Experience with modelling of processes with heavy flavor in ATLAS analyses (including calibrations) - What are the key issues?

Open Day of Hbb/FTAG Workshop

Stony Brook University

Tuesday 5th September 2017

Paul Thompson

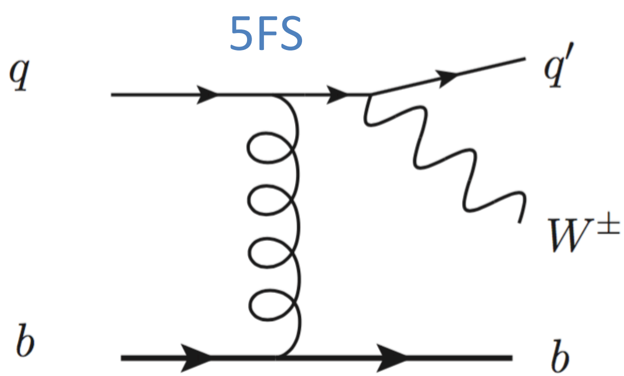
University of Birmingham

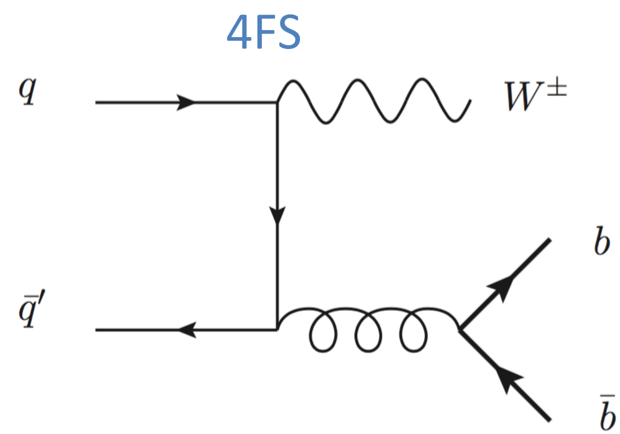
On behalf of the ATLAS Collaboration

UNIVERSITY OF BIRMINGHAM

ATLAS EXPERIMENT

Heavy Flavour at the LHC





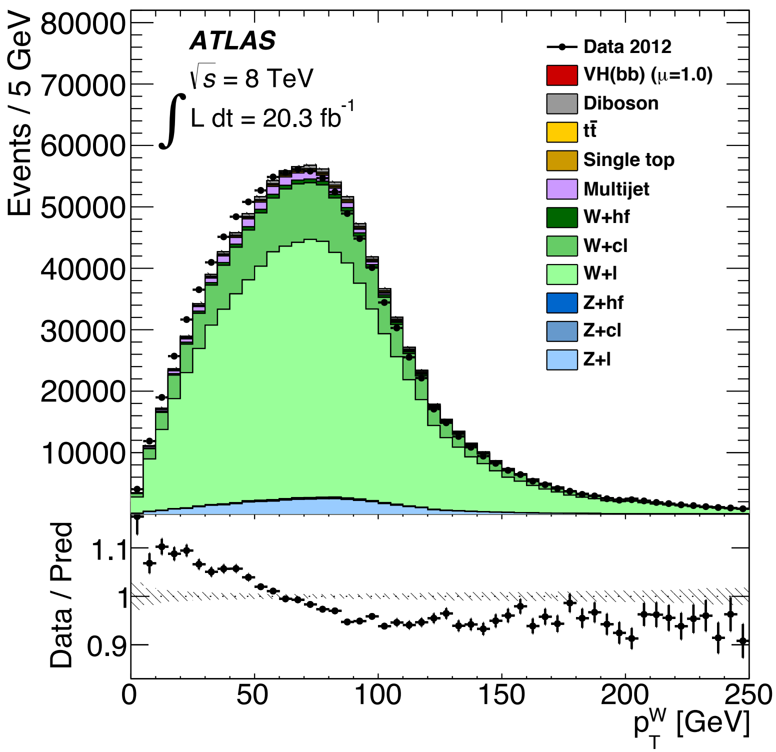
Modelling events using Monte Carlo (MC) at the LHC involves:
Matrix elements to leading order (LO) or next-to-leading order (NLO)
Higher multiplicities through parton shower (PS) or multi-leg with each multiplicity at LO.
Need to avoid double counting using matching schemes (CKKW-L, MLLM, FxFx etc.)

