

# 1. Study of associated production of vector bosons and b-jets in pp collisions at the LHC<sup>1, 2, 3,</sup> 4

## 1.1 Introduction

The vector boson production in association with one and two b jets at the CERN Large Hadron Collider is important for many different experimental and theoretical reasons. Bottom quarks have a peculiar signature which allows one to easily identify them thanks to a displaced decay vertex. The associated production with vector bosons is an important background to VH production with the Higgs boson decaying to b quarks, and many new physics searches. Theoretically, it offers an interesting testing ground for predictions involving heavy quarks.

There are two possible options for the calculation of processes with b-quarks in the final state at hadron colliders. In the four-flavour scheme (4F) b-quarks are not present in the parton density of the incident protons. They can only be generated in the final state and they are usually massive. In the five-flavour scheme (5F) the b-quark mass is considered small with respect to the scale of the process  $Q$  and powers of logarithms of the type  $\log(Q^2/m_b^2)$  are resummed into a b parton density function. The b-quark is therefore massless in this approach, though higher order mass effects can be included in the calculation. A critical review of the different flavour number schemes and of the status of theoretical calculations is available in Ref. [1]. To all orders in perturbation theory the two approaches give identical results up to power suppressed mass terms. At finite order, however, they may give different results. In the 4F scheme the computation is more complicated, but the full kinematics of the heavy quarks is taken into account. Furthermore it can be easily interfaced to parton showers, even at NLO using the MC@NLO [2] or the POWHEG [3] formalisms. On the other hand logarithms in the initial state are not resummed and could lead to large discrepancies in the inclusive quantities like the total cross-section. In the 5F approach, on the opposite, calculations for the inclusive quantities are highly simplified and generally more accurate, but differential distributions and exclusive observables are technically more involved.

The goal of this study is to compare the most recent measurements with the predictions of the state of the art generators using 4F and 5F scheme. The report is organised as follows. In Section 1.2 we provide a short description of the ATLAS and CMS measurements, available in the Rivet framework, for  $V + b + X$  and  $V + b\bar{b} + X$ , where  $V$  is a  $Z$  or a  $W$  boson. In Section 1.3 we describe the generator setups used to obtain the predictions, which are compared to the measurements in Section 1.4 for the  $Z$  and 1.5 for the  $W$ , before conclusions are drawn in Section 1.6.

## 1.2 Rivet Routines

Results in this study were produced using three Rivet routines to compare to published ATLAS and CMS data.

**ATLAS Z+b(b)** Measurement of differential production cross-sections for a  $Z$  boson in association with b-jets in proton-proton collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector [4] (Rivet routine ATLAS\_2014\_I1306294). A pair of opposite-sign charge dressed leptons<sup>5</sup> (i.e. electrons or muons) with  $p_T > 20$  GeV and  $|\eta| < 2.5$  are required, with a dilepton mass between 76 and 106 GeV. Anti- $k_t$  0.4 jets are reconstructed from all final state particles, and required to have  $p_T > 20$  GeV,  $|y| < 2.4$  and not overlap with the leptons used to make the  $Z$  candidate ( $\Delta R(jet, l) > 0.5$ ). Jets are labelled as b-jets based on matching with  $\Delta R < 0.3$  to a weakly decaying b-hadron with  $p_T > 5$  GeV.

---

<sup>1</sup>Section coordinator: V. Ciulli

<sup>2</sup>Contributing authors: M. Bell, J. Butterworth, G. Hesketh, F. Krauss, G. Luisoni, G. Nail, D. Napoletano, C. Oleari, L. Perrozzi, S. Platzer, C. Reuschle, B. Waugh

<sup>3</sup>S. Platzer acknowledges support by a FP7 Marie Curie Intra European Fellowship under Grant Agreement PIEF-GA-2013-628739.

<sup>4</sup>The work of C. Reuschle is supported in part by the U.S. Department of Energy under grant DE-FG02-13ER41942.

<sup>5</sup>Leptons are dressed by adding the four-vectors of all photons within  $\Delta R < 0.1$  to the lepton 4-vector

Distributions include the  $p_T$  and rapidity of  $b$ -jets and of the  $Z$ -boson, and for each  $b$ -jet, the  $y_{boost}$  of the  $b$ -jet and  $Z$ . For events with  $Z p_T > 20$  GeV, the  $\Delta R$ ,  $\Delta\phi$ , and  $\Delta y$  between the  $Z$  and all  $b$ -jets are plotted. For events with at least two  $b$ -jets, the  $\Delta R$  and di- $b$ -jets mass for the two leading  $b$ -jets, along with the  $Z p_T$  and rapidity are plotted.

**CMS Z+BB** Cross-section and angular correlations in  $Z$  boson with  $b$ -hadrons events at  $\sqrt{s} = 7$  TeV [5] (Rivet routine CMS\_2013\_I1256943). A pair of opposite-sign charge dressed lepton with  $p_T > 20$  GeV and  $|\eta| < 2.4$  are required, with dilepton mass between 81 and 101 GeV. Exactly two weakly decaying  $b$ -hadrons with  $p_T > 15$  GeV and  $|\eta| < 2$  are then required.

Distributions include the  $Z p_T$ , the  $\Delta R$  and  $\Delta\phi$  between  $b$ -hadrons,  $\Delta R$  between the  $Z$  and closest  $b$ -hadron, and the asymmetry of the  $\Delta R$  between the  $Z$  and closest  $b$ -hadron, and the  $Z$  and the furthest  $b$ -hadron. The angular distributions are repeated with a requirement of  $Z p_T > 50$  GeV.

**ATLAS W+b** Measurement of the cross-section for  $W$  boson production in association with  $b$ -jets in pp collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector [6] (Rivet routine ATLAS\_2013\_I1219109). A dressed lepton with  $p_T > 25$  GeV and  $|\eta| < 2.5$  and a same-flavour neutrino with  $p_T > 25$  GeV are used to form a  $W$  candidate, which is required to have a transverse mass greater than 60 GeV. Anti- $k_t$  0.4 jets are reconstructed from all final state particles, and required to have  $p_T > 25$  GeV,  $|y| < 2.1$  and not overlap with the charged lepton used to make the  $W$  candidate ( $\Delta R(jet, l) > 0.5$ ). Events with more than two selected jets are vetoed, and the at least one of the selected jets is required to be labelled as  $b$ -jet, based on matching with  $\Delta R < 0.3$  to a weakly decaying  $b$ -hadron with  $p_T > 5$  GeV.

Distributions include the number of jets, and the  $b$ -jet  $p_T$  in events containing exactly one or two selected jets.

### 1.3 Event generators

#### SHERPA

In this section we present the setups that are used in this study for the SHERPA event generator [7]. In particular for Z+b(b) we consider three different classes of samples: 4F MC@NLO, 5F MEPS and a 5F MEPS@NLO one.

**4F MC@NLO:** This first set of results is obtained in the four-flavour scheme, and based on the MC@NLO technique [2], as implemented in SHERPA [8]. In a four-flavour scheme calculation,  $b$ -quarks can only be produced as final state massive particles. They are, therefore, completely decoupled from the evolution of the strong coupling,  $\alpha_s$  and that of the PDFs. In this scheme the associated production at tree-level starts from processes such as  $jj \rightarrow b\bar{b}Z$  where  $j$  can be either a light quark or a gluon. No specific cuts are applied on the  $b$ -quarks, their finite mass regulates collinear divergences that would appear in the massless case. In most cases, therefore, a  $b$ -jet actually originates from the parton shower evolution and hadronization of a  $b$ -quark produced by the matrix element.

**5F MEPS:** In a 5F scheme  $b$ -quarks are treated as massless partons. Collinear logs are resummed into a  $b$ -PDF and they can appear as initial state particles as well as final state ones. In order to account for 0 and 1  $b$ -jets bins as well as to cure the collinear singularity that would arise with a massless final state parton, we use multi-jet merging. In SHERPA, the well-established mechanism for combining into one inclusive sample towers of matrix elements with increasing jet multiplicity at tree-level is the CKKW [9]. For this sample we merge together LO samples of  $jj \rightarrow Z$ ,  $jj \rightarrow Z + j$ ,  $jj \rightarrow Z + jj$ ,  $jj \rightarrow Z + jjj$  where now  $j$  can be a light quark, a  $b$ -quark or a gluon, and all these samples are further matched to the SHERPA parton shower CSS [10]. Merging rests on a jet-criterion, applied to the matrix elements. As a result, jets are being produced by the fixed-order matrix elements and further evolved by the parton shower. As a consequence, the jet criterion separating the two regimes is typically chosen such that the jets produced by the shower are softer than the jets entering the analysis. This is realised here by a cut-off of  $\mu_{jet} = 10$  GeV.

**5F MEPS@NLO:** In this scheme we use the extension to next-to leading order matrix elements, in a technique dubbed MEPS@NLO [11]. In particular, we merge  $jj \rightarrow Z$ ,  $jj \rightarrow Z + j$ ,  $jj \rightarrow Z + jj$  calculated with NLO accuracy and we further merge this sample with  $jj \rightarrow Z + jjj$  at the LO. As in the previous case matching criterion has to be chosen, and this is realised by a cut-off of  $\mu_{\text{jet}} = 10 \text{ GeV}$ .

In SHERPA, tree-level cross sections are provided by two matrix element generators, AMEGIC++ [12] and COMIX [13], which also implement the automated infrared subtraction [14] through the Catani–Seymour scheme [15, 16]. For parton showering, the implementation of [10] is employed with the difference that for  $g \rightarrow b\bar{b}$  splittings the invariant mass of the  $b\bar{b}$  pair, instead of their transverse momentum, is being used as scale. NLO matrix elements are instead obtained from OPENLOOPS [17, 18].

The SHERPA W+b sample is the same one used for the study presented in Section ?? . It is generated with AMEGIC at LO in the 5F scheme, using NNPDF [19] library. The  $b$ -quark is massive, with the mass value set to 4.75 GeV, and the  $W$  boson is treated through the narrow width approximation. The order of the electroweak couplings is fixed to 2, therefore removing single-top contribution. Multi-parton interactions (MPI) are switched on/off to estimate this contribution.

## HERWIG7

In this section we present the setup for those results obtained with the HERWIG7 event generator [20, 21].

