Measurements of Z boson production in association with jets at ATLAS

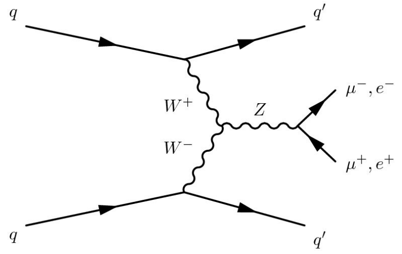
Federico Sforza (UniGE & INFN), on behalf of the ATLAS Collaboration

DIS 2021 -13/04/2021

Virtual Event @ Stony Brook University

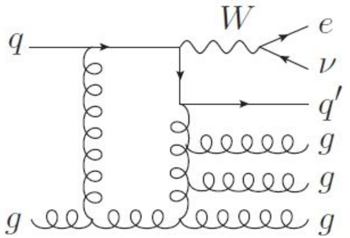
Key role of W or Z plus jet production at LHC

W and Z boson plus jets production at LHC is abundant and has a clear experimental signature ⇒ crucial probe of QCD and EWK mechanics

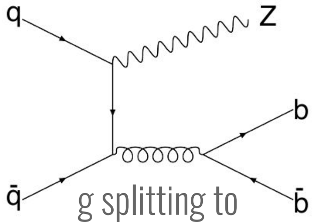


Z+jets EWK production

→ Test of state-of-the-art predictions: fixed-order in perturbative QCD (pQCD) and MC simulations

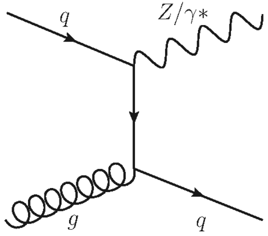


production of W+ n-partons



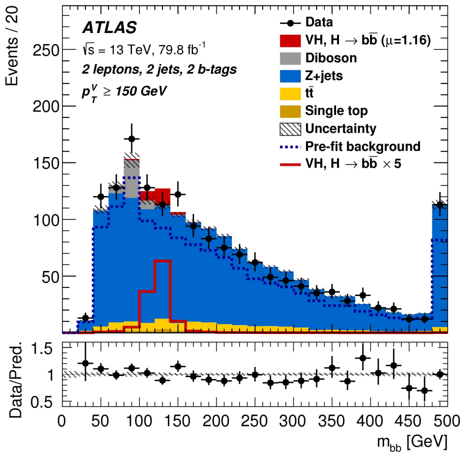
g splitting to heavy-quarks

→ Sensitive to various aspects of the proton Parton Density Functions (PDFs)



qg initial state

→ Large and irreducible background to BSM searches and Higgs measurements



VH(bb) obs. [PRL B 786 (2018) 59]

The presented Z plus jet ATLAS measurements

A vast set of inclusive and differential measurements at different √s is available!

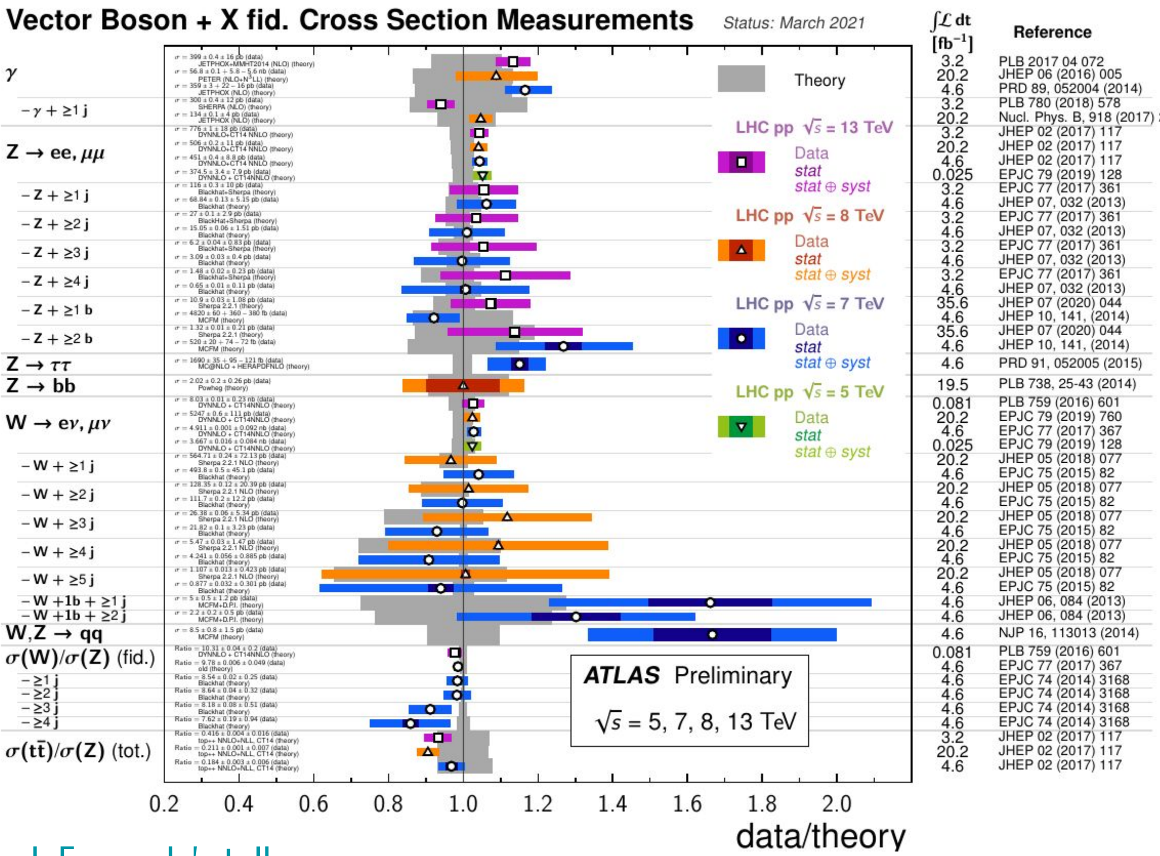
Here mainly discussing two recent results:

- ***Double differential Z+jets @ 8 TeV***
  → Important new constraint for PDF fits [[EPJC 79 (2019) 847](#)]
- ***Differential Z+b(b) jets @ 13 TeV***
  → Test of MC (ME+PS) and proton flavour structure [[JHEP 07 (2020) 44](#)]

Note also these interesting/related talks:

