Measurement of associated production of Z boson with b-jets in the ATLAS experiment

Results released in arXiv:2003.11960, submitted to JHEP

Semen Turchikhin

Dzhelepov Laboratory of Nuclear Problems, JINR

Dzhelepov Laboratory of Nuclear Problems seminar
8 April 2020

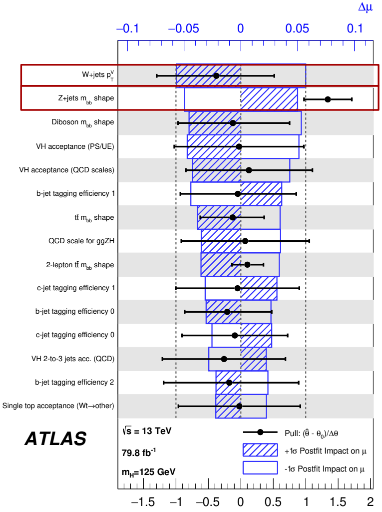
V+heavy-flavour (HF) jets production measurements: motivation

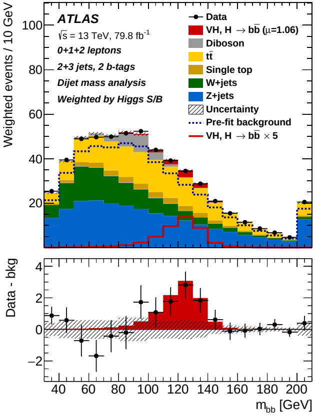
- Test of perturbative QCD predictions
  - Available at *NLO precision* for a while
  - Calculations performed within 4- or 5-flavour number scheme
    - 4FNS: b quark appears in (massive) final state from $g \to b\bar{b}$
    - 5FNS: allow b quark density in initial state, typically massless

  - Can constrain heavy quark PDF
    - In case of c quark – sensitivity to *intrinsic charm* component
    - Beyond the scope of the current measurement, but very much anticipated
- Benchmark for Monte Carlo (MC) generators
  - Commonly used for background modelling in Higgs studies and BSM searches

Introduction

- V+HF jets is an important background for
  - $VH(b\bar{b})$ study
  - BSM searches with leptons and HF jets
- $VH(b\bar{b})$ analyses systematically limited, V+HF background modelling is a dominant one
  - estimated from MC-to-MC comparison or data-driven
  - unfolded measurement could suggest a better strategy





PLB 786 (2018) 59, ATLAS observation of $H \to b\bar{b}$ and VH production

Introduction

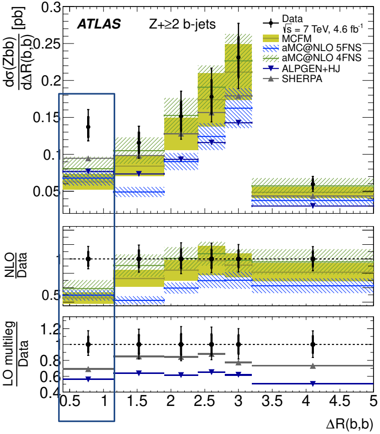
Earlier measurements

- First measurements of $Z + b$-jets and $W + b$-jets in pp collisions at Tevatron (CDF, D0)
- ATLAS: $W + b$-jets and $Z + 1, 2b$-jets at $\sqrt{s} = 7$ TeV
- CMS: same processes and $Z + 1, 2b$-jets at $\sqrt{s} = 8$ TeV
- CMS also measured $Z + c$-jets production at $\sqrt{s} = 7$ TeV and $\sigma(Z + c)/\sigma(Z + b)$ ratio at $\sqrt{s} = 13$ TeV

Measurements limited both statistically and systematically

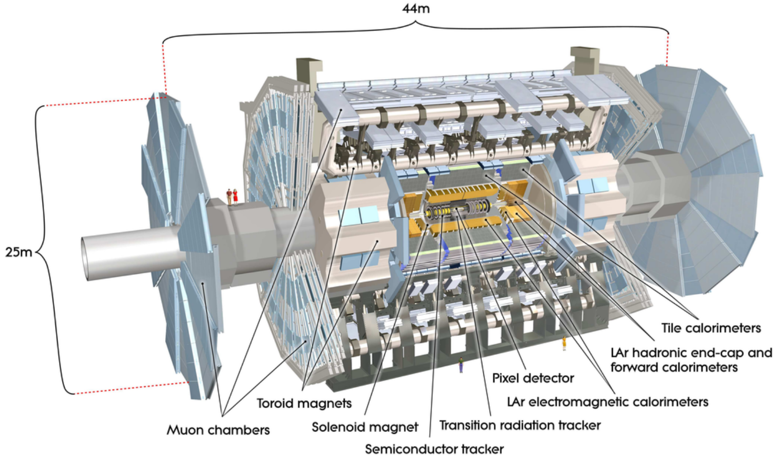
Goals of this measurement

- Inclusive cross-sections for $Z+ \geq 1b$-jet, $Z+ \geq 2b$-jets
- Differential cross-sections:
  - $Z+ \geq 1b$: p_{T} and $|y|$ of leading b-jet and Z, $\Delta\phi_{Zb}$, Δy_{Zb}, ΔR_{Zb}
  - $Z+ \geq 2b$: $p_{\mathrm{T}}(Z)$, m_{bb}, $\Delta\phi_{bb}$, Δy_{bb}, ΔR_{bb}, $p_{\mathrm{T}\,bb}$, $p_{\mathrm{T}\,bb}/m_{bb}$
- Use Run-2 data of 2015–16, **35.6 fb^{-1} @ 13 TeV**
- Consider $Z \to \mu^+\mu^-$ and $Z \to e^+e^-$ channels



JHEP 10 2014) 141, ATLAS $Z + b$-jets measurement at $\sqrt{s} = 7$ TeV

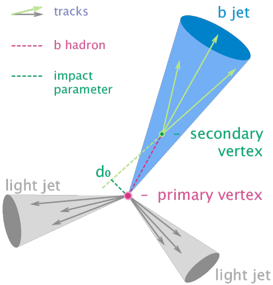
Noticeable disagreements between data and all predictions at small ΔR_{bb}, m_{bb}



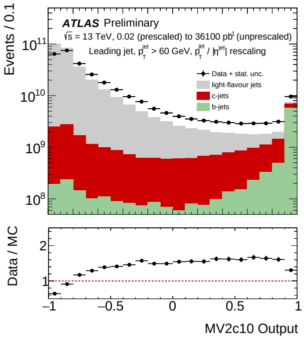
- Inner Detector coverage $|\eta| < 2.5$
- Insertable B-Layer installed for Run-2, improved *b*-jet tagging performance (factor ∼4 light jet rejection)

