ICHEP 2022
BOLOGNA

Measurements of Z boson production in association with heavy flavor at ATLAS

Benedetto Giacobbe

On behalf of the ATLAS Collaboration

6–13 07 2022

outline

- Vector Bosons with Heavy Flavors in ATLAS:
  - From the resolved to the boosted regimes
- A new measurement: Z production in association with (b-tagged) large-radius jets
  - The physics case: inclusive and 2 b-tag selection
  - Objects selection and detector level results
  - Going to particle level
  - Results and discussion
- Conclusions

arXiv:2204.12355

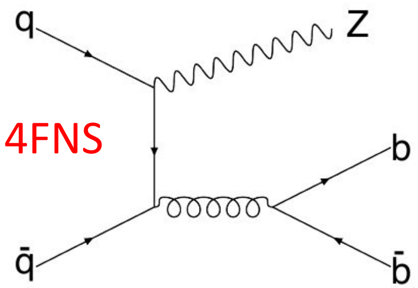
The ATLAS Collaboration - "Measurement of cross-sections for production of a Z boson in association with a flavor-inclusive or doubly b-tagged large-radius jet in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS experiment"

Z+Heavy Flavors: why exploring the boosted regime

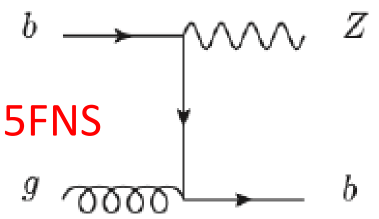
- Relatively larger production of boosted population at 13 TeV wrt 7 TeV
- Boosted jet production in association with Z-boson to reduce QCD background
- Boosted-jet topology sensitive to New Physics: heavy resonances decaying into boosted jets
- Understand background for H->bb (and other heavy resonances) in boosted regime
- Test of our understanding of hard collinear gluon splitting (small bb separation)
  - Description of gluon-splitting into Heavy Quarks (HQ) from Parton Shower (PS) models
  - Slicing of HQ production between ME and PS
- The initial state: how b-quark participate to hard scattering ?
  - 4 Flavor Number Scheme: bb only from gluon-splitting in ME
  - 5 Flavor Number Scheme: b-content in the proton PDF (massless approximation)
  - Fusing scheme (4+5FNS): in principle more accurate scheme in all kinematic regions

