PROCEEDINGS OF SCIENCE

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

Benedetto Giacobbe*on behalf of the ATLAS Collaboration

Istituto Nazionale di Fisica Nucleare Sezione di Bologna,
Viale Carlo Berti Pichat 6/2, 40127 Bologna, Italy

E-mail: benedetto.giacobbe@bo.infn.it

Standard Model measurements involving the production of Z and W bosons in association with jets is an important part of the ATLAS physics program at the LHC. The measurement of integrated and differential production cross sections constitutes a powerful tool to test our understanding of the Monte Carlo generators and of the hadronization processes, as well as a way to explore the initial state of the proton-proton interaction. In these proceedings, a recent ATLAS measurement of the associated production of the Z-boson with a high-p_{T} (boosted) jet is described, both flavor-inclusive and with b-quark tagging of the jet. The measurement is performed using proton-proton collisions at $\sqrt{s}$ = 13 TeV at the LHC with an integrated luminosity of 36 fb^{-1} acquired by the ATLAS detector during Run-2. A discussion of the results and their complementarity with a previous measurement of the associated production of the Z-boson with b-tagged jets in the resolved regime is given. Particular emphasis is devoted to the discussion of the modeling issues affecting the flavor-inclusive selection and of the ability of the experimental results obtained with the b-tag selection to discriminate among models implementing different flavor-number schemes to describe the initial state of the interaction.

41st International Conference on High Energy physics - ICHEP2022
6-13 July, 2022
Bologna, Italy

*Speaker

PoS(ICHEP2022)798

1. Introduction

Jets produced by the decay of high-momentum particles can be detected as a single large-radius (large-R) *merged* jet with an internal sub-structure. This is also referred to as *boosted* jet. The deconvolution and the flavor-tagging of the small-R jets contained into it is a key technique [1] in ATLAS for the search of the hadronic decays of the Higgs-boson but also for the search of Beyond Standard Model heavy resonances and for the study of the dynamics in the high-p_{T} regime which is very sensitive to possible deviations from the SM predictions. All these studies can be best performed exploiting the pp collisions produced at the LHC collider. For these searches the correct understanding of the hard collinear parton splitting in QCD is crucial, as it constitutes their main background, and the approach of the MC generators, including factorization scales and splitting between matrix elements and parton-shower is best validated in the boosted regime. Also the impact of the heavy-quark masses to the calculations and to the modeling of the gluon-splitting into light or heavy quarks can be tested in the high-p_{T} regime exploiting efficient techniques for the tagging of the heavy-quark content of the large-R jets.

Another subject where the measurement of the production of high-p_{T} jets with heavy-quark tagging is crucial is the understanding of the initial state of the proton-proton interaction, by assessing whether the source of the heavy-quarks present in the final state has to be identified in the partonic cross-section only, or also in the proton's PDF, which would imply the contribution from heavy-quarks to the proton's structure. The first case is described by MC generators based on the *4 Flavor Number Scheme* (4FNS) models, where the b-quark density is neglected in the PDF and the initial-state b-quarks, with mass effects, are only due to gluon-splitting. The second case is encoded into the *5 Flavor Number Scheme* (5FNS) models which include the initial-state b-quarks into the proton's PDF, usually with massless approximation, and therefore allow also for a single b-quark to participate to the hard-scatter. In the truncated calculations the two approaches lead to different predictions for inclusive and differential b-quark production cross sections, both having pro and cons depending on the process scales. Recent computations such as the NLO Sherpa *fusing* scheme combine the strong aspects of both approaches and are predicted to approach the 5FNS scheme predictions in the boosted regime. Complementary measurements in both the resolved and boosted jet regimes are crucial to test the predictivity of the various approaches in a large momentum range and to provide a wider understanding of the initial state of the pp interaction. ATLAS has performed a measurement of the inclusive and differential production cross sections of a Z-boson in association with ≥ 1 and ≥ 2 b-jets in the resolved regime, compared to both 4FNS, 5FNS and fusing models with LO and NLO approximations [2]. NLO calculations with 5FNS and fusing scheme well reproduce both inclusive cross sections, which is not the case for the 4FNS largely underestimating the ≥ 1 b-jet cross section. Hints of mis-modeling are observed in the differential cross section at small jet-jet distance ($\Delta R(b\bar{b}) < 1$) which can be probed with high resolution in the boosted regime described in this report.