Based on extensions of the previously developed MATCHBOX module [22], HERWIG7 facilitates the automated setup of all ingredients necessary for a full NLO QCD calculation in the subtraction formalism: an implementation of the Catani–Seymour dipole subtraction method [15, 16], as well as interfaces to a list of external matrix–element providers – either at the level of squared matrix elements, based on extensions of the BLHA standard [23, 24, 25], or at the level of color–ordered subamplitudes, where the color bases are provided by an interface to the COLORFULL [26] and CVOLVER [27] libraries.

For this study the relevant tree-level matrix elements are taken from MADGRAPH5\_aMC@NLO [28, 29] (via a matrix–element interface at the level of color–ordered subamplitudes), whereas the relevant tree-level/one-loop interference terms are provided by OPENLOOPS [17, 18] (at the level of squared matrix elements).

Fully automated NLO matching algorithms are available, henceforth referred to as subtractive ( $NLO\oplus$ ) and multiplicative ( $NLO\otimes$ ) matching – based on the MC@NLO [2] and POWHEG [3] formalism respectively – for the systematic and consistent combination of NLO QCD calculations with both shower variants (the angular–ordered *QTilde* shower [30] and the *Dipole* shower [31]) in HERWIG7.

We consider four different classes of samples, for varying combinations of matching and shower algorithms (a selection of plots can be found in sections 1.4 and 1.5):

**4F, Zbb** For this set we consider the subtractive and multiplicative matching together with the *QTilde* shower. The tree-level process of the underlying hard sub-process in this case is  $pp \rightarrow e^+e^-b\bar{b}$ . For this sample the  $b$  quark is considered massive and  $p$  only consists of light quarks or a gluon, not a  $b$  quark.

**5F, Zbb** For this set we consider the subtractive and multiplicative matching together with the *QTilde* and *Dipole* shower. The tree-level process of the underlying hard sub-process in this case is  $pp \rightarrow e^+e^-b\bar{b}$ . For this sample the  $b$  quark is treated as massless, and  $p$  may also include a  $b$  quark. Generator-level cuts on the  $b$  quarks have thus been applied. Only in the shower evolution of the *QTilde* shower is the  $b$  quark assumed massive.

**5F, Zb** For this set we consider the subtractive and multiplicative matching together with the *QTilde* and *Dipole* shower. The tree-level process of the underlying hard sub-process in this case is  $pp \rightarrow e^+e^-j_b$ , where  $j_b \ni \{b, \bar{b}\}$ . For this sample the  $b$  quark is treated as massless, and  $p$  may also include a  $b$  quark. Generator-level cuts on the  $b$  quark have thus been applied. Only in the shower evolution of the *QTilde* shower is the  $b$  quark assumed massive. For single  $b$ -quark

production only one  $p$  must contribute a  $b$  quark at a time, at the level of the hard sub-process at hand.

**4F, Wbb** For this set we consider the subtractive and multiplicative matching together with the  $QTilde$  shower. The tree-level process of the underlying hard sub-process in this case is  $pp' \rightarrow Wb\bar{b} \rightarrow l\nu_l b\bar{b}$ , where  $l \in \{e^+, e^-, \mu^+, \mu^-\}$  and  $\nu_l$  the associated (anti-)neutrino. For this sample the  $b$  quark is considered massive and  $p, p'$  only consist of light quarks or a gluon, not a  $b$  quark;  $p'$  simply denotes the contribution of pair-wise different quark flavours in the initial state, as a result from the  $Wdu$  or  $Wsc$  vertex.

In all samples the uncertainty bands are purely from scale variations by simultaneously varying all scales in the hard sub-process and in the shower by factors of two up and down, i.e. factorization and renormalization scale in the hard sub-process, as well as scales related to  $\alpha_s$  and the PDFs in the shower, as well as the hard shower scale. The central scale choice is always a fixed scale (the  $Z$  mass in associated  $Z$  production; the  $W$  mass in associated  $W$  production).

The PDF sets being used are MMHT2014lo68cl and MMHT2014nlo68cl [32], i.e. the default PDF sets to which the showers are currently tuned. An internal study showed that using different PDF sets (a different  $n_f=5$  PDF set for the 5F runs or  $n_f=4$  PDF sets for the 4F runs) results in only minor differences, within the scale variation uncertainties.

In case the  $b$  quark is assumed massive, its mass is set to the default value in HERWIG7. All other relevant parameters, like  $W$  and  $Z$  mass and width, etc., are set to their respective default values in HERWIG7 as well.

For the 5F, Zbb sample we cut on the final state  $b$  quarks by including them into the jet definition and requiring at least two jets, with a min.  $p^\perp$  of 18 GeV and 15 GeV for the first and second jet respectively. The statistics for this sample are 100k unweighted events.

For the 5F, Zb sample we cut on the final state  $b$  quark (similarly to above) by requiring at least one jet, with a min.  $p^\perp$  of 18 GeV for the first jet. The statistics for this sample are 100k unweighted events.

For both 5F samples we apply generator level cuts on the invariant mass of the charged-lepton pair, with a min. invariant mass of 60 GeV and a max. invariant mass of 120 GeV.

For the 4F, Zbb sample we apply no cuts on the  $b$  quarks. However, we require the same generator level cuts as for the 5F, Zbb and Zb samples on the invariant mass of the charged-lepton pair again. In addition we cut slightly on the charged leptons, with a min.  $p^\perp$  of 5 GeV and a rapidity range between -4 and 4. The statistics for this sample are 100k unweighted events.

For the 4F, Wbb sample we also apply no cuts on the  $b$  quarks. However, we require a slight generator level cut on the transverse mass of the  $W$ , with a min. transverse mass of 20 GeV. We also cut slightly on the charged lepton, with a min.  $p^\perp$  of 5 GeV and a rapidity range between -4 and 4. The statistics for this sample are 100k unweighted events.

#### POWHEG BOX

The results obtained with the POWHEG BOX framework are based on the generators presented in ref. [33]. The tree-level amplitudes, which include Born, real, spin- and colour-correlated Born amplitudes, were automatically generated using an interface [34] to MADGRAPH4 [35, 36], whereas the one-loop amplitudes were generated with GOSAM [37, 38] via the Binoth-Les-Houches (BLHA) interface [23, 24], presented for the POWHEG BOX and GOSAM in [39]. The version of GOSAM is the 2.0: it uses QGRAF [40], FORM [41] and SPINNEY [42] for the generation of the Feynman diagrams. These diagrams are then computed at running time with NINJA [43, 44], which is a reduction program based on the Laurent expansion of the integrand [45], and using ONELOOP [46] for the evaluation of the scalar one-loop integrals. For unstable phase-space points, the reduction automatically switches to GOLEM [47], that allows to compute the same one-loop amplitude evaluating tensor integrals.

Further details can be found in ref. [33]. Here we briefly summarize the most important features.

1. We have used a mixed renormalization scheme [48], generally known as decoupling scheme, in which the  $n_{\text{lf}}$  light flavours are subtracted in the usual  $\overline{\text{MS}}$  scheme, while the heavy-flavour loop is subtracted at zero momentum. In this scheme, the heavy flavour decouples at low energies. To make contact with other results expressed in terms of the  $\overline{\text{MS}}$  strong coupling constant, running with 5 light flavours, and with pdfs with 5 flavours, we have switched our scheme using the procedure discussed in ref. [49].
2. We have generated  $Wb\bar{b}j$  events using the `MINLO` [50] prescription, that attaches a suitable Sudakov form factor to the  $Wb\bar{b}j$  cross section at NLO, and subtracts its expansion (not to have double counting of the Sudakov logarithms), in order to get a finite cross section down to small transverse momentum of the hardest jet. The scales of the primary process (i.e. the process obtained by the attempt to cluster a  $Wb\bar{b}j$  event with a procedure similar to CKKW [9]) have been chosen as follows:

- (a) if there has been a clusterization, then the scales are set to

$$\mu_{\text{R}} = \mu_{\text{F}} = \mu \equiv \frac{\sqrt{\hat{s}}}{4}, \quad \hat{s} = (p_{\text{W}} + p_{\text{b}} + p_{\bar{\text{b}}})^2, \quad (1)$$

where  $p_{\text{W}}$ ,  $p_{\text{b}}$  and  $p_{\bar{\text{b}}}$  are the momenta of the  $W$ ,  $b$  and  $\bar{b}$  in the primary process

- (b) If the event has not been clustered by the `MINLO` procedure, i.e. if the underlying Born  $Wb\bar{b}j$  process is not clustered by `MINLO`, we take as scale the partonic center-of-mass energy of the event.

The bands in the plots of Figures 13 and 14 of this section are the envelope of the distributions obtained by varying the renormalization and factorization scales by a factor of 2 around the reference scale  $\mu$  of eq. (1), i.e. by multiplying the factorization and the renormalization scale by the scale factors  $K_{\text{F}}$  and  $K_{\text{R}}$ , respectively, where

$$(K_{\text{R}}, K_{\text{F}}) = (0.5, 0.5), (0.5, 1), (1, 0.5), (1, 1), (2, 1), (1, 2), (2, 2). \quad (2)$$

These variations have been computed using the POWHEG BOX reweighting procedure, that recomputes the weight associated with an event in a fast way.

## 1.4 Z+b(b) production

### *Z+b(b) with SHERPA*

Figures 1 and 2 show a selection of the plots comparing Sherpa predictions to data. There is overall a good agreement, but for the normalization. The 5F LO order predictions are generally below the data, though compatible within the large scale uncertainty. For NLO predictions this uncertainty is smaller and some patterns can be observed. Both the 5F and the 4F NLO are in good agreement with distributions for events with two b-tagged jets. But when a single b-jet is tagged, the 5F and 4F results have an opposite behaviour: the 5F is 20% above the data (except for high  $Z$   $p_T$ ), while 4F is 20% below.

It is nevertheless remarkable that the ratio of 4F NLO predictions to data is flat for all the observables. This is particularly interesting, since it is more efficient to generate a sample of  $Z + b\bar{b}$  events with the 4F scheme than with the 5F. The reason why an overall normalization factor is needed could lie in the large logarithms, that in the 5F scheme are resummed in the  $b$  parton distribution function. However they might not affect the shape of the distributions. To check this hypothesis the 4F NLO predictions have been rescaled to the integrated cross-sections calculated with MCFM [51]. Depending on the observable and the applied selection, four different cross-sections are defined, as explained in [4]. The value is corrected for QED final-state radiation, hadronisation, underlying event and multi-parton interactions. The uncertainty is given by the envelope of the results obtained with several PDFs, taking for each the sum in quadrature of all theory uncertainties. A selection of the plots is shown in Figure 3. The results are very encouraging but further studies are needed to understand if this approach fails for other observables, e.g. those related to the presence of additional light-quark jets.