→ discussion of V+jj EWK production in M. Goblirsch-Kolb's talk

→ PDF interpretation using W, Z + jets data in J. Ferrando's talk

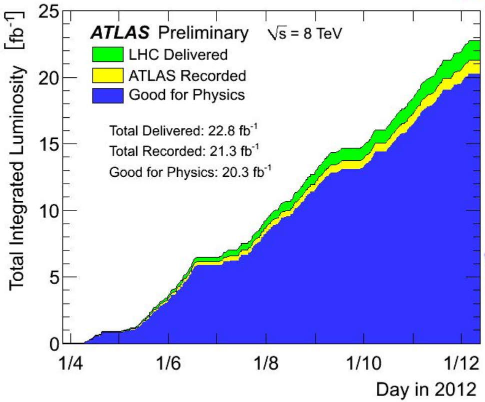


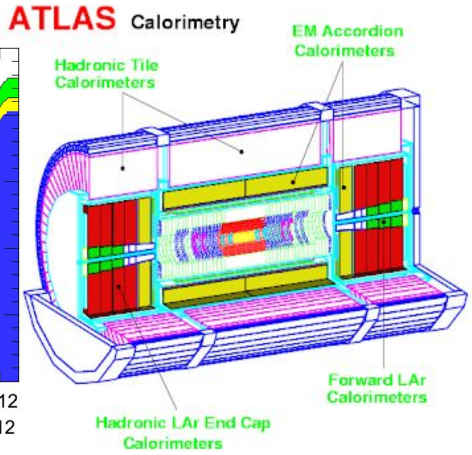
Measurements with different datasets and different challenges

Z+jets @ 8 TeV [EPJC 79 (2019) 847]

→ 20 fb^{-1} of data collected during LHC Run 1

→ ATLAS calorimeter capability of forward jet reconstruction crucial in this analysis for the exploration of PDFs at high-x values

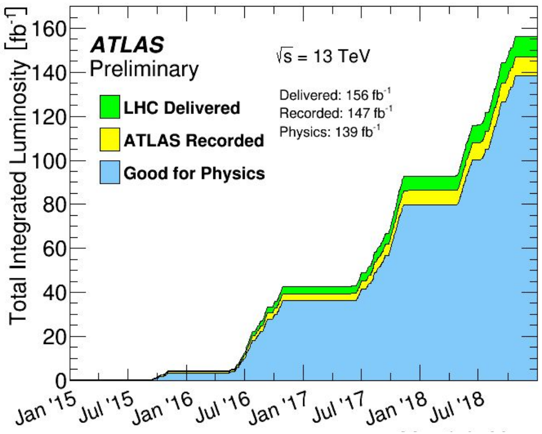


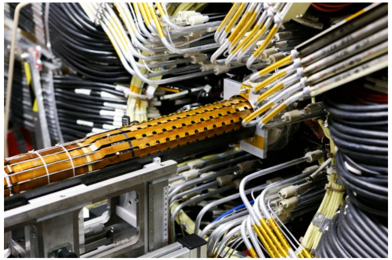


Z + b(b) jets @ 13 TeV [JHEP 07 (2020) 044]

→ 36 fb^{-1} of data collected during LHC Run 2

→ Identification of heavy-hadrons (m_b~*4.2* GeV) decay at O(100 μm) w.r.t. primary vertex (PV) improved with new Run 2 pixel detector (IBL) very close to beamline





IBL installation at ~3 cm from beamline

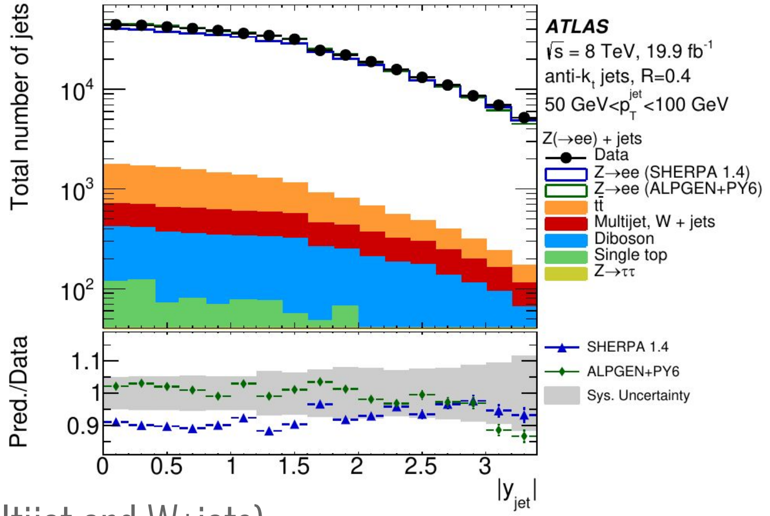
Z plus inclusive jets @8 TeV : Selection of signal candidates

Z→ee candidate selection very efficient because of low p_T cuts during LHC Run 1:

- Trigger online selection: single-electron (p_T>24 GeV) and di-electrons (p_T>12 GeV)
- Offline selection : 2 opposite-sign (OS) electrons of p_T > 20 GeV, $|\eta|$ < 2.47, $66<m_{ll}<116$ GeV

Jet reconstruction and inclusive selection:

- Anti-k_t clustering algorithm (ΔR=0.4), selection of ⩾ 1 jet with |y|<3.4 and p_T > 25 GeV (→ extends to ~ 1 TeV)
- Pile-up jets suppressed requiring high fraction of jet tracks associated to PV... but no tracking coverage for $|\eta|$>2.4
  ⇒ p_T > 50 GeV for jets up to |y|~3.4



Non-Z event background in selection is almost negligible:

→ estimated with MC (top, diboson) and data-driven techniques (multijet and W+jets)

Z plus inclusive jets @8 TeV : Unfolding of detector effects

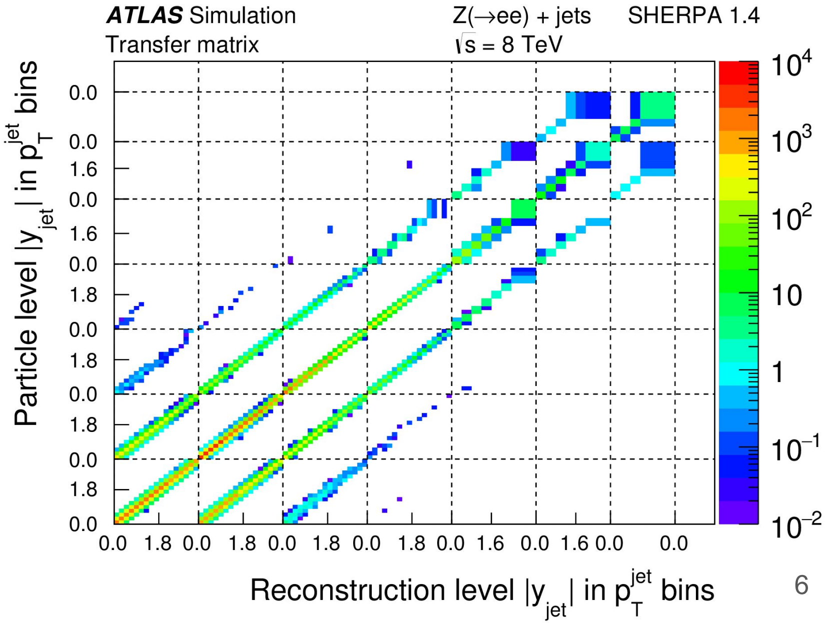
The analysis goal: $$\frac{d^2\sigma}{dp_T^{jet} d|y_{jet}|} = \frac{1}{\mathcal{L}} \frac{N_i^{\mathcal{P}}}{\Delta p_T^{jet} \Delta|y_{jet}|}$$ Double differential cross section measurement → strong constraint in PDF fits