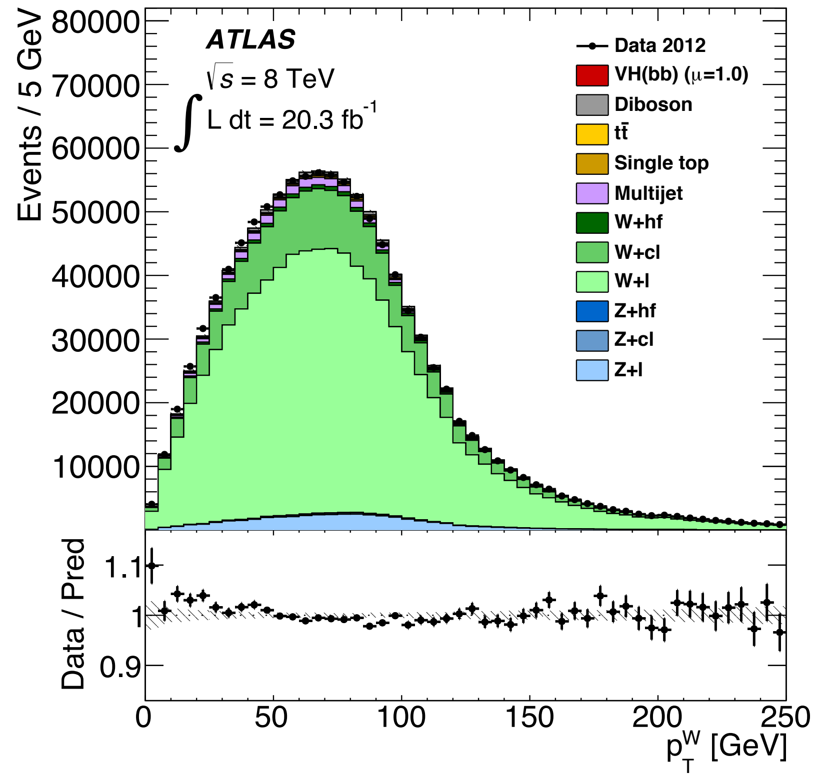
Heavy Flavour provides testing ground for QCD, multi-scale m_Z, m_W, p_T^{jet} and challenge how to deal with the heavy quark mass:
5 Flavour scheme (5FS) allows to resum large logarithmic terms into in the b-quark PDF
4 Flavour scheme (4FS) allows to take care of mass effects but does not resum any possible large logs. *Challenging to produce inclusive flavour MCs avoiding overlaps of HF in ME and PS*

In this short introduction just focus on SM VHbb analysis and two areas namely V+HF jet production and single top Wt

V+Jets Modelling in Run 1

JHEP01(2015)069





Used leading order 4FS Sherpa v1.4

For both W+jets and Z+jets had mis-modellings in p_T^V and $\Delta\phi jj$.

Applied various reweights depending on boson type and number of b-tags

Example shown is p_T^W with 0 b-tags after applying reweight as a function of $\Delta\phi jj$

Each mis-modelling led to reweight plus sizeable systematic uncertainties in the corrections

ATLAS MC Notes

ATLAS (via its Physics Modelling Group – PMG) continues to provide up-to-date documentation of the ATLAS MC status – existing/planned models/systematic prescriptions. Generally aim to be in time for the various experimental-theoretical workshops throughout the year so that can get feedback from theoretical community.
https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MCPublicResults

Recent examples of MC notes :

