Heavy flavour production: experimental perspective

Andy Buckley

School of Physics and Astronomy,
The University of Glasgow

ATLAS/CMS MC workshop,
CERN, 4 May 2017

Heavy flavour production at the LHC

Heavy flavour (*b* and *c* quark) production is key to many important analyses at LHC Run 2

- Relatively unconstrained from theory: ambiguities in scale choices, resummation, mass treatment, ME/shower interface
- b production can be a crucial irreducible search background, e.g.
  - $V + b\bar{b}$ in $VH \to b\bar{b}$; and
  - $t\bar{t} + b\bar{b}$ in $t\bar{t}H \to b\bar{b}$
- Experimentally also important to understand tagging behaviour: flavour fractions and feed-in to b tagging from c jets

Experimental issues often = MC issues!
Quite a technical/bookkeeping minefield...

HF@LHC workshop in April 2016 a good forum to discuss – valuable talks and discussion, some resulting studies underway

$\Rightarrow$ **HF@LHC2 in Durham, UK from 6–8 Sept 2017!**

In this talk

Inputs from HF@LHC, new ATLAS notes, and Feb $t\bar{t} + b\bar{b}$ meeting
Many measurements, lots of MC/data, no overwhelming conclusion
$\Rightarrow$ necessarily incomplete summary

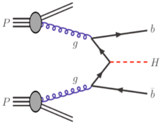
- $V + b(b)$
  ATL-PHYS-PUB-2017-006
  Measurement/constraint prospects
- **$t\bar{t}$ and $t\bar{t} + b\bar{b}$**
  ATL-PHYS-PUB-2016-016 $t\bar{t} + b\bar{b}$
  ATL-PHYS-PUB-2017-007 Sherpa and MG5_aMC@NLO
  Common meeting on tt+b-jet backgrounds to ttH(bb), 6 Feb
- $g \to b\bar{b}$ and other experimental constraints

Recurring issues:

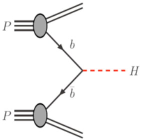
- Source of initial-state HF? 4- vs. 5-flavour matched ME/PS event simulation
  - "5F for rate/stability; 4F for kinematics" $\Rightarrow$ norm vs. shape.
  - Complicated by NLO, mass effects, and $+1 \times b/c$ bins.
- Combination/HFOR and MC systematics/disagreements

4- vs. 5-flavour

4F and 5F schemes



NNLO correction in the 5FS



4F scheme

✘ It does not resum possibly large logs, yet it has them explicitly
✘ Computing higher orders is more difficult
✔ Mass effects are there at any order
✔ Straightforward implementation in MC event generators at LO and NLO

5F scheme

✔ It resums initial state large logs into b-PDFs leading to more stable predictions
✔ Computing higher orders is easier
✘ p_T of bottom enters at higher orders
✘ Implementation in MC depends on the gluon splitting model in the PS

5

from Maria Ubiali / Fabio Maltoni

4F scheme requires event vetoing to eliminate HF double-counting by parton shower emissions. Built-in in Sherpa, *ad hoc* for MG5 (& Alpgen)

Practicalities

HFOR: 4F requires some HF Overlap Removal $\Rightarrow$ usual cut forces "shower narrow / ME wide" in ΔR. Ad hoc, incomplete, fragile! $\Rightarrow$ ATLAS nearly completely using 5F. Similar needed for combination of 5F $t\bar{t} + X$ with 4F $t\bar{t} + b\bar{b}$

Fitting flavour fractions: requires separation of samples to allow normalization floating.
For Sherpa and MG5_aMC needs *particle-level filtering.*
In all cases $\Rightarrow$ slicing. . . kinematics *and* flavour: bookkeeping!!