Need to extract $N_i^{\mathcal{P}}$ = number of jets in data at particle-level (after background subtraction)

- Estimate of detector efficiency and resolution using Sherpa v1.4 MC (LO ME+PS) → 2D unfolding matrix:

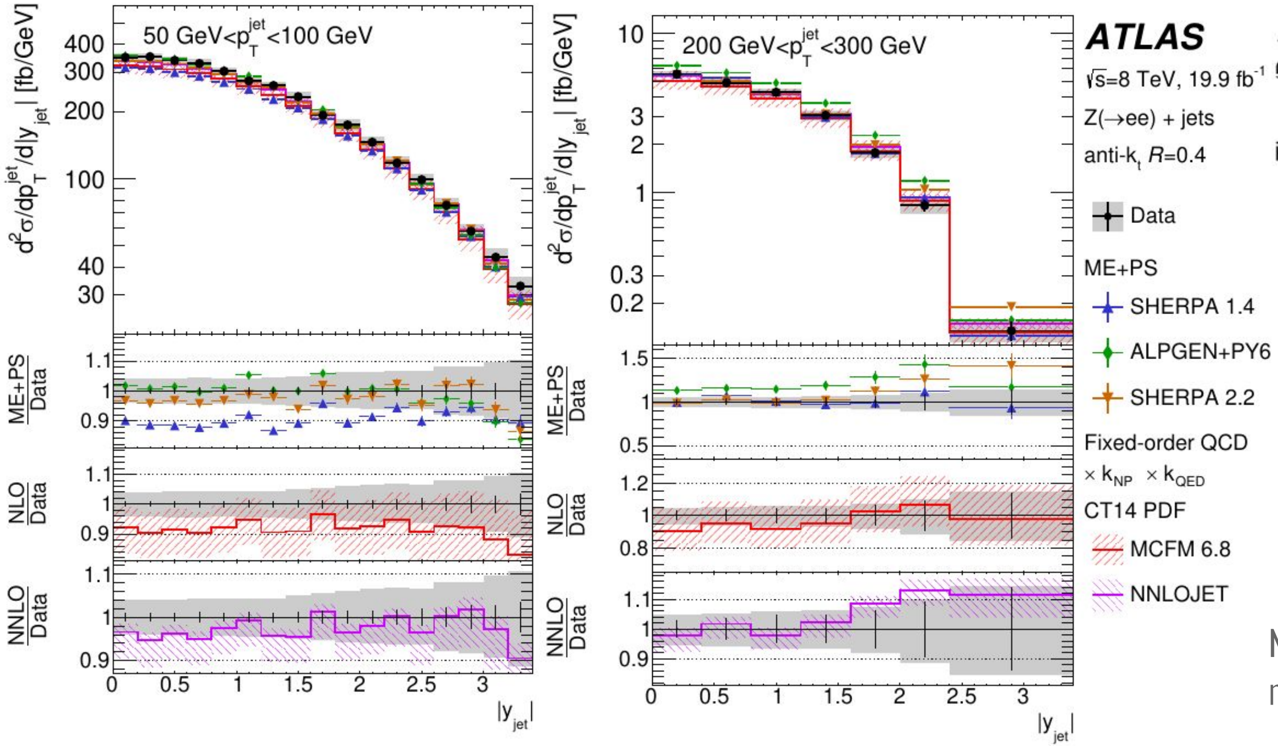
$$N_i^{\mathcal{P}} = \frac{1}{\mathcal{E}_i^{\mathcal{P}}} \sum_j U_{ij} \mathcal{E}_j^{\mathcal{R}} N_j^{\mathcal{R}}$$

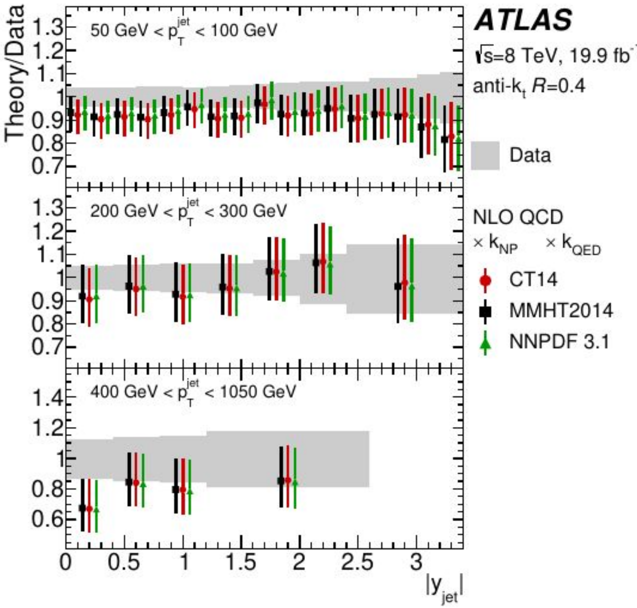
- Iterative Bayesian unfolding allows to obtain particle-level quantities



Z plus inclusive jets @8 TeV : Comparison with predictions

ME+PS MC predictions with LO ME (Alpgen+Py6 and Sherpa 1.4), NLO ME (Sherpa 2.2), and fixed-order NLO QCD using different PDFs ⇒ good agreement within uncertainties (dominated by JES and statistics at high-p_T)





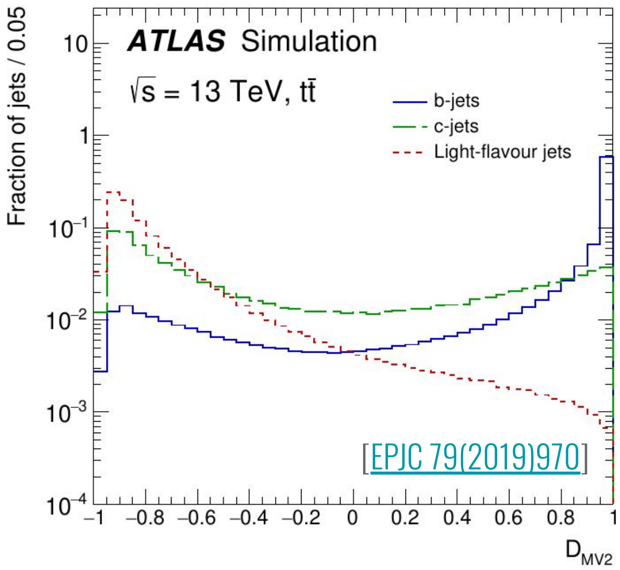
Note: fixed-order calculation corrected for non-perturbative effects before comparison

