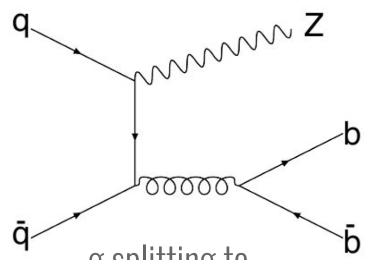
Measurements of Z boson production in association with jets at ATLAS

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

DIS 2022 -03/05/2022

Santiago de Compostela, Spain

Key role of W or Z plus jet production at LHC

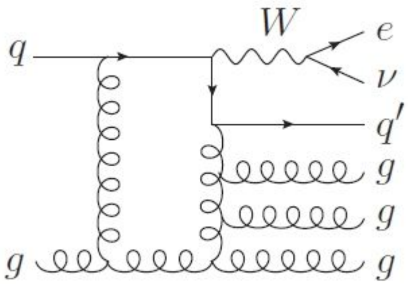


g splitting to heavy-quarks

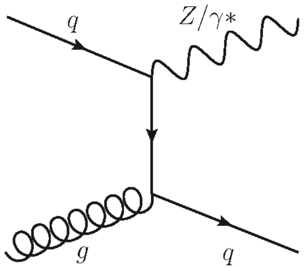
Abundant production of W or Z + jets & clear experimental signature

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

→ Large & irreducible background to BSM searches or Higgs measurements: *crucial for experiments to rely on precise models affected by small uncertainties!*

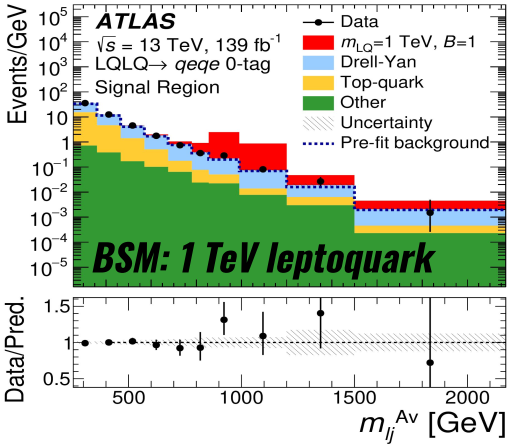


production of W+4-partons with NLO contribution

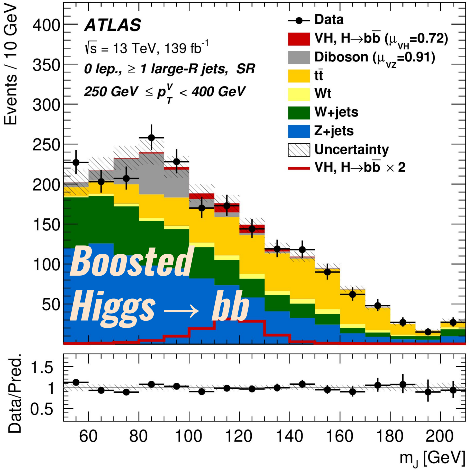


Z+1 parton @LO from qg initial state

[JHEP 10 (2020) 112]



[PLB816 (2021) 136204]



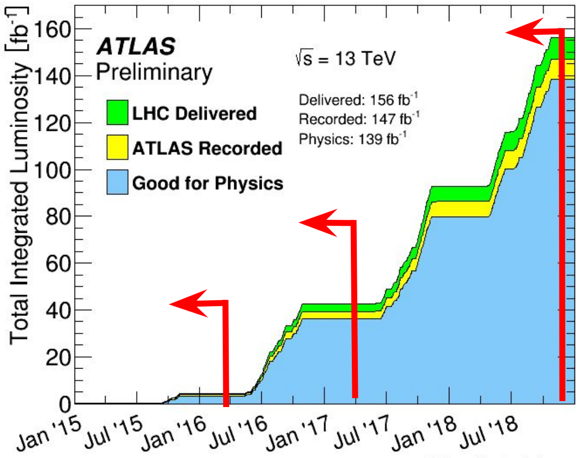
→ Clearly, also sensitive of proton PDFs:

Check new results in [F. Giuli's talk tomorrow!]

Challenging new measurements of Z+jets with ATLAS experiment

Several key measurements have been obtained using ATLAS Run 2 *pp* collision data at 13 TeV, a relevant past examples:

- *Early Z+jets measurements with 3.2 fb^{-1}* [EPJC 77 (**2017**) 361]
  → important for first benchmark of V+jets MC @13 TeV
- *First Z+b(b)-jets measurement with 36 fb^{-1}* [JHEP 07 (**2020**) 44]
  → probing MC modeling of heavy flavour jets, substantial differences depending of Flavour Number Scheme of simulation

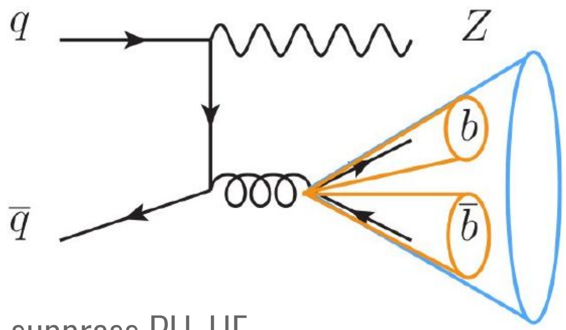


Today's focus on results accessible thanks to improved reconstruction/full dataset large statistics:

- *Z + large-radius jet, flavour inclusive and with 2 b-tags, 36 fb^{-1}* [arXiv:2204.12355] **NEW!**
  → first probe of jet sub-structure in boosted regime, also testing kinematics of close-by b-quarks
- *Z + high-p_T jets in extreme phase-space, with 139 fb^{-1}* [arXiv:2205.02597] **NEW!**
  → probe of cutting edge MC and pQCD predictions in scarcely explored areas (e.g. collinear Z-jet)

Z plus a large-radius jet and identification of *bb* final state

Investigating Z production in association with: ***high-p_T, boosted, coloured object, w/ or w/o 2 b-jets***



Z→ll (l=e/mu) selection: efficient for online/offline identification, final state with low background from other SM processes

Large Jet (J) reconstruction and double-b-tagging:

- **"J" definition**: Anti-k_t ΔR=1.0 & ***trimming*** ΔR=0.2 subjets if p_T<0.05 p_T(J) → suppress PU, UE
- **Boosted object phase space definition:** p_T(J) > 200 GeV , $|\eta| < 2$
- **Double b-tagging:** 2 small track-jets (Anti-k_t ΔR=0.2) matched to **J** and passing a b-hadron MVA selection (MV2c10) with eff_b~70%, eff_c~14% , eff_{light}~0.8% (on tt simulation)

Signal MC estimate and syst. using ME+PS matched/merged simulations:

Sherpa 2.2.1: with ME 0-2p NLO + 3-4p LO

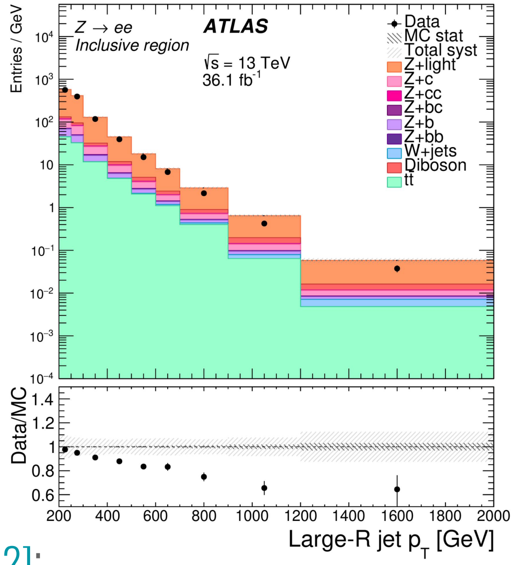
MG5_aMC+Py8: with ME 0-4p LO (CKKW-L)

$Z + J$: sample composition and correction to particle level

Analysis of selected data:
→ *flavour-inclusive*
→ *double b-tag*

Sherpa 2.2.1 signal MC shows overestimate in inclusive vs under-estimate in 2-b-tag

