(multilepton, $\gamma\gamma$): MadGraph5_aMC@NLO(NLO)+MadSpin+Pythia8 [as used for tt+W/Z]

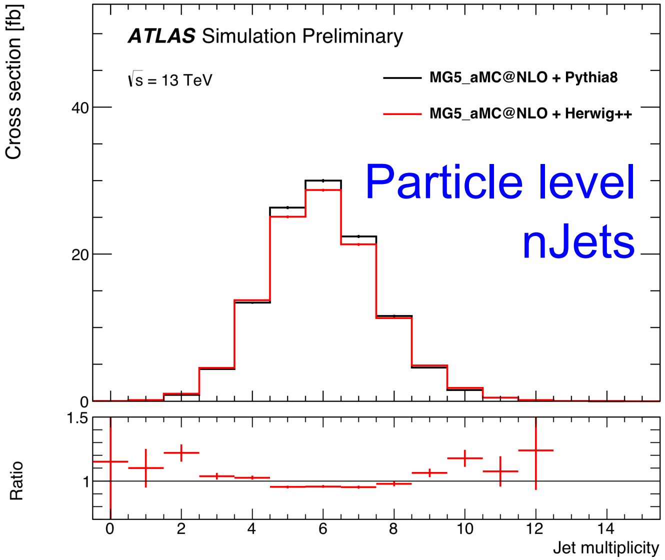
- Scale choice & PDF set: using multiple event weights

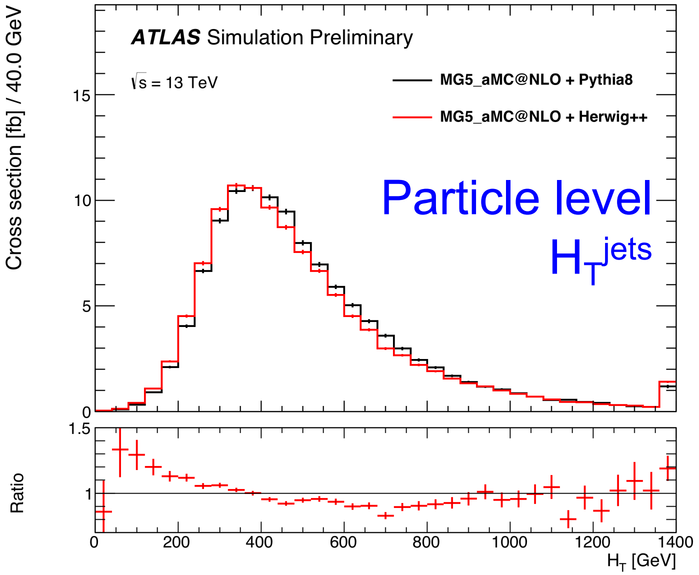
* Caveat of MadGraph5_aMC@NLO (NLO mode): ~25% of events having negative weights

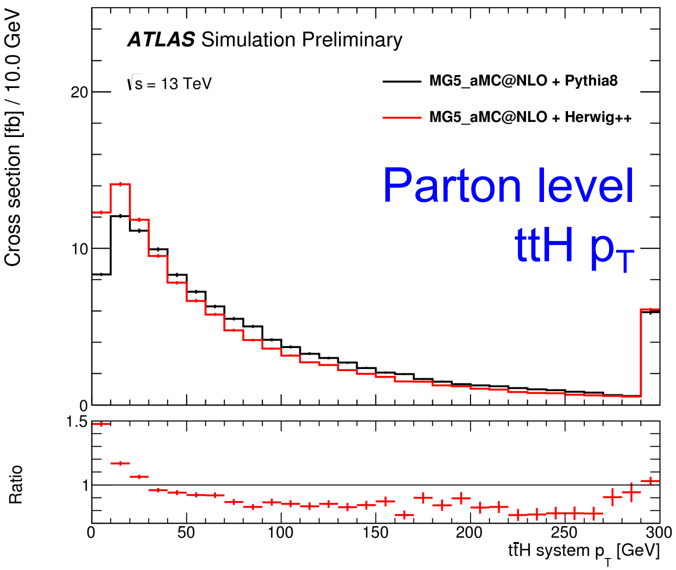
tt+H modelling: studies at particle/parton level (ATLAS)