Large weights! (Not specific to HF, but. . .) Instability particularly with Sherpa

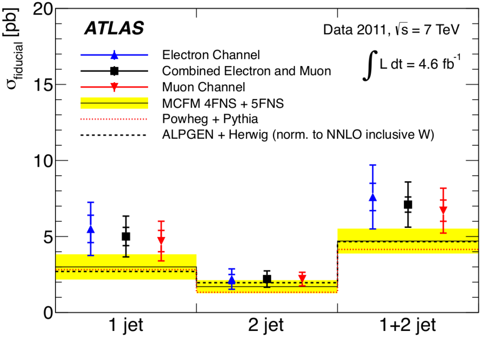
CPU usage: multileg, esp. Sherpa & NLO, are massive CPU hogs. ATLAS high-p_{T} b/c-filtered samples are slower per-event than full detector simulation!

$$V + \mathrm{HF}$$

$W + b(b)$

$V + b(b)$ is irreducible background for $VH(\to b\bar{b})$ searches; single-b important for control regions. Modelling is largest search uncertainty: lack of control from measurements. c fraction affects b-tagging.

ATLAS 7 TeV $W + b(b)$, JHEP 06 (2013) 084

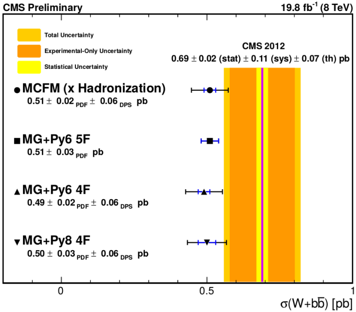


Total cross-section – 4F mismodelling in 1b bin

$W + b(b)$

$V + b(b)$ is irreducible background for $VH(\to b\bar{b})$ searches; single-b important for control regions. Modelling is largest search uncertainty: lack of control from measurements. c fraction affects b-tagging.

CMS 8 TeV $W + b\bar{b}$, CMS-PAS-SMP-14-020

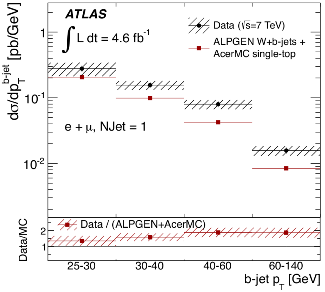


Similar inclusive result and MC/data to ATLAS

$W + b(b)$

$V + b(b)$ is irreducible background for $VH(\to b\bar{b})$ searches; single-b important for control regions. Modelling is largest search uncertainty: lack of control from measurements. c fraction affects b-tagging.

ATLAS 7 TeV $W + b(b)$, JHEP 06 (2013) 084

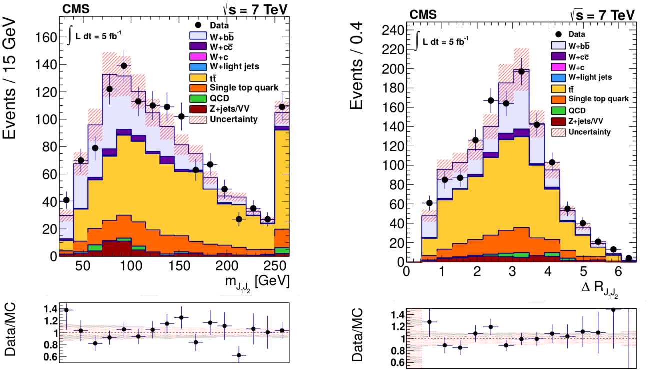


Alpgen (4F) discrepancy increasing at large p_{T}

$W + b(b)$

$V + b(b)$ is irreducible background for $VH(\to b\bar{b})$ searches; single-b important for control regions. Modelling is largest search uncertainty: lack of control from measurements. c fraction affects b-tagging.