Z+b-jets @13 TeV : Selection of candidate events

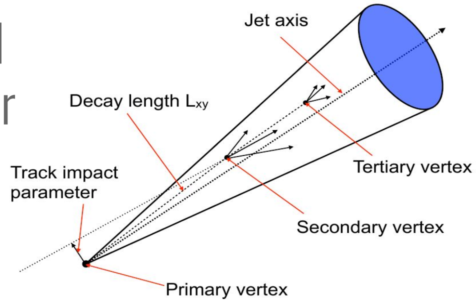
$Z \rightarrow ll$ + jets candidate selection similar to Run 1 but relies on higher p_T cuts in Run 2:

- Single lepton trigger of p_T>25 GeV, 2 OS leptons (ee/$\mu\mu$), with p_T>27 GeV, $|\eta|$<2.5, 76<m_{ll}<106 GeV
- ≥1 or ≥ 2 jets reconstructed with Anti-k_t algorithm (ΔR=0.4) with p_T>20 GeV and $|\eta|$<2.5

b-jet candidate selection relies on long lifetime, secondary vertices, decay pattern, etc.



- Tracking & jet information condensed using multivariate (MV) algorithms for separation of *b*-jets vs different flavour jets - c or light-flavour



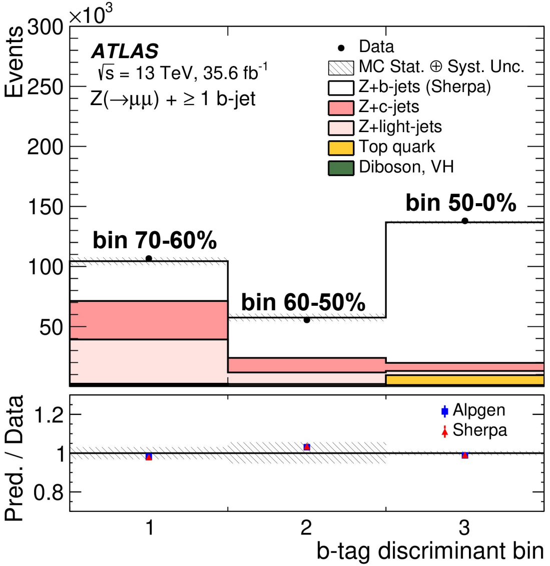
- Require 1 or 2 jets passing a “cut” on MV *b*-tagging algorithm corresponding of 70% efficiency for b-jets (vs mistag of ~10% c-jets and ~0.4% for light-jets)

Z+b-jets @13 TeV : data-driven background of different flavour

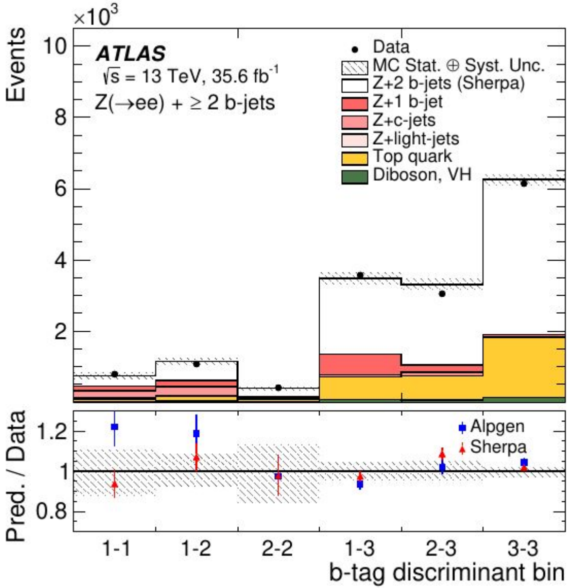
Z+jets non-b-jet background sizable and affected by large uncertainties:

⇒ extract *Z+c*-jets and *Z+light*-jets with data-driven template fit using MV discriminant binned according to specific *b*-tag purity cuts (with efficiency calibration for each bin and for each flavour)

Z+⩾ 1 b-jet background fit



Z+⩾ 2 b-jet background fit



Notes on Z+⩾2 b-jets:

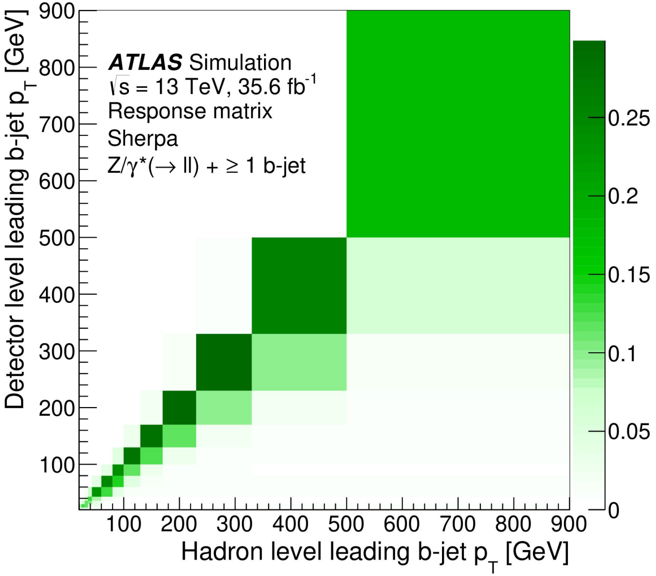
→ the presence of two jets in the selection allows multiple combination of b-tag MV bins

→ sizable background from top-quark estimate using MC and validated using high-stat. dedicated control region

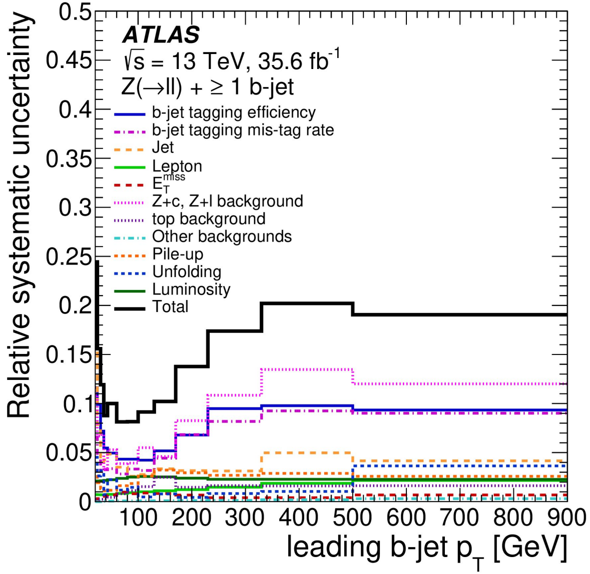
Z+b-jets @13 TeV : Unfolding and systematic errors

Similar approach to Z+jets:
background subtracted data events brought to particle-level correcting for efficiency and detector resolution effects using *Z+b*-jets MC (*NB:* b-jet efficiency calibrated using data control)

E.g. unfolding matrix on *b*-jet p_T :