What do we Know from the Resolved regime ?
See next slide

4FNS

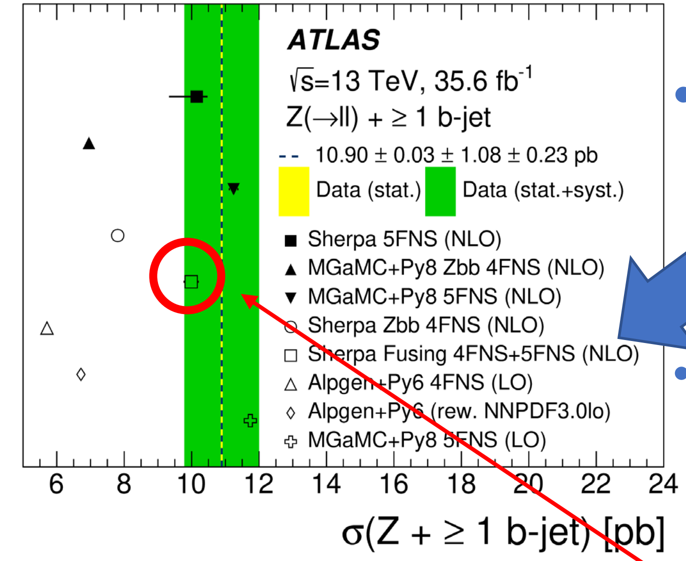


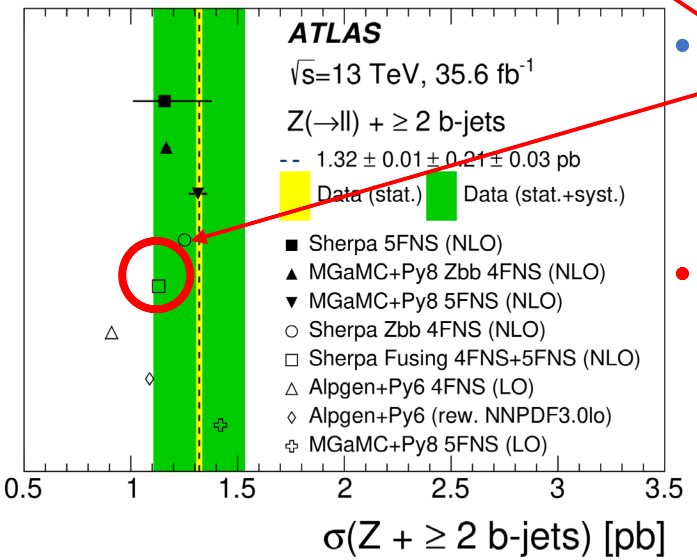
5FNS



Z+b(b): from resolved to boosted regimes

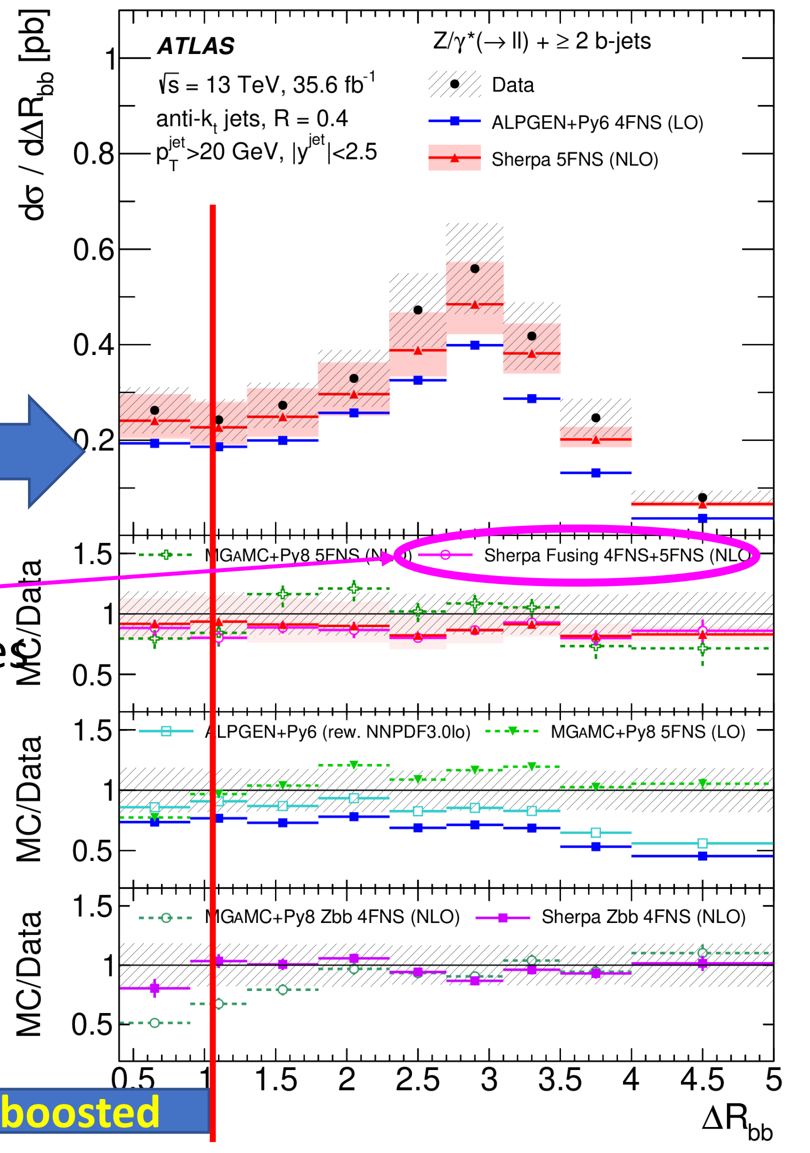
JHEP 07 (2020) 44



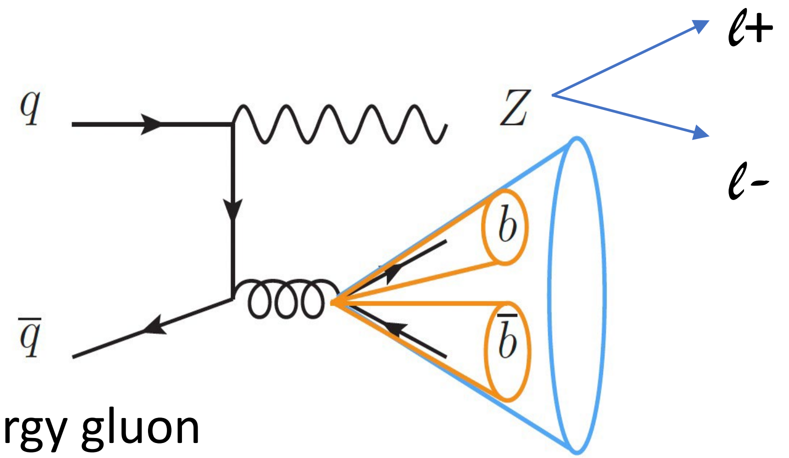


- Total cross sections well described by NLO calculations
  - 4-FNS largely underestimate ≥1 b-jet production
  - Both 4 and 5-FNS describe ≥ 2 b-jet production