Two-level trigger system

- Hardware **Level-1**
  - Uses fast calorimetry and muon chambers information
  - 100 kHz output
- Software **High-level trigger**
  - Full detector information in *Regions-of-Interest* available
  - ∼1000 Hz average output



- Jet flavour tagging is based on *b* hadron decay signatures: displaced vertex, high impact parameter tracks, semileptonic decays
- Various algorithms combined in a multivariate classifier
- MV2c10 algorithm (JINST 11 (2016) 04008) is used in the measurement
- Specific selections based on MV2c10 output – working points (WP)
  - Have calibrated *b*-tagging and *c*, light jet mis-tagging efficiency

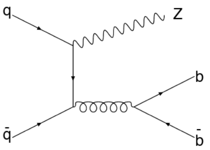


WP	Cut value X	b-jet efficiency (ε_b)	c-jet mistag rate (ε_c)	LF-jet mistag rate ($\varepsilon_{\mathrm{LF}}$)
85%	0.1758	85%	32%	2.9%
77%	0.6459	77%	16%	0.77%
70%	0.8244	70%	8.3%	0.26%
60%	0.9349	60%	2.9 %	0.065%
50%	0.9769	50%	0.94 %	0.017%

Strategy and event selection

Analysis flow

- Select events with 2 leptons and $\geq 1, 2$ *b*-tagged jets
- Evaluate and subtract background contributions
- Run unfolding to particle-level, in *fiducial volume* close to the detector-level selection
- Compare with theoretical predictions



} isolated electrons or muons
p_T>27 GeV, |η|<2.5 (2.47 in e-channel, no crack region)

} b-jets p_T>20 GeV, |y|<2.5
ΔR(jet, lep)>0.4

	Regions			
	Pre-tag region	Signal regions	Z+jets Validation Region	$t\bar{t}$ Validation Region
Leptons	2 same-flavour, opposite-charge			1 e, 1 μ, opposite-charge
$m_{\ell\ell}$	76 GeV $< m_{\ell\ell} <$ 106 GeV			
E_T^{miss}	**to reduce ttbar** $E_T^{miss} <$ 60 GeV if $p_T^{\ell\ell} <$ 150 GeV			
Jets	≥ 1 or ≥ 2 jets			
b-tagging efficiency working point selection	-	70%	≥ 1 b-jet at 77%–70%	70%
Number of b-jets	-	≥ 1 b-jets (1-tag region) ≥ 2 b-jets (2-tag region)	≥ 1 b-jets	≥ 2 b-jets

Background sources

- $Z + c$- or light jets
  - Shapes from MC (validated with data), normalization from flavour fit to data
- Di-leptonic $t\bar{t}$ (+ single-top) events (dominant in 2-tag region)
  - Use MC, validate with data control region ($e^{\pm}\mu^{\mp}$)
- Di-boson, $V + H$ production, $Z \to \tau^{+}\tau^{-}$, W+jets
  - Small contribution, estimate with MC
- QCD multi-jet production
  - Templates derived in enriched control regions (loose lepton ID requirement)
  - Fit templates for $m_{\ell\ell}$ distribution $\to$ negligible contribution found

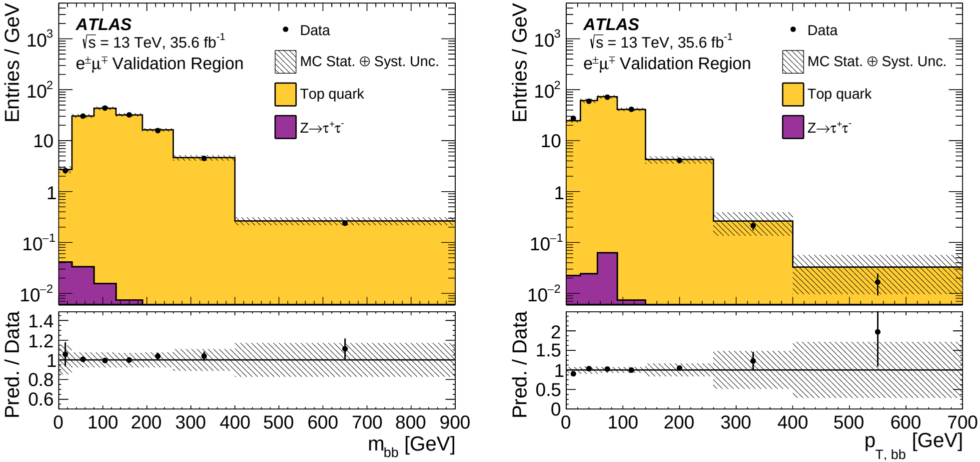
1-tag region	
Signal	
$Z + b$, $Z + bb$	59%
Backgrounds	
$Z + c$	18%
$Z + l$	18%
Top	4%
Diboson, VH	1%
Others	<1%
Total predicted	470 000 ± 650
Data	499 645

2-tag region	
Signal	
$Z + bb$	60%
Backgrounds	
$Z + b$	9%
$Z + c$	5%
$Z + l$	<1%
Top	23%
Diboson, VH	2%
Others	1%
Total predicted	33 070 ± 180
Data	36 548

Process	Generator	Order of cross-section calculation	Reference normalisation	Normalisation cross-section uncertainty
$Z \to \ell\ell$ ($\ell = e, \mu, \tau$) with $66 < m_{\ell\ell} < 116$ GeV	Sherpa	NNLO	[44–47]	5%
$W \to \ell\nu$ ($\ell = e, \mu, \tau$)	Sherpa	NNLO	[44–47]	5%
$t\bar{t}$	Powheg-Box	NNLO + NNLL (m_{top} = 172.5 GeV)	[55–61]	6%
Single top (t-, Wt-, s-channel)	Powheg-Box (m_{top} = 172.5 GeV)	NLO		6%
Dibosons $Z(\to \ell\ell) + Z(\to qq)$, $W(\to \ell\nu) + W(\to qq)$)	Sherpa	NLO	[69]	5%
Higgs ($qq \to Z(\to \ell\ell) + H(\to b\bar{b})$ $gg \to Z(\to \ell\ell) + H(\to b\bar{b})$ $qq \to W(\to \ell\nu) + H(\to b\bar{b})$)	Powheg-Box	NNLO QCD + NLO EW NLO + NLL NNLO QCD + NLO EW	[73–75]	3%