### *Z+b(b) with HERWIG7*

A selection of results obtained with HERWIG7 is shown in Figures 4, 5 and 6 for the 5F, Zbb setup, in Figures 7, 8 and 9 for the 5F, Zb setup, and in Figures 10, 11 and 12 for the 4F, Zbb setup. We refer to section 1.3 for the process setups of the 5F, Zb and Zbb samples, as well as for the 4F, Zbb sample.

Regarding the 5F, Zbb setup we note that we should expect somewhat large uncertainties in observables which are also sensitive to events with 1  $b$  jet, since at the level of the hard sub-process events with only 1  $b$  jet are only described by the real emission in this sample. Within those uncertainties, the prediction describes the data, with the exception of the  $\Delta\phi(Z, b)$  and  $\Delta R(Z, b)$  observables in Figure 5, where the prediction shows a slight tendency to be systematically above the data towards the lower values of  $\Delta\phi(Z, b)$  and  $\Delta R(Z, b)$ . Looking at the CMS data comparisons the *Dipole* shower together with the subtractive matching seems to undershoot the data in the  $\Delta R_{BB}$  and  $\Delta\phi_{BB}$  observables, which, however, seems not to be the case in the corresponding observables (for  $Z + \geq 2b$ -jets) in the ATLAS data comparisons (not shown here); however, here the combinations with the multiplicative matching seem to overshoot the data.

A brief internal study with the  $NLO \oplus Q\tilde{T}ilde$  combination showed that a 5F, Zbj setup (where at the level of the hard sub-process events with only 1  $b$  jet are already produced at the Born level) yields the expected reduction in the uncertainty bands.

The predictions from the 5F, Zb setup describes the data overall well. There seems to be a slight tendency, though, to systematically overshoot the data in the  $b$ -jet  $|y|$  and  $y_{\text{boost}}(Z, b)$  observables for  $Z + \geq 1 b$ -jet in Figure 8 (the same holds for the  $\Delta y(Z, B)$  observable, not shown here).

With the 4F, Zbb setup we see that the predictions are generally below the data. However, also here we notice a similar behaviour as already pointed out for the SHERPA results, in that the ratio of this 4F NLO prediction to data is flat – except for the  $\Delta R_{BB}$  and  $\Delta\phi_{BB}$  observables in Figure 12, where we notice that the predictions fall below the data at the lowest end of the distributions.

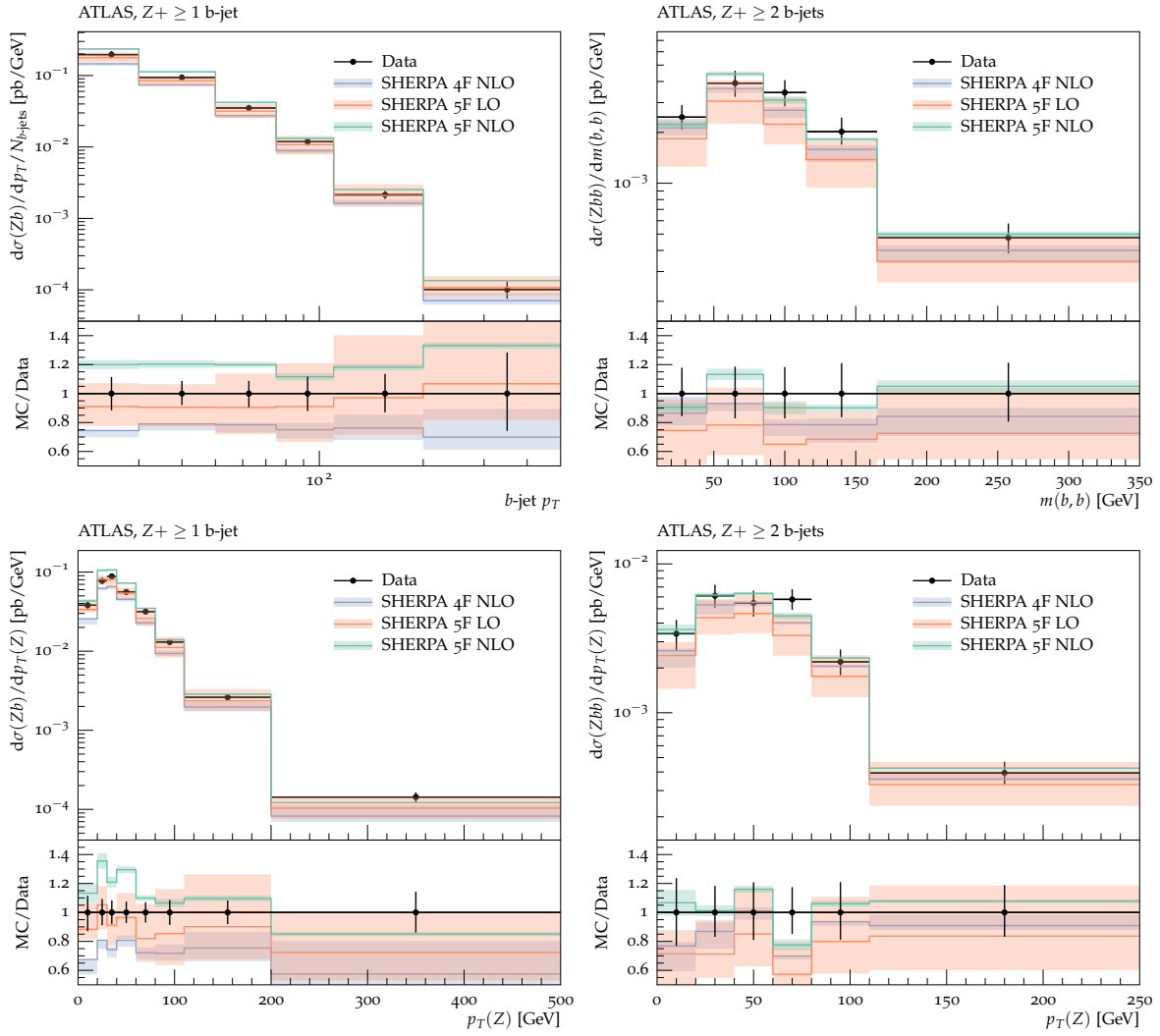


Fig. 1: A selection of the plots comparing Sherpa predictions to ATLAS results.

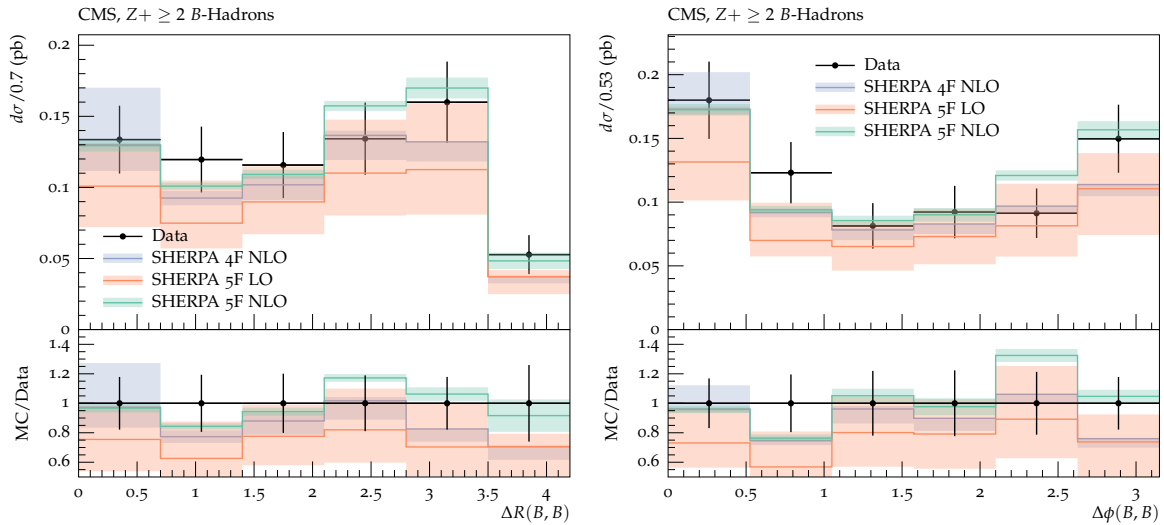


Fig. 2: A selection of the plots comparing Sherpa predictions to CMS results.

