

# Performance of various event generators in describing multijet final states at the LHC

**Stefan von Buddenbrock**

School of Physics, University of the Witwatersrand, Johannesburg 2050, South Africa.

E-mail: `stef.von.b@cern.ch`

**Abstract.** At the Large Hadron Collider (LHC), the most abundant processes which take place in proton-proton collisions are the generation of multijet events. These final states rely heavily on phenomenological models and perturbative corrections which are not fully understood, and yet for many physics searches at the LHC, multijet processes are an important background to deal with. It is therefore imperative that the modelling of multijet processes is better understood and improved. For this reason, a study has been done with several state-of-the-art Monte Carlo event generators, and their predictions are tested against ATLAS data using the RIVET framework. The results display a mix of agreement and disagreement between the predictions and data, depending on which variables are studied. Several points for improvement on the modelling of multijet processes are stated and discussed.

## Introduction

The biggest challenges to deal with in proton-proton ( $pp$ ) collisions arise from multijet processes, as far as Standard Model (SM) backgrounds are considered. Due to the nature of quantum chromodynamics (QCD), multijet production processes have the largest cross sections at the Large Hadron Collider (LHC). In addition to this, their partial reliance on non-perturbative QCD makes them difficult to deal with from a theoretical perspective. This is because simulation of fragmentation and hadronisation depend on a non-perturbative calculations, these often being done using phenomenological models. It is therefore of importance to study the performance of event generators in describing multijet final states, since certain combinations of matrix element (ME) calculations, parton shower (PS) and hadronisation models do not always provide an accurate description of the data.

In ATLAS, a number of generators are used to model multijet processes. These are discussed in detail in the next section. The predictions of these generators can be compared both to each other and to data corrected for detector effects (unfolded datasets). The simplest way of doing this is by using the RIVET analysis system [1], which has a large set of built in analyses and distributions of unfolded data from various experiments. This short paper will present a subset of distributions relating to multijet processes, and compare the current set of ATLAS Monte Carlo (MC) multijet samples to unfolded data. From these results, information can be extracted about how to improve the modelling of the generators for future generation of samples in ATLAS.

## Multijet event generators in ATLAS

A variety of MC event generators are used for studying multijet topologies in ATLAS. These involve different combinations of ME and PS programs. For a general review of event generators currently used in LHC physics, the reader is encouraged to look at Ref. [2]. Below is a list of the event generators considered in this study, as well as a few notes about their set up:

- PYTHIA 8 [3]: The prediction by PYTHIA 8 is sliced up by jet  $p_T$  using filters. The lowest  $p_T$  filtered samples use the PYTHIA 8 built in diffractive scattering processes (`SoftQCD`) to generate events. The rest of the slices use the elastic scattering processes (`HardQCD`). The chosen tune for the PYTHIA 8 samples is the A14 tune [4], which assumes the NNPDF23L0 parton density function (PDF).
- SHERPA [5]: The official SHERPA samples make use of a  $2 \rightarrow 3$  ME calculation,<sup>1</sup> matched with a CKKW scheme to a default SHERPA PS that use the CT10 tune. This sample has known issues with forward jets. The SHERPA prediction is also sliced in jet  $p_T$ .
- POWHEG+PYTHIA 8 [6]: The POWHEG ME is generated using the `Dijet` code that is provided with version 2 of the POWHEG-BOX. It is passed to the PYTHIA 8 PS, with the A14 tune. The sample is also sliced in jet  $p_T$ .
- HERWIG++ [7]: Like PYTHIA 8, the HERWIG++ sample makes use of the built-in `MEMinBias` process to simulate diffractive scattering for the lowest two slices in jet  $p_T$ , and the `MEQCD2to2` for the remaining slices. These samples make use of the UE-EE5 tune, and therefore the CTEQ6L1 PDF.
- MG5\_aMC@NLO+PYTHIA 8 [8]: The MG5\_aMC@NLO+PYTHIA 8 samples use a  $2 \rightarrow 4$  ME matched with a PYTHIA 8 PS using the CKKW-L scheme. The ME makes use of the NNPDF30NLO PDF, while the PS uses the A14 tune as described above. These samples are sliced at the ME level in parton  $H_T$ .