$t\bar{t}$ background

- MC validated in a control region
  - opposite-charge electron+muon, ≥ 2 b-tagged jets
- Perfect agreement found within the uncertainties of $t\bar{t}$ production modelling



Z+jets flavour fit

Flavour fit: maximum-likelihood fit to data based on a flavour-sensitive distribution – b-tagging discriminant (MV2c10) output, to extract normalization for Z+jets background

- Done separately for 1-tag and 2-tag regions
- 1-tag
  - Discriminating variable: leading b-tagged jet MV2c10 output
    - Signal template: $Z+\geq 1b$-jet
    - Background – single template: $Z+c$ and Z+light jets
- 2-tag
  - Discriminating variable: combination of MV2c10 outputs for two leading b-jets
    - Signal template: $Z+\geq 2b$-jet
    - Background – single template: $Z+1b$, $Z+c$ and Z+light jets

1-tag

Generator	Signal SF	Z+jets background SF	Signal post-fit yield	Z+jets background post-fit yield	Signal + Z+jets post-fit yield
SHERPA	1.109 ± 0.003	0.861 ± 0.004	309 650 ± 810	166 640 ± 650	476 290 ± 750
ALPGEN	1.480 ± 0.004	1.015 ± 0.002	297 670 ± 740	178 100 ± 400	475 810 ± 480

2-tag

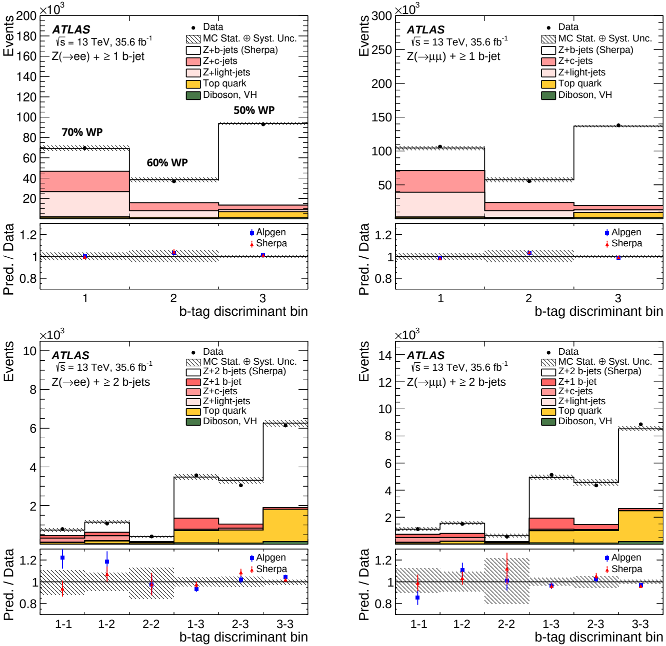
Generator	Signal SF	Z+ jets background SF	Signal post-fit yield	Z+ jets background post-fit yield	Signal + Z+jets post-fit yield
SHERPA	1.18 ± 0.01	1.08 ± 0.04	23 440 ± 250	4780 ± 180	28 220 ± 200
ALPGEN	1.18 ± 0.01	1.30 ± 0.05	23 650 ± 240	4550 ± 180	28 200 ± 200

Z+jets flavour fit

- Simultaneous fit of electron and muon channels
- Use binning of MV2c10, corresponding to the calibrated working points

1-tag

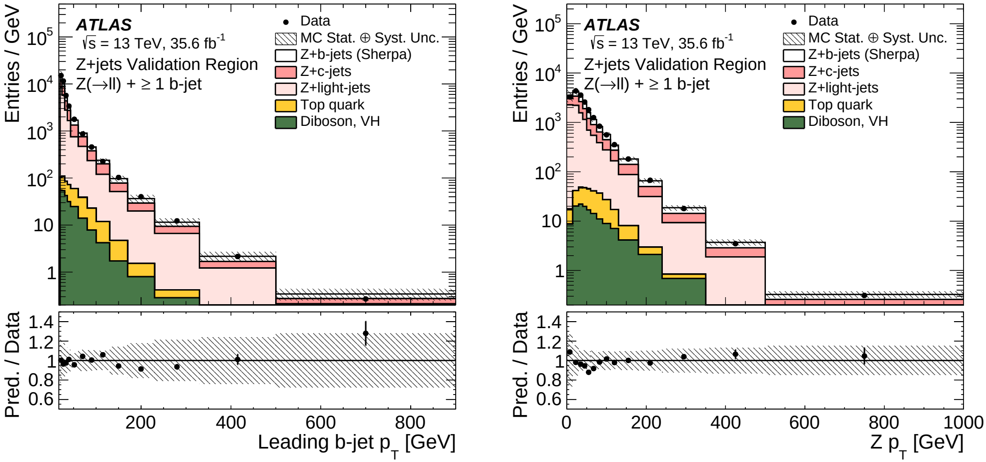
2-tag



- Split the background templates to study systematics
- 1-tag: into $Z + c$ and Z+light
- 2-tag: into $Z + 1b$ and $Z + c$,light

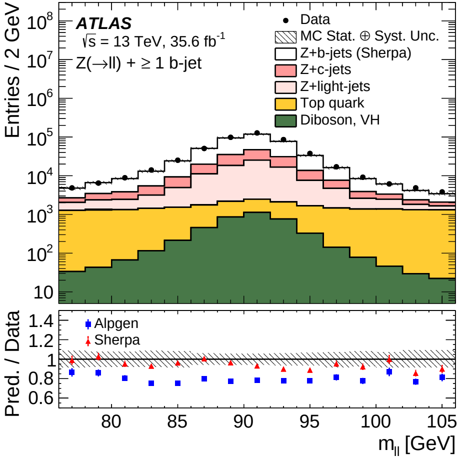
Z+jets validation region

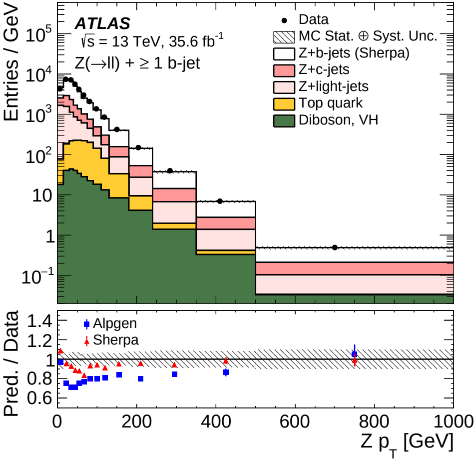
- Define the region enriched with $Z + c$, Z+light jets
  - Require ≥ 1 b-tagged jets passing the MC2c10 cut between 77% and 70% efficiency WPs
  - c- (light) jets mis-ID rate is 7.7% (0.51%)
- $Z + c$ and Z+light jets constitute 50% and 28% of the sample, respectively
- Perfect agreement found within the flavour-tagging uncertainty