	Inclusive		2-tag	
	ee	$\mu\mu$	ee	$\mu\mu$
$Z{+}b\bar{b}$	324 ± 4	305 ± 4	163.8 ± 2.6	157.2 ± 2.5
$Z{+}c\bar{c}$	536 ± 10	530 ± 9	12.3 ± 1.8	19.3 ± 2.0
$Z{+}bc$	89 ± 2	81 ± 2	14.6 ± 1.2	12.1 ± 0.9
$Z{+}b$	2588 ± 13	2423 ± 12	14.8 ± 1.1	12.4 ± 1.3
$Z{+}c$	5073 ± 32	4862 ± 39	5.5 ± 1.3	6.9 ± 1.7
$Z{+}l$	53 808 ± 164	51 206 ± 145	9.4 ± 1.1	11.1 ± 1.5
$t\bar{t}$	5960 ± 46	5204 ± 43	82.7 ± 5.3	75.4 ± 5.6
W+jets	73 ± 4	7 ± 1	0.4 ± 0.1	< 0.1
Diboson	2042 ± 17	1834 ± 16	21.5 ± 1.4	20.7 ± 1.4
MC total	70 493 ± 175	66 452 ± 158	324.9 ± 6.8	315.1 ± 7.2
Data	66 481	65 034	391	384



Correction at particle level: *using Fully Bayesian Unfolding method* [arXiv:1201.4612]:

→ Likelihood of data d given signals σ, nuisances Λ
→ Use pdf priors: flat for signal, Gauss for sys. nuisances
→ Compute posterior by MCMC sampling from pdf(σ, Λ)

$$\mathcal{L}(d|\sigma,\Lambda) = \prod_{i \,\in\, \text{recobins}} \text{Poiss}(d_i|x_i(\sigma,\Lambda))$$

$$x_i(\sigma,\Lambda) = L(\Lambda) \times (b_i(\Lambda) + M_{ij}(\Lambda)\,\sigma_j)$$

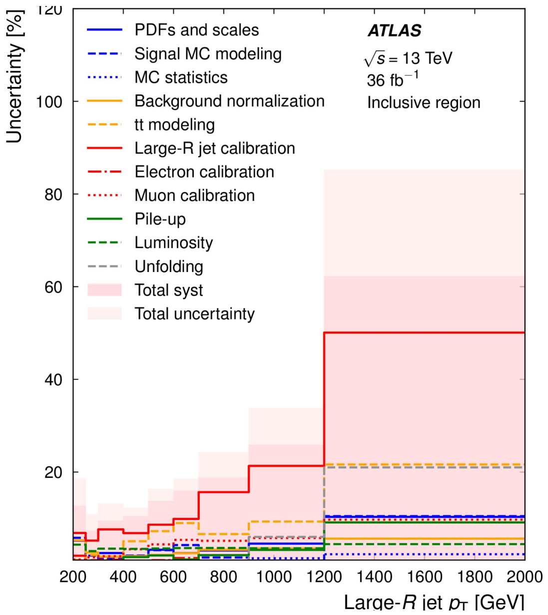
luminosity — background — response matrix

$Z + J$: measurement in flavour-inclusive phase-space

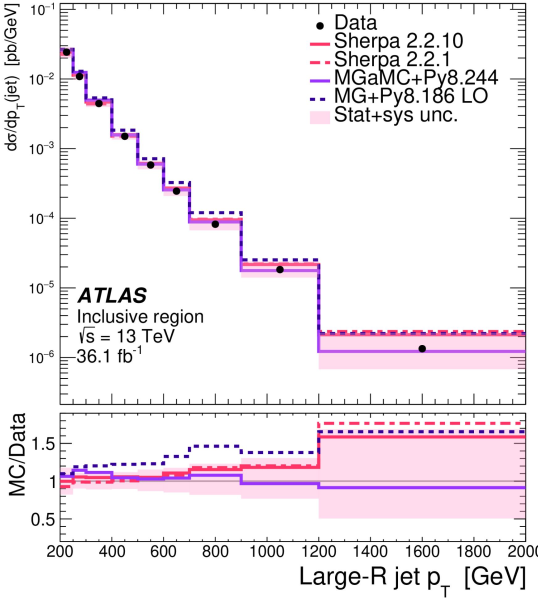
Differential measurements in flavour-inclusive boosted phase-space focusing on large-R jets

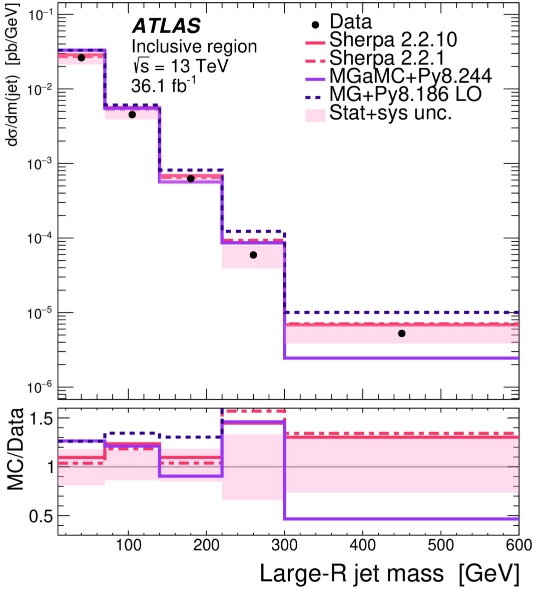
Main uncertainty:

→ large-R J reco. and calib.



Predictions test of new Sherpa 2.2.10 and MGaMC+Py8 NLO (FxFx)





All MCs overestimate cross-section, but ok within uncertainties, although slightly high large-R jet p_T - best description from FxFx

$Z + J$: results in the double b-jet tagged region

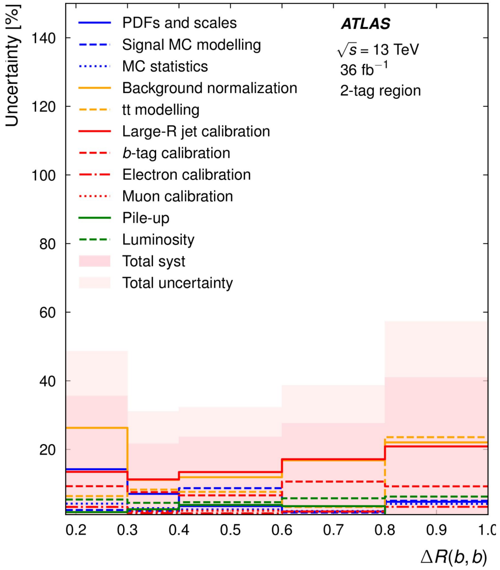
Differential measurements in double b-tagged boosted phase-space focusing on *bb* subjets

Main uncertainties:

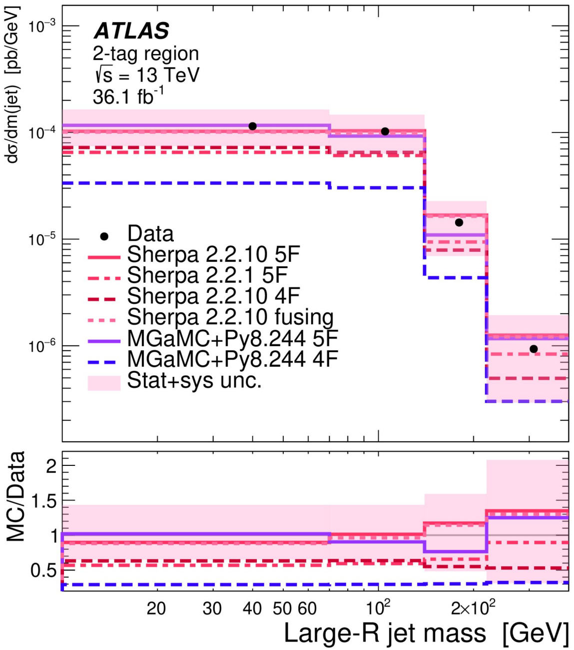
→ large-R J reco. and calib.

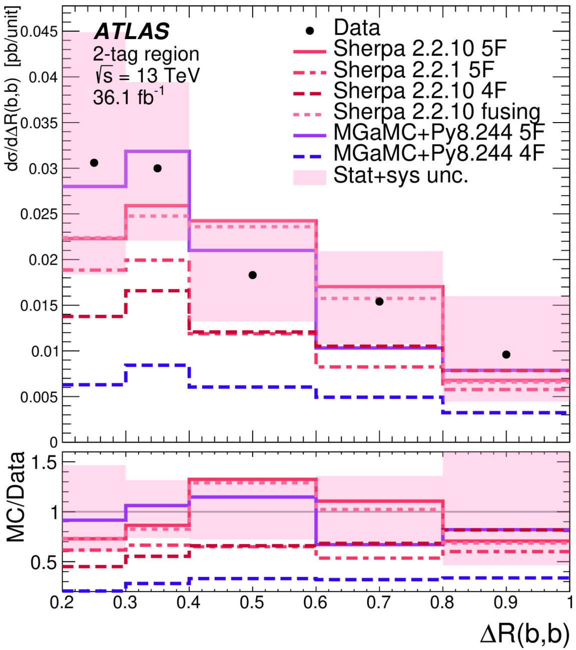
→ residual top bkg.