ATL-PHYS-PUB-2016-005

tt+H (*tt*→lep+jets, *H*→*bb*), Parton shower and hadronisation





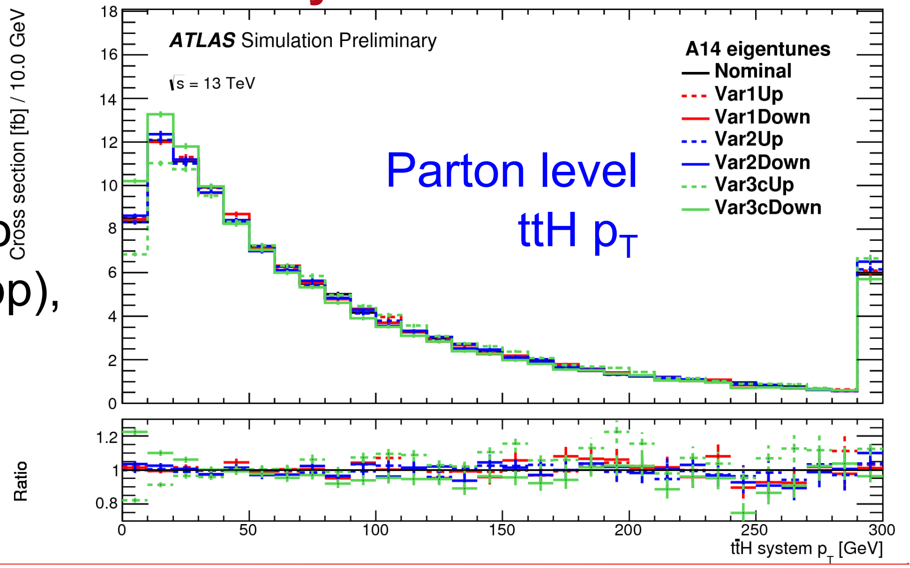


* MG5_aMC@NLO+Pythia8 prediction: slightly more events with six jets (number of expected jets for the selected channel *tt+H* with *tt*→lep+jets, *H*→*bb*). In addition, jets transverse momenta is harder.

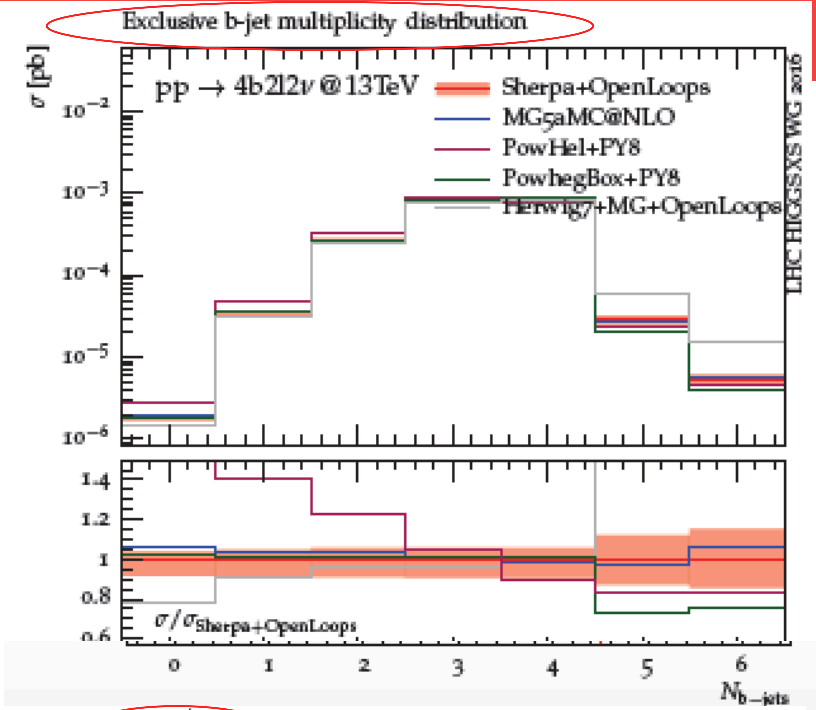
* Visible effects in low region of *tt+H* p_T spectrum due to different showering and hadronisation model (Py8/HWpp), larger than A14 Var3c (ISR) variations.