Reconstruction-level distributions for $Z+\geq 1b$-jet

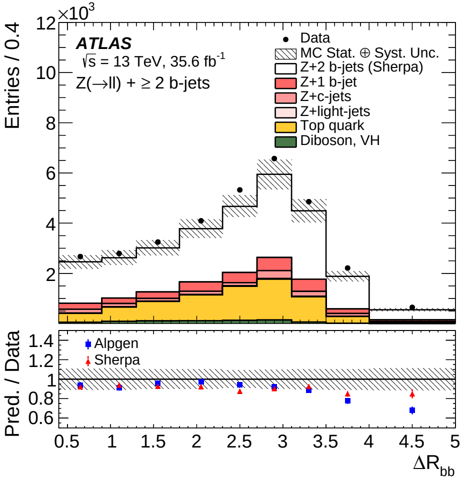
- Normalization of Z+jets background MC corrected according to the flavour fit scale factors
- Signal $Z+\geq 1b$ MC not corrected
- Electron and muon channels combined

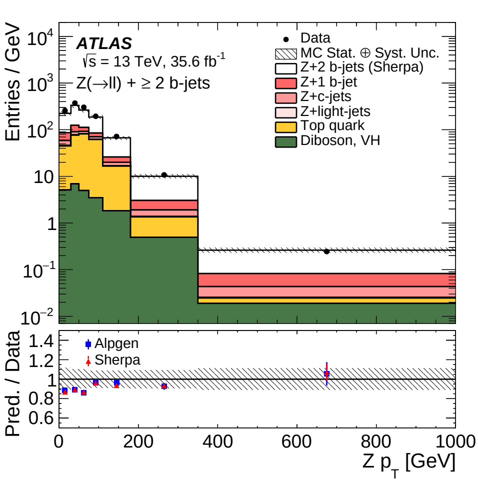




Reconstruction-level distributions for $Z+\geq 2b$-jets

- Normalization of Z+jets background MC corrected according to the flavour fit scale factors
- Signal $Z+\geq 2b$ MC not corrected
- Electron and muon channels combined





Correction to particle level

- Detector-level background-subtracted data distributions are corrected to the fiducial phase space at particle level

Kinematic variable	Acceptance cut
Lepton p_{T}	$p_{\mathrm{T}} > 27$ GeV
Lepton η	$\lvert\eta\rvert < 2.5$
$m_{\ell\ell}$	$m_{\ell\ell} = 91 \pm 15$ GeV
b-jet p_{T}	$p_{\mathrm{T}} > 20$ GeV
b-jet rapidity	$\lvert y\rvert < 2.5$
b-jet–lepton angular distance	$\Delta R(b\text{-jet}, \ell) > 0.4$

- Inclusive $Z+\geq 1b$-jet and $Z+\geq 2b$-jets cross-sections
  - Corrected by reconstruction efficiency from MC
- Differential cross-sections
  - Unfolding using *Bayesian iterative method*
- Electron and muon channels are combined at reconstruction level
  - Individual results cross-checked, agree at $\sim 1.5\sigma$ within statistical + uncorrelated systematics uncertainties

Iterative Bayesian unfolding as implemented in RooUnfold package

$$\sigma_i = \frac{1}{\epsilon_i L} \sum U_{ij} f_j N_j^{bsD}$$

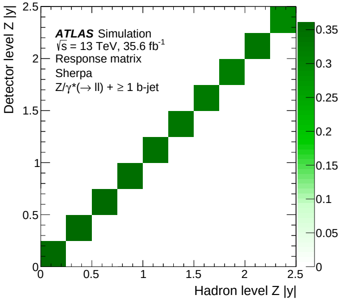
- **Differential cross-section in bin *i*** (σ_i)
- **(i,j) element of the unfolding matrix after two iterations of the Bayesian Unfolding** (U_{ij})
- **Data-bkg events** (N_j^{bsD})
- **Reconstruction efficiency in bin *i*** (ϵ_i)
- **Luminosity** (L)
- **Correction factor for the unmatched events in bin *j*.** (f_j)
  ➡It represents the number of detector-level events not passing the particle-level selection

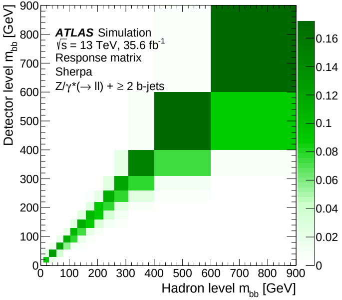
The unfolding matrix U_{ij} is evaluated through the Bayesian unfolding as:

$$U_{ij} = \frac{U_{ji} \cdot P_{0,i}}{\sum_{l=1}^{N} U_{jl} \cdot P_{0,l}}$$

U_{ji} = matrix filled with events that pass both detector- and particle-level selections (matched)

P_0 = prior, corresponding to particle-level distribution in the 1st iteration and to the result of step *n-1* for iteration *n*

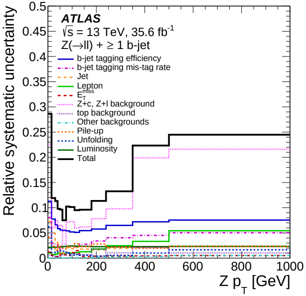


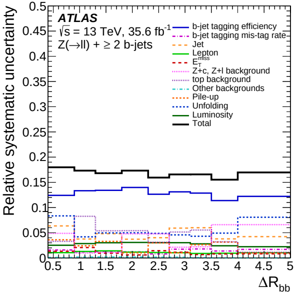


Dominant sources

- *b-jet tagging efficiency* (less mis-tag rate)
- *Z+jets background* – affects inclusive cross-sections and extreme phase space regions for $Z+\geq 1b$
- *$t\bar{t}$ modelling* – main background uncertainty in $Z+\geq 2b$
- *Unfolding procedure*