→ b-tag efficiency calib.



Predictions test also include 4 vs 5 Flavour Number Scheme MCs





Underestimate in 4FNS generators. Similar performance for Fusing 4+5FNS

Z + high-p_T jets in extreme phase-space with full Run 2 dataset

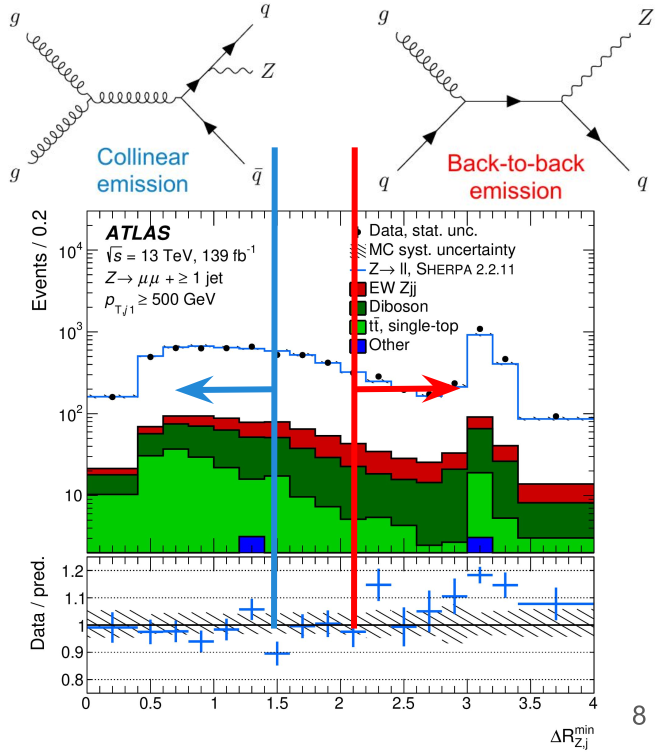
*Also here, investigating Z production in association with high p_T jets, but full 13 TeV dataset of **139 fb^{-1}** follows to pin down precisely areas of the phase-space challenging to model for QCD predictions*

Selection: $Z \rightarrow ll$ ($l=e/\mu$) plus **≥ 1 jet** (Anti-k_t $\Delta R=0.4$), **p_T>100 GeV**

Inclusive analysis further divided in multiple measurements:

- **High-p_T:** presence of $p_T(jet_1)$ > 500 GeV
- **Collinear:** High-p_T + ΔR(Z,closest-jet) < 1.4
- **Back-to-back:** High-p_T + ΔR(Z,closest-jet) > 2.0
- **High-S_T:** i.e. large jet activity with $S_T \equiv \sum_{jets} p_T^{jet}$ > 600 GeV

Sizable SM background at high p_T originates from $t\bar{t}$ contamination:
⇒ data-driven estimate with eμ to reduce top modeling uncertainties

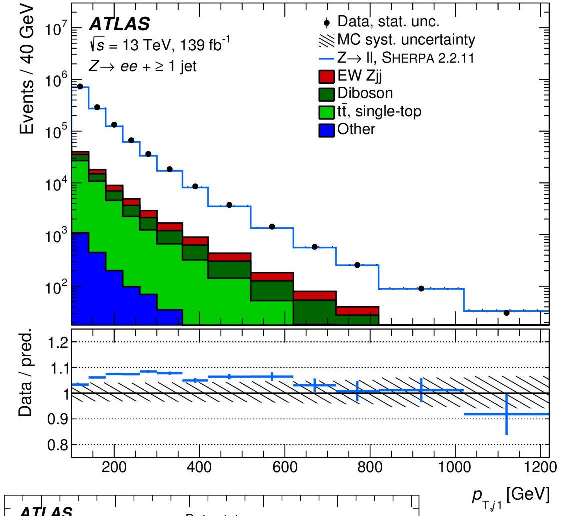


Z + high-p_T jets : reco, unfolding, systematic uncertainties

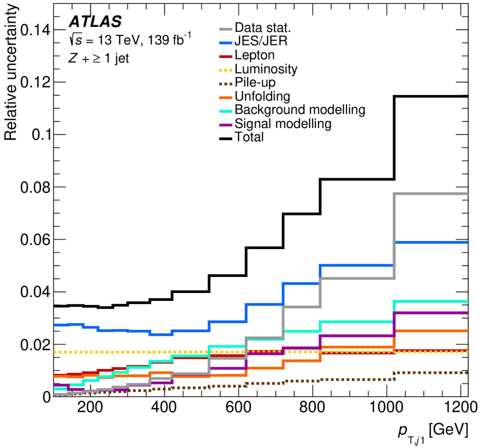
→ Newest Sherpa v2.2.11 (0-2p NLO + 3-5p LO) developed for ATLAS full Run 2 analyses and towards Run 3 [arXiv:2112.09588] used for nominal signal modeling and correction to particle level

→ Iterative unfolding (standard) method with 2 iterations

→ Latest JER/JES calibration give good precision up to high-p_T



Uncertainty source [%]	*Inclusive*	*High-p_T*	*Collinear*	*Back-to-back*	*High-S_T*
JER/JES	2.6	3.2	2.8	3.6	2.8
Lepton	0.9	1.6	1.4	2.0	1.1
Luminosity	1.7	1.7	1.7	1.7	1.7
Pile-up	0.1	0.4	0.4	0.4	0.4
Unfolding	0.5	1.0	1.1	1.4	0.8
Background modelling	0.5	2.0	2.0	1.9	1.7
Signal modelling	0.5	1.2	1.1	1.1	1.1
Total syst. uncertainty	3.4	4.8	4.4	5.3	4.2
Data stat. uncertainty	0.1	2.1	2.9	2.7	1.2
Total uncertainty	3.4	5.3	5.3	5.9	4.4



Z + high-p_T jets : Comparison with theoretical predictions

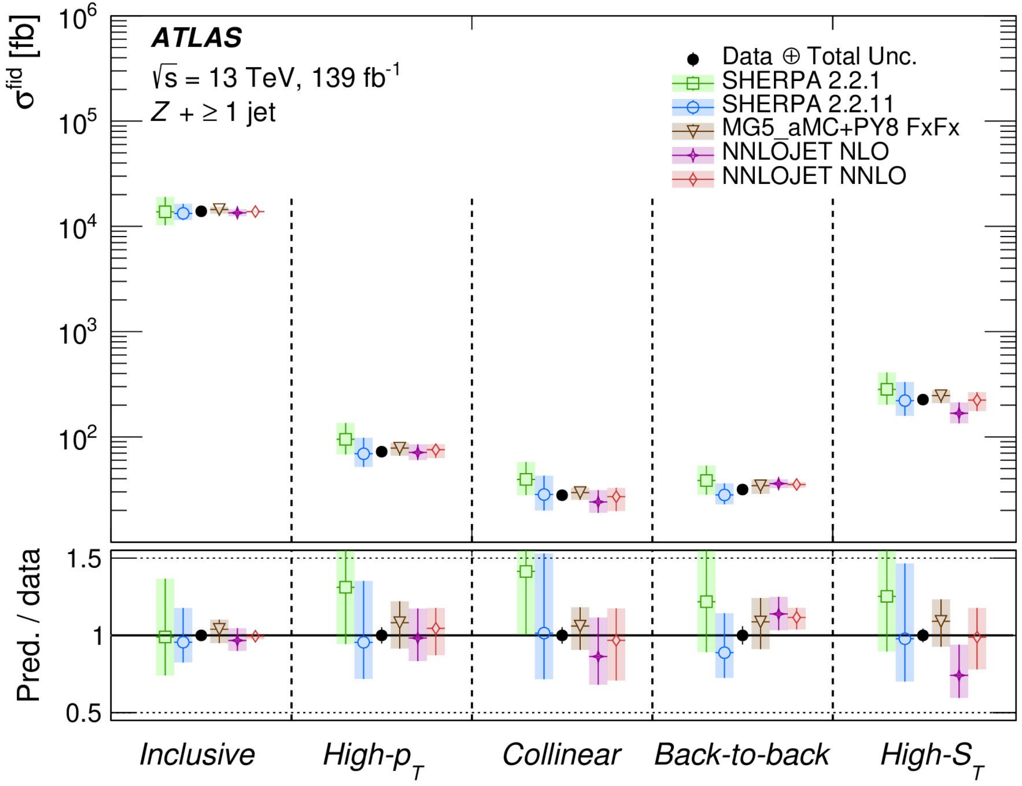
Theoretical predictions:

State-of-the-art NLO ME+PS QCD (sh2.2.11 includes also NLO EW) and pQCD in NNLO:

SHERPA v.2.2.11	0-2p NLO, 3-5p LO
MG5_aMC+PY8 FxFx	0-3p NLO
SHERPA v.2.2.1	0-2p NLO, 3-4p LO
MG5_aMC+PY8 CKKWL	0-4p LO
NNLOJET@NNLO	1p NNLO
NNLOJET@NLO	1p NLO

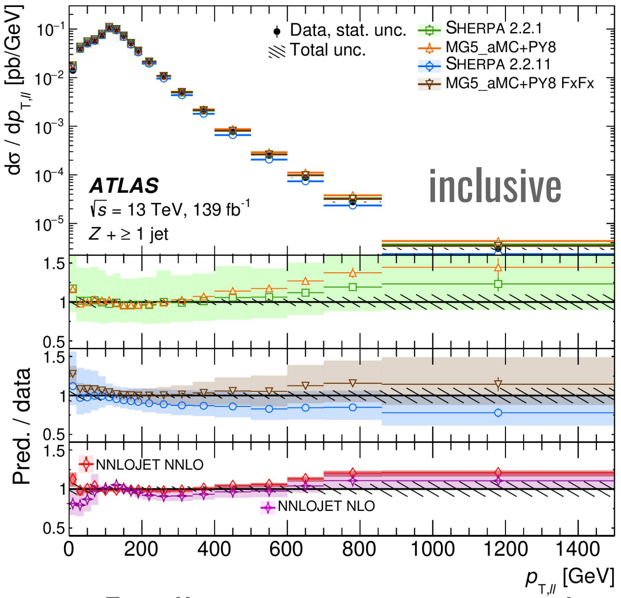
→ ***Data cross section always more precise than predictions!***

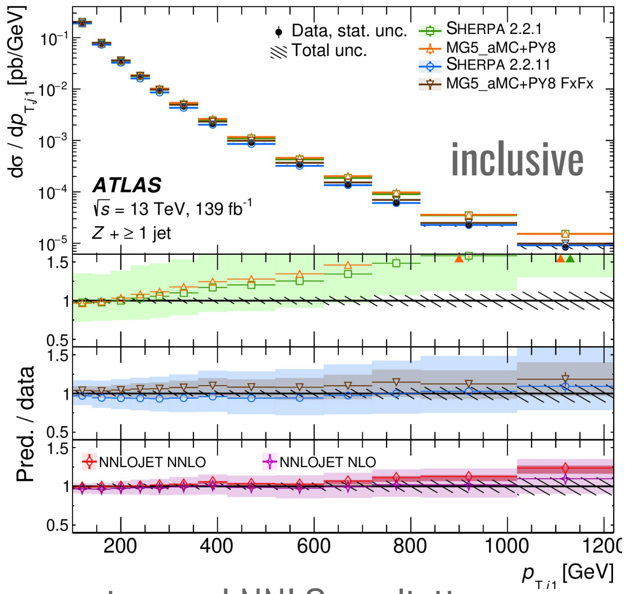
→ 7-point scale uncertainties show relevant variations between predictions

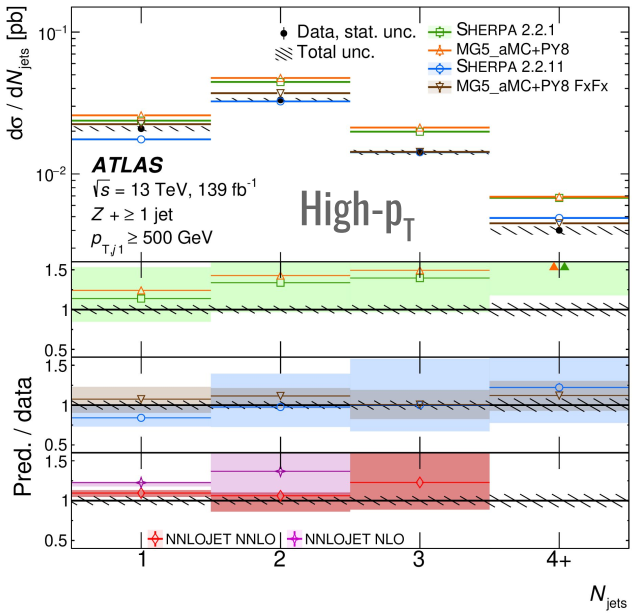


Z + high-p_T jets : inclusive and high-p_T results

Zp_T, leading jet p_T, N-jets: key variables in pQCD predictions but also crucial for searches/measurements because correct modeling with MC simulation allows to place selection cuts with reduced uncertainties







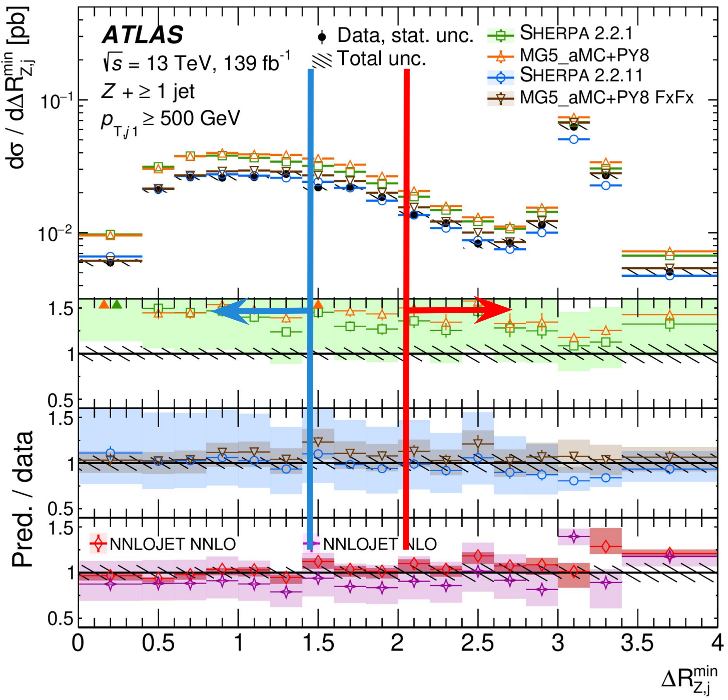
Excellent agreement from modern generators and NNLO predictions

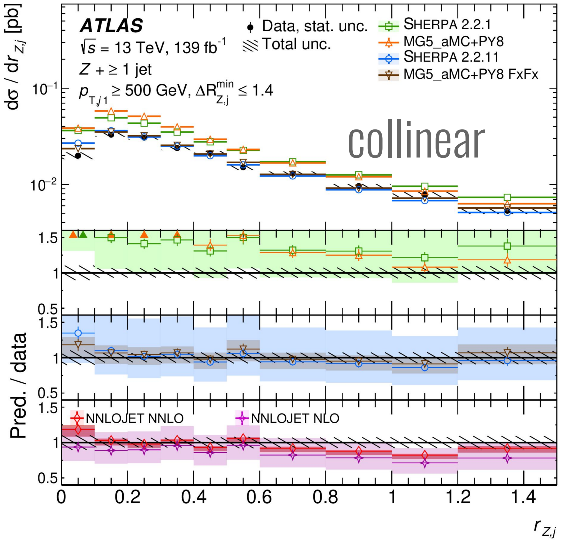
NB: sizable scale uncertainty differences: sh2.2.1 ~ sh2.2.11 **>** MGaMG FxFx ~ NLO **>>** NNLO

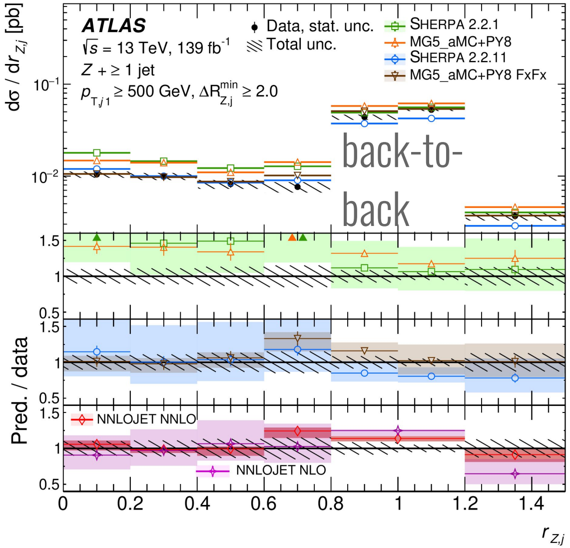
Z + high-p_T jets : results in collinear & back-to-back phase- space

Study of $r_{Z,j}$ variable interesting in back-to-back and collinear regions:

$$r_{Z,j} = \frac{p_{T,\ell\ell}}{p_{T,\text{closest-jet}}}$$







→ *Collinear:* low $r_{Z,j}$, as expected. Good agreement in new predictions

→ *back-to-back:* $r_{Z,j}$~1, confirms Z+1-hard jet, sensitive to EW corrections

Historical modeling issues in collinear region solved in new generators!