CMS 7 TeV $W + b\bar{b}$, PLB 735 (2014) 204

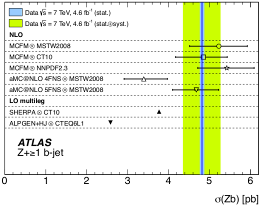


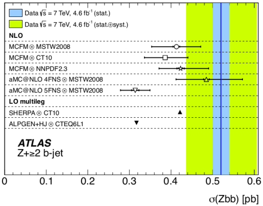
No systematic discrepancies

$Z + b(b)$

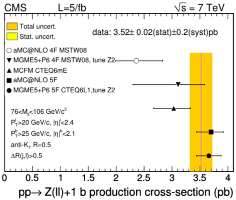
Again, key background for $VH(\to b\bar{b})$ searches

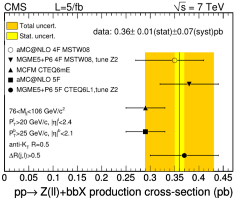
ATLAS 7 TeV
$Z + b(b)$
JHEP 10 (2014) 141





CMS 7 TeV
$Z + b(b)$
JHEP 06 (2012) 126





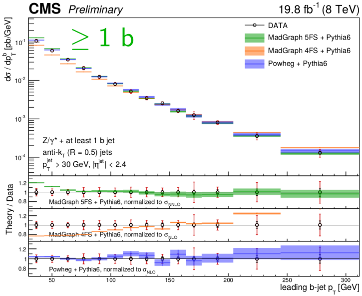
ATLAS/CMS consistent: 4F poor for $1b$, 5F dubious for $2b$
(ATLAS deviations more extreme)

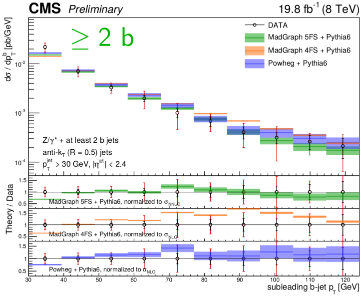
$Z + b(b)$

CMS 8 TeV $Z + b(b)$,
CMS-PAS-SMP-14-010

MC normalised: 4F to NLO, 5F to NNLO

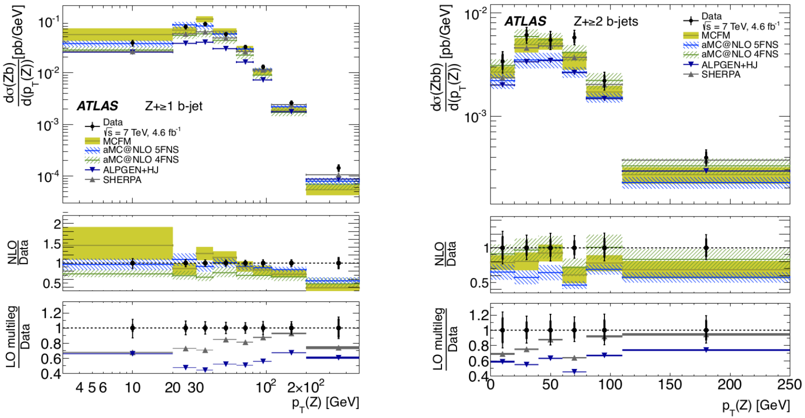
Opposite low-p_T deviations for MG+PY6 4/5F, high/mid-p_T deviation for MG 4F. POWHEG+Py8 describes well





$Z + b(b)$

ATLAS 7 TeV $Z + b(b)$, JHEP 10 (2014) 141

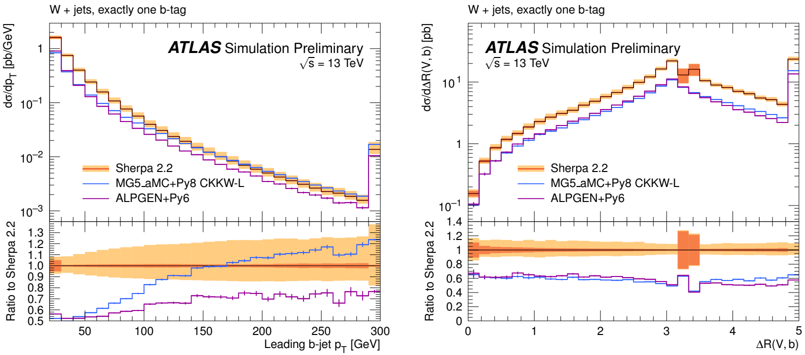


NLO deviations at high Z p_T– add more legs in Run 2

$V + b(b)$ MC comparisons

From ATL-PHYS-PUB-2017-006 MC note:

$W + b$



All samples are normalised to NNLO:

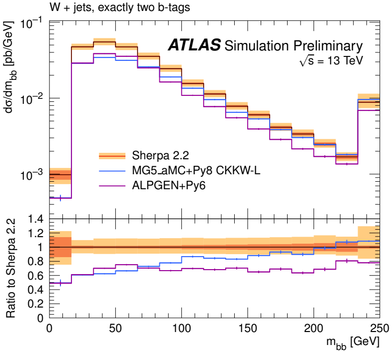
$\Rightarrow$ *factor ~ 2 normalisation difference is pure acceptance!*

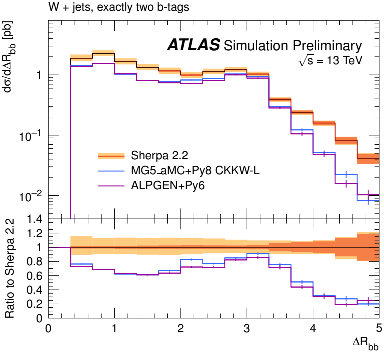
MG5_aMC@NLO much harder p_{T} than others. Large-weight artifacts!

$V + b(b)$ MC comparisons

From ATL-PHYS-PUB-2017-006 MC note:

$W + b\bar{b}$





All samples are normalised to NNLO:

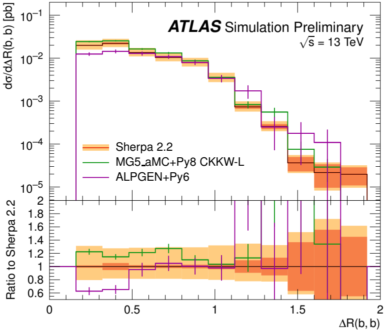
$\Rightarrow$ *factor* ~ 2 *normalisation difference is pure acceptance!*

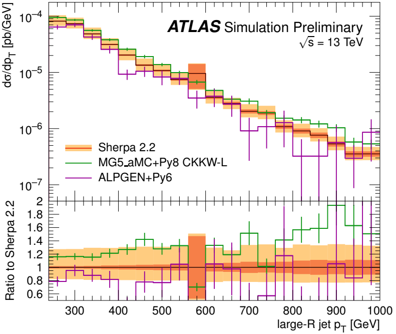
Significant shape difference from MG5 in m_{bb}, Sherpa in ΔR_{bb}

$V + b(b)$ MC comparisons

From ATL-PHYS-PUB-2017-006 MC note:

Z+b-tagged fat-jet with $R = 0.2$ matched & tagged track-jets





Low stats, but Alpgen deviates for small ΔR. Large-weight artifacts!

More V + HF

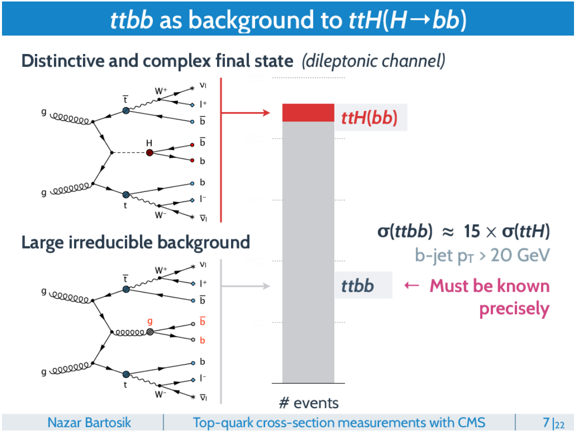
Not time here to talk about $W + c$ – sorry. Important input (with $Z + b(b)$) for PDF fits. Asymm c content?