- Differential cross section wrt ΔR(bb) well described by NLO SHERPA while large mismodelling in the g-splitting region by MadGraph 4FNS
- SHERPA fusing scheme (4+5 FNS) well describes some (not all) variables:
  - Interesting to test in the boosted regime
- Boosted regime test low-ΔR(bb) region dominated by g-splitting

JHEP 07 (2020) 44

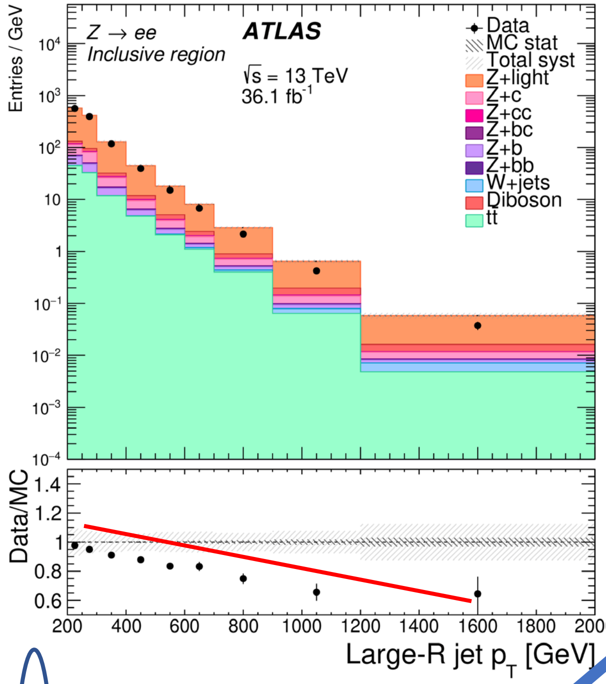


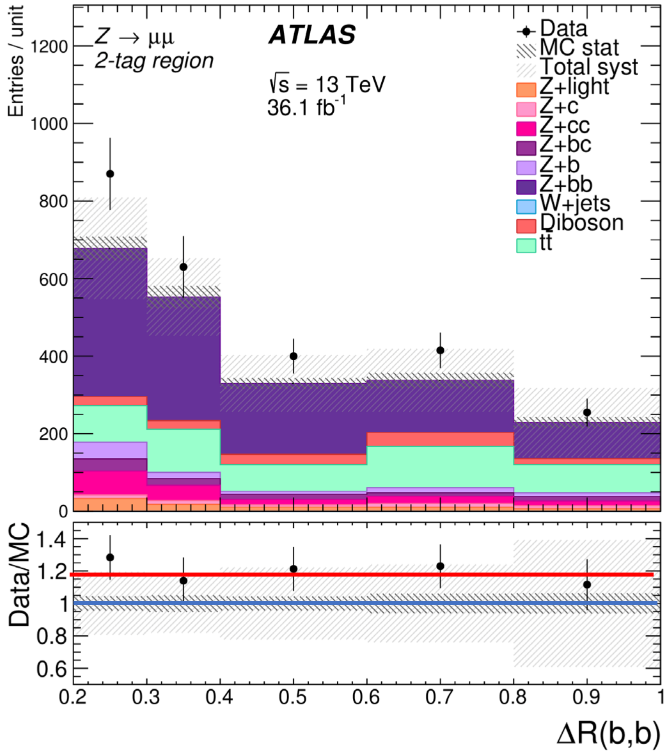
Z+large-R jets: object selection and phase-space regions