Dominant systematic uncertainties due to b-jet efficiency calibration and background subtraction:

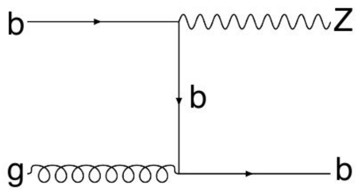


Z+b-jets @13 TeV : Comparison with theoretical predictions

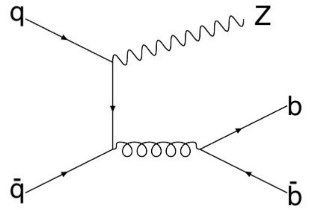
Measurement challenges theoretical predictions for vector boson (Z here) production+b quarks:

- Different Flavour Scheme (FS) calculations ⇒ 4FS "easy" use of b-quark mass in ME, while 5FS allows resummation of large logarithm in the PDFs
- Different levels of ME accuracy ⇒ *n-partons, LO, NLO, etc.*
- Various approaches for inclusion of b-quark mass in PS and gluon splitting

A long standing discussion:



5FS: Z+*b* @LO with b from the PDF and m_b=0



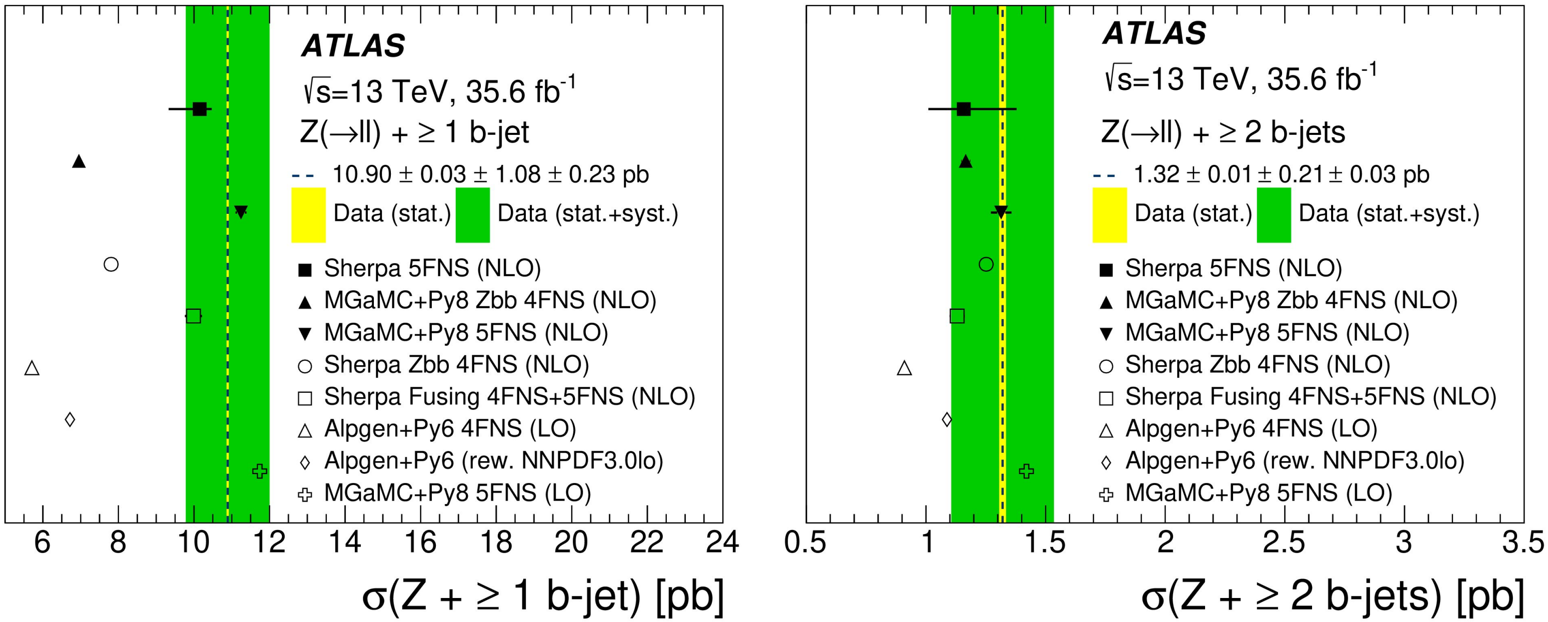
Z+b(b) @LO with g⟶ bb and $m_b \neq 0$ (***4FS***) or m_b=0 (***5FS***)

Tested MC predictions (ME+PS):

Generator	$N_{max}^{partons}$ NLO	$N_{max}^{partons}$ 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

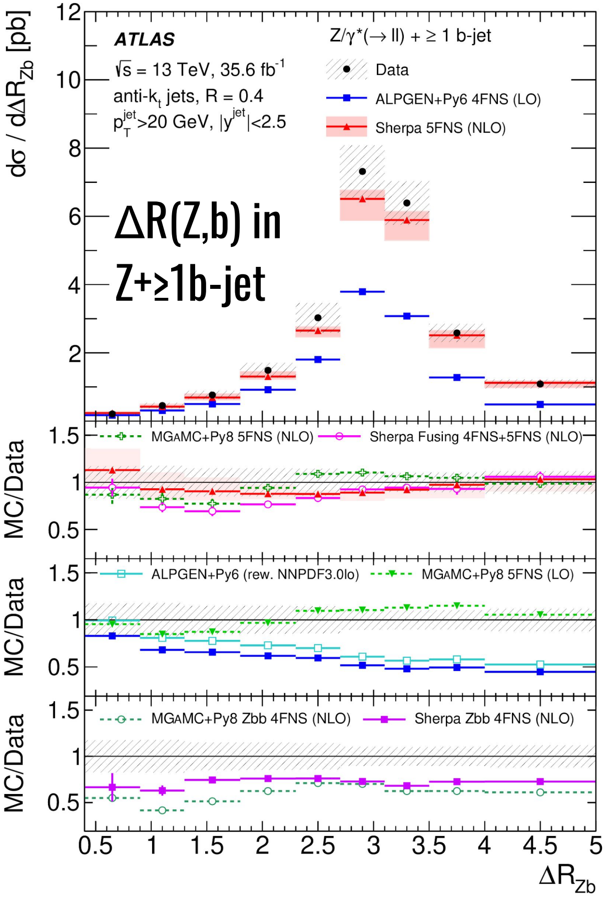
NB: V+HF main background for V+Higgs measurements ⇒ crucial for experiments to have reliable MC