Source of uncertainty	$Z(\rightarrow \ell\ell)+\geq 1$ b-jet [%]	$Z(\rightarrow \ell\ell)+\geq 2$ b-jets [%]
b-jet tagging efficiency	7.0	14
b-jet mistag rate	2.4	1.1
Jet	2.4	5.0
Lepton	0.8	1.2
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.6	1.3
$Z+c$ and $Z+l$ backgrounds	4.5	1.1
Top background	0.5	3.8
Other backgrounds	<0.1	0.1
Pile-up	1.7	2.6
Unfolding	3.8	4.1
Luminosity	2.3	2.9
Total [%]	10	16





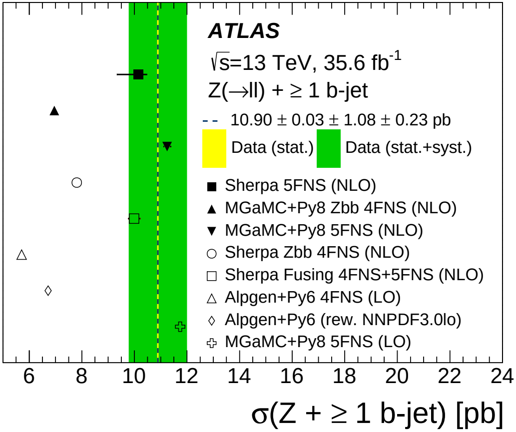
Theoretical predictions

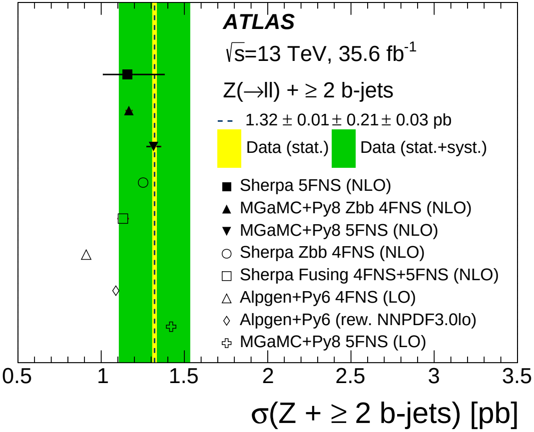
- Totally 8 predictions compared to the unfolded results
  - **LO** vs **NLO** matrix elements
  - **4FNS** vs **5FNS** calculations

Generator	$N_{\max}^{\text{partons}}$ NLO	$N_{\max}^{\text{partons}}$ LO	FNS	PDF set	Parton Shower
		Z+jets (including $Z+b$ and $Z+bb$)			
Sherpa 5FNS (NLO)	2	4	5	NNPDF3.0nnlo	Sherpa
Sherpa Fusing 4FNS+5FNS (NLO)	2	3	5 (*)	NNPDF3.0nnlo	Sherpa
Alpgen + Py6 4FNS (LO)	-	5	4	CTEQ6L1	Pythia v6.426
Alpgen + Py6 (rew. NNPDF3.0lo)	-	5	4	NNPDF3.0lo	Pythia v6.426
MGaMC + Py8 5FNS (LO)	-	4	5	NNPDF3.0nlo	Pythia v8.186
MGaMC + Py8 5FNS (NLO)	1	-	5	NNPDF3.0nnlo	Pythia v8.186
		$Z+bb$			
Sherpa Zbb 4FNS (NLO)	2	-	4	NNPDF3.0nnlo	Sherpa
MGaMC + Py8 Zbb 4FNS (NLO)	2	-	4	NNPDF3.0nnlo	Pythia v8.186

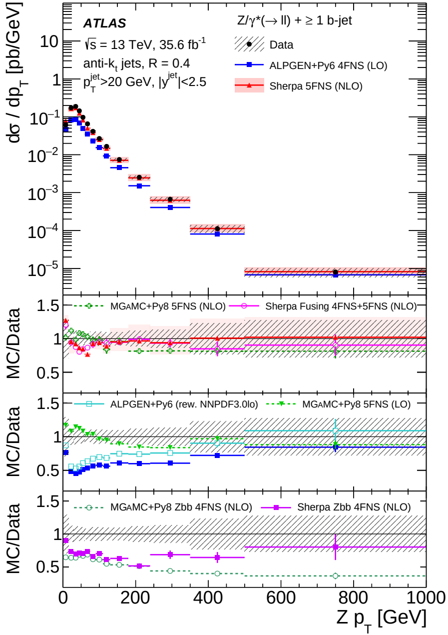
Inclusive cross-section results

- 4FNS predictions
  - Systematically lower than data in $Z + \geq 1b$ region
    - Both LO (ALPGEN+PY6) and NLO (SHERPA ZBB and MGAMC ZBB)
  - Agree with data for $Z + \geq 2b$
    - Except ALPGEN+PY6 showing 2σ discrepancy which improves with a newer PDF
- 5FNS predictions describe well both $Z + \geq 1b$ and $Z + \geq 2b$ data



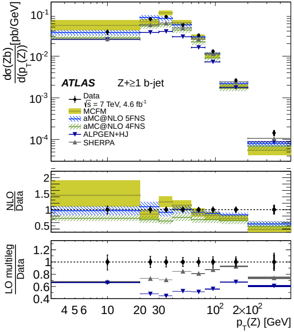


Differential cross-sections for $Z+\geq 1b$-jet

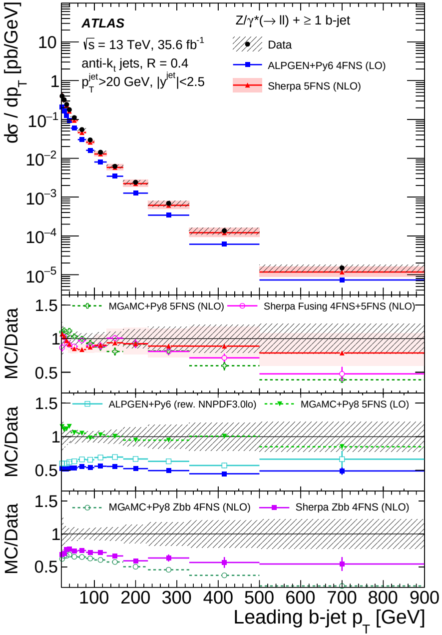


p_T of Z and b-jets test pQCD over a wide range of scales and provide input to background predictions for other processes

- All predictions show a trend at low $p_T < 100\,\text{GeV}$, except for MGAMC+PY8 5FNS (NLO)
  - soft radiation plays a role
- Best agreement by SHERPA 5FNS and SHERPA FUSING 4FNS+5FNS
- Harder $p_T(Z)$ spectrum in ALPGEN+PY6 already seen in Run-1