## Key comparisons to data

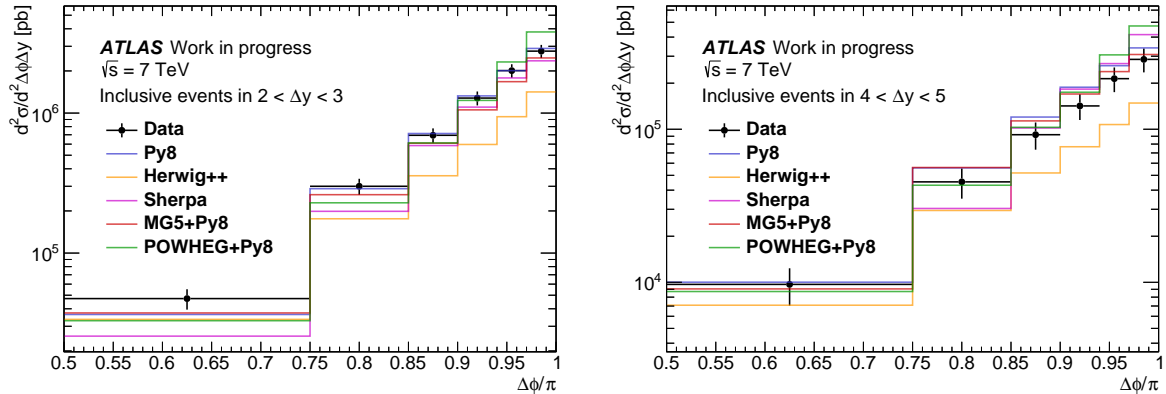
As mentioned above, RIVET is used to compare the predictions of these variables against each other and unfolded data. In this short paper, the predictions are compared in the context of three different aspects of jet physics, namely azimuthal decorrelations, jet fragmentation and jet shapes. Note that all jets considered in the following analyses are constructed using the anti- $k_T$  algorithm [9] with a radius parameter of  $R = 0.6$ .

### *Azimuthal decorrelations*

Purely elastic scattering of QCD partons most often results in a dijet event – that is, exactly two well separated jets in the final state. In such a case, the azimuthal separation between the two jets should be  $\pi$  radians. However, in theory, one expects to see more QCD interactions in the elastic scattering of quarks and gluons. This extra activity can produce more jet activity in multijet events. Depending on how much more activity is found in the event, the azimuthal angle between the two leading jets will deviate from  $\pi$ . This is known as an azimuthal decorrelation.

In order to study azimuthal decorrelations, one typically looks at the azimuthal angle between the two leading jets in multijet events. ATLAS performed a differential cross section measurement of azimuthal decorrelation variables with the Run 1 7 TeV dataset [10]. The corresponding RIVET routine for this analysis is `ATLAS_2014_I11307243`. In Figure 1, some comparison plots are shown from this analysis. The different generators mostly perform well against the data, although discrepancies arise in different regions of the distributions, particularly for HERWIG++.

<sup>1</sup> That is, up to three partons can be generated in the final state.



**Figure 1.** Differential cross section measurements as a function of the azimuthal angle between the two leading jets in multijet events [10]. The plots are made for inclusive multijet events in bins of rapidity separation, with  $2 < \Delta y < 3$  on the left and  $4 < \Delta y < 5$  on the right.

### Jet fragmentation

The behaviour of the fragmentation function used in different PS models is most commonly studied by looking at the densities of jet constituents in selected jets. The simplest RIVET routine to use when studying jet fragmentation is `ATLAS_2011_I929691`, which is a 7 TeV measurement of charged jet constituent densities as a function of three different variables [11]. Firstly, the variable  $z$  is scanned, which is the fraction of longitudinal momentum carried by a jet constituent:

$$z = \frac{\vec{p}_{\text{jet}} \cdot \vec{p}_{\text{ch}}}{|\vec{p}_{\text{jet}}|^2}. \quad (1)$$

Here,  $\vec{p}_{\text{ch}}$  denotes the 3-momentum of the charged jet constituent, and  $\vec{p}_{\text{jet}}$  is the 3-momentum of the jet. Secondly, the distance between the jet axis and the jet constituent in units of  $\phi$  and  $y$  is scanned over (denoted by  $r$ ). And thirdly, the jet constituent's momentum transverse to the jet axis,  $p_{\text{T}}^{\text{rel}}$ , is scanned over:

