- you mention (page 5, second line of page) the ”mixed event distribution”. Could you explain what is the mixed event distribution for readers not familiar ? Is this distribution obtained in data or in full simulation for the determination of the pair acceptance ?

R: Added the following sentence to explain the mixed event: “The mixed event distribution is constructed by correlating trigger and associated particles from different events in the data sample.”

•For the study with Pythia 8 described on page 5, is it a study done at generator level or after full event simulation (for example to obtain Fig.2) ? Is jT computed exactly as for the analysis with data ? (You explain how jT is computed in data only later, on page 6, second paragraph so was not sure if it meant that the computation was different than for the Pythia 8 study)

R: The study is generator level. Added now “generator level” also to the text to be more explicit. Also moved the sentence describing how leading and associated particles in xlong bins are paired to the beginning of page 5 to be clearer that the determination of jT is the same in simulations and data.

• Could you explain more what is ”this scaling” (bottom of second paragraph of page 5). I guessed you meant that it is the factor ”0.63” mentioned on Fig. 2 that is needed to have both distributions coincide at low jT but it may be better to write it explicitly.

R: You are correct. Changed the text now in such a way that the factor 0.63 in Figure 2 is explicitly mentioned.

• Equations 5 and 6: they are obtained with two-jet events. Can multijet or multi-parton interactions modify significantly their shapes ?

R: They are obtained from all events with leading particle requirement. In ALICE acceptance(|\eta|<0.8), we do see most of the time single-jet events. The ratio of single-jet to two-jet events is about few percent and multi-jet events are negligible. Since we are looking at the leading particle in an event, the relative contributions should be negligible. It is similar for multi-parton interactions. If there were contributions from these, it should contribute to the background template.

• Equation 6: perhaps it is good to define Γ as the gamma function also.
R: Defined Γ as the gamma function

• Figures 3: I think it would be clearer to draw the pull value in stead of distribution/fit or signal/fit for the bottom panels of the jT distributions.
R: We think that in this particular case having the ratio plotted instead of the pull value is better. They both give an idea about the goodness of the fit, but in the ratio you can also easily see the small bump around jT = 0.4. It is important to show it, since it represents the extra correlation component originating from the decay of short lived particles in addition to the jet correlation. Handling this component is one of the main systematic uncertainties of the measurement.

• Figures 3: The wide component seems to decrease much faster as a function of jT than for the Pythia 8 study. Is there a reason why ?
R: This has to do with the selected mass (Mjj) for the resonance. When doing the Pythia 8 study with the artificial decaying resonance, we tried different values for the resonance mass and it turn out that the mass has a strong dependence on how quickly the wide component dies out. All slopes were well described by the inverse gamma function, but just the decrease speed changed. You can see the effect of changing the Mjj from the attached Figure 1. The smaller the value for

Mjj, the faster the decrease of the wide component. Thus the difference between the slopes in this simulation and data is most likely due to the fact that the data corresponds to somewhat smaller Mjj as is shown in the paper.

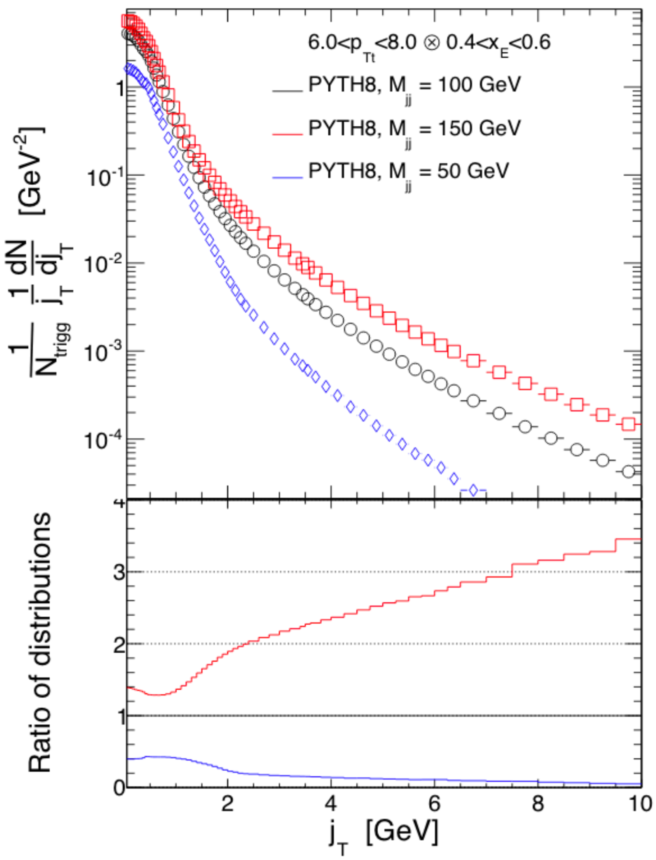


Figure 1: Pythia 8 study with different mass Mjj given to the decaying resonance.

• Page 7, end of second paragraph: you write that you obtain the uncertainty on the fit comparing the fit to the total distribution to the fit to the background subtracted distribution. But these two distributions are not independent since the background subtracted distribution is obtained after the total fit (which must be done to determine the background contribution) so I am not sure what is really checked by this method.

R: This is a bit technical and it tells us about the stability of the signal and background fitting. We have two different fitting methods in the analysis code, the first approach is to do a single fit (narrow component + wide component + background) to the whole distribution. This will give us the background and signal at the same time. Note that for the background the shape is given by the eta-gap method and here we only fit the scale. The second approach is to normalize the background to the tail of the jT distribution, subtract it and then do a two-component fit (narrow component + wide component) to the background subtracted distribution to fit the signal. These two methods should give the same answer and there is no strong reason to prefer either of the approaches. Thus we assign the difference as a systematic uncertainty coming from the signal fitting method.

• Table 1: I suppose this table gives the uncertainties for the data, not for the Herwig or Pythia simulations. Could you add this information to the caption ?

R: Added to the caption: “for data” so that this is completely clear.

•Page 8, First paragraph of Section 5: Could you move the definition of RMS earlier in the article ? (For example you mention it already on page 7 before defining it

R: Moved it now to the first mention of RMS.