- Lepton and Z-boson selection:
  - Pairs $e^{\pm}$ ($\mu^{\pm}$) in the calorimeter (muon system) acceptance
    - PID (e/μ), Isolation, P_T> 27 GeV
    - $M_{\ell\ell}$> 50 GeV
- Jet selections (phase spaces):
  1. **Inclusive** Large-R jet as a definition of a hadronically decaying high-energy gluon
     - Anti-k_t algorithm with R=1, trimming to suppress PU and UE, p_T>200 GeV (boosted), |η|<2 for tracker coverage
  2. Large-R jet with **2 b-tagged** sub-jets: leading b-quarks in the jet from *g->bb* splitting
     - Anti-k_t algorithm with R=0.2, matching with tracks pointing to IP and p_T>10 GeV, |η|<2.5
     - b-tagging with MV2c10 algorithm at 70% b-tag efficiency (large rejection of c- and light-jets)
- differential cross section: what we do measure and test
  - Overall kinematics of the large-R jet: M^J and P_T^J (both inclusive and 2-tag)
  - Z-Jet properties and correlations: P_T(Z+J) and Δϕ(Z,J) (inclusive)
  - g->bb splitting properties: ΔR(bb) (2-tag selection)
  - Test of overall scale: total cross-sections from integration of differential ones

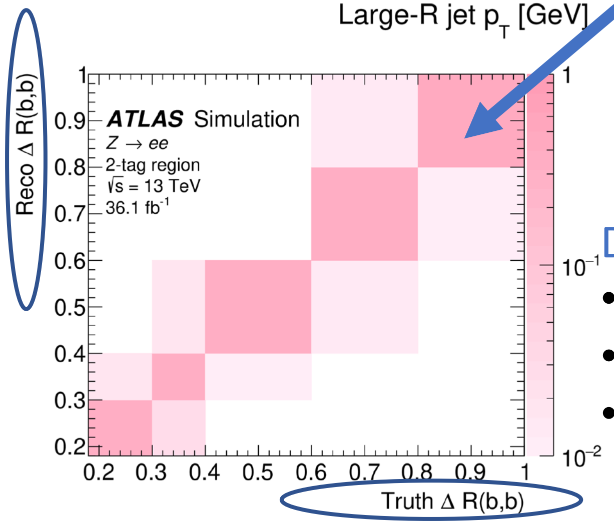
From detector to particle level





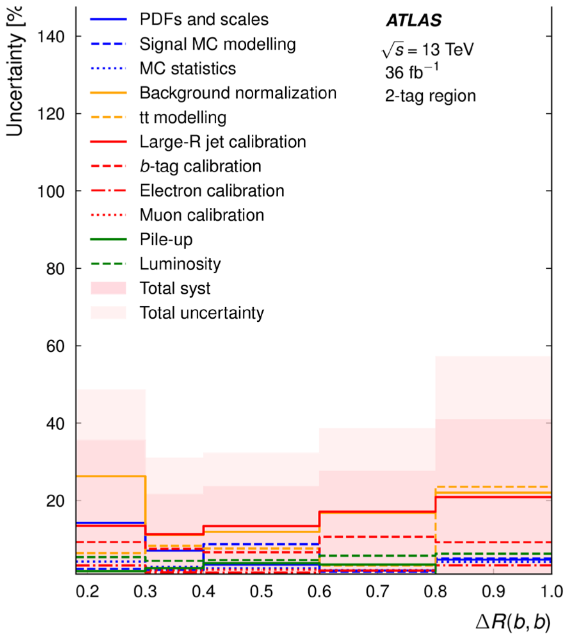
- Dominant background due to tt followed by Z+light-jets
- Detector level comparison to MC:
  - Overestimation and mis-modelling from SHERPA in various kinematic distributions for inclusive selection. Overall underestimation but better modelling in 2-tag selection