- ATLAS simulation of boson plus jets processes in Run 2

  https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-063/

- Multi-Boson Simulation for 13 TeV ATLAS Analyses

  https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-005/

- Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO

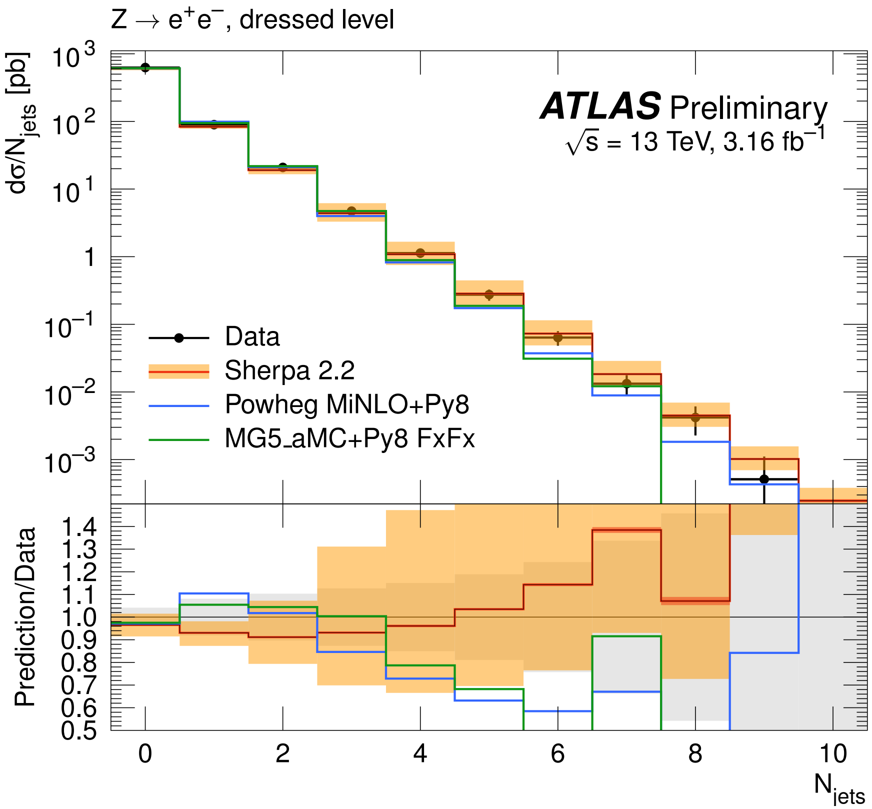
  https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2017-007/

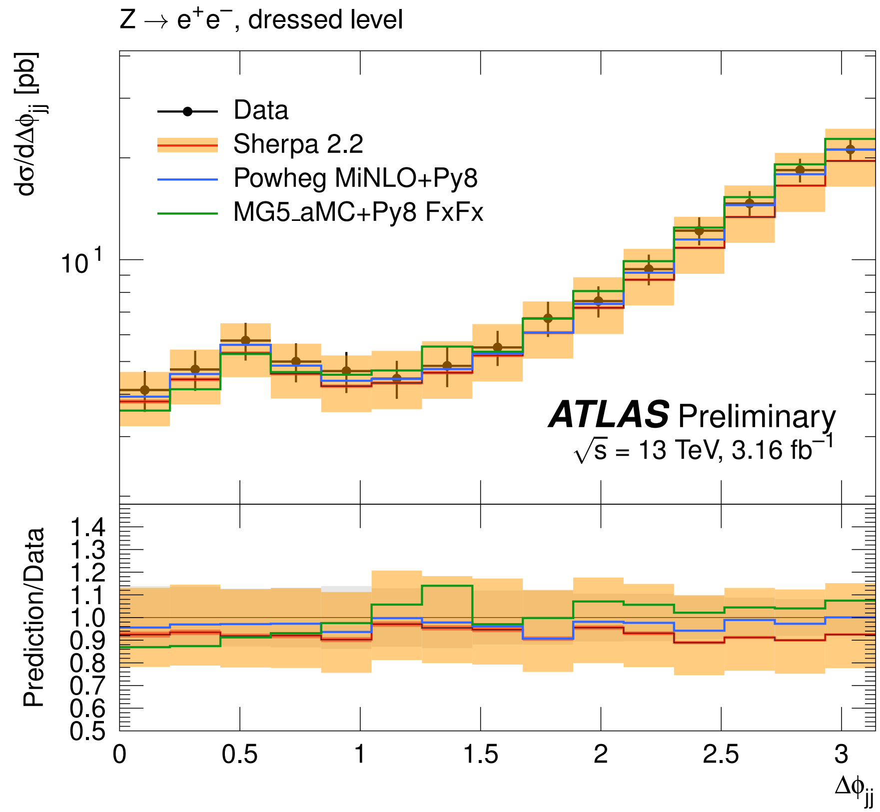
- Modelling of the $t\bar{t}H$ and $t\bar{t}V$ ($V = W, Z$) processes for $\sqrt{s}$ = 13 TeV ATLAS analyses

  https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-005/

ATLAS MC Note

V+Jets Modelling in Run 2





Comparison with recent 13TeV 2015 Z+jets data (arXiv: 1702.05725)

Reasonable description of data by latest NLO models:

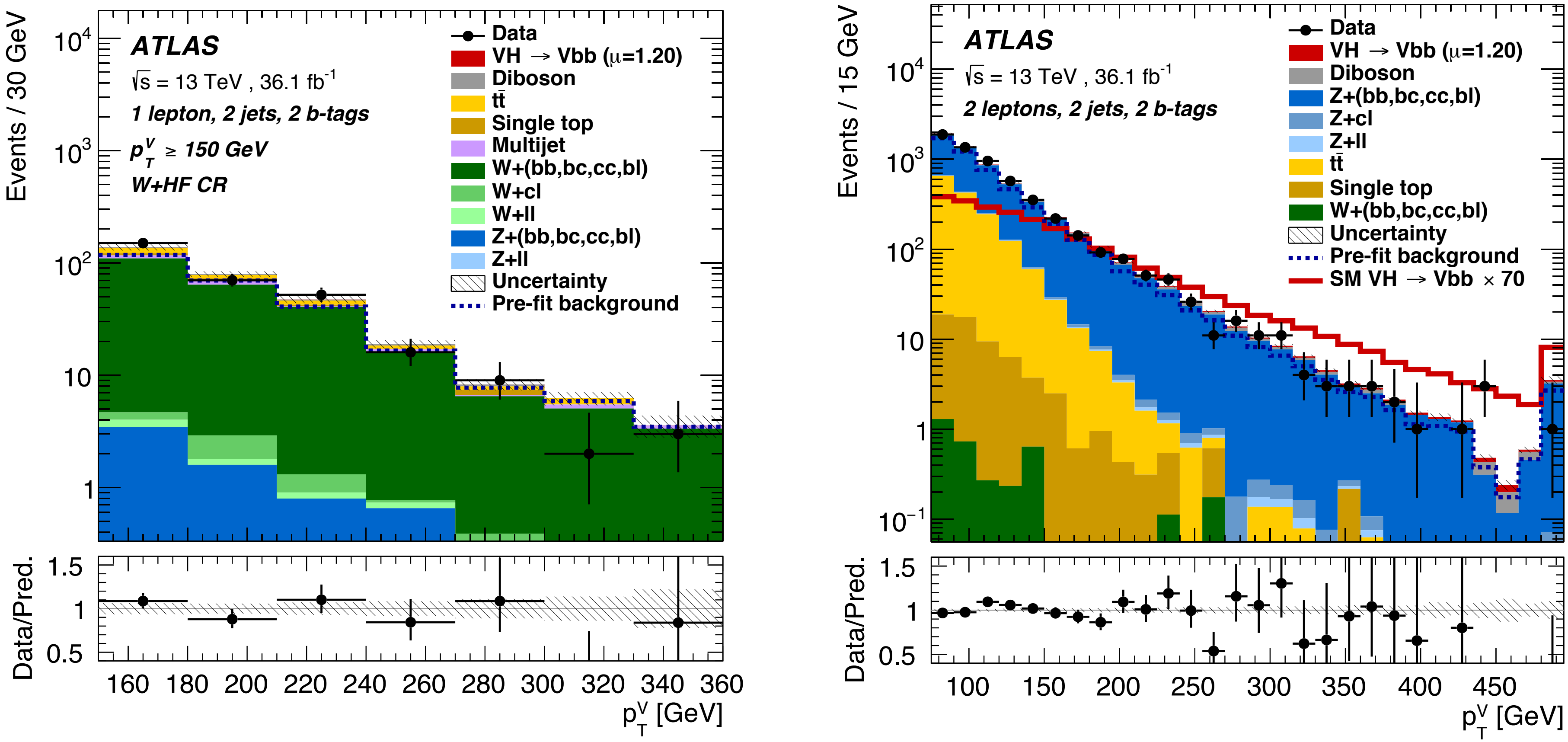
Sherpa 2.2 MEPS@NLO, NLO for V+0,1,2 jets, LO multi-leg up to 4. NNLO PDF

Powheg MinNLO+Pythia8, with NLO accuracy for Z+1 jet

MadGraph5_aMC@NLO+Pythia8 using NLO up to 2 jets and FxFx merging with PS

Run 2 V+Jets HF Modelling

https://arxiv.org/abs/1708.03299



Can obtain relatively pure control regions for W+HF and Z+HF

Sherpa v2.2 does a good job of describing shapes (no need for reweights). Fit requires adjustment of predicted normalisations (see over)

Systematic uncertainties: Zbb from data, Wbb from MC (data CR is at low m_{BB})

W/Z Background Model

Free Floating Normalisations

Process	Normalisation factor
W + HF 2-jet	1.22 ± 0.14
W + HF 3-jet	1.27 ± 0.14
Z + HF 2-jet	1.30 ± 0.10
Z + HF 3-jet	1.22 ± 0.09

Data show need for better theoretical understanding - or agrees within full theoretical error?
– Central value difference is the choice of scale, scheme (5FS vs 4FS), order of pQCD or choice of mass or PDF or fragmentation parameters or…?