Analyses in the pipeline: ATLAS $W/Z + b$ (resolved) and $Z + b\bar{b}$ (boosted) analyses on-going at 13 TeV

ATLAS $W + c$ (13 TeV) and $W/Z + D$-mesons (8 TeV) also in progress but on longer timescales

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

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

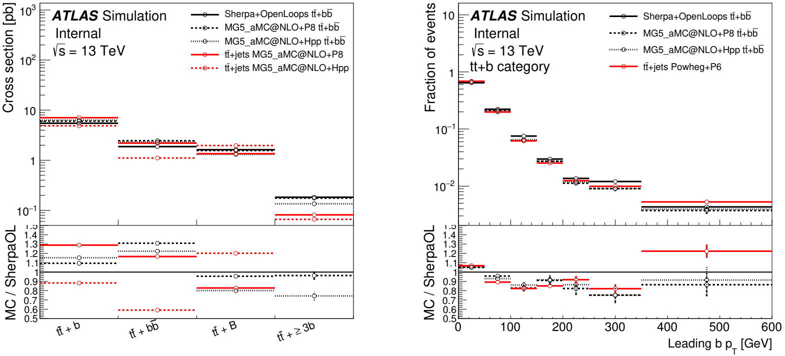


- 4F used since m_b important and all current 5F is massless.
- As search background: differences between different 4F generators larger than single-generator systematics. Under control?
- Uncertainty correlations: now agreed (? cf. S. Pozzorini talk at $t\bar{t}b\bar{b}$ meeting) to correlate within b/light categories, but uncorrelated between.

$t\bar{t} + b\bar{b}$ cross-sections

$\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj} = 1.2$–$2.2\%$ measurement vs. theory have similar uncertainties from both ATLAS and CMS:

ATLAS top-quark production PUB note, ATL-PHYS-PUB-2016-016:



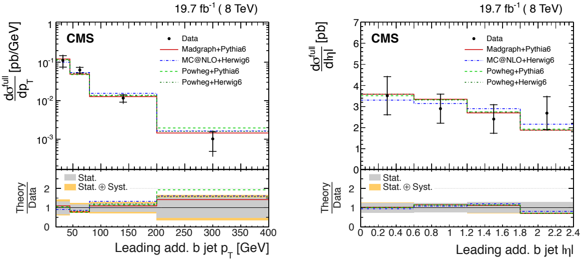
MC norm uncertainty 30%, shape uncertainties 20%.
Significant 5F/4F & shower sensitivity

$t\bar{t} + b\bar{b}$ cross-sections

CMS differential $t\bar{t} + b\bar{b}$ measurements:

Differential tt + bb: results | dileptonic

Leading additional *b* jet: $p_T > 20$ GeV, $|\eta| < 2.4$ — arXiv:1510.03072



MC predictions normalised to Data

Limited by statistical uncertainty

Well described by the considered MC predictions

Nazar Bartosik | Top-quark cross-section measurements with CMS | 17|22

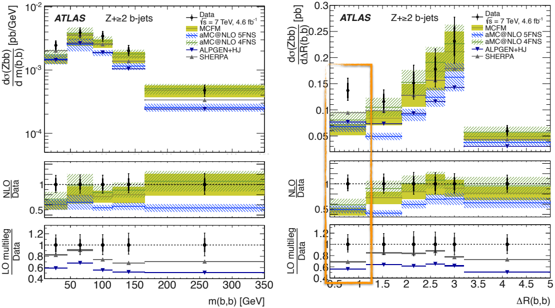
Agreement, but highly limited by stat uncertainty ⇒ Run 2

Other aspects of b production & decay

$g \to b\bar{b}$ splitting

ATLAS 7 TeV $Z + b(b)$, JHEP 10 (2014) 141

Z+b(b) – dσ/dm$_{bb}$ and dσ/dΔR(b,b)



Low ΔR(b,b) large discrepancy!