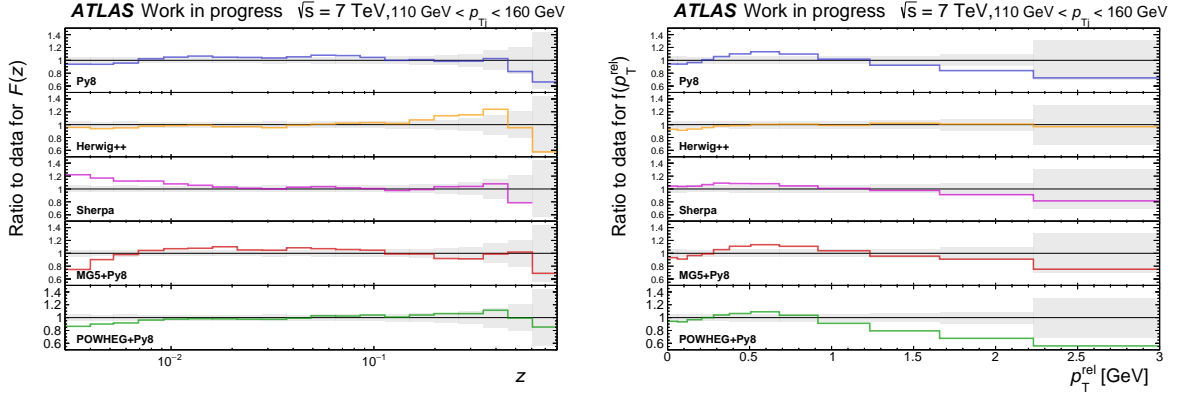
$$p_{\text{T}}^{\text{rel}} = \frac{|\vec{p}_{\text{ch}} \times \vec{p}_{\text{jet}}|}{|\vec{p}_{\text{jet}}|}. \quad (2)$$

In Figure 2, plots are shown for jet constituent densities as a function of two of these variables. In this case, HERWIG++ arguably performs the best compared with the data, and SHERPA tends to perform the poorest. This is most probably due to the old version of SHERPA used by ATLAS in the official samples. Recent studies on newer SHERPA samples in ATLAS have seen the problems with jet fragmentation fixed, although these results could not be shown in this short paper.

### Jet shapes

Jet algorithms can tell us about the geometry of the constituents of a jet. But to understand how energy is distributed in the average jet, it is more instructive to look at jet shapes. Similarly to jet fragmentation measurements, jet shapes are studied through looking at jet constituent densities. These are distributed as a function of the distance away from the axis of a jet,  $r$ . Typically, we look at the jet  $p_{\text{T}}$  weighted density in bins of annulus areas in the jet cone,

$$\rho(r) = \frac{1}{\Delta r N_{\text{jet}}} \sum_{\text{jets}} \frac{p_{\text{T}}(r - \Delta r/2, r + \Delta r/2)}{p_{\text{T}}(0, R)}, \quad (3)$$



**Figure 2.** Ratio plots of the different event generators compared to the data in measurements of jet constituent densities as a function of  $z$  (left) and  $p_T^{\text{rel}}$  (right) [11]. These plots are both shown in the same bin of jet  $p_T$ .