Summary and conclusions

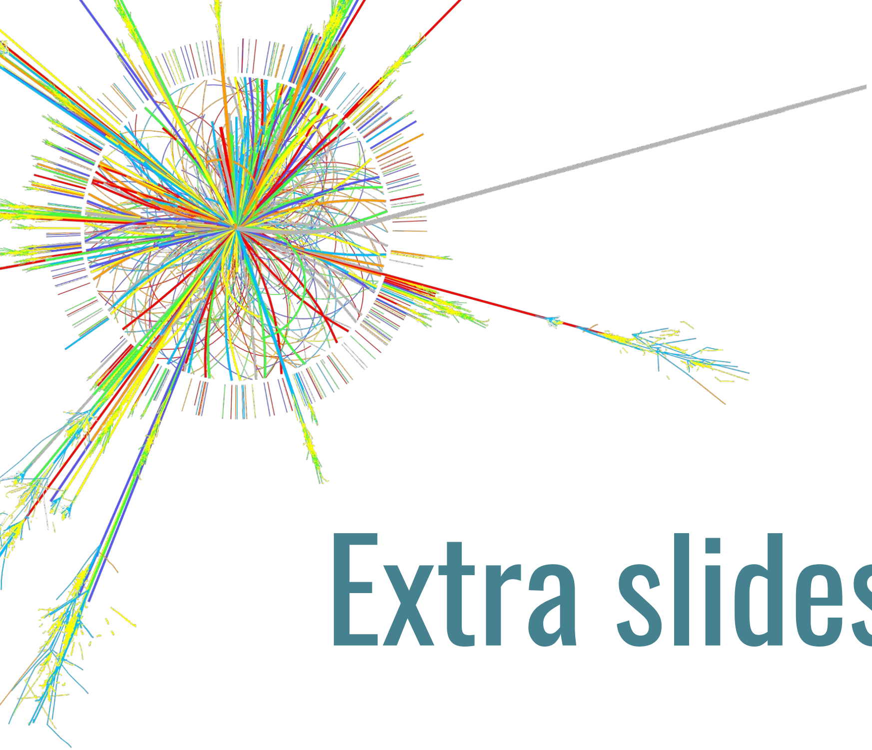
V+jets extreme phase-space became accessible thanks to improved reconstruction techniques and to the large stat. of full LHC Run 2 ⇒ crucial for benchmarking QCD predictions, test of MC modeling, and relevant for BSM searches or new Higgs measurements

Presented two new results comparing extreme phase-space data to state-of-the-art predictions:

- *Z + large-radius jet, flavour inclusive and with 2 b-tags* [arXiv:2204.12355]
- *Z + high-p_T jets in extreme phase-space with full Run 2 dataset* [STDM-2018-49]

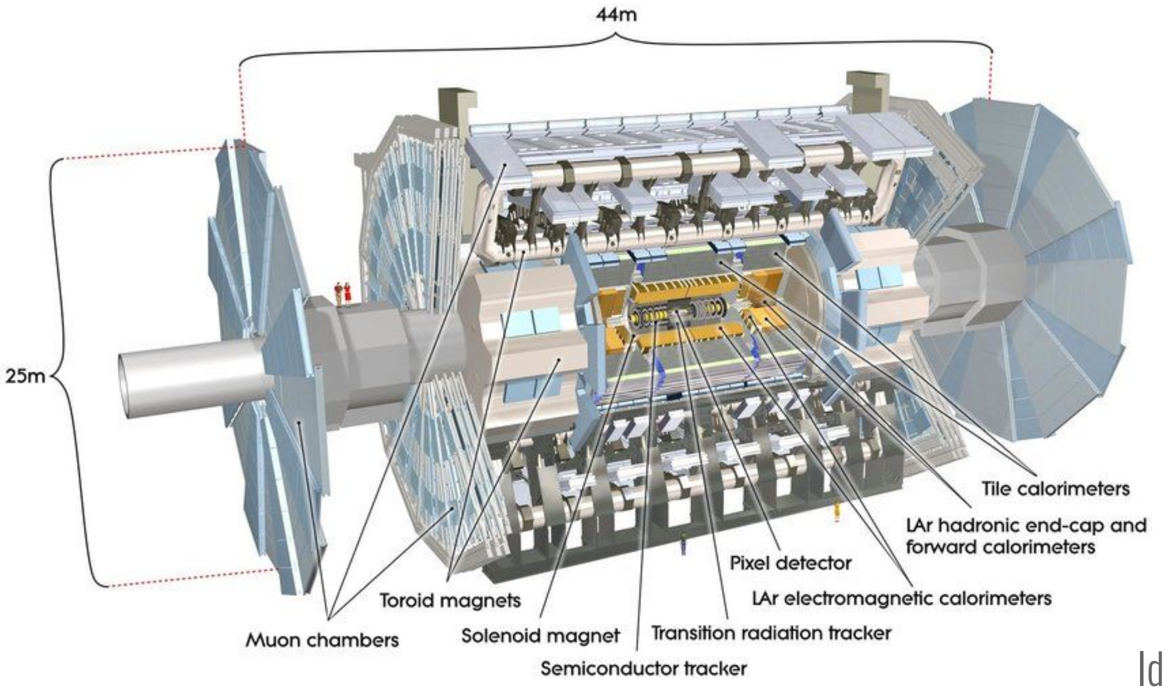
Good agreement with data within (still large) uncertainties in boosted regime, while agreement is excellent in collinear and back-to-back topologies when comparing to new NLO ME+PS matched/merged status and to NNLO predictions

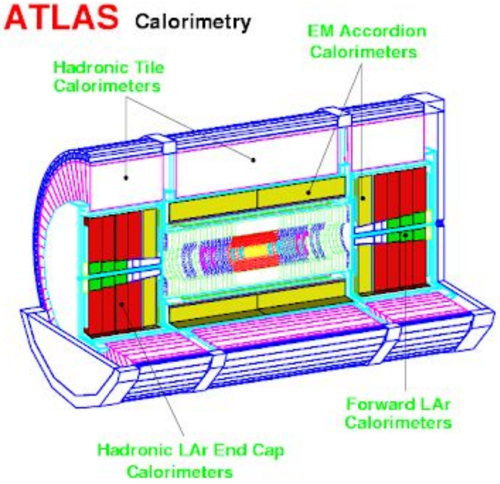
Results have been already partially used to test next generation of ATLAS V+jets MC simulation [arXiv:2112.09588], e.g. Sherpa 2.2.11 (0-2p NLO +3-5 LO) and MGaMC FxFx (0-3p NLO)