- Unfolding procedure based on *Fully Bayesian Unfolding method* [arXiv:1201.4612]
  - correct for detector's effects (from detector to particle level)
  - obtain detector-independent distributions within acceptance to compare to other experiments & calculations
  - measurement systematics propagated through unfolding



Dominant systematics:

- large-R Jet calibration (both inclusive & 2-tag)
- residual top background (modelling and normalization)
- b-tag efficiency calibration (only 2-tag)

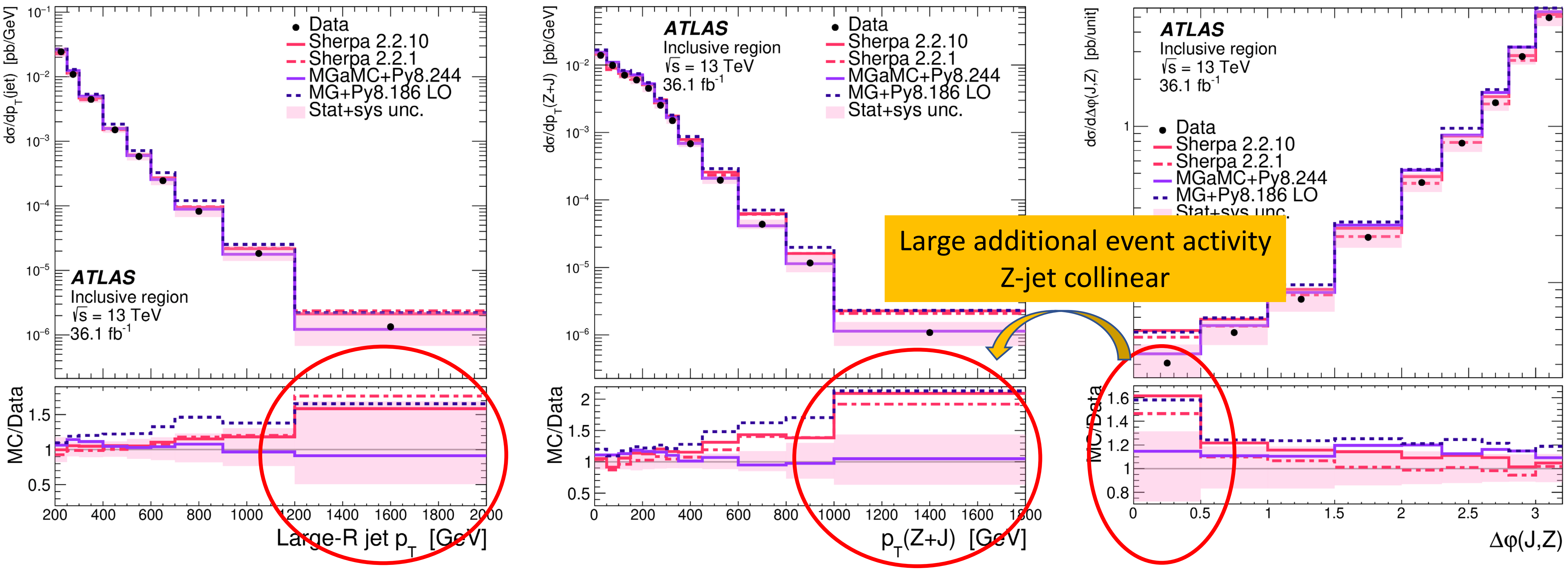


The total cross sections

σ [fb]	Measured	Sherpa 2.2.10	Sherpa 2.2.1	MadGraph5 +Pythia8	MadGraph2.2.2 +Pythia8
MC features		NLO – 4/5/fusing	NLO - 5FNS	NLO – 4/5FNS	LO – 5FNS
σ^{incl} ($x10^3$)	2.37±0.28	2.53±1.25	2.37	2.68±0.67	2.84
$\sigma^{2\text{-tag}}$	14.6±4.6	9.4±3.1 (4FNS) 14.9±4.2 (5FNS) 14.3±4.8 (fusing)	9.1	4.4±1.1 (4FNS) 14.4±1.9 (5FNS)	
R= $\sigma^{2\text{-tag}}/\sigma^{incl}$ [%]	0.62±0.12	0.59±0.39	0.42	0.54±0.21	0.38