2. Object selection and phase space regions

Proton-proton collisions at $\sqrt{s} = 13$ TeV acquired by the ATLAS experiment [3] during Run-2 for a total integrated luminosity of 36.1 fb^{-1} are used in this measurement [4], using single-electron and single-muon triggers. Candidate events were required to have a primary vertex, charged leptons from Z-decay and large-R jets, with a reconstructed sub-structure of small-R track-jets, optionally b-tagged. At reconstruction level, electrons and muons are selected within the acceptance of the calorimeters and muon system with $p_{\mathrm{T}} > 27$ GeV with particle identification criteria and isolated from activity in the calorimeters and tracking system to enhance the purity of the signal. Jets, both large-R and the contained small-R ones, are the other crucial ingredients in this measurement, representing respectively any high-momentum hadronically decaying particle and their leading decay partons. Large-R jets are reconstructed from calibrated topological clusters of calorimeter cells using the anti-k_t algorithm with radius parameter $R = 1.0$ and must have $p_{\mathrm{T}} > 200$ GeV and $|\eta| < 2$. The contained small-R jets ($R = 0.2$), reconstructed starting from tracks pointing to the primary vertex to ensure a high angular resolution, are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$. The possible b-jet nature of the small-R jets is assigned based on the ATLAS MV2c10 multivariate b-tagging algorithm [5] which ensures a high rejection power against $c-$ and light-jets at 70% b-tagging efficiency.

Based on this definition of the basic objects, the *inclusive* event selection, both in data and MC, is defined as containing events with two same-flavor leptons with an invariant mass > 50 GeV (for the Z-selection) and at least one large-R jet. A subset of these events in which the large-R jet contains exactly two b-tagged subjets constitute the *2-tag* (or *exclusive*) selection. For both selections, the total cross sections are measured, as well as their ratio which profits from the cancellation of several systematic uncertainties. Differential cross section measurements aim at determining the properties of:

- the large-R jet: mass (m_J) and transverse momentum (p_T^J), for both the inclusive and 2-tag selections;
- the Z-jet system: transverse momentum of Z-jet vector sum (p_T^{Z+J}) and their angular separation ($\Delta\phi(Z, J)$), for the inclusive selection only;
- the internal sub-structure of the large-R jet (2-tag selection only): angular separation ($\Delta R(b, \bar{b})$) between the two b-tagged subjets.

Reconstruction-level distributions are corrected to the particle-level fiducial phase-space using the unfolding procedure to deconvolute the effects and biases introduced by the detector and the reconstruction. In order to minimize extrapolations and provide a more straightforward comparison to theoretical calculations and other measurements, all particle-level cross-sections are measured in a fiducial region. The Fully Bayesian Unfolding (FBU [6]) technique is used which performs a likelihood fit in the parameter space of signal cross-sections plus a set of nuisance parameters representing the background compositions and the systematic uncertainties, all discussed hereafter.

The main background processes considered are Z+jets, W+jets, and dibosons, simulated with NLO Sherpa 2.2.1, and $t\bar{t}$ simulated with Powheg-Box v2 HVQ, and they are used both to estimate the background levels, to unfold the data and for comparison to the unfolded distributions. An alternative LO Z+jets sample was simulated using MadGraph 2.2.2 and was used for the evaluation of the systematic uncertainties on the main background. In addition to the MC generator versions used for Z+jets, the signal was also generated at particle level with the Sherpa 2.2.10 generator with both 5FNS, 4FNS and *fusing* schemes for comparison with the unfolded data. All the details about the used MC generators and parton-shower algorithms can be found in [4]. The systematic uncertainties in the measurement are mainly due to detector effects and reconstruction, MC modeling, and from the unfolding procedure and they are dominated by: the leptons energy/momentum scale and resolution and their reconstruction, identification, isolation and trigger efficiencies; the energy scale, mass scale, energy resolution, and mass resolution of the large-R jet; the flavor tagging of sub-jets (for the 2-tag selection), including b-tagging efficiency and mis-tag rates for c- and light-flavor jets; the uncertainty arising from the reweighting of the MC predictions to the pile-up distribution measured in the data; the modeling of both the signal and background processes. The impact of each uncertainty to the final results depends on the specific measured cross-section and varies typically from $\sim 10\%$ to up to more than 50% in the extreme regions of some of them.