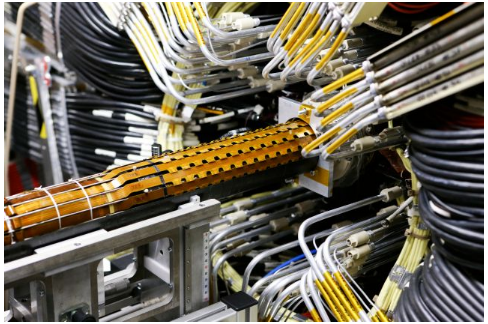


Extra slides & Backup Material

ATLAS Detector

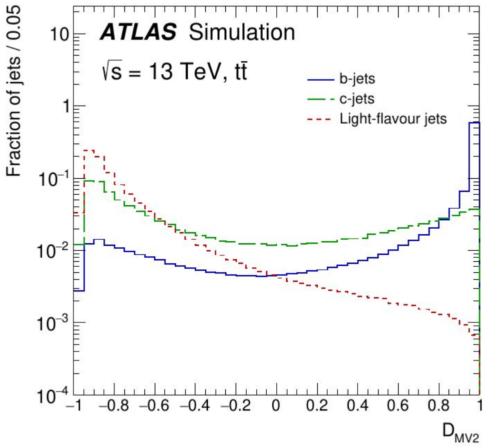




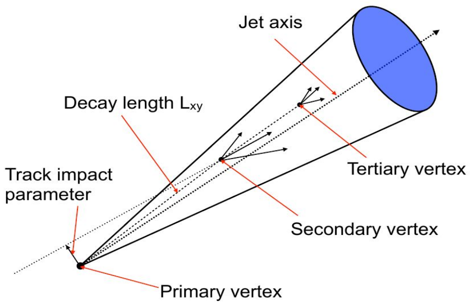


Identification of heavy-hadrons (m_b~4.2 GeV) decay at O(100 μm) w.r.t. primary vertex (PV) improved with new Run 2 pixel detector (IBL) very close to beamline

b-jet identification: Selection and Efficiency Calibration

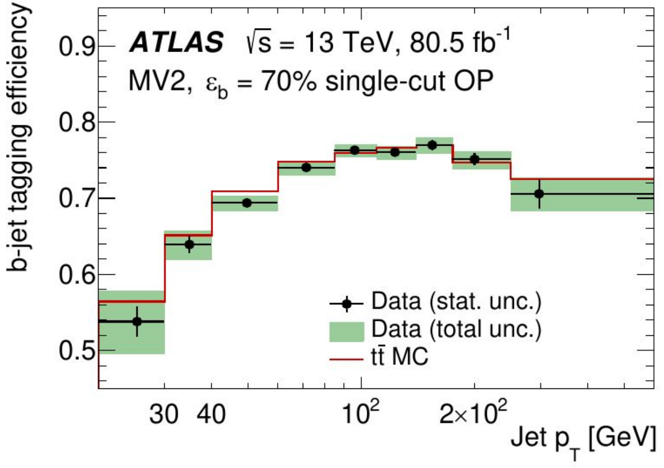


- Tracking & jet information condensed using multivariate (MV) algorithms for separation of *b*-jets vs different flavour jets - c or light-flavour



- *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)

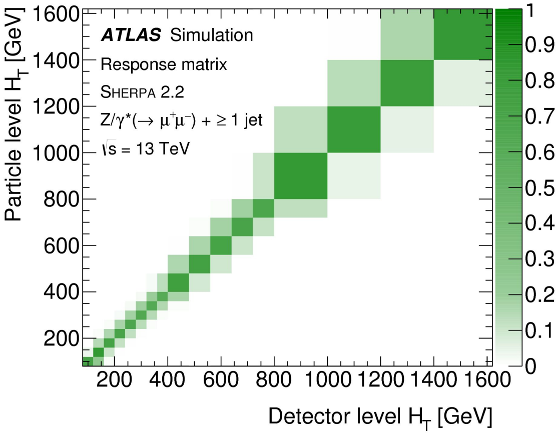


Iterative Unfolding with Bayesian method

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:

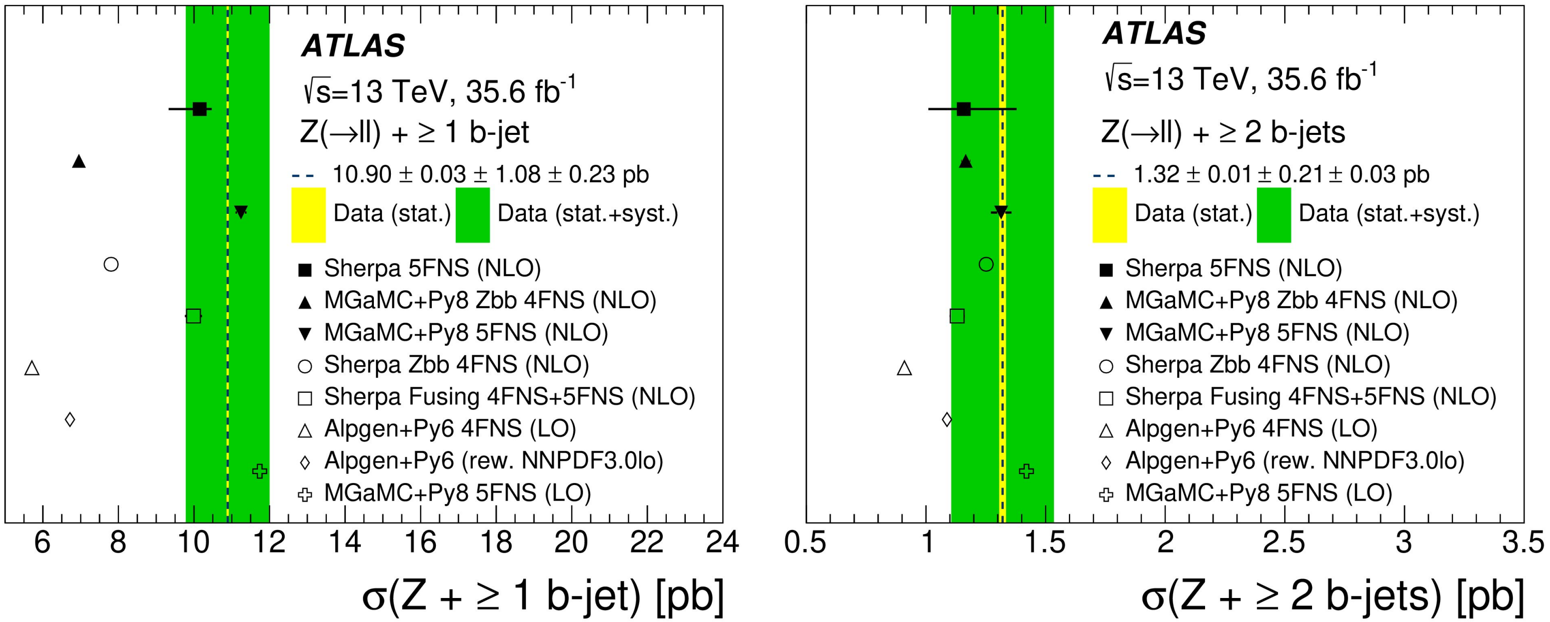
$$M(T_i|R_j) = M(R_i|T_j)\, P_0(T_j) \,/\, \mathrm{Sum}_l \; M(R_i|T_l)\, P_0(T_l)$$

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

$$T_j = \mathrm{Sum}_i \; M(T_j \mid R_i)R_i$$

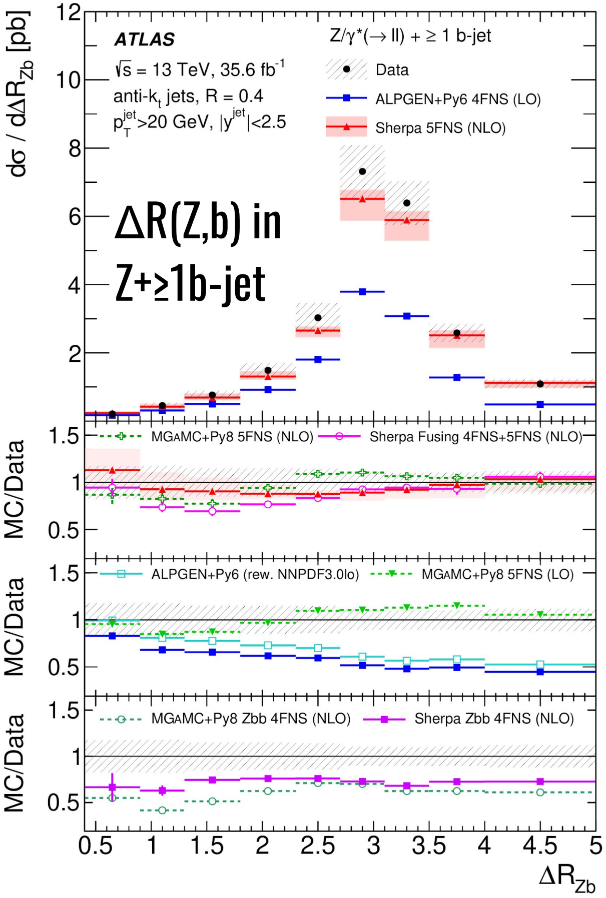
- In each further iteration the estimator of the unfolded distribution from previous iteration is used as a new prior