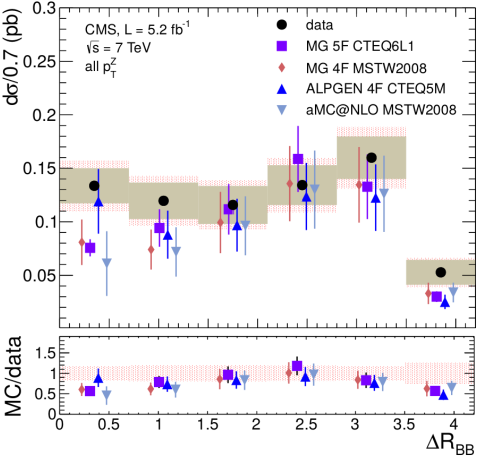
Testing here gluon splitting to bb!!! → very interesting to be followed up in Run 2!!

C. Debenedetti, UCSC/SCIPP - HF@LHC workshop, Durham, 20/4/2016 18/24

from Chiara Debenedetti

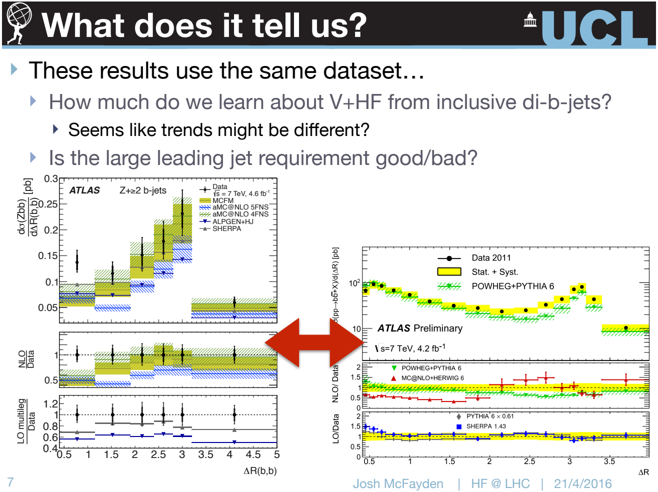
$g \to b\bar{b}$ splitting

CMS 7 TeV $Z + b\bar{b}$, JHEP 12(2013) 039



CMS sees a similar trend with Alpgen(+Herwig?) best,
but shower rather than 4F effect since 4F MG5 is poor

$g \to b\bar{b}$ splitting



from Josh McFayden

Tension between ATLAS $Z + b\bar{b}$ and di-b-jet?

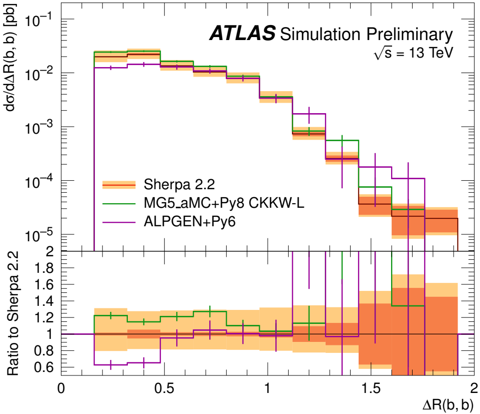
ATLAS gluon splitting measurement via $J/\psi + \mu$ not *quite* ready for this workshop! Get below jet-R limit.

$g \to b\bar{b}$ splitting measurement prospects

ATLAS gluon splitting measurement via $J/\psi + \mu$ not *quite* ready for this workshop

⇒ No jet radius: break through the below jet-R resolution limit

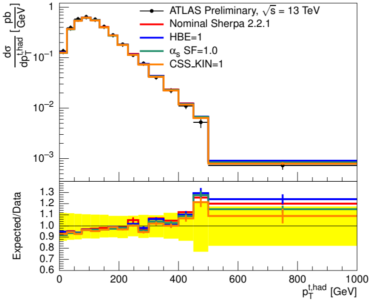
Also:

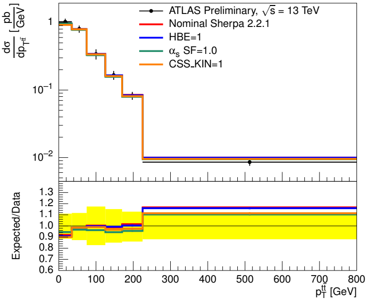


This $\Delta R_{b\bar{b}}$ in $Z + b\bar{b}$ with fat jet and $R = 0.2$ tagged subjets shown earlier is underway on 13 TeV ATLAS data.