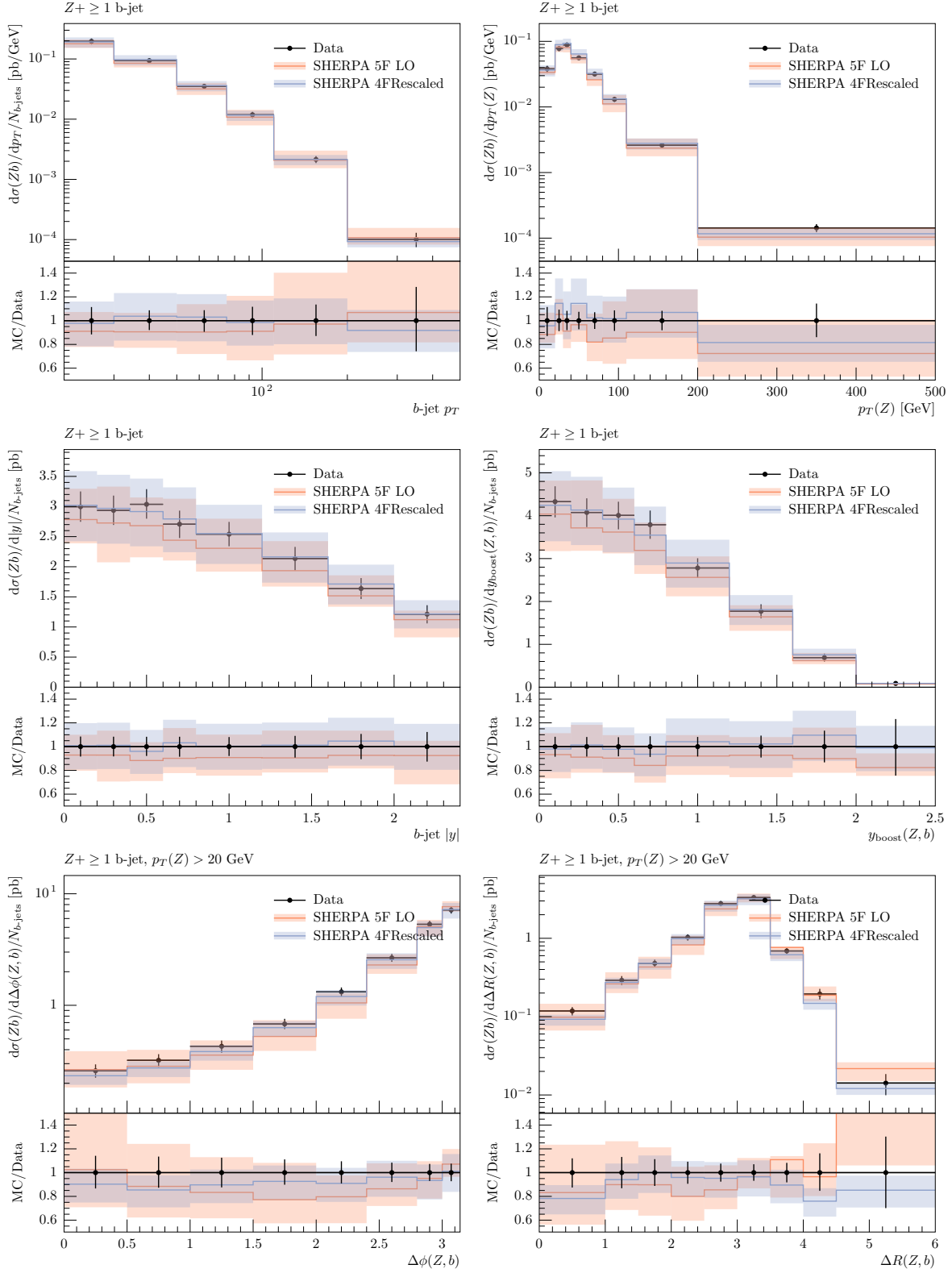


Fig. 3: A selection of the plots comparing rescaled Sherpa 4F NLO predictions to ATLAS results.



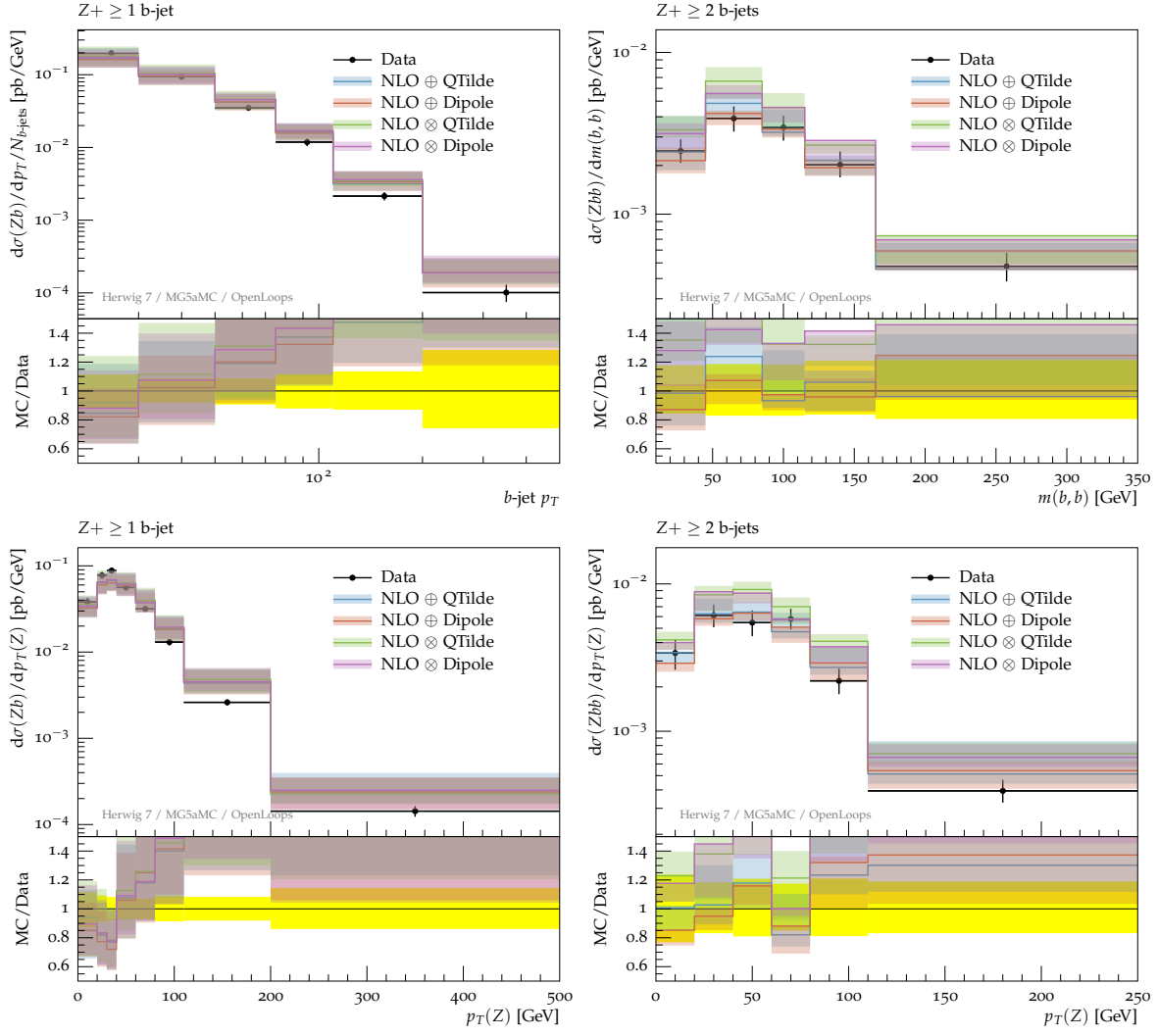


Fig. 4: A selection of the plots comparing HERWIG7 5F, Zbb predictions to ATLAS results.

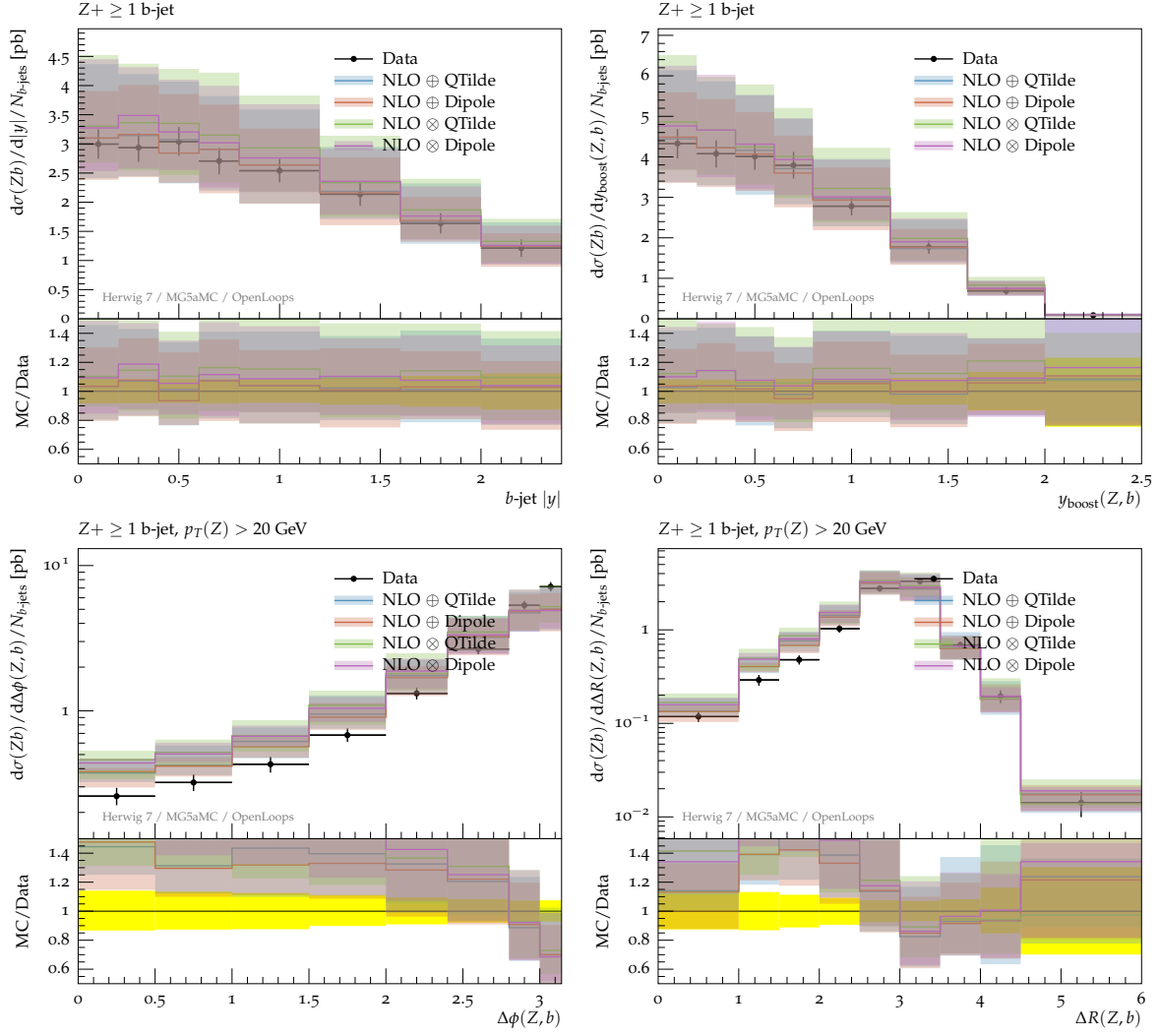


Fig. 5: A selection of the plots comparing HERWIG 7 5F, Zbb predictions to ATLAS results; together with the plots in figure 4 a comparison to the rescaled SHERPA 4F NLO predictions in Figure 3 can be made.