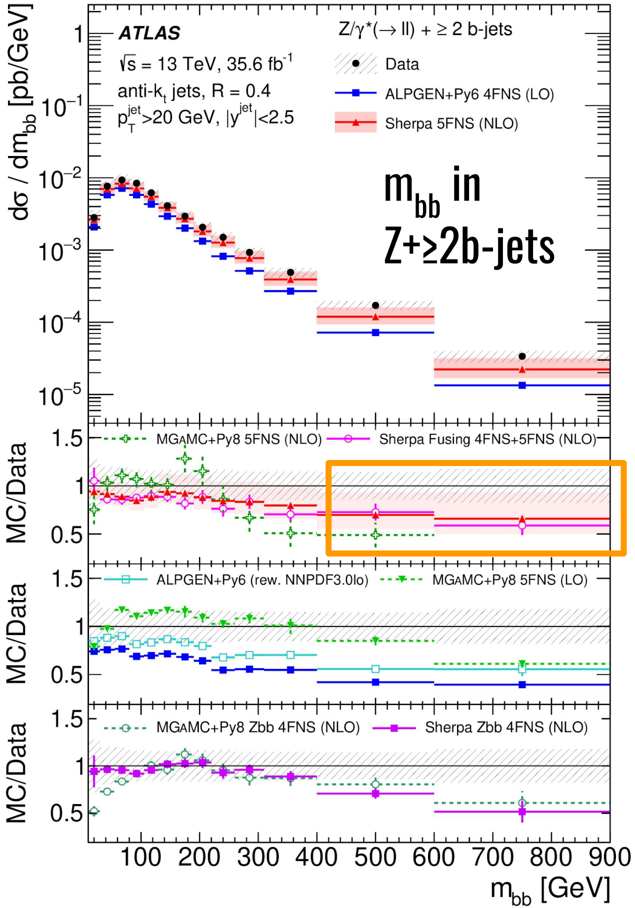
Z+≥1 b-jet and Z+≥2 b-jets inclusive cross-section: Results



- 4FS largely undershoots *Z+≥1 b*-jets cross-section in all configuration
- *Z+≥2 b*-jets uncertainties still too large to favour any of the more recent predictions

Differential cross-section measurements & challenges





Z+≥1 b-jet & Z+≥2 b-jets phase space (mostly) well described by 5FS, while 4FS shows deficits in Z+≥1 b-jet

Some tensions with data at high m_{bb} (and high jet-p_T) but large errors in both theory and measurement:

⇒ *challenge for searches and test of other process in such phase space...*

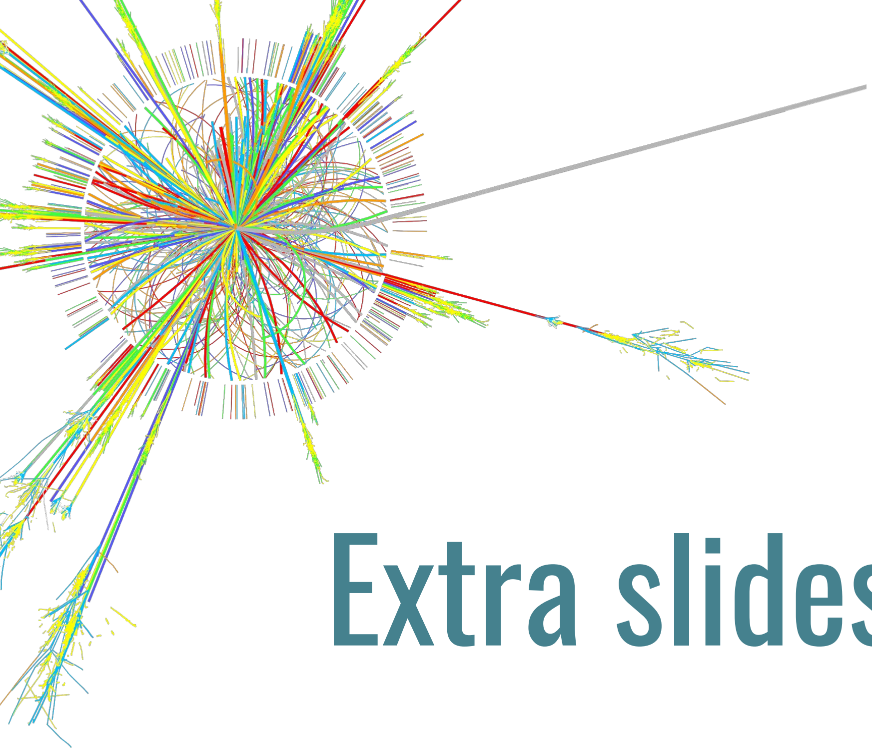
Summary and prospects

Presented today two recent results:

- *Double differential Z+jets @ 8 TeV* [EPJC 79 (2019) 847]
- *Differential Z+b(b) jets @ 13 TeV* [JHEP 07 (2020) 44]

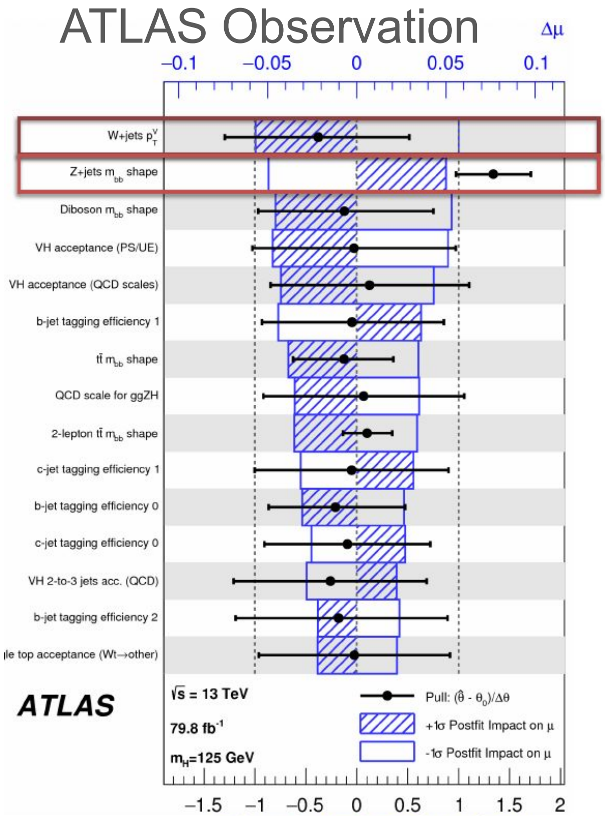
They show the importance of V+jets measurements at LHC for the deep understanding of QCD theory, pQCD predictions, PDFs, MC (ME+PS) technology, and also for the general advancement of LHC fundamental physics in all its aspects

Future measurements will use the much larger statistics of the full LHC Run 2 (and maybe even the first Run 3 data), however the main benefit to V+jets precision measurements will come from the ongoing improvements in analysis techniques and systematic reduction ⇒ *stay tuned!*