Z+jets	
$Z + ll$ normalisation	18%
$Z + cl$ normalisation	23%
$Z + bb$ normalisation	Floating (2-jet, 3-jet)
$Z + bc$-to-$Z + bb$ ratio	30-40%
$Z + cc$-to-$Z + bb$ ratio	13-15%
$Z + bl$-to-$Z + bb$ ratio	20-25%
0-to-2 lepton ratio	7%
p_T^V, m_{bb}	S
W+jets	
$W + ll$ normalisation	32%
$W + cl$ normalisation	37%
$W + bb$ normalisation	Floating (2-jet, 3-jet)
$W + bl$-to-$W + bb$ ratio	26% (0-lepton) and 23% (1-lepton)
$W + bc$-to-$W + bb$ ratio	15% (0-lepton) and 30% (1-lepton)
$W + cc$-to-$W + bb$ ratio	10% (0-lepton) and 30% (1-lepton)
0-to-1 lepton ratio	5%
W+HF CR to SR ratio	10% (1-lepton)
p_T^V, m_{bb}	S

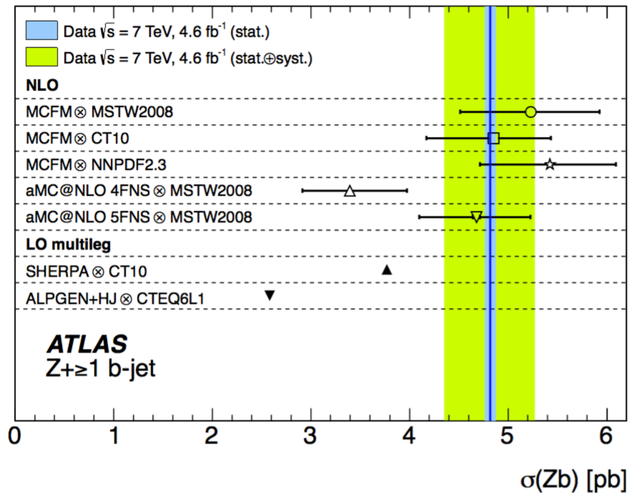
Free floating uncertainties

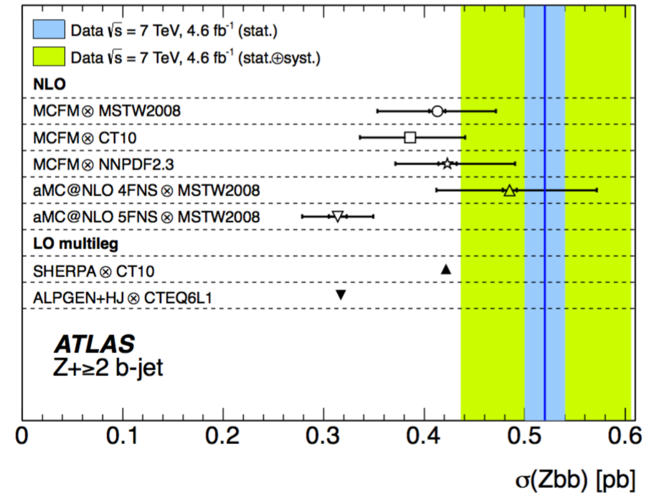
Extrapolation to 0-lepton

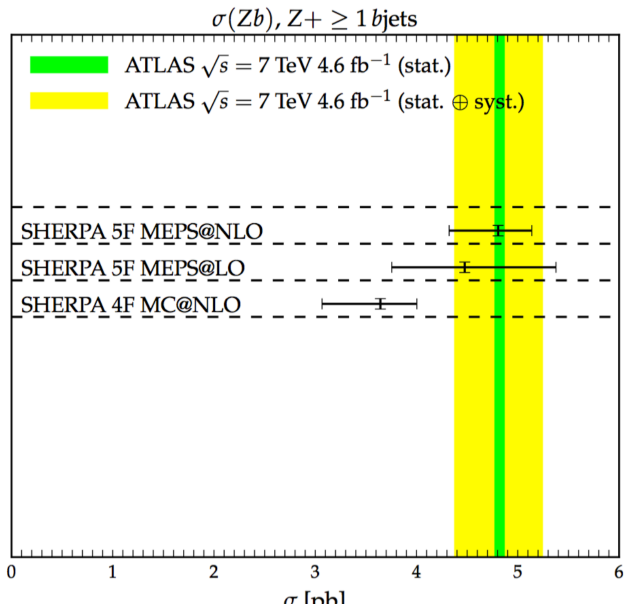
Errors on flavour ratios taken from comparison of models
As is p_T^V,m_{bb} shape (S) uncertainty in W+jets