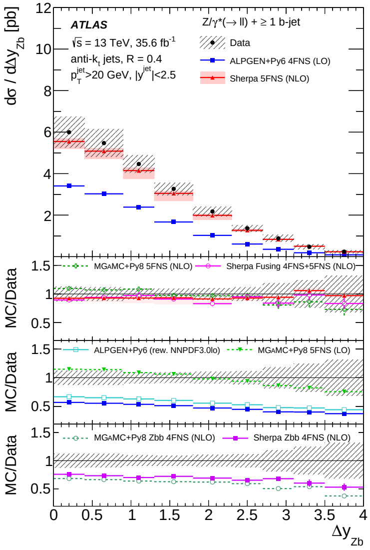


Differential cross-sections for $Z + \geq 1b$-jet



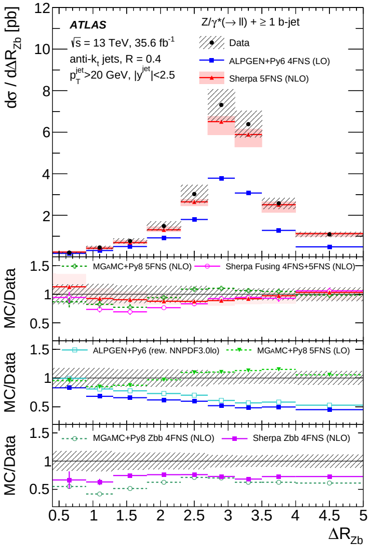
p_T of Z and b-jets test pQCD over a wide range of scales and provide input to background predictions for other processes

- Best agreement by SHERPA 5FNS, not confirmed by SHERPA FUSING 4FNS+5FNS at high p_T
- MGAMC+PY8 5FNS LO (4 partons in ME) better than NLO (1 parton in ME only), where additional hard radiation is simulated only via PS
- 4FNS *Zbb* predictions of SHERPA and MGAMC+PY8 give softer spectrum than data
  - Although inclusive ALPGEN+PY6 4FNS describes the shape well

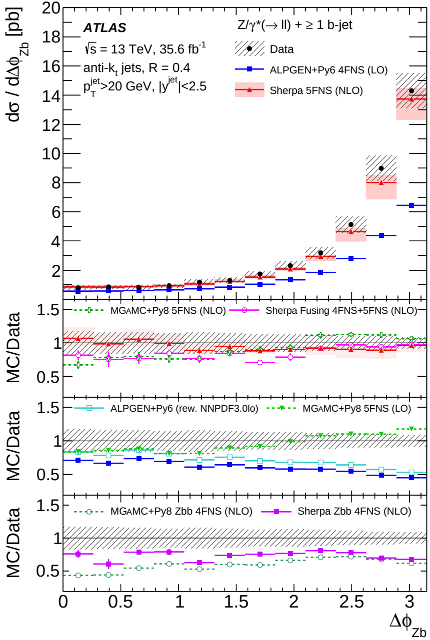


Sensitive to b quark PDFs and higher order diagram contributions

- Good description by SHERPA 5FNS and SHERPA FUSING 4FNS+5FNS
- Other predictions give *smaller rapidity separation*
- Use of different PDFs in ALPGEN show only small effect

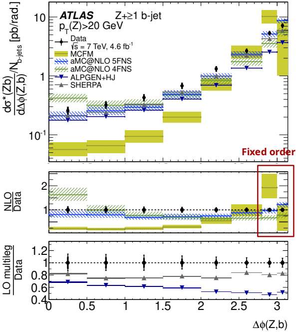


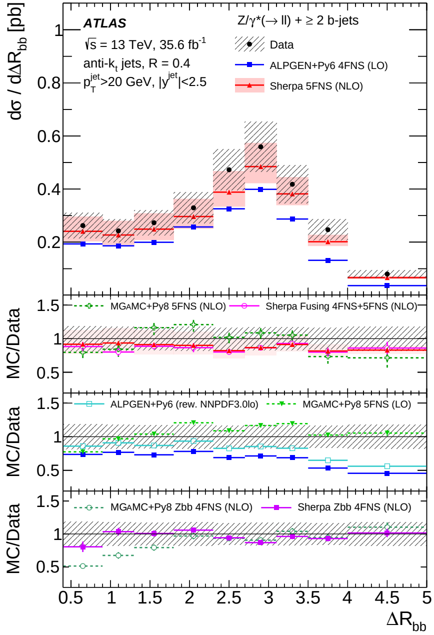
Differential cross-sections for $Z+ \geq 1b$-jet



Sensitive to additional radiation: LO gives only $\Delta\phi_{Zb} = \pi$, NLO is first order populating $\Delta\phi_{Zb} < \pi$ $\to$ ME+PS are better to describe that region

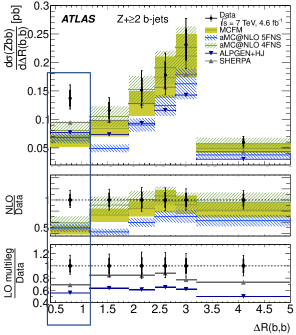
- Best agreement by SHERPA 5FNS
- SHERPA FUSING 4FNS+5FNS a little worse in for low $\Delta\phi_{Zb}$
  - Correlated with effect seen in leading b-jet p_{T}
  - These scheme needs further investigation for collinear Zb production in high-p_{T} regime
- MGAMC+PY8 5FNS NLO is slightly worse than LO



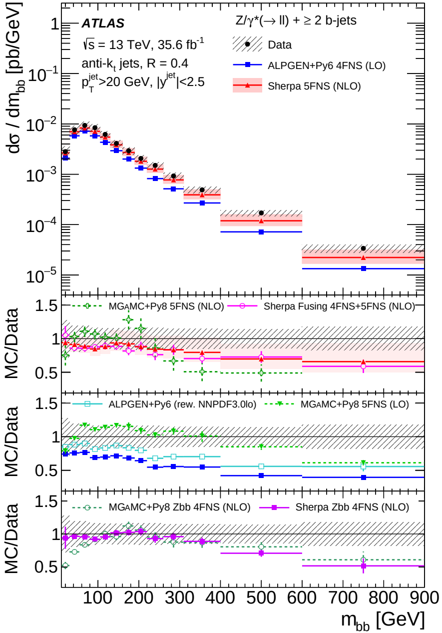


Sensitive to different production mechanisms of *Zbb* final state, e.g. gluon splitting dominant at low ΔR_{bb}