3. Measurement of differential cross sections

The electron and muon channels were first unfolded independently to verify their mutual consistency and then a global unfolding was performed to obtain the final results. Some of the obtained differential cross-sections unfolded at particle-level as a function of various kinematic variables are presented and discussed hereafter.

Figures 1 shows examples of global properties of the large-R jet and of the Z-jet system in the inclusive selection, namely the cross-sections as a function of the large-R jet p_{T} (left) and of $\Delta\phi(Z, J)$ (right), respectively. Data are compared to the particle-level predictions from NLO Sherpa 2.2.1 and 2.2.10, NLO MadGraph5 and LO MadGraph, all normalized to their own calculated cross-sections. The MadGraph5 predictions well describe the data in the full spectra, while excesses in both Sherpa and LO MadGraph are present in the extreme phase-space regions (high jet p_{T} and low angular Z-jet separation), due to a mismodeling of the extra radiation in these generators affecting the region at high jet-p_{T} and collinear Z-jet emission. This represents a very useful input from the present analysis for the tuning of the MC generators in the boosted phase-space.

In Figure 2 (left and center), two examples of global properties of the large-R jet with 2-tag selection are given, namely the cross-sections as a function of the large-R jet p_{T} and mass. Data are compared to NLO Sherpa and MadGraph5 predictions, using both the 4FNS and the 5FNS (for Sherpa 2.2.10 also the *fusing* scheme). The study of the predictions of the various flavor schemes can be indeed probed by the exclusive selection where the b-tag is aimed to select jets originated from the b-quark hadronization. Within the larger uncertainties (both systematic and statistical) all models qualitatively well describe the shape of the distributions, also in the extreme regions of the phase-space, where discrepancies were observed in the inclusive selection. More solid conclusions about the data description by the various flavor schemes can be drawn from the comparison of the total cross-sections (see next section).

Figure 2 (right) shows the differential cross-section as a function of the distance between the two sub-jets for the 2-tag selection. This variable describes the internal dynamics of the b-tagged large-R jet and therefore is a very valuable observable as input for analyses searching for heavy resonances decaying into $b\bar{b}$ in the boosted regime for which the QCD background needs a precise modeling. Moreover this measurement is complementary to the resolved analysis already published by ATLAS [2] as it probes the low separation region. The shapes of the differential cross-sections in the exclusive selection are well described by all generators, both 4 and 5FNS, in their full range, although a better agreement in absolute value is provided by the 5FNS and *fusing* scheme MC, while 4FNS provide a clear underestimation (see next section).

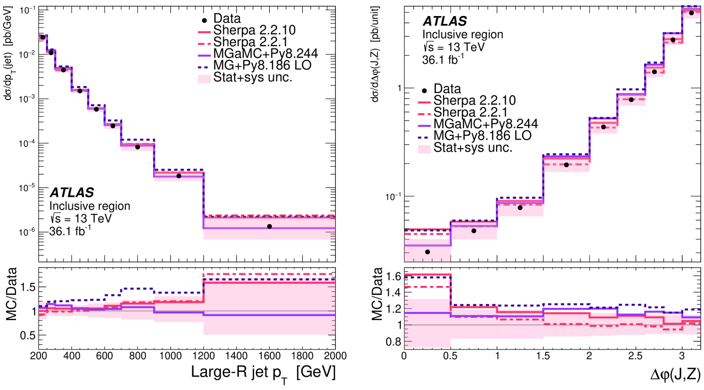

Figure 1: Inclusive selection: differential cross-sections unfolded at particle-level as a function of the large-R jet transverse momentum (left) and of the Z-jet angular separation [4]