- NLO generators well describe the inclusive cross-section
  - Measurement have competitive/smaller uncertainty than calculations (dominated by scale variations)
- Exclusive phase space measurement well reproduced by 5FNS schemes, not by 4FNS. Fusing scheme close to 5FNS
  - Confirm expectation that 5FNS is the best choice for the boosted HF production
- R measurement characterized by cancellation of some uncertainties (2-tag could benefit from full Run-2 statistics)
  - Discriminating power among pQCD models

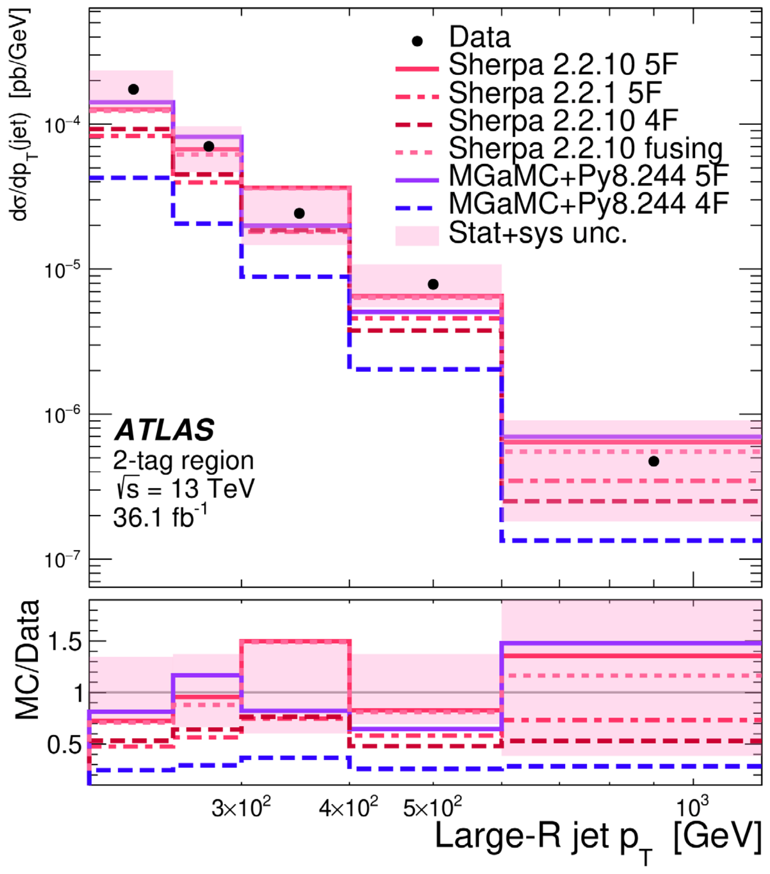
Inclusive properties of the large-R jet and Z-J system