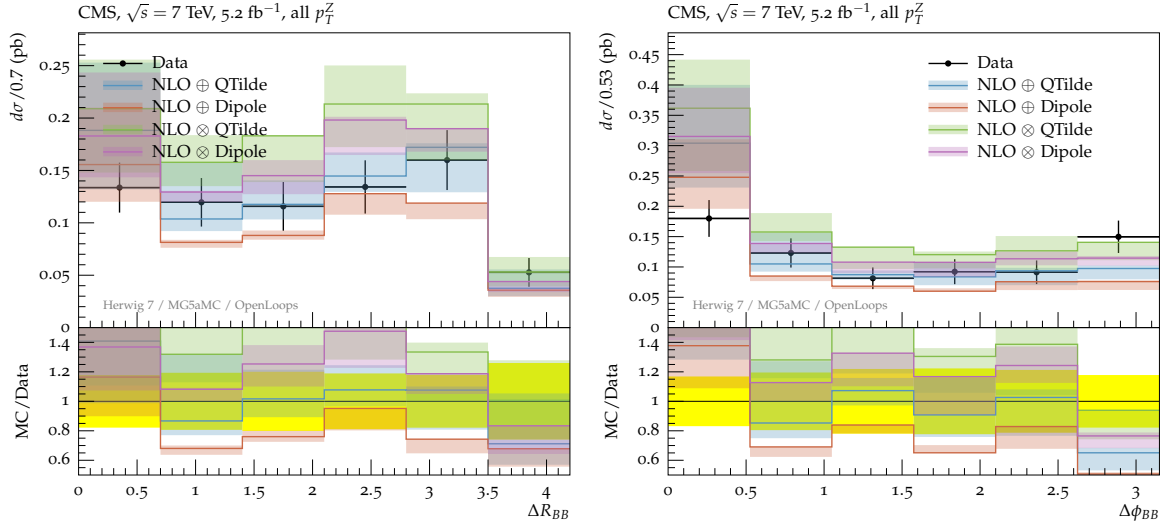


Fig. 6: A selection of the plots comparing HERWIG 7 5F, Zbb predictions to CMS results.

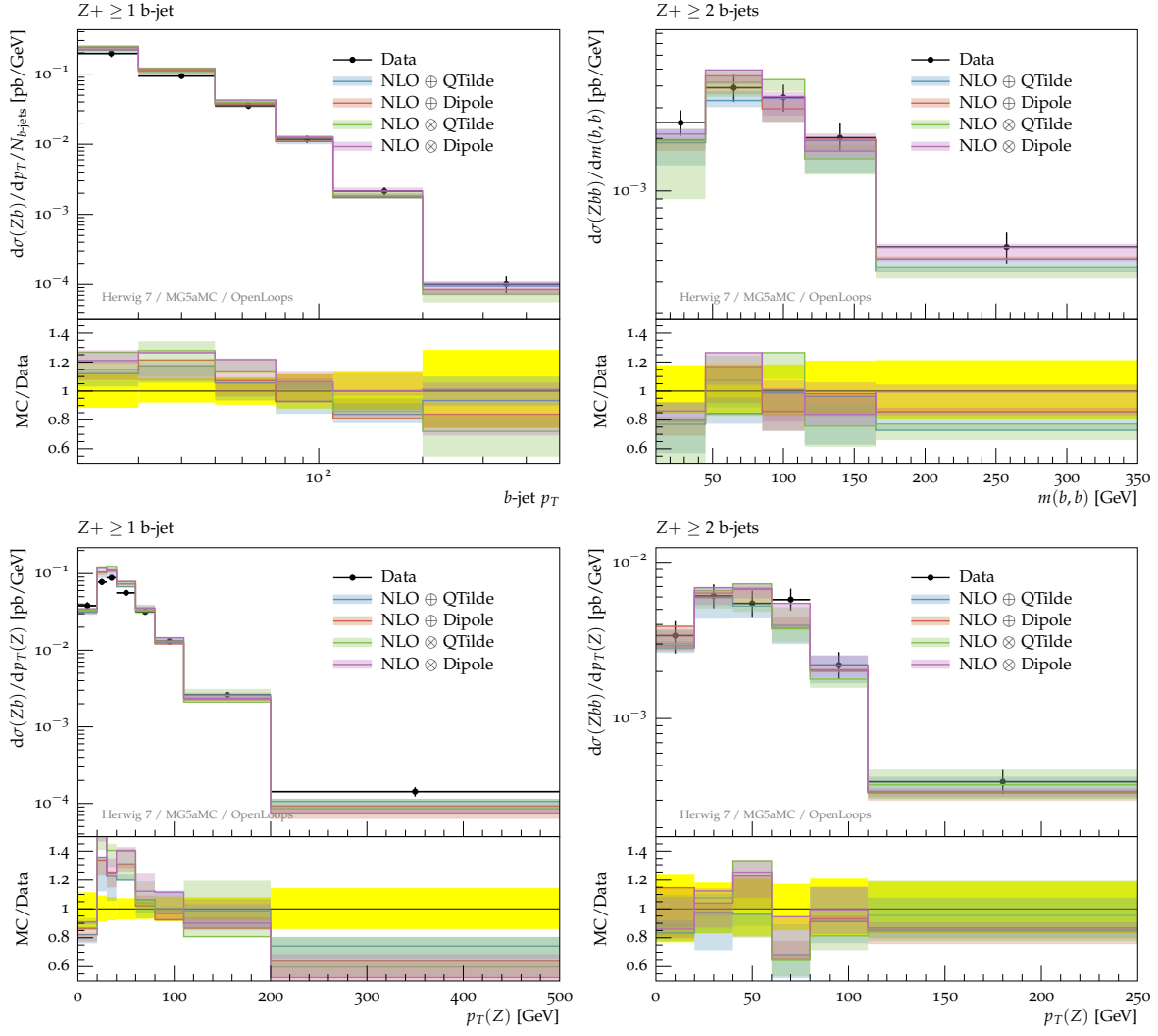


Fig. 7: A selection of the plots comparing HERWIG7 5F, Zb predictions to ATLAS results.

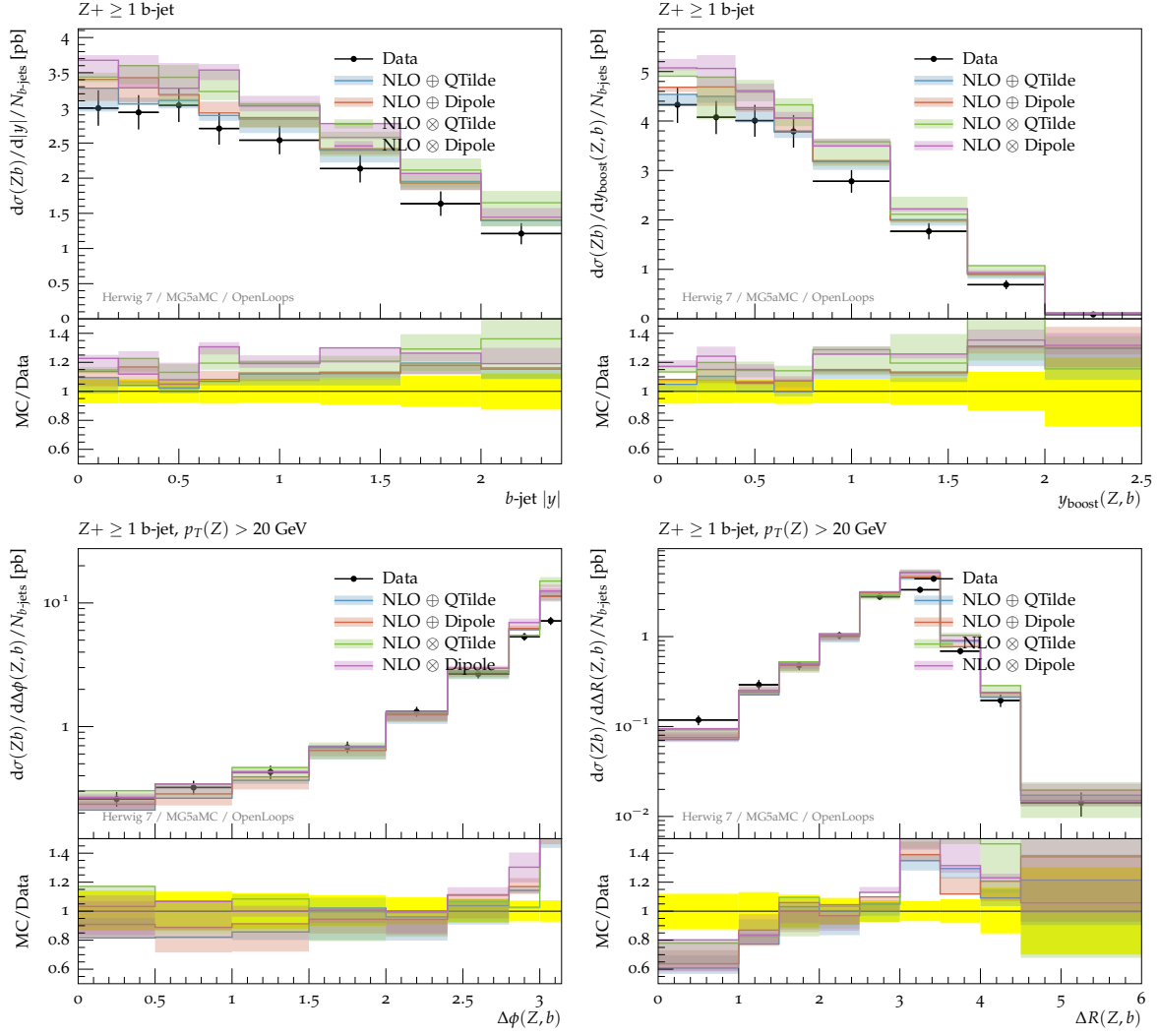


Fig. 8: A selection of the plots comparing HERWIG 7 5F, Zb predictions to ATLAS results; together with the plots in Figure 7 a comparison to the rescaled SHERPA 4F NLO predictions in Figure 3 can be made.