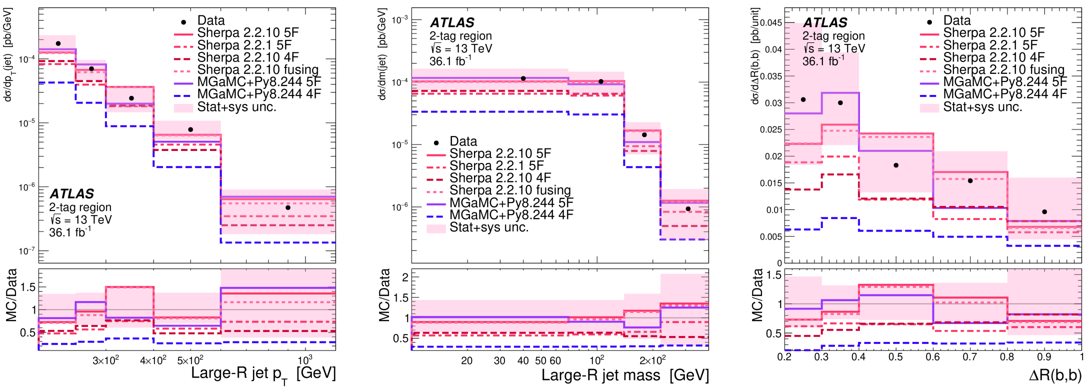

Figure 2: Exclusive selection: differential cross-sections unfolded at particle level as a function of the large-R jet transverse momentum (left), jet mass (center) and jet-jet separation [4].

4. Measurement of total cross sections

The total fiducial cross sections for both the inclusive and exclusive selections are obtained through the integration of the respective angular cross-sections in the fiducial phase-space. The measured values are respectively $\sigma^{incl} = 2.37 \pm 0.28$ pb and $\sigma^{2-tag} = 14.6 \pm 4.6$ fb.

The measured inclusive cross-section is well reproduced by all NLO predictions (Sherpa 2.2.1 and 2.2.10 and MadGraph5) with only a slight overestimation still well compatible with the data, while the LO MadGraph central configuration provides a substantial overestimation [4]. As for the exclusive 2-tag cross section, Sherpa 2.2.10 (both 5FNS and *fusing* schemes) and MadGraph5 (5FNS) reproduce very well the measured values, while all generators using the 4FNS largely underestimate the data, confirming the expectation that 5-flavor and *fusing* schemes are the more appropriate choice for heavy-quark production in the boosted phase-space.

A measured ratio $R = \sigma^{2-tag}/\sigma^{incl} = 0.62 \pm 0.12\%$ between the two selections is obtained. Both Sherpa 2.2.10 and MadGraph5 NLO predictions provide consistent values with the measured data, while the older Sherpa 2.2.1 and the LO MadGraph provide a large underestimation. Thanks to the reduced systematic uncertainty due to the cancellation of the systematics common to the two selections, this ratio represents a new experimental discriminator between pQCD models of high-p_T heavy-flavor production, which is also the case for the total exclusive cross-section despite the larger measurement uncertainties.

5. Conclusions

The measurements of total and differential cross-sections for the production of a leptonically decaying Z-boson in association with a large-radius boosted jet, both with and without requirement of its b-quark content, represent a fundamental piece of the ATLAS SM program at the LHC, and complement a similar measurement in the resolved regime. These measurements provide an important test of perturbative QCD, with particular emphasis on the production rates and kinematics of bottom quarks. These are a significant background to several important Higgs or New Physics searches, and are affected by significant theory and modeling uncertainties. The observation of significant mismodelling in the NLO MC description of the inclusive selection by some of the used generators calls for a tuning of the used MC. The same does not happen in the exclusive b-tag selection, which is not affected by the same problems, at least in the limit of the available statistics. A clear indication of the superior power of the 5-flavor and *fusing* schemes to predict the total cross sections with respect to the 4FNS was clearly obtained. This analysis would greatly benefit from the addition of more data, in particular with the addition of the Run-3 statistics being collected.

References

[1] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet Substructure as a New Higgs-Search Channel at the Large Hadron Collider*, Phys. Rev. Lett. 100 (2008) 242001, arXiv:0802.2470 [hep-ph].

[2] ATLAS Collaboration, *Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP 07 (2020) 044, arXiv:2003.11960 [hep-ex].

[3] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08003.

[4] 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*, arXiv:2204.12355 [hep-ex].

[5] ATLAS Collaboration, *Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC*, Eur. Phys. J. C 79 (2019) 375, arXiv:1808.07858 [hep-ex].

[6] G. Choudalakis, *Fully Bayesian Unfolding*, arXiv:1201.4612 [physics.data-an]