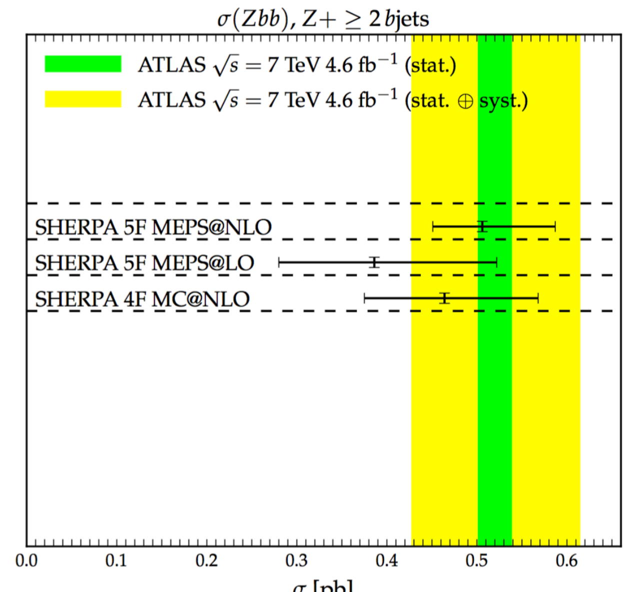
Extrapolation from W+HF to SR

V+Jets HF Modelling Run 1 Data (revisited)





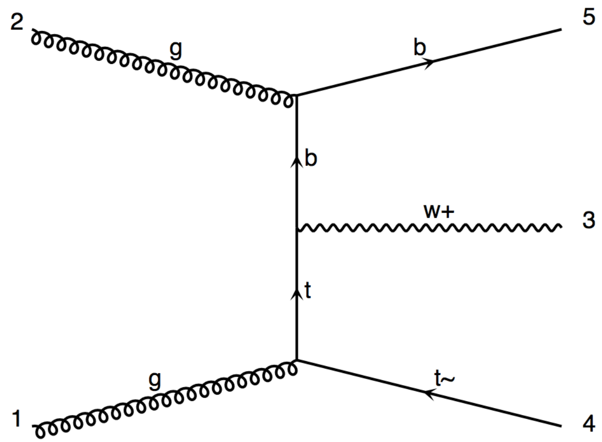




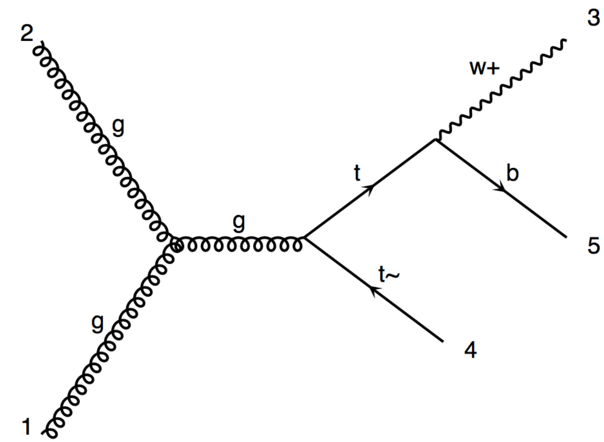
Models from original paper:Zbb central values lower, although agree within uncertainties
Comparison with Sherpa MEPS@NLO (http://arxiv.org/abs/arXiv:1612.04640) also reasonable.
V+Heavy Flavour data analyses from Run 2 will be vital to continue studies…

Single Top (Wt)

Singly resonant (sr)



Doubly resonant (dr)