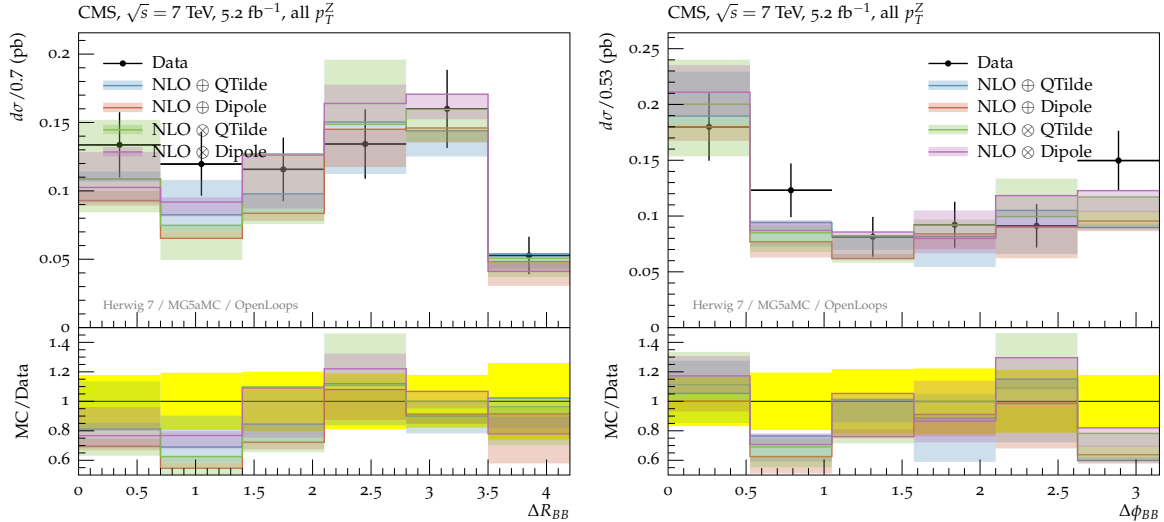


Fig. 9: A selection of the plots comparing HERWIG 7 5F, Zb predictions to CMS results.

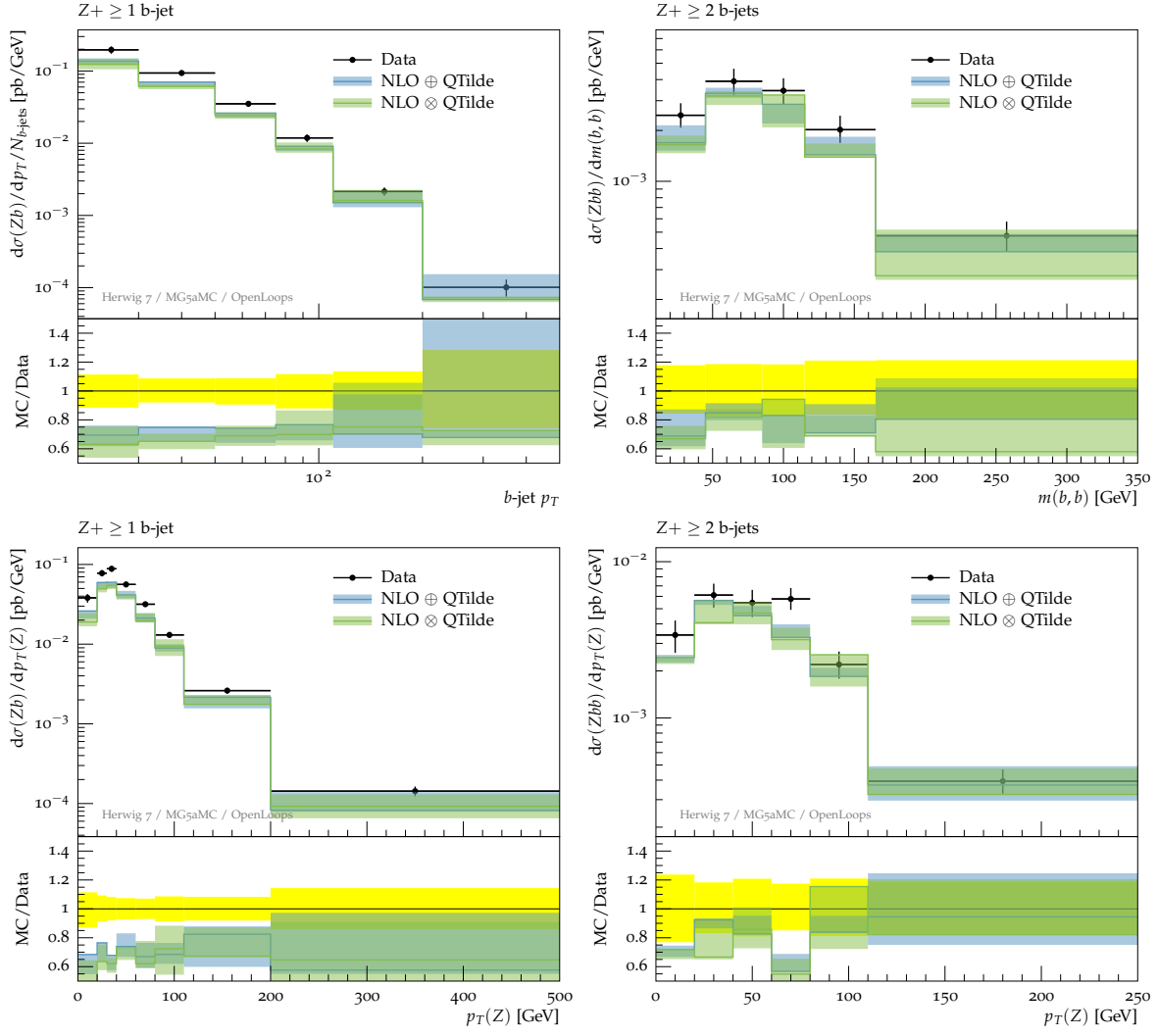


Fig. 10: A selection of the plots comparing HERWIG7 4F, Zbb predictions to ATLAS results.

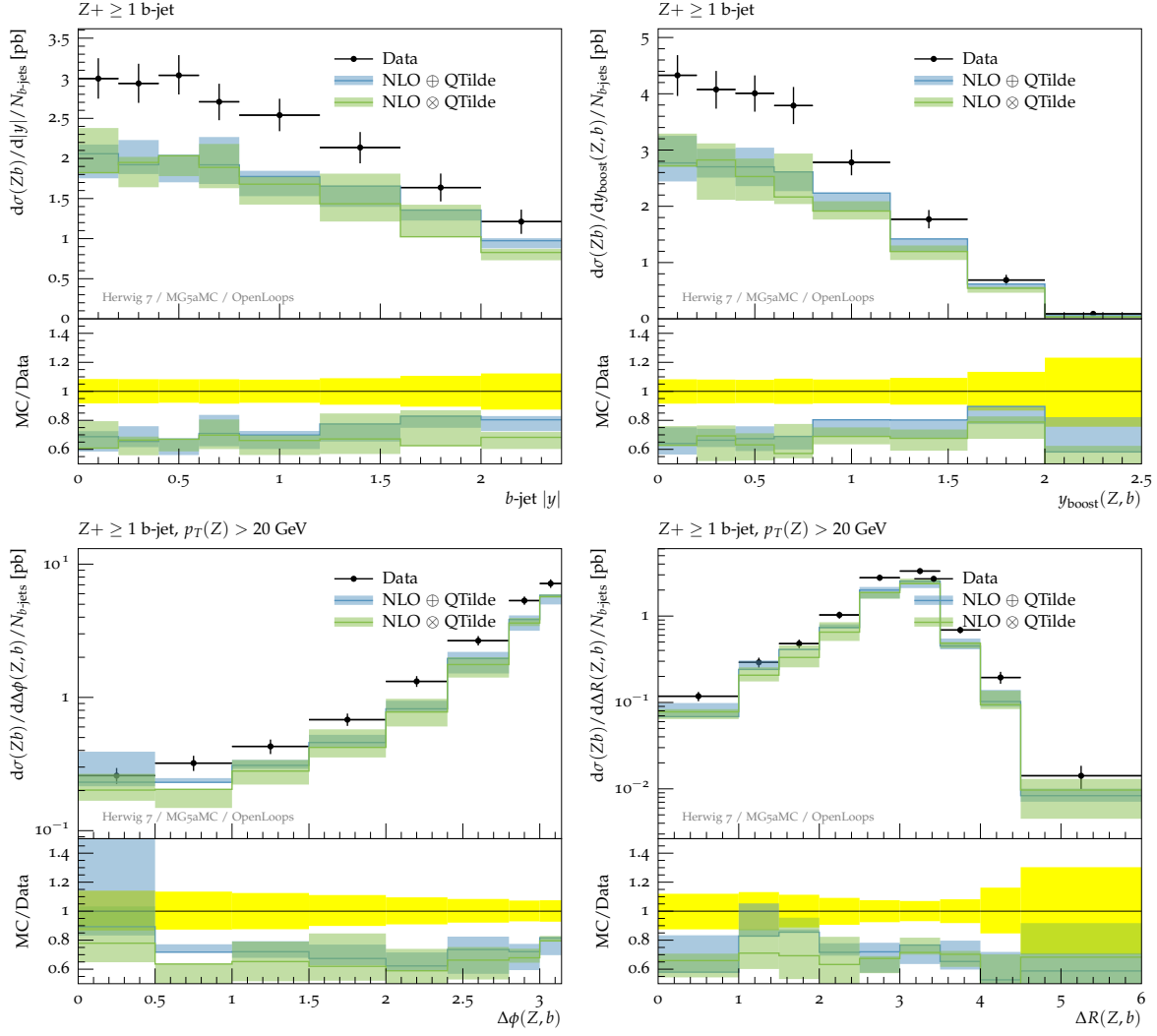


Fig. 11: A selection of the plots comparing HERWIG7 4F, Zbb predictions to ATLAS results; together with the plots in Figure 10 a comparison to the rescaled SHERPA 4F NLO predictions in Figure 3 can be made.

