Dear Editor,
We would like to thank the referees for reviewing this paper and furnishing this report.
We have carefully considered all comments, and we have applied several changes to the original version of the paper to address the issues raised. Detailed responses to all the comments can be found below.
We are at your disposal for any further clarifications and/or additional information.

Sincerely,
H. Yang, A. Bermudez Martinez, L.I. Estevez Banos, F. Hautmann, H. Jung, M. Mendizabal, K. Moral Figueroa, S. Prestel, S. Taheri Monfared, A.M. van Kampen, Q. Wang, K. Wichmann

Reviewer(s)' Comments to Author (if there no comments below, please check the attachments):
Referee: 1

Comments to the Author
In this letter the authors consider the distribution of Z+jet LHC events with respect to the azimuthal Zj correlation DPhi, and compare the results to an analogous distribution in dijet events. They propose the experimental measurement of the DPhi -> pi region in order to observe factorisation-breaking effects.

I find the content of the article as it is now of quite limited scope, although the study may be interesting: the authors propose results on a single distribution of a very well-studied process at NLO+PS; the novel part, which should be the proposal to use this distribution for the study of factorisation-breaking effects, is also described too briefly in my opinion.
ANSWER: Although the measurement of the azimuthal correlation of $\Delta\phi_{Zj}$ has been measured and studied in NLO+PS predictions, the back-to-back region of $\Delta\phi_{Zj}$ has only been measured in dijet event, but was never investigated in Zj events. In addition the focus here is on the very high p_T region, where a small $\Delta\phi_{Zj}$ deviation from 180 degrees is caused by one or more jets with measurable transverse momenta.

To recommend the letter for publication in EPJC, I think the text needs revision, addressing the following points.

1) I find it inconvenient when a method is described only by means of references to previous publications (see in particular section 2). In order to help the reader, I would suggest to spend some sentences trying to elucidate more clearly the salient features of the method, and make this section self-contained.
ANSWER: We have now extended the discussion and moved figures of the transverse momentum distribution for quarks and gluons at two different scales.

We change the order of question 2 and 3 for easier argumentation, and first answer question 3
3) Still on the matching, in the CASCADE3 reference [34] one can read that final-state radiation (FSR) is performed using the relevant PYTHIA6 routines. The subtraction terms for the MC@NLO matching should then be the HERWIG6

ones for initial-state radiation (ISR) (assuming CASCADE3 = HERWIG6 in that case), and the PYTHIA6 ones for FSR. Have the authors implemented such a mixed set of subtraction terms, or have they activated special options to this aim in MadGraph5_aMC@NLO?
I urge the authors to discuss this point explicitly in the letter: from the current text, it seems the HERWIG6 subtraction terms are applied to FSR as well, which would spoil NLO accuracy in presence of FSR.
ANSWER: We have investigated in detail the contribution of final state radiation in MCatNLO+CAS3 and compared it with MCatNLO+H6. We found that using the PYTHIA6 final state shower with the angular ordering veto condition, as used in CASCADE, agrees very well within uncertainties coming from shower parameter variations with the ones obtained from MCatNLO+H6.
We have included an appendix to the paper draft, showing this comparison.

2) Regarding the matching to NLO, for initial-state radiation the MC@NLO method is applied by means of subtraction terms that are conceived for the HERWIG6 shower. Although CASCADE3 and HERWIG6 share the same kind of angular-ordered branching algorithm, in order to be allowed to use the HERWIG6 subtraction terms without spoiling NLO accuracy, one should make sure that the actual first emission in HERWIG6 coincides with that in CASCADE3, including for instance so called dead-zones. Could the authors show that the first emission of CASCADE is identical as that of HERWIG6?
ANSWER:
After having shown, that the final state radiation agrees between MCatNLO+CAS3 and MCatNLO+H6, we have compared the contribution of initial and final state parton shower. We observe a significant dependence on the parton shower parameters from herwig6. One has to note, that even the transverse momentum and rapidity of the first emission is affected by subsequent emissions.
Given the differences in the parameter settings of the PB TMD shower and the ones which can be set in H6, we conclude, that the H6 subtraction terms can be consistently used with CASCADE3.
The comparison plots are also shown in the appendix of the new draft.

The dead zones come essentially from a region of $z<1$ for heavy quarks due to mass effect. In the PB-TMDs, all masses are treated massless, and $z \to 1$, therefore no dead zones are treated. This is different in the final state shower, where the dead zone effect is treated. As has been shown in the answer to question 2, the predictions from CASCADE3 and H6 using final state showers agree very well in the phase space region investigated here.

4) The dominance of the qq channel over the qg one at pt(leading) > 1 TeV is arguable. From figure 4 this happens almost at 2 TeV for dijet, and beyond the displayed range for Z+j, so I would suggest to rephrase.
ANSWER: We have reformulated this sentence to: “At high $\ptmax > 1000$ \GeV\ the qq channel becomes important for both \Zjet\ and multijet final states, ... “

5) Although the analysis of ISR justifies the stronger correlation at high pt(leading) with respect to low pt(leading), in the letter there seems to be no conclusive argument explaining the similarity between the dijet and Z+j distributions at large DPhi. Could the authors add some statements on this, perhaps before discussing the matching scale?
ANSWER: At large pt the gg channel becomes less important compared to the qq channel for dijets.

The decorrelation is essentially driven by initial state radiation, therefore it becomes similar for jj and Zj, if the initial states are similar, and this happens at large pt.
We have added a sentence:" We conclude, that the main effect of the Δφ decorrelation comes from initial state radiation, and the shape of the Δφ decorrelation in the back-to-back region becomes similar between Z+jet and dijet processes at high pleading where similar initial partonic states are important. "

6) I don't understand the discussion on the interpretation of the matching scale. In pt-ordered showers the matching scale limits the hardness (pt) of the first emission. Being pt-ordered, the subsequent emissions are limited by the first, and in turn by the matching scale. Could the authors clarify why this would be different for angular and pt showers?
I'