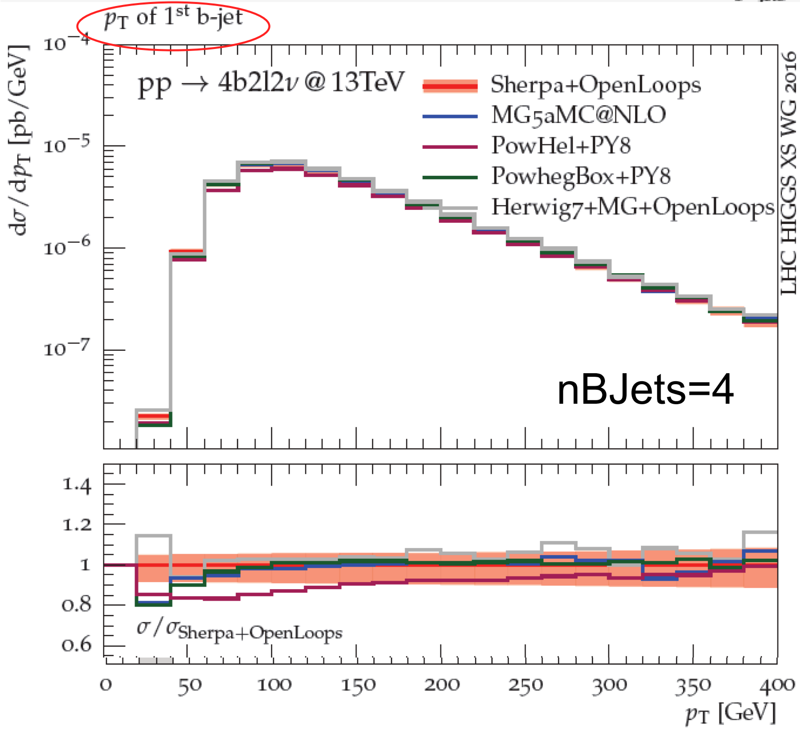
* Scale choice: main effect from μ_R, cross-section varies 9%, shape effect <1%

A14 Pythia8 tune variations



tt+H modelling: YR4 studies





nBJets=4

Yellow Report4, arXiv:1610.07922

Five NLO QCD+PS setups were compared:

- S- MC@NLO.: Sherpa(NLO)+OpenLoops +Sherpa PS
- MG5_aMC@NLO (fixed NLO)+Pythia8
- PowHel(fixed NLO)+Pythia8
- Powheg(fixed NLO)+Pythia8
- HERWIG7 using OpenLoops+MG5_aMC@NLO+Herwig7

using

5F scheme

$\mu_R=\mu_F=$ $\mu_0 = (E_T(t)E_T(\bar{t})E_T(H))^{1/3}$, where $E_T = \sqrt{M^2+p_T^2}$

$\mu_Q=H_T/2$ with $H_T = E_T(t)+E_T(\bar{t})+E_T(H)$

for samples with Sherpa and MG5_aMC@NLO

for Powheg $h = H_T/2$ in the definition of $h_{\text{damp}} = h^2/(h^2+p_T^2)$

PDF set: NLO PDF4LHC15_30

Uncertainty band: scale variations (factor 2 up/down)

Plots for pp→4b2l2v

- → Discrepancies in PowHel for nBjets<4
- → Discrepancies for nBJets>4 mainly of parton-shower origin
- → Kinematic distributions are quite compatible for nBJets=4