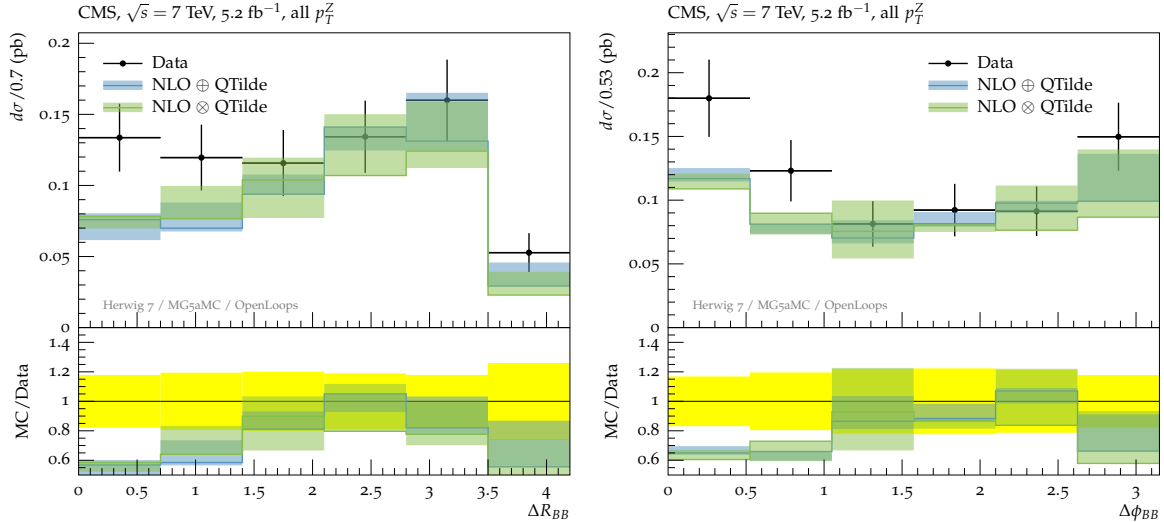


Fig. 12: A selection of the plots comparing HERWIG7 4F, Zbb predictions to CMS results.

## 1.5 W+b production

Figures 13 and 14 present a comparison between the data of the ATLAS W+b measurement and 4F NLO predictions obtained with the POWHEG BOX generator for  $Wb\bar{b}j$  presented in sec. 1.3 and the HERWIG7 generator for  $Wb\bar{b}$  discussed in sec. 1.3. Figures 15 and 16 present the same comparison but with 5F LO predictions obtained with the SHERPA generator as explained in sec. 1.3.

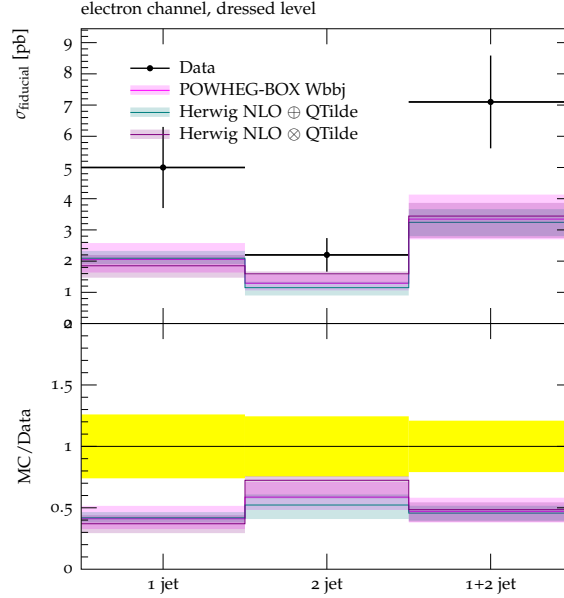


Fig. 13: ATLAS measured cross sections for  $Wb$  production with only a b-tagged jet (“1 jet”), one b-tagged and at least an additional jet (“2 jet”), or both (“1+2 jet”). The theoretical results are at the full shower+hadron level. No DPI corrections are included.

The first Figure present results for the total cross section in the case of only one b-tagged jet (“1 jet”), one b-tagged and at least an additional jet (“2 jet”), or both (“1+2 jet”). Despite the underlying generators consider two different underlying partonic processes,  $Wb\bar{b}j$  in the POWHEG BOX and  $Wb\bar{b}$  in the case of HERWIG7, the agreement between the two 4F NLO theoretical predictions is very good and the uncertainties due to scale variations are very similar. The general good agreement between the  $Wb\bar{b}$  and the  $Wb\bar{b}j$  predictions is not trivial a priori due to the different choices of scales in the two codes, and the fact that in the  $Wb\bar{b}j$  code the transverse momentum of the jet goes to zero and the divergent behavior is regulated by the `MinLO` prescription [33]

The agreement of these 4F NLO predictions with the experimental data is however not so good. As shown in [33] this can only partially be compensated by the missing double-parton interaction (DPI) corrections, which are estimated and applied as an additive correction factor in the original ATLAS publication [6]. This is confirmed by the results obtained with SHERPA with and without MPI. Compared to the estimate of DPI corrections given in the paper, the MPI contribution determined by SHERPA is smaller for the “1 jet” case, but not incompatible given the large uncertainty:  $0.2 \text{ pb}$  vs  $1.0^{+0.4}_{-0.3} \text{ pb}$ ; while a similar result is obtained for the “2 jet” case:  $0.4$  vs  $0.3 \pm 0.1$ .

The 5F LO prediction for the cross section agrees very well with data for the “2-jet” final state, while overshoot the data for “1 jet” by more than one standard deviation (taking into account the experimental uncertainty only). Previous studies [52] have shown that the contribution from  $qb \rightarrow Wbq'$  can be sizeable in the 5F scheme, while in the 4F scheme it appears at tree level only in the NLO corrections to  $q\bar{q}' \rightarrow Wb\bar{b}$ . Nevertheless it is quite surprising that the 5F predictions at LO are in better agreement

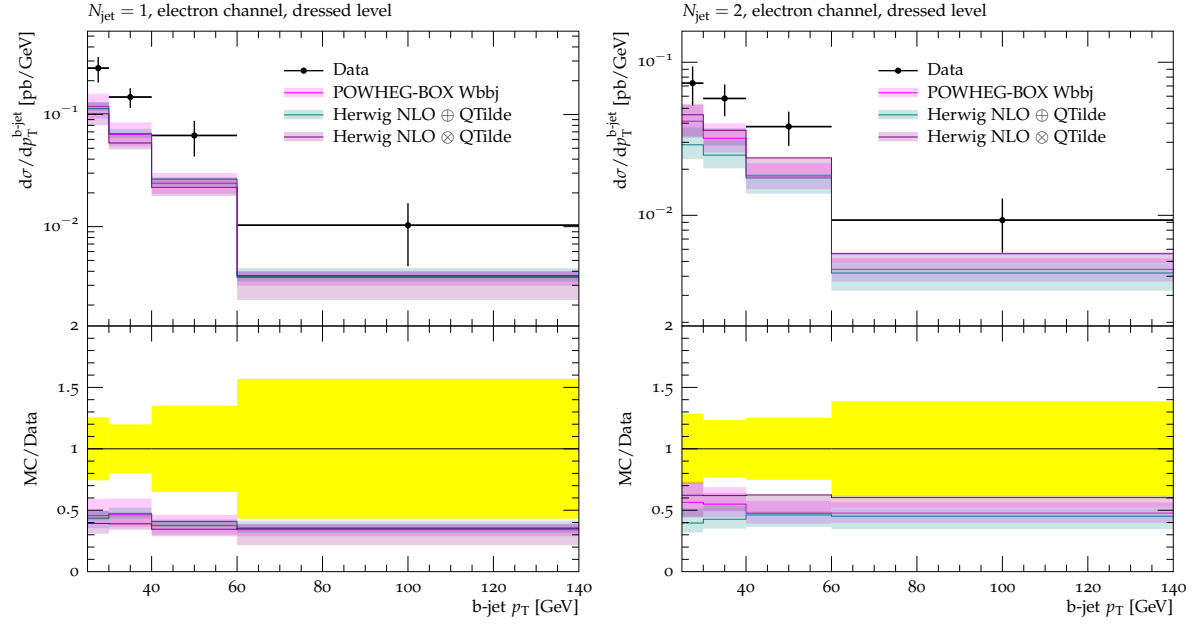


Fig. 14: ATLAS measured differential  $p_T$  distribution of the  $b$ -tagged jet in  $W + b$  events with a single jet (left) or with at least one additional jet (right). The theoretical results are at the shower+hadron level. No DPI corrections are included.

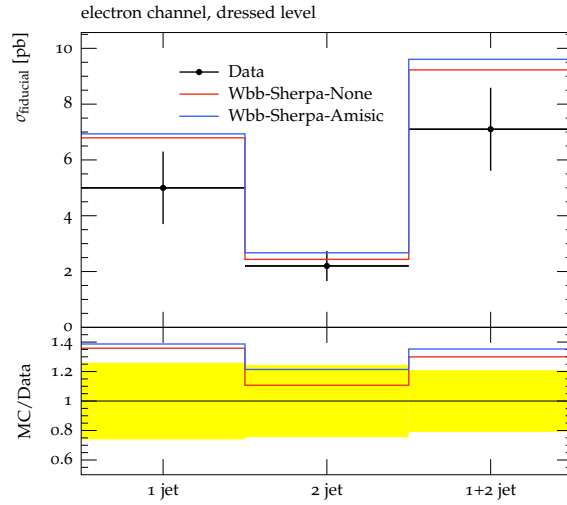


Fig. 15: ATLAS measured cross sections for  $Wb$  production with only a  $b$ -tagged jet (“1 jet”), one  $b$ -tagged and at least an additional jet (“2 jet”), or both (“1+2 jet”). Superimposed are shown the theoretical results obtained with SHERPA at LO in the 5F scheme, both with and without MPI.