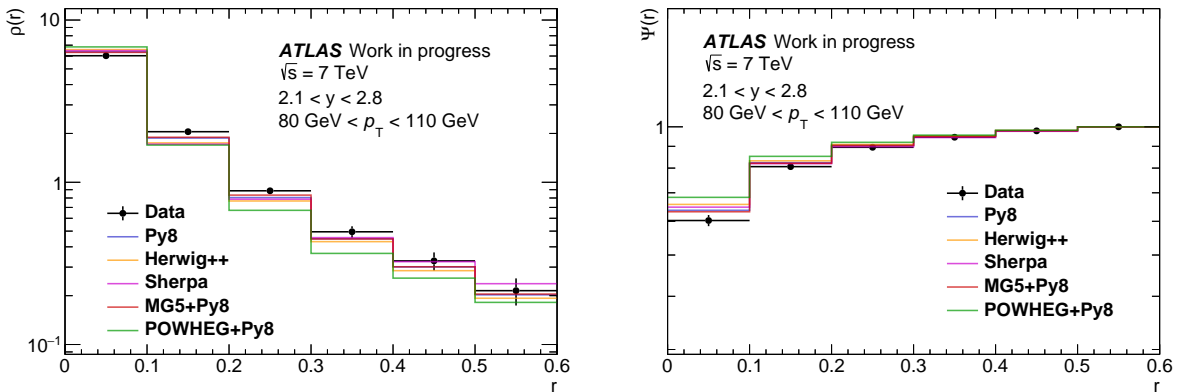
where  $p_T(r_1, r_2)$  is the sum of the jet constituent  $p_T$  between  $r_1$  and  $r_2$  away from the jet cone axis. In addition to this, we measure the integrated  $p_T$  weighted density,

$$\Psi(r) = \frac{1}{N_{\text{jet}}} \sum_{\text{jets}} \frac{p_T(0, r)}{p_T(0, R)}. \quad (4)$$

The RIVET routine `ATLAS_2011_S8924791` contains a large set of doubly differential jet shapes corresponding to an ATLAS 7 TeV measurement [12]. In Figure 3, some plots are shown in a single bin of the ATLAS analysis. Here, most of the generators considered agree relatively well with the data. It should be noted that POWHEG+PYTHIA seems to predict a different jet shape than what is seen in the data.

### Summary

Using the ATLAS multijet samples, comparisons have been made to unfolded data using RIVET. Three different measurements have been considered in this short paper. In each, it can be seen that the different generators tend to perform better in some regions of the phase space than



**Figure 3.** Jet shape measurements in terms of differential  $p_T$  density (left) and integrated  $p_T$  density (right) [12].

others, while there is no clear choice for one generator performing systematically better than any of the others.

However, the information from these comparisons is still useful for the ATLAS collaboration to improve the modelling of multijet processes by knowing where the current predictions fail. There are yet many more measurements that can be considered in this study, and in future these studies will be extended to a more comprehensive study.

## References

- [1] Buckley A, Butterworth J, Lonblad L, Grellscheid D, Hoeth H, Monk J, Schulz H and Siegert F 2013 *Comput. Phys. Commun.* **184** 2803–2819 (arXiv:1003.0694)
- [2] Buckley A *et al.* 2011 *Phys. Rept.* **504** 145–233 (arXiv:1101.2599)
- [3] Sjöstrand T, Ask S, Christiansen J R, Corke R, Desai N, Ilten P, Mrenna S, Prestel S, Rasmussen C O and Skands P Z 2015 *Comput. Phys. Commun.* **191** 159–177 (arXiv:1410.3012)
- [4] ATLAS Collaboration 2014 ATLAS Pythia 8 tunes to 7 TeV data ATL-PHYS-PUB-2014-021 URL <https://cds.cern.ch/record/1966419>
- [5] Hoeche S, Krauss F, Schonherr M and Siegert F 2013 *JHEP* **04** 027 (arXiv:1207.5030)
- [6] Alioli S, Nason P, Oleari C and Re E 2010 *JHEP* **06** 043 (arXiv:1002.2581)
- [7] Bahr M *et al.* 2008 *Eur. Phys. J.* **C58** 639–707 (arXiv:0803.0883)
- [8] Alwall J, Frederix R, Frixione S, Hirschi V, Maltoni F, Mattelaer O, Shao H S, Stelzer T, Torrielli P and Zaro M 2014 *JHEP* **07** 079 (arXiv:1405.0301)
- [9] Cacciari M, Salam G P and Soyez G 2008 *JHEP* **04** 063 (arXiv:0802.1189)
- [10] ATLAS Collaboration 2014 *Eur. Phys. J. C* **74** 3117 (arXiv:1407.5756)
- [11] ATLAS Collaboration 2011 *Eur. Phys. J. C* **71** 1795 (arXiv:1109.5816)
- [12] ATLAS Collaboration 2011 *Phys. Rev. D* **83** 052003 (arXiv:1101.0070)