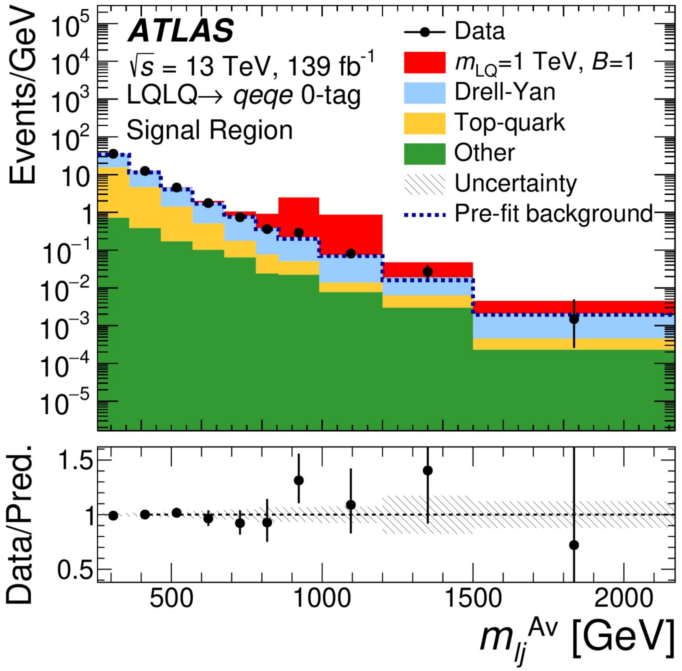


Extra slides & Backup Material

Examples of impact of V+jets on Higgs and BSM searches

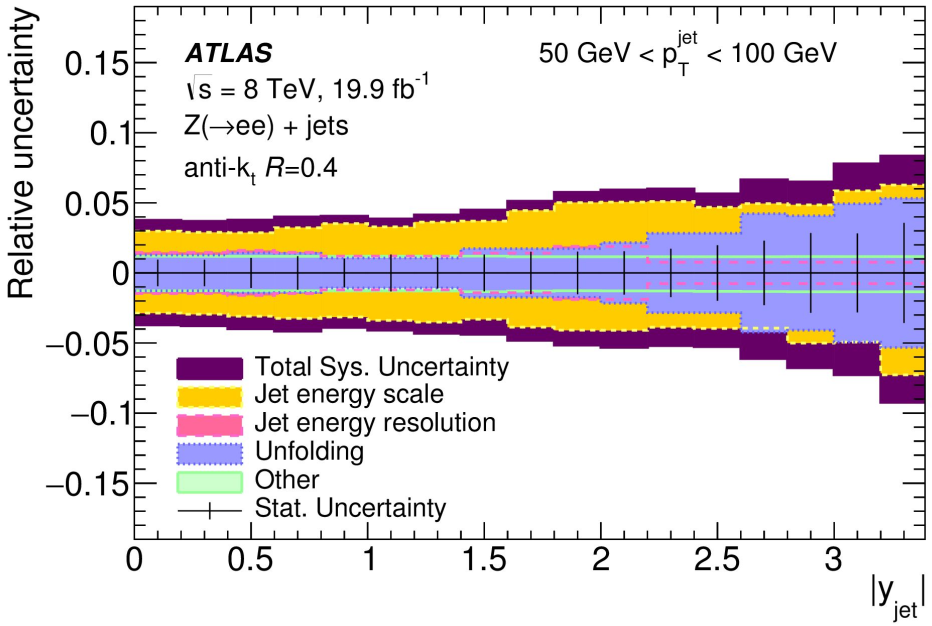


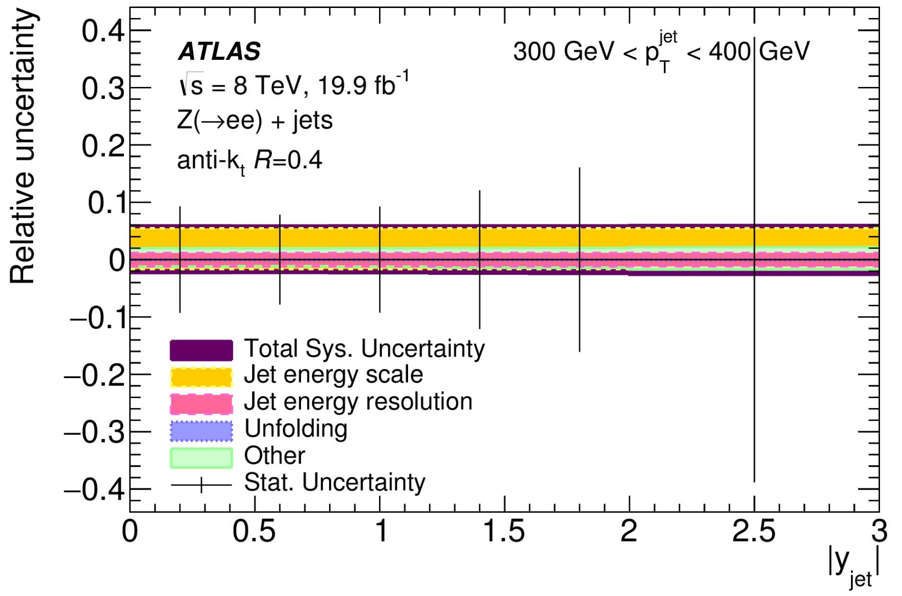
https://arxiv.org/abs/1808.08238



https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EXOT-2019-13/

Uncertainties in Z plus inclusive jets double differential measurement



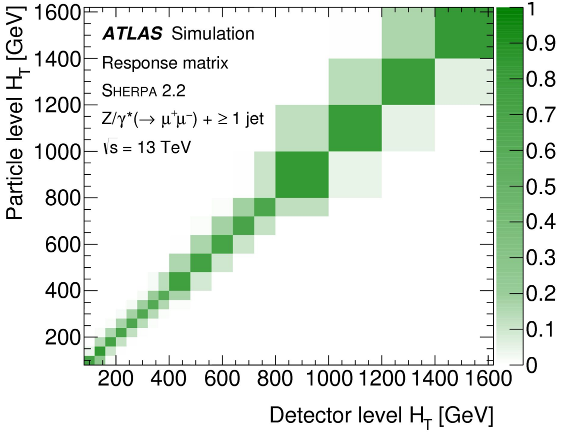


Iterative Unfolding with Bayesian method

Response matrix accounts for migrations using MC simulation:

$$M_{ij}=M(R_i|T_j)$$

Conditional probability that the effect R_i is produced by the cause T_j



How to extract *"prediction-unbiased"* probability using iterative Bayesian unfolding:

- Bayes theorem:

$$M(T_i|R_j) = M(R_i|T_j)\, P_0(T_j) \,/\, \mathrm{Sum}_l\ M(R_i|T_l)\, P_0(T_l)$$

- Particle level MC used as initial prior, $P_0(T_j)$, to determine a first estimate of the unfolded data distribution:

$$T_j = \mathrm{Sum}_i\ M(T_j\,|\,R_i\,)R_i$$

- In each further iteration the estimator of the unfolded distribution from previous iteration is used as a new prior