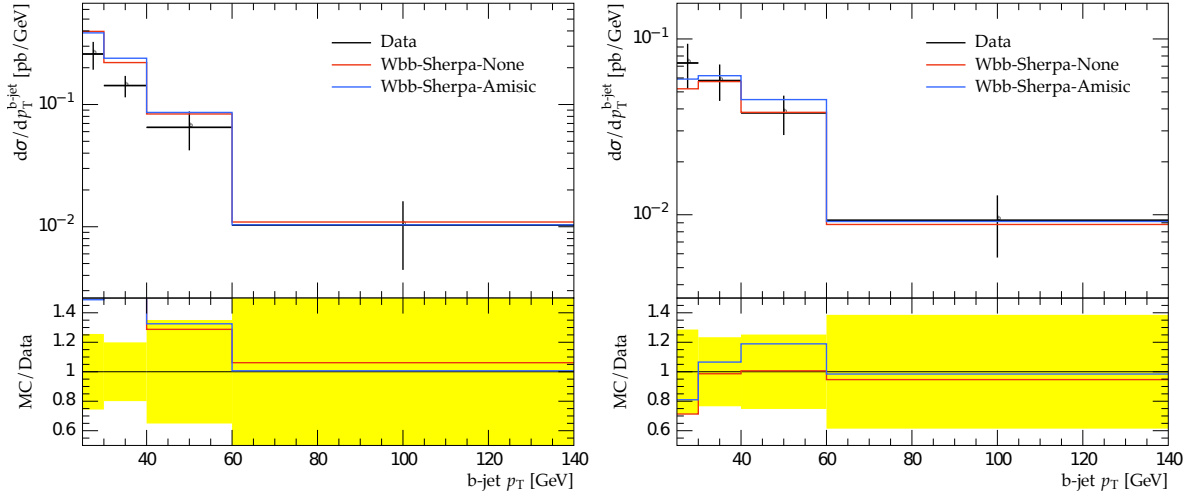


Fig. 16: ATLAS measured differential  $p_T$  distribution of the  $b$ -tagged jet in  $W + b$  events with a single jet (left) or with at least one additional jet (right). Superimposed are shown the theoretical results obtained with SHERPA at LO in the 5F scheme, both with and without MPI.

with the data with respect to 4F at NLO. As pointed out in the introduction, this might arise from the presence of large logarithms in the initial state, which can be properly resummed in the  $b$ -quark PDF.

Turning to less inclusive distributions, the ATLAS analysis presented above also contains measurements for the transverse momentum spectrum of the  $b$ -tagged jet in the “1 jet” and “2 jet” samples. A comparison of the data with the POWHEG BOX and HERWIG7 predictions for these observables is shown in the plots of Figure 14. Similarly to the inclusive case, also for these distribution the different theoretical predictions are in very good agreement among each other, but roughly a factor of two smaller than the measurements. It is likely that also in this case, only part of the difference can be explained with missing DPI corrections (not given in the paper for these differential distributions). It is interesting to note that, as for the  $Zb(b)$ , the results are off by a constant factor, i.e. the shape is correct and only the normalization seems to be wrong.

The  $b$ -jet  $p_T$  spectrum predicted by the 5F LO calculation is in very good agreement with data in the “2 jet” case, but it is steeply falling in the “1 jet” sample, substantially overshooting the data for low  $p_T$  values.

## 1.6 Conclusions

We presented a comparison of generators predictions using 4F and 5F scheme to most recent measurements of vector boson production in association with  $b$ -jets at the LHC. In the 4F scheme a good agreement is found among the different generators at NLO accuracy, and among different matrix-element to parton-shower matching algorithms. The agreement with data however is good only when two  $b$ -jets are tagged in the final state or, when one  $b$ -jet only is required, if a rescaling to the 5F integrated cross-section is applied. For Wbb, even taking into account the contribution from MPI, predictions seem to significantly undershoot the data. The  $Zb(b)$  production has been compared with predictions obtained in the 5F scheme with different setups, i.e. explicitly requiring one or two  $b$ -jets in the final state or a  $b$ -quark in the incoming proton when calculating the matrix element; or with no requirements on  $b$ -jets (treating them as light quarks) and combining final states with additional jets with a merging technique at LO and, where possible, also at NLO. Pro and cons of the different approaches are more difficult to pin down. In some case the scale uncertainty is quite large and not all distributions show a nice agreement with data,

especially if two b-tagged jets are present in the final state. For the Wb final state results have been also compared to a 5F calculation at LO: even if the precision is limited, because higher order corrections are missing, the cross section agrees better with data than for the 4F NLO calculation, as observed for the Zb(b) process and as previous NLO studies at parton level suggested.

Overall these results show that the associated production of vector bosons and b-jets is still an important benchmark for perturbative QCD at hadron colliders. More measurements and additional theoretical studies are definitely needed.

## References

- [1] F. Maltoni, G. Ridolfi, and M. Ubiali, *JHEP* **07** (2012) 022, [1203.6393]. [Erratum: *JHEP*04,095(2013)].
- [2] S. Frixione and B. R. Webber, *JHEP* **06** (2002) 029, [hep-ph/0204244].
- [3] P. Nason, *JHEP* **11** (2004) 040, [hep-ph/0409146].
- [4] G. Aad *et. al.*, **ATLAS** Collaboration *JHEP* **10** (2014) 141, [1407.3643].
- [5] S. Chatrchyan *et. al.*, **CMS** Collaboration *JHEP* **12** (2013) 039, [1310.1349].
- [6] G. Aad *et. al.*, **ATLAS** Collaboration *JHEP* **06** (2013) 084, [1302.2929].
- [7] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, *JHEP* **02** (2009) 007, [0811.4622].
- [8] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *JHEP* **09** (2012) 049, [1111.1220].
- [9] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *JHEP* **11** (2001) 063, [hep-ph/0109231].
- [10] S. Schumann and F. Krauss, *JHEP* **03** (2008) 038, [0709.1027].
- [11] S. Hoeche, F. Krauss, M. Schonherr, and F. Siegert, *JHEP* **1304** (2013) 027, [1207.5030].
- [12] F. Krauss, R. Kuhn, and G. Soff, *JHEP* **02** (2002) 044, [hep-ph/0109036].
- [13] T. Gleisberg and S. Höche, *JHEP* **12** (2008) 039, [0808.3674].
- [14] T. Gleisberg and F. Krauss, *Eur. Phys. J.* **C53** (2008) 501–523, [0709.2881].
- [15] S. Catani and M. H. Seymour, *Nucl. Phys.* **B485** (1997) 291–419, [hep-ph/9605323].
- [16] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, *Nucl. Phys.* **B627** (2002) 189–265, [hep-ph/0201036].
- [17] F. Cascioli, P. Maierhofer, and S. Pozzorini, *Phys. Rev. Lett.* **108** (2012) 111601, [1111.5206].
- [18] F. Cascioli, S. Hche, F. Krauss, P. Maierhfer, S. Pozzorini, and F. Siegert, *J. Phys. Conf. Ser.* **523** (2014) 012058.
- [19] R. D. Ball *et. al.*, **NNPDF** Collaboration *JHEP* **04** (2015) 040, [1410.8849].
- [20] J. Bellm *et. al.*, 1512.01178.
- [21] M. Bahr *et. al.*, *Eur. Phys. J.* **C58** (2008) 639–707, [0803.0883].
- [22] S. Platzer and S. Gieseke, *Eur. Phys. J.* **C72** (2012) 2187, [1109.6256].

- [23] T. Binoth *et. al.*, *Comput. Phys. Commun.* **181** (2010) 1612–1622, [1001.1307]. [,1(2010)].
- [24] S. Alioli *et. al.*, *Comput. Phys. Commun.* **185** (2014) 560–571, [1308.3462].
- [25] J. R. Andersen *et. al.*, 1405.1067.
- [26] M. Sjodahl, *Eur. Phys. J.* **C75** (2015), no. 5 236, [1412.3967].
- [27] S. Pltzer, *Eur. Phys. J.* **C74** (2014), no. 6 2907, [1312.2448].
- [28] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *JHEP* **07** (2014) 079, [1405.0301].
- [29] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *JHEP* **06** (2011) 128, [1106.0522].
- [30] S. Gieseke, P. Stephens, and B. Webber, *JHEP* **12** (2003) 045, [hep-ph/0310083].
- [31] S. Platzer and S. Gieseke, *JHEP* **01** (2011) 024, [0909.5593].
- [32] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Eur. Phys. J.* **C75** (2015), no. 5 204, [1412.3989].
- [33] G. Luisoni, C. Oleari, and F. Tramontano, *JHEP* **04** (2015) 161, [1502.01213].
- [34] J. M. Campbell, R. K. Ellis, R. Frederix, P. Nason, C. Oleari, *et. al.*, *JHEP* **1207** (2012) 092, [1202.5475].
- [35] T. Stelzer and W. F. Long, *Comput. Phys. Commun.* **81** (1994) 357–371, [hep-ph/9401258].
- [36] J. Alwall *et. al.*, *JHEP* **09** (2007) 028, [0706.2334].
- [37] G. Cullen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, *et. al.*, *Eur.Phys.J.* **C72** (2012) 1889, [1111.2034].
- [38] G. Cullen, H. van Deurzen, N. Greiner, G. Heinrich, G. Luisoni, *et. al.*, *Eur.Phys.J.* **C74** (2014), no. 8 3001, [1404.7096].
- [39] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, *JHEP* **1310** (2013) 083, [1306.2542].
- [40] P. Nogueira, *J.Comput.Phys.* **105** (1993) 279–289.
- [41] J. Kuipers, T. Ueda, J. Vermaseren, and J. Vollinga, *Comput.Phys.Comm.* **184** (2013) 1453–1467, [1203.6543].
- [42] G. Cullen, M. Koch-Janusz, and T. Reiter, *Comput.Phys.Comm.* **182** (2011) 2368–2387, [1008.0803].
- [43] H. van Deurzen, G. Luisoni, P. Mastrolia, E. Mirabella, G. Ossola, *et. al.*, *JHEP* **1403** (2014) 115, [1312.6678].
- [44] T. Peraro, *Comput.Phys.Comm.* **185** (2014) 2771–2797, [1403.1229].
- [45] P. Mastrolia, E. Mirabella, and T. Peraro, *JHEP* **1206** (2012) 095, [1203.0291].
- [46] A. van Hameren, *Comput.Phys.Comm.* **182** (2011) 2427–2438, [1007.4716].
- [47] G. Cullen, J. P. Guillet, G. Heinrich, T. Kleinschmidt, E. Pilon, *et. al.*, *Comput.Phys.Comm.* **182** (2011) 2276–2284, [1101.5595].

- [48] J. C. Collins, F. Wilczek, and A. Zee, *Phys. Rev.* **D18** (1978) 242.
- [49] M. Cacciari, M. Greco, and P. Nason, *JHEP* **05** (1998) 007, [[hep-ph/9803400](#)].
- [50] K. Hamilton, P. Nason, and G. Zanderighi, *JHEP* **10** (2012) 155, [[1206.3572](#)].
- [51] J. M. Campbell and R. K. Ellis, *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10–15, [[1007.3492](#)].
- [52] J. M. Campbell, F. Caola, F. Febres Cordero, L. Reina, and D. Wackeroth, *Phys. Rev.* **D86** (2012) 034021, [[1107.3714](#)].