Again, get past the calo jet resolution limit.

Heavy baryon modelling (in $t\bar{t}$)

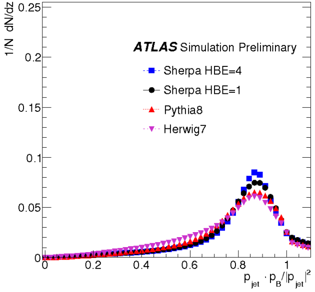




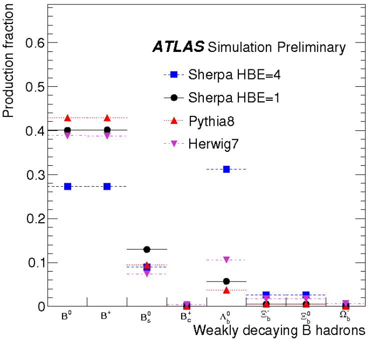
Heavy baryon modelling: hadronisation detail.

Won't affect main event kinematics, but affects decay topologies and kinematics (e.g. frag function) ⇒ tagging?

ATLAS measurement of *b*-track-jet frag functions in pipeline



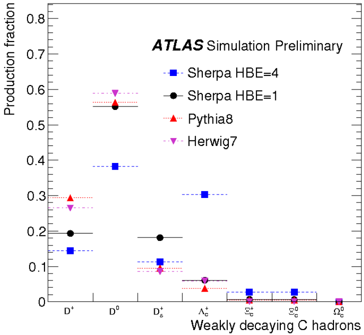
Heavy baryon modelling (in $t\bar{t}$)



Species	SHERPA v2.2 HBE=4	SHERPA v2.2 HBE=1	PYTHIA8	HERWIG7	World Average[24]
B^+	27.3	40.1	42.9	38.8	40.4 ± 0.6
B^0	27.2	40.1	42.9	38.7	40.4 ± 0.6
B_s^0	9.0	13.0	9.4	7.4	10.3 ± 0.5
Baryons	36.5	6.8	4.8	15.1	8.8 ± 1.2

Sherpa *b*-baryon enhancement is too strong by default.

Heavy baryon modelling (in $t\bar{t}$)



Species	SHERPA v2.2 HBE=4	SHERPA v2.2 HBE=1	PYTHIA8	HERWIG7	World Average[25]
D^+	14.5	19.3	29.3	26.5	22.56 ± 0.77
D^0	38.5	55.1	56.4	58.9	56.43 ± 1.51
D_s^0	11.3	18.1	9.5	8.5	7.97 ± 0.45
Baryons	35.9	7.5	4.8	6.1	10.8 ± 0.91

Sherpa *c*-baryon enhancement also too strong by default.

Summary

- **Great progress in MC/data modelling of HF since Run 1**
  - Worth remembering that we're lucky that this can be modelled at all
  - It's come at a substantial CPU price, though: MC is no-longer "free"
- **Theory errors still large for searches:** profiling helps, but…
- **Need Run 2 high-stats measurement analyses** in $V + b\bar{b}$ and $t\bar{t} + b\bar{b}$ (and more) to constrain models & squeeze systematics. Rivet analyses have been very useful for both experiment and theory.
- No big breakthroughs to announce now, but some significant improvements, and new Run 2 analyses + MC developments on their way
- **A reminder!** HF@LHC2, Durham, UK, 6-8 Sept