- All SHERPA predictions describe well the entire distribution
  - Substantial improvement w.r.t. LO predictions of SHERPA used for the Run-1 measurement
- Large mismodelling by MGAMC+PY8 ZBB 4FNS (NLO) in the $g \to b\bar{b}$ dominated region



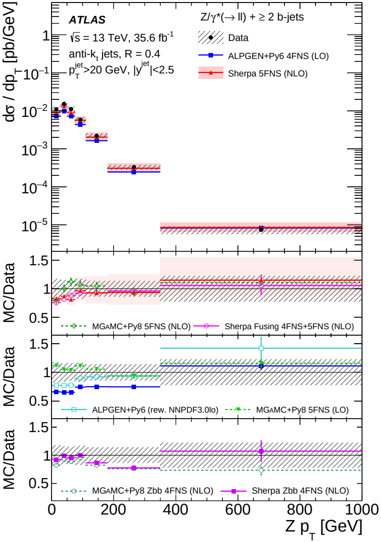
Differential cross-sections for $Z+ \geq 2b$-jets



Important variable for $VH(bb)$ studies and BSM searches

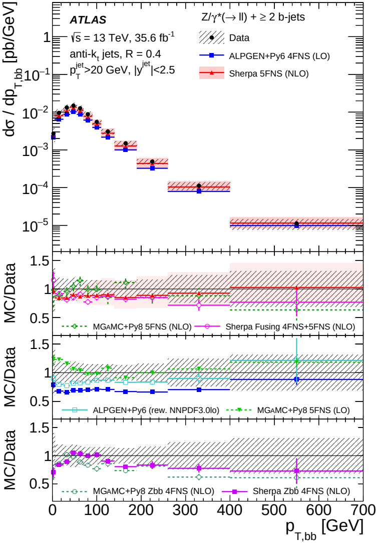
- Good modelling by all SHERPA predictions for $m_{bb} < 300$ GeV
- Others are worse, particularly MGAMC+PY8 ZBB 4FNS (NLO) which shows the discrepancy consistent with seen for ΔR_{bb}
- All predictions underestimate data at high m_{bb}

Differential cross-sections for $Z+ \geq 2b$-jets

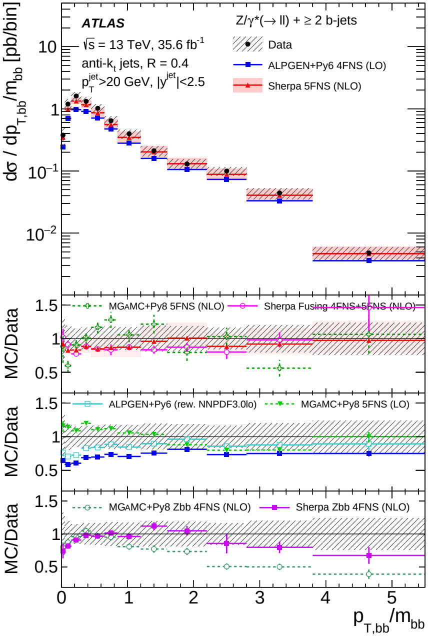


Probe pQCD in a wide scale range

- Most of predictions agree with data within large experimental uncertainties
- Alpgen shows harder p_T spectra, as in $Z+ \geq 1b$ case
- 4FNS NLO predictions (SHERPA ZBB, MGAMC+PY8 ZBB agree better than in $Z+ \geq 1b$ case, but still not perfect



Differential cross-sections for $Z+\geq 2b$-jets



Sensitive to gluon splitting: low (high) values correspond to hard (soft) splitting

- Best agreement by SHERPA 5FNS and SHERPA FUSING 4FNS+5FNS
- Again large mismodelling by MGAMC+PY8 ZBB 4FNS (NLO)

Conclusions

ATLAS $Z+1,2b$-jets results obtained using partial Run-2 dataset ($35.6\,\text{fb}^{-1}$) are compared to a wide range of predictions

- **Inclusive production cross-sections**
  - NLO 5FNS Sherpa and Madgraph predictions describe data well
  - LO 4FNS MC largely underestimate data
  - 4FNS *Zbb* NLO predictions agree with data only for $Z+2b$-jets
- **14 differential cross-sections**
  - Sherpa 5FNS provides the best description of data overall
    - The only sizeable mismodelling is high m_{bb} region
  - NLO Sherpa Fusing 4FNS+5FNS predictions generally agree with Sherpa 5FNS
    - Merging technique effects are minor at scales of the measurement
    - Small additional discrepancies at high b-jet p_{T} and small $\Delta\phi_{Zb}$
  - MGaMC+Py8 5FNS LO is good in most cases
    - Sometimes better than NLO, due to larger number of partons in matrix element
  - *Zbb* 4FNS NLO predictions of Sherpa and MGaMC+Py8 demonstrate large discrepancies, even for $Z+2b$-jets

The measurement provides an important input for quantitative understanding of pQCD, improvement of predictions and MC modelling

Results available at arXiv:2003.11960, submitted to JHEP

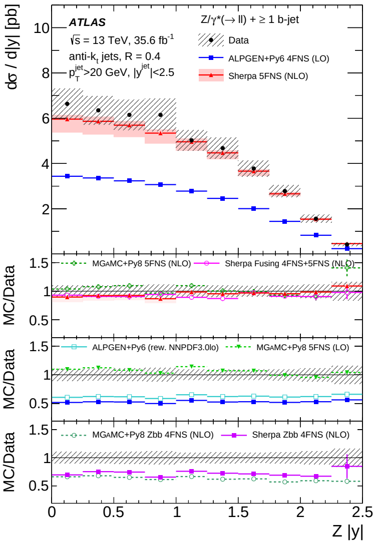
Backup slides

Full selection criteria

	Electron channel	Muon channel
Trigger	Single electron	Single muon
Leptons	Tight Isolated PV association: $\|d_0/\sigma_{d_0}\| < 5$, $\|z_0 \sin\theta\| < 0.5$ mm $p_T > 27$ GeV $\|\eta\| < 1.37$ or $1.52 < \|\eta\| < 2.47$	Medium Isolated PV association: $\|d_0/\sigma_{d_0}\| < 3$, $\|z_0 \sin\theta\| < 0.5$ mm $p_T > 27$ GeV $\|\eta\| < 2.5$
Jets	$p_T > 20$ GeV and $\|y\| < 2.5$ $\Delta R(\text{jet}, \ell) > 0.4$	
b-jet	$p_T > 20$ GeV and $\|y\| < 2.5$	