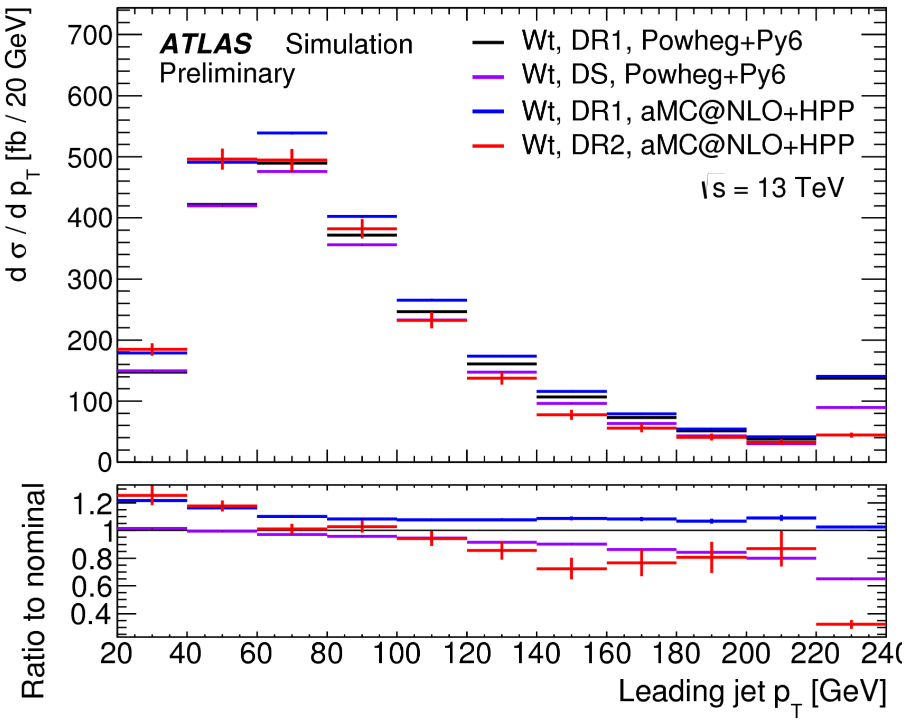


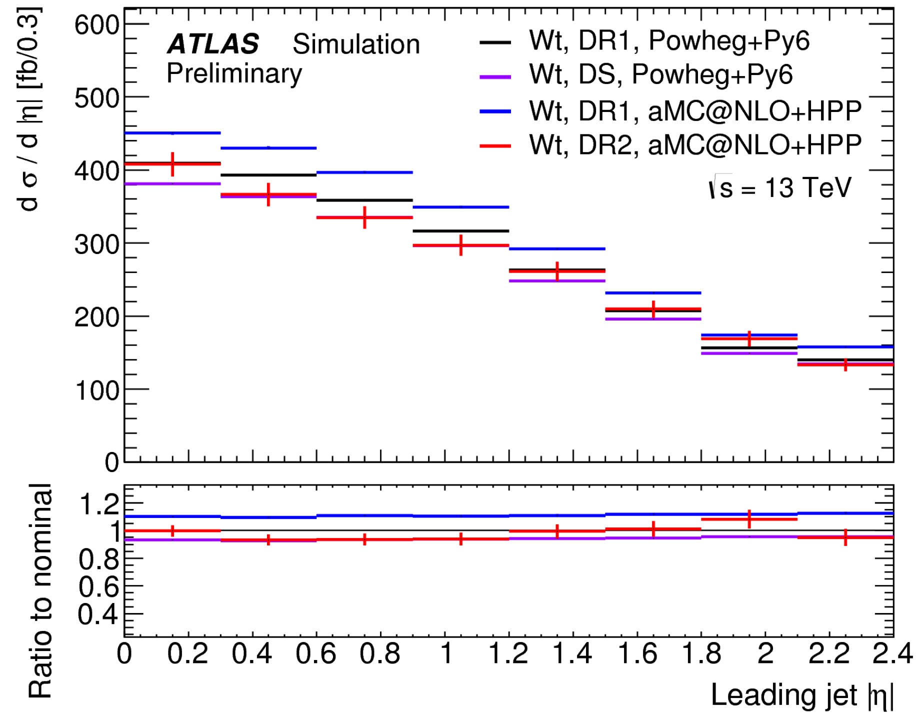
$$|\mathcal{M}_{\text{tot}}|^2 = |\mathcal{M}_{\text{sr}}|^2 + 2\text{Re}(\mathcal{M}_{\text{sr}} \cdot \mathcal{M}_{\text{dr}}^*) + |\mathcal{M}_{\text{dr}}|^2$$

Single top Wt and t-channel contributes background to the 1-lepton channel:
Systematic errors on m_{bb} and p_T^V shapes estimated from comparing MC models with different implementations of ME, parton shower modelling/fragmentation
Wt at NLO has singly-resonant and doubly resonant contributions (interferes with LO ttbar)
Powheg+Pythia6 used in VHbb analysis with Diagram Removal 1 (DR1) which sets M_{dr}=0
Interference estimated using Diagram Subtraction (DS)
Uncertainty taken from largest differences and for Wt comes from DS
MadGraph5 aMC@NLO also implements DR1 and assesses interference using alternative DR2

Single Top (Wt)

ATLAS MC Note https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2016-020/





Generator comparison for dilepton selection:

For leading jet p_T similar prediction for *DR1 Powheg+Pythia6* and DR1 MG5_aMC@NLO+HPP

Also similar interference estimate between DS Powheg+Pythia6 and DR2 MG5_aMC@NLO+HPP