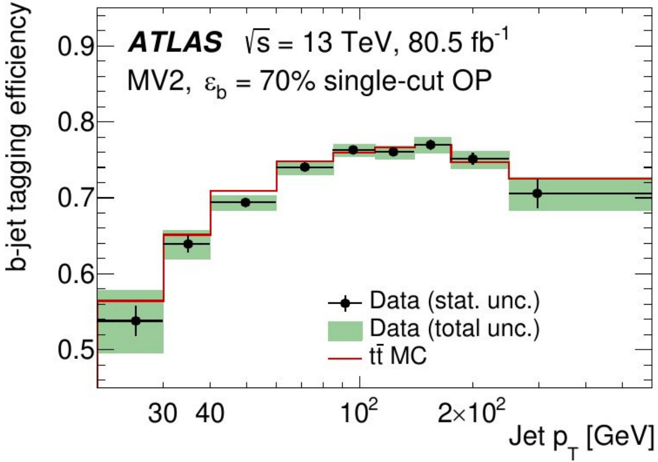
b-jet identification: Selection and Efficiency Calibration

- In practical terms the *b*-jet identification proceeds with a *"cut"* on the MV algorithm output

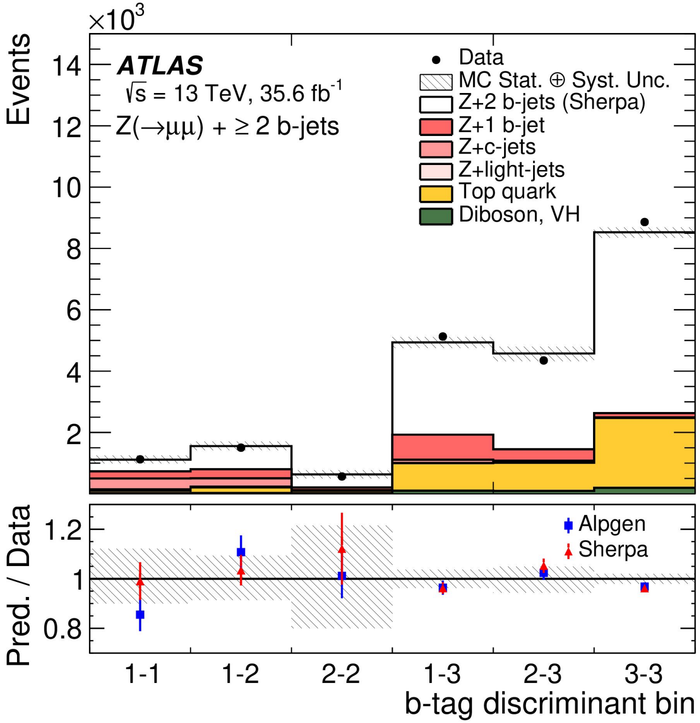
WP	Cut value X	b-jet efficiency (ε_b)	c-jet mistag rate (ε_c)	LF-jet mistag rate (ε_{LF})
85%	0.1758	85%	32%	2.9%
77%	0.6459	77%	16%	0.77%
70%	0.8244	70%	8.3%	0.26%
60%	0.9349	60%	2.9 %	0.065%
50%	0.9769	50%	0.94 %	0.017%

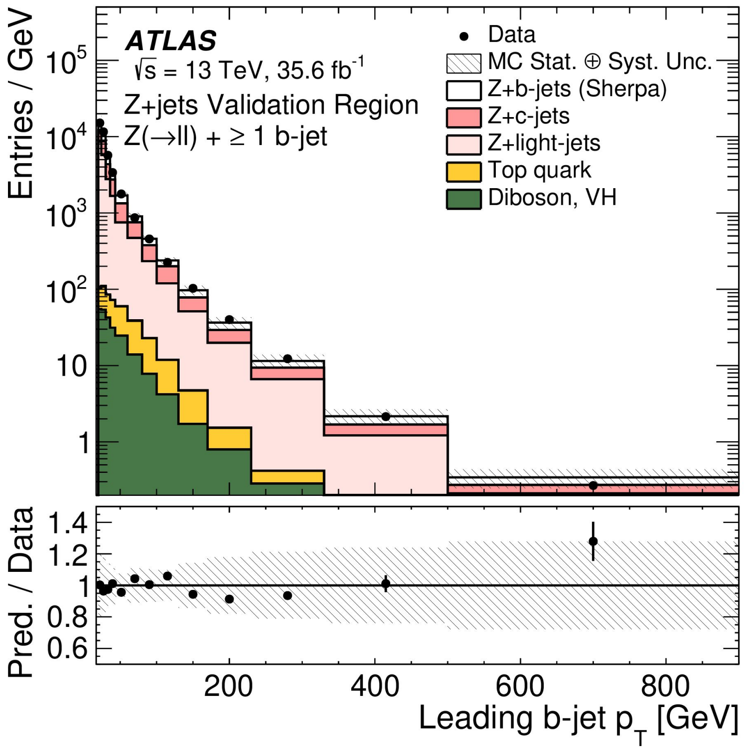
- *b*-jet identification performance evaluated on MC but crucial to calibrate it with reference candles in *Data* ⇒ inaccuracy in detector description or QCD simulation of *b/c/light*-jets

- Example of $t\bar{t}$ events used for the extraction of the efficiency correction (EPJC 79(2019)970)



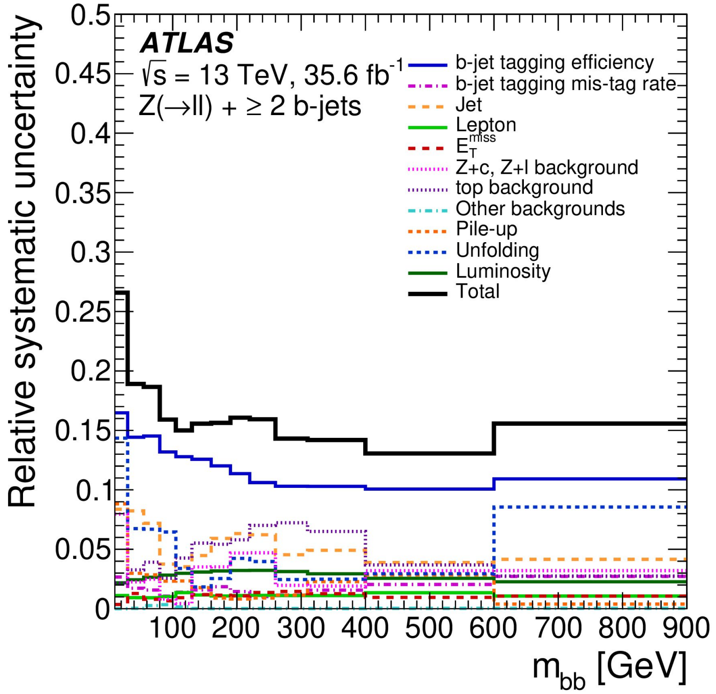
Data driven Z+jets background and validation





Z+b jets systematic uncertainties

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_T^{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



Other Z+b-jets differential distributions