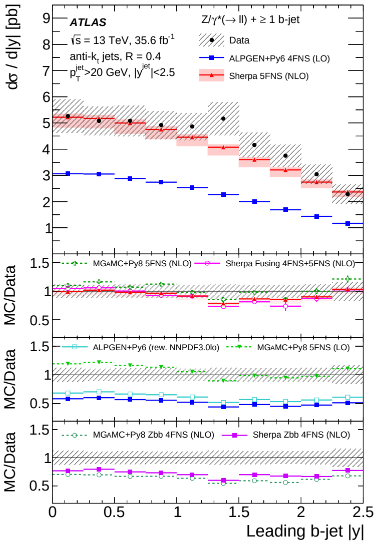
Regions

	Pre-tag region	Signal regions	Z+jets Validation Region	$t\bar{t}$ Validation Region
Leptons	2 same-flavour, opposite-charge			1 e, 1 μ, opposite-charge
$m_{\ell\ell}$	$76\text{ GeV} < m_{\ell\ell} < 106\text{ GeV}$			
E_T^{miss}	$E_T^{miss} < 60$ GeV if $p_T^{\ell\ell} < 150$ GeV			
Jets	≥ 1 or ≥ 2 jets			
b-tagging efficiency working point selection	-	70%	≥ 1 b-jet at 77%–70%	70%
Number of b-jets	-	≥ 1 b-jets (1-tag region) ≥ 2 b-jets (2-tag region)	≥ 1 b-jets	≥ 2 b-jets

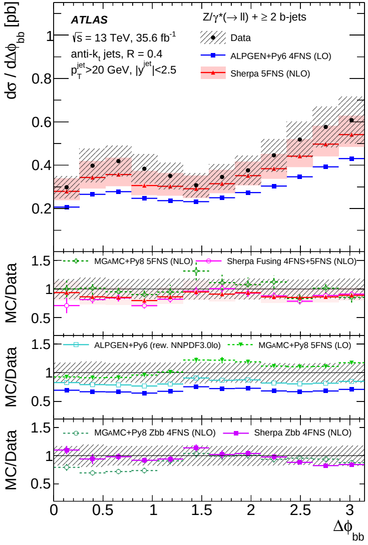


Sensitive to *b* quark PDFs and higher order diagram contributions

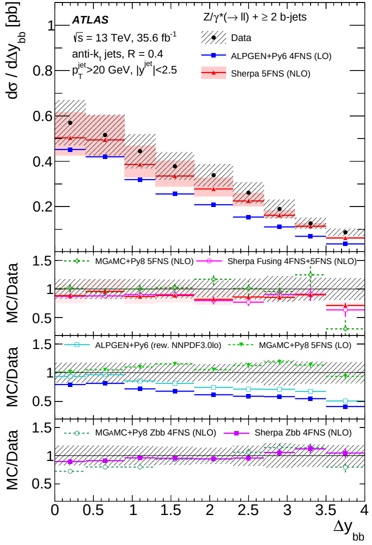
- All MC predictions provide satisfactory description
- Some modulation w.r.t. data in leading *b*-jet p_T, sometimes beyond the experimental uncertainty



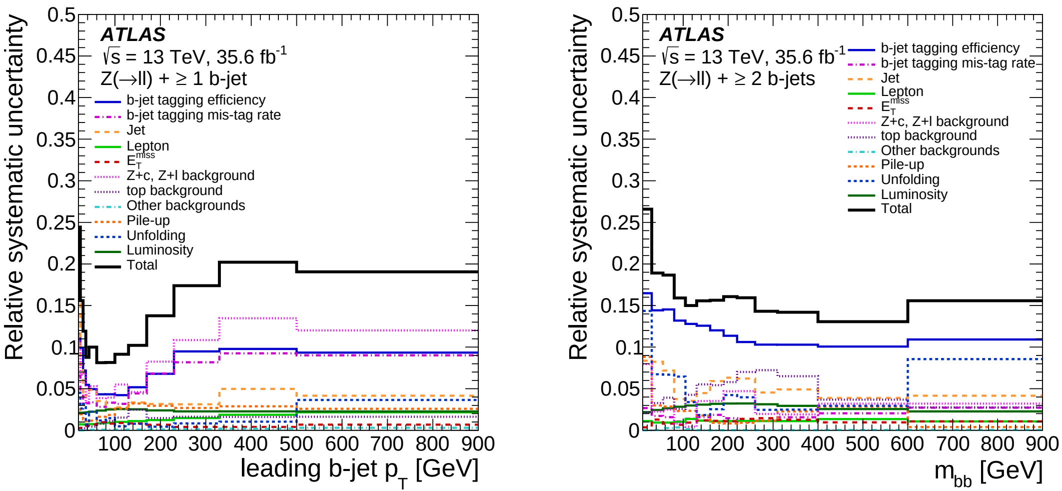
Differential cross-sections for $Z+\geq 2b$-jets



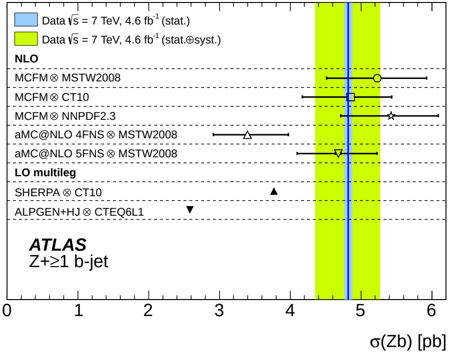
- Most of predictions provide satisfactory description within large experimental uncertainties
- Disagreement at low $\Delta\phi_{bb}$ for MGAMC+PY8 ZBB 4FNS NLO
- Mismodelling of Δy_{bb} by ALPGEN
  - Small effect of PDF

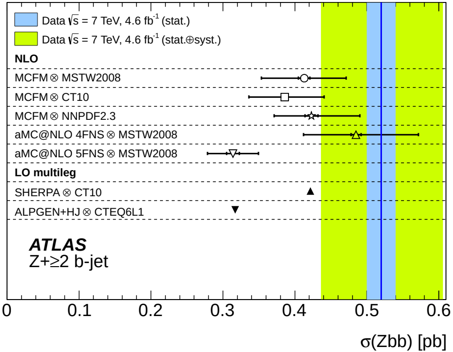


Systematic uncertainties



7 TeV cross-sections





JHEP 10 (2014) 141, ATLAS $Z + b$-jets measurement at $\sqrt{s} = 7\,\text{TeV}$