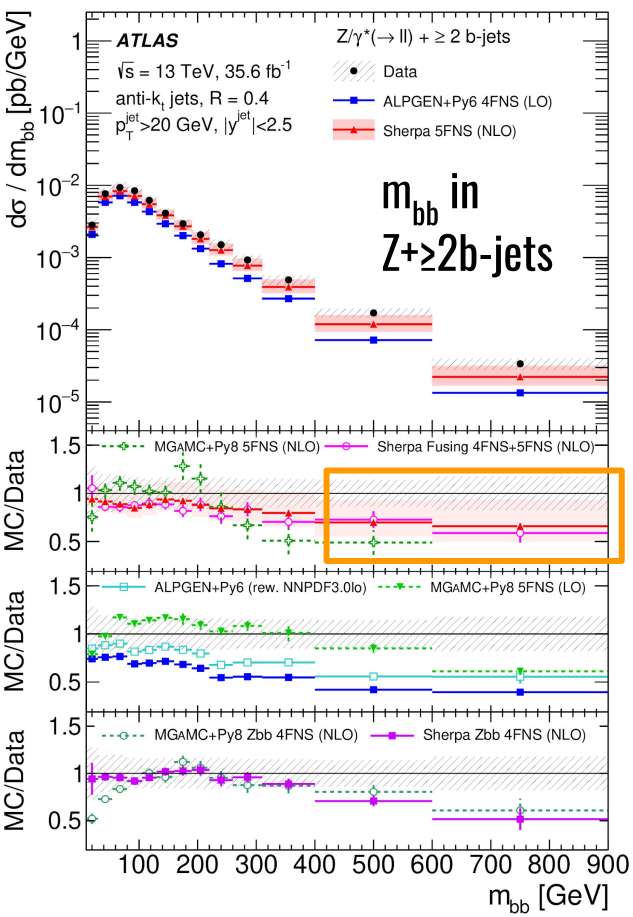
Z+≥1 b-jet and Z+≥2 b-jets inclusive cross-section: Results



- 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 in resolved Z+b(b)





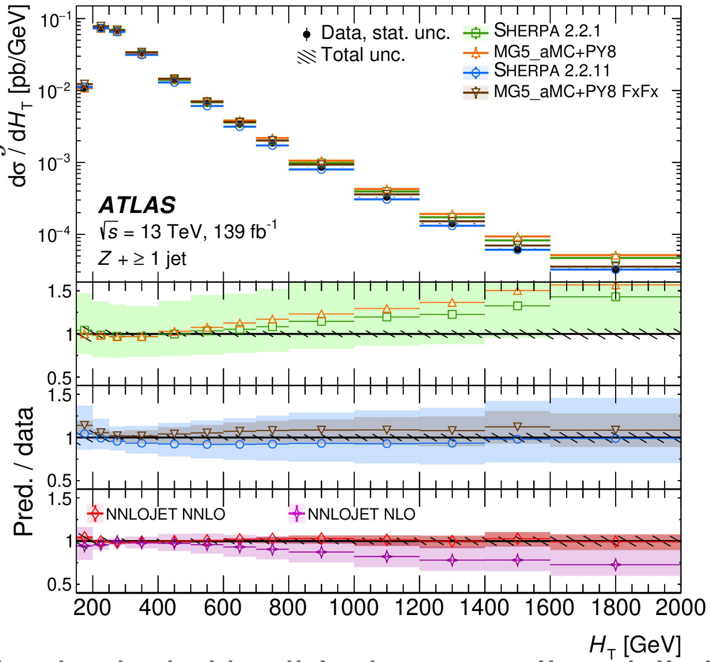
Z+≥1 b-jet & Z+≥2 b-jets phase space (mostly) well described by 5FS, while 4FS shows deficits in Z+≥1 b-jet

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…*

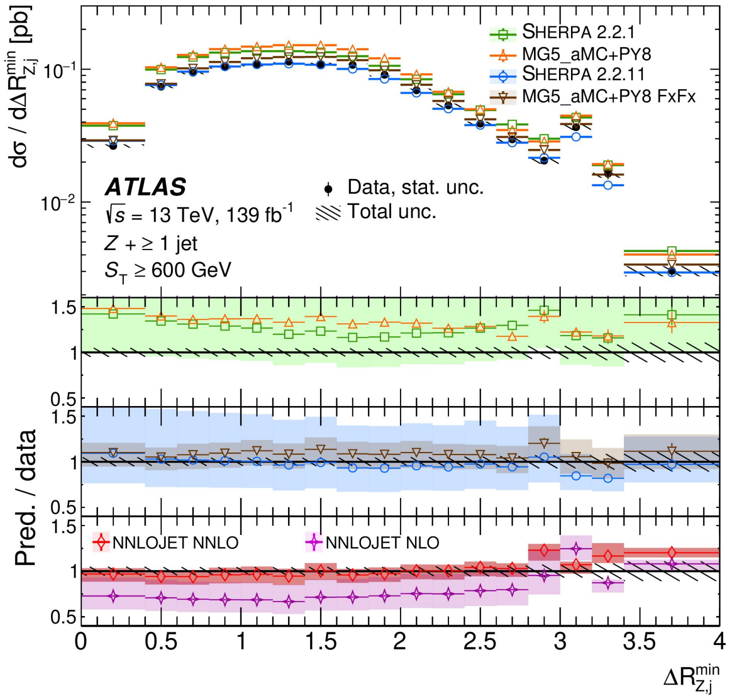
Z + high-p_T jets : results vs jet activity and at high S_T

$$H_T \equiv p_T^{\ell_1} + p_T^{\ell_2} + \sum_{jet} p_T^{jet}$$



S_T > 600 GeV shapes ΔR(Z,closest-jet)

→ less back-to-back, dominated by di-jet events

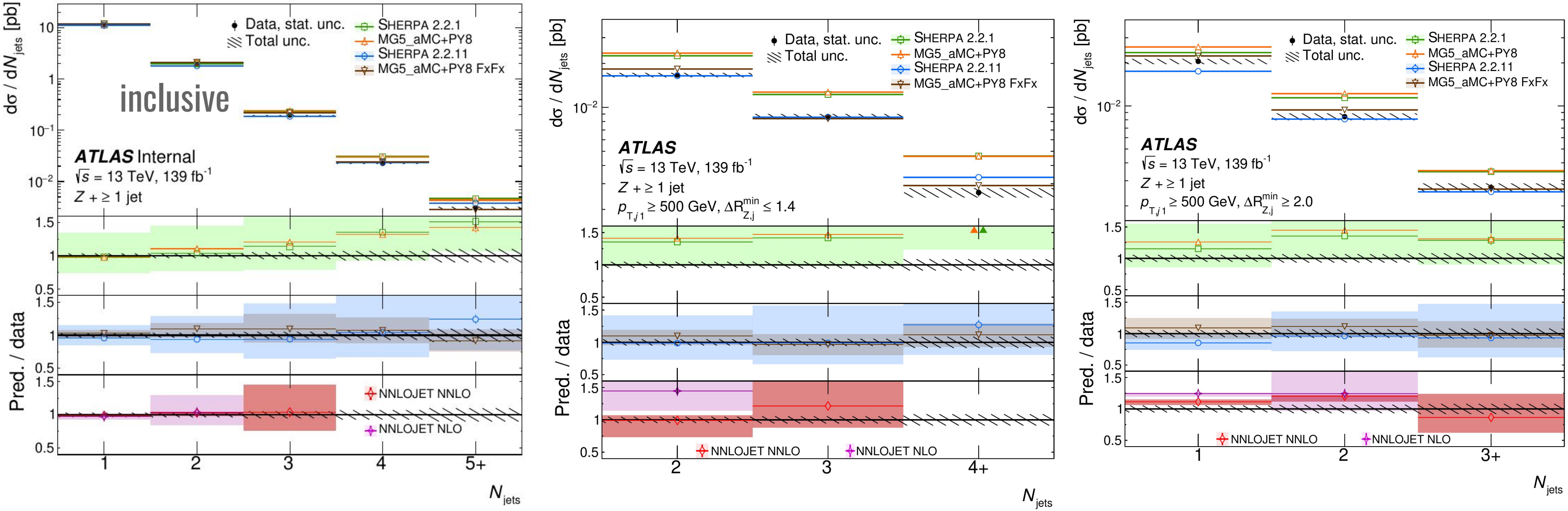


Low H_T , dominated by di-lepton p_T , well modelled

⇒ therefore disagreement at higher H_T from jets!