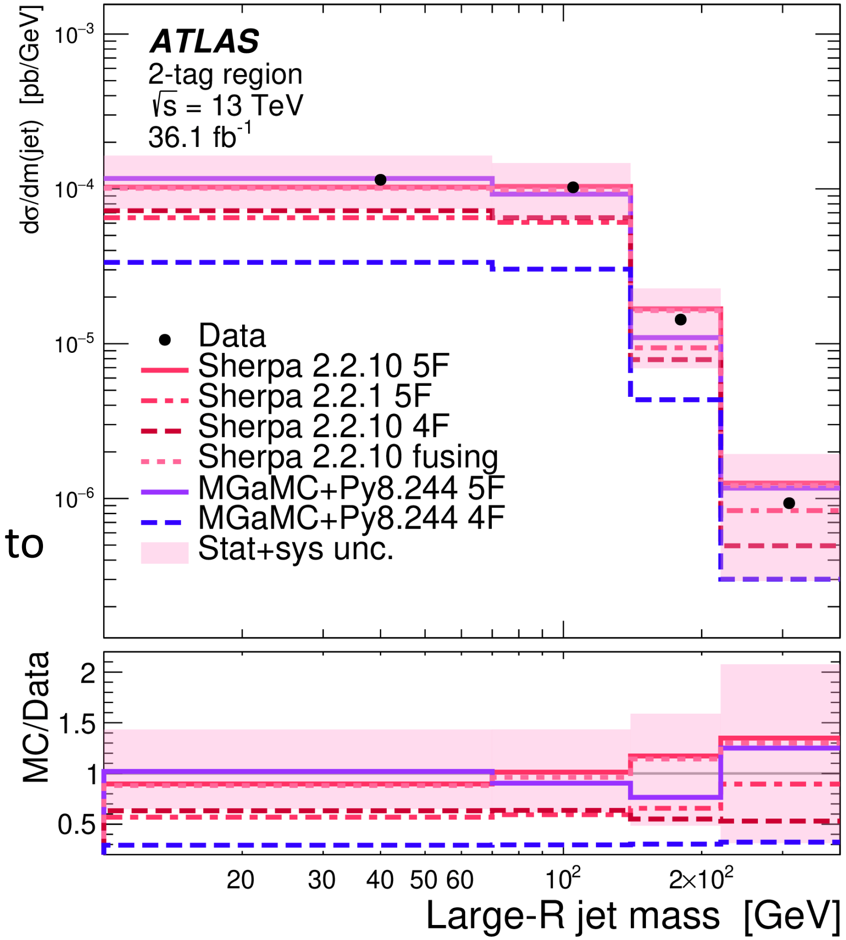
- Serious mis-modelling of NLO Sherpa's and LO MadGraph in the extreme phase space regions
  - poor description of extra-radiation - resulting in larger transverse recoil of jet and J+Z - calls for MC tuning
- NLO MadGraph well describes jet kinematics and Z+J correlations

Global properties of 2 b-tag large-R jet

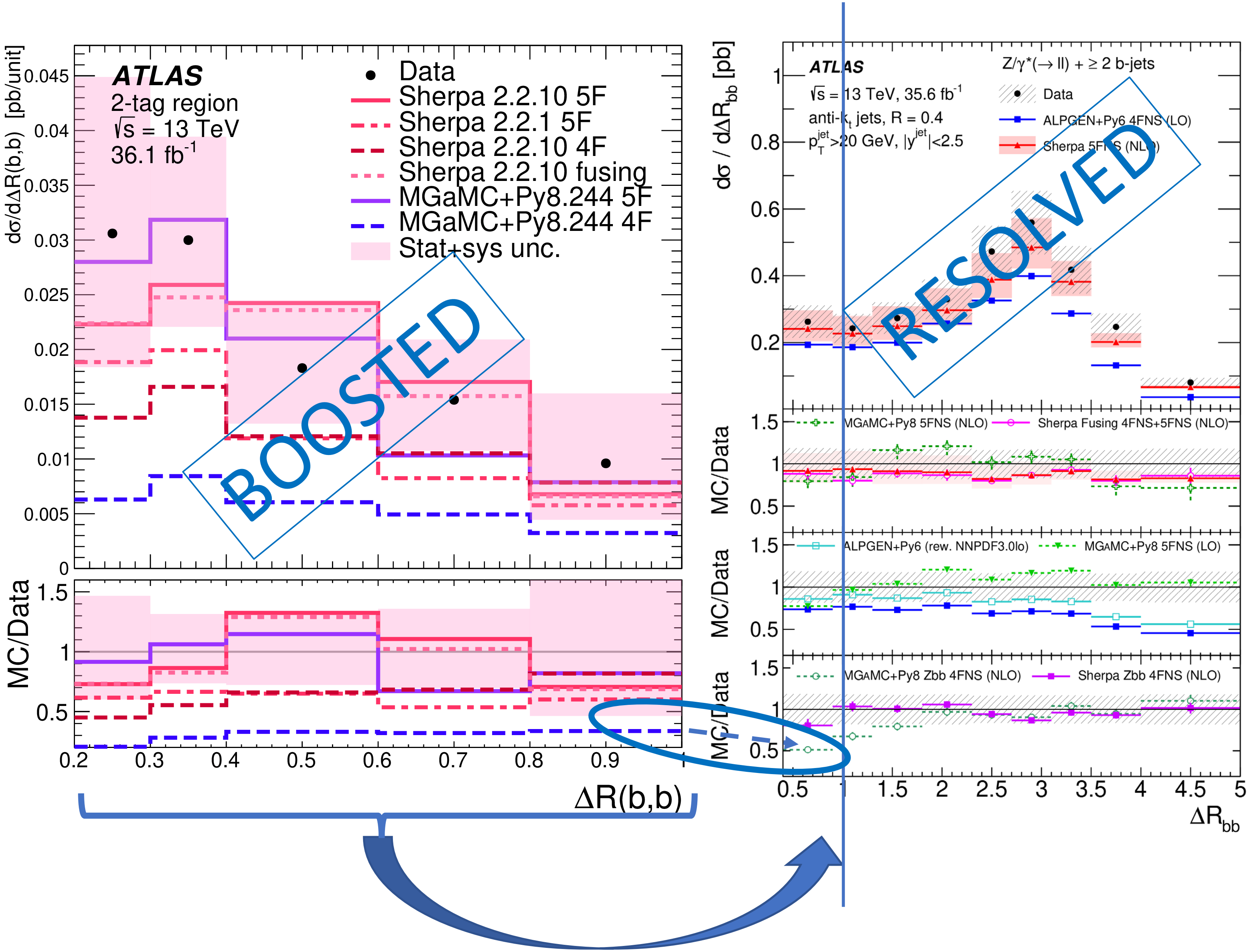
- Statistically limited measurements prevents from strong conclusions
  - Interesting to explore full Run-2 data-set to possibly provide stronger discrimination among models