Interference effects visible in leading jet p_T but not leading jet eta

Key Issues (summary)

- Monte Carlo generators vital for understanding modelling of important backgrounds to new physics and improving our understanding of underlying QCD as well as parameter tuning
- Focussing on VHbb issues: V+HF and single top systematics as key issues

- *Do we understand the likely V+HF scale factor difference of ~1.2? Should we be concerned or is it reasonable within the theoretical uncertainties (Note: ttbar SF~1)?*
- We use multi-leg 5FS up to 2jets at NLO with B-hadron filter (3 jets@NLO takes too long)
- For the W+HF systematic uncertainties we use MC model comparisons
- Can developments in 4FS including Wbb+j @NLO be used to help to reduce this error?
- *Can we boost our HF statistics more directly and/or improve our Vbbj precision?*

- The single top systematic error contributes as much as errors from more dominant backgrounds
- Largest uncertainty is from Wt and ttbar interference effects
- *Is it possible to have the interference calculation (and therefore smaller errors) accounted for in our simulations?*

Back up

V+jets Modelling in Run 2

- Updated Run 2 models originally tuned on published 7/8 TeV Run-I data, see how they compare at 13TeV

5FS Models used by ATLAS:

- MadGraph5+Pythia8. Leading order in ME up to 4 jets , PS beyond. Different final state parameter tunings "A" and "B" and move from LO to NLO PDF
- Sherpa v2.2 NLO for V+0,1,2 jets, LO multi-leg up to 4 jets, PS for higher multiplicities.

Sample	W, 7 TeV		W, 13 TeV		Z, 7 TeV			Z, 13 TeV		
	σ [pb]	k-fac	σ [pb]	k-fac	σ [pb]	k-fac	f	σ [pb]	k-fac	f
NNLO	10455	-	20080	-	964	-	-	1906	-	-
Sherpa 2.1	11270	0.928	22108	0.9083	1150	0.912	1.08	2290	0.9013	1.08
Sherpa 2.2	10600	0.986	-	-	1080	0.981	1.08	-	-	-
MadGraph+Pythia8 A	8290	1.26	16707	1.21	827	1.287	1.10	1710	1.239	1.10
MadGraph+Pythia8 B	8960	1.17	17880	1.123	907	1.17	1.10	1840	1.15	1.10

- Also used Alpgen in 4FS with overlap between HF and LF samples removed
- Also investigated aMC@NLO with FxFx merging. NLO for V+0,1,2 jets and PS beyond
- And also Powheg MiNLO with NLO for V+1jet

Now have Run 2 ATLAS SM V+jets measurement (inclusive - not yet HF)…

Impact of Systematics on μ

Source of uncertainty		σ_μ
Total		0.39
Statistical		0.24
Systematic		0.31
Experimental uncertainties		
Jets		0.03
E_T^{miss}		0.03
Leptons		0.01
b-tagging	b-jets	0.09
	c-jets	0.04
	light jets	0.04
	extrapolation	0.01
Pile-up		0.01
Luminosity		0.04
Theoretical and modelling uncertainties		
Signal		0.17
Floating normalisations		0.07
Z+jets		0.07
W+jets		0.07
$t\bar{t}$		0.07
Single top-quark		0.08
Diboson		0.02
Multijet		0.02
MC statistical		0.13

Limiting factors
- Signal modelling
- Monte Carlo statistics
- Flavour tagging
- Background modelling

Systematically Limited

Signal modelling dominated by extrapolation from high p_T^V to full phase space and showering Pythia 8 vs Herwig 7. Doesn't affect significance

Monte Carlo Stats despite flavour filtering and p^{TV} slicing

Background model, all contribute similar level
Will improve as we collect more data but theoretical progress crucial too
Surprising that single top contributes (see over)

Top Background Model

Free floating uncertainties

Process	Normalisation factor
$t\bar{t}$ 0- and 1-lepton	0.90 ± 0.08
$t\bar{t}$ 2-lepton 2-jet	0.97 ± 0.09
$t\bar{t}$ 2-lepton 3-jet	1.04 ± 0.06

$t\bar{t}$ (all are decorrelated between the 0+1 and 2-lepton channels)	
$t\bar{t}$ normalisation	Floating (0+1 lepton, 2-lepton 2-jet, 2-lepton 3-jet)
0-to-1 lepton ratio	8%
2-to-3-jet ratio	9% (0+1-lepton only)
W+HF CR to SR ratio	25%
p_T^V, m_{bb}	S

Note: Theory agrees with data within errors

Extrapolation uncertainties analogous to W/Z