Again, excellent modeling from recent NLO ME+PS generators (Sh2.2.11, MGaMC FxFx), and NNLO pQCD

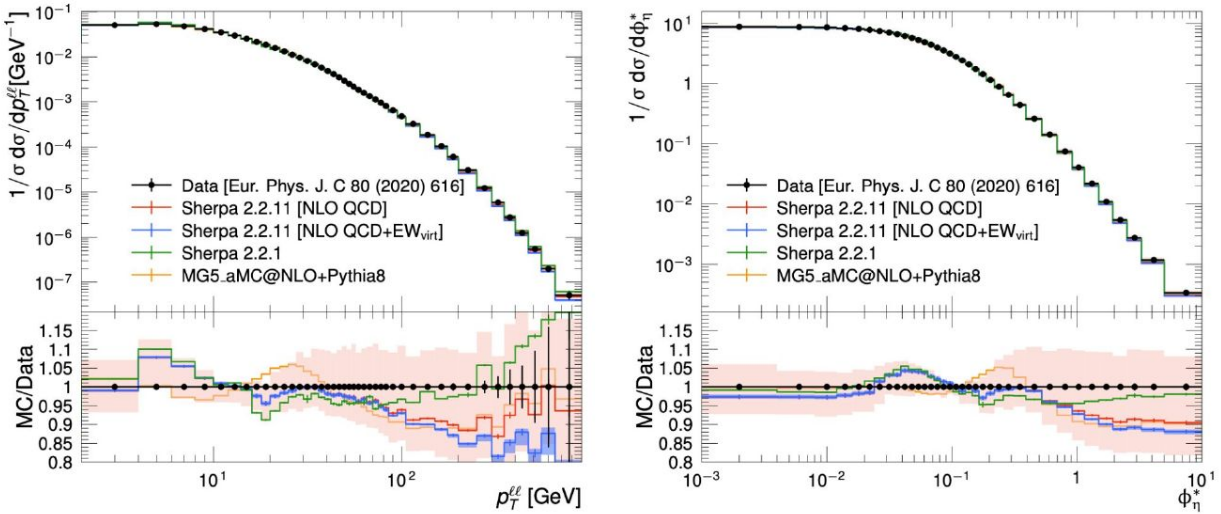
$Z + high\text{-}p_T\ jets$: jet multiplicities



Modelling and computational improvements to the simulation of single vector-boson plus jet processes for the ATLAS experiment [arXiv:2112.09588]

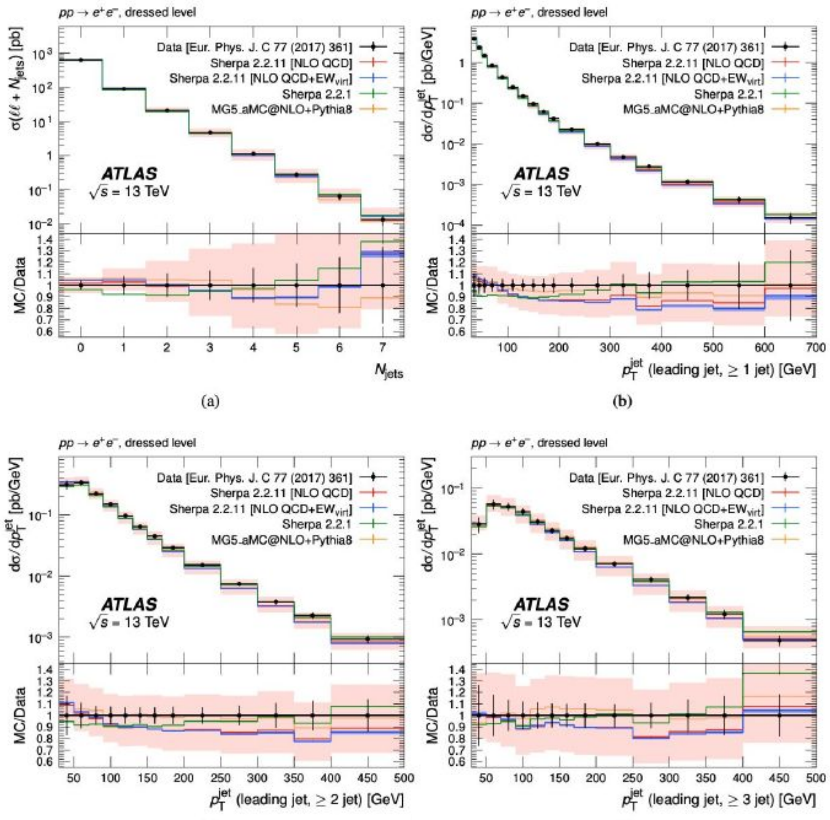
→ Sh. 2.2.11: 0-2p NLO, 3-5LO stat. enhancing function: **log(max H_T, Vp_T)**

→ MGaMC FxFx: 0-3p NLO FxFx stat. enhancing function: **H_T^2**

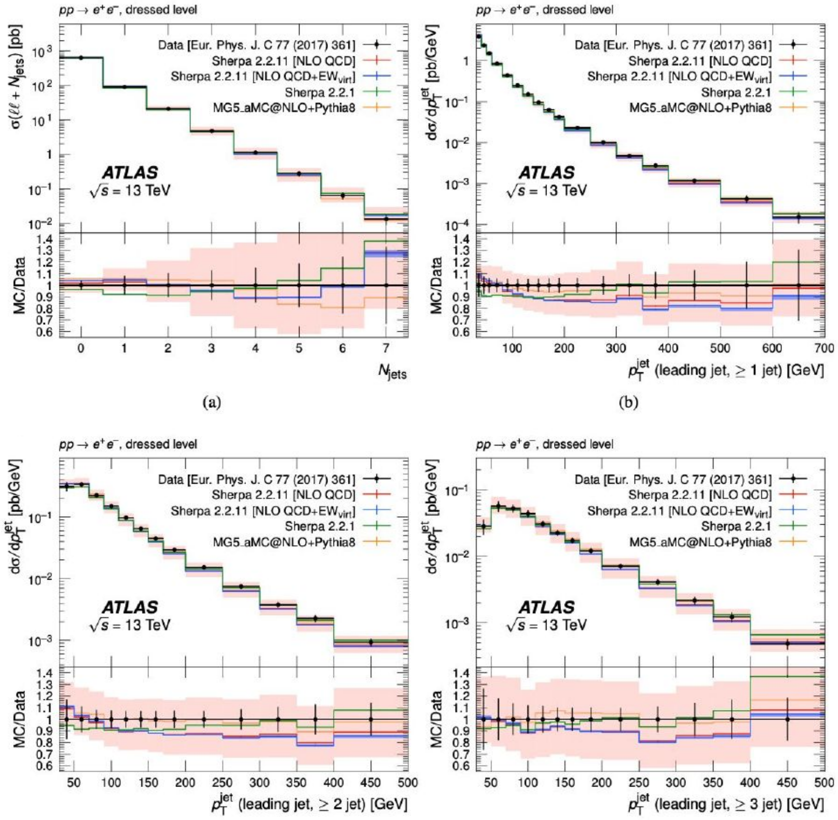


	$\sigma(pp \to \ell\ell)$ [pb]	$\sigma(pp \to \mu\nu)$ [pb]
Sherpa 2.2.1	2160 ± 86 (scale) ± 53 (PDF+α_S)	20697 ± 828 (scale) ± 507 (PDF+α_S)
Sherpa 2.2.11	2221 ± 155 (scale) ± 47 (PDF+α_S)	21781 ± 1525 (scale) ± 462 (PDF+α_S)
MG5_aMC@NLO+Pythia	2268 ± 92 (scale) ± 23 (PDF)	22163 ± 898 (scale) ± 225 (PDF)
NNLO QCD	$2067\ ^{+10}_{-14}$ (scale) $^{+52}_{-39}$ (PDF)	$20080\ ^{+100}_{-161}$ (scale) $^{+522}_{-401}$ (PDF)

Modelling and computational improvements to the simulation of single vector-boson plus jet processes for the ATLAS experiment [arXiv:2112.09588]



(a) (b) (c) (d)



(a) (b)