- 5FNS provide better normalization
  - See total cross section
- Fusing scheme close to 5FNS also in differential distributions
- Rather flat D/MC ratios indicate no major mismodelling issues contrary to inclusive selection



Internal properties of the 2 b-tag large-R jet



- **2 b-tag boosted jets explore low ΔR(bb) region** (extreme phase space In resolved analysis)
- Overall normalization discrepancy observed in total cross section for 4FNS (in particular MGaMC)
- **All MC show flat ratios to data (within uncertainties),** including MGaMC (4FNS) which gave largest disagreement in resolved analysis

conclusions

- Z production in association with HF is a relevant part of the ATLAS physics program
  - Part of the vast Z/W+jets (or photons) ATLAS program at all energies
- First measurement of Z+jet in the boosted regime and with *b*-content substructure
  - Explore pQCD in extreme phase-space regions, study *b*-content in the proton, compare to a variety of state-of-art MC calculations
- Measurements already competitive with MC predictions despite limited statistics (2 b-tag selection): ability to discriminate between models
  - Interesting to explore full Run-2 statistics to boost discrimination power
- NLO calculations together with 5FNS scheme best describe total cross sections and most of measured differential cross sections with some exceptions in extreme regions of the inclusive distributions
  - data-to-MC comparison valuable input for further development of MC tunings

backup

Data and Montecarlo simulations

- pp collisions acquired in early Run-2 (36.1 fb^{-1} ± 2.1%) @ 13 TeV
  - Single-lepton trigger + DQ + basic event selection
- MC generators/tunings with 5FNS
  - 5FNS Sherpa 2.2.1: with ME 0-2p NLO + 3-4p LO
  - 5FNS MG5_aMC+Py8: with ME 0-4p LO (CKKW-L)

→ Signal and background simulation, Detector level comparison, unfolding

  - NLO SHERPA 2.2.10 with different tunings: 5FNS, 4FNS, fusing variation
  - NLO MG5_aMC+Py8 with different tunings: 5FNS, 4FNS

→ Particle level only: comparison with Unfolded data

unfolding

- Correction from detector to particle level with Fully Bayesian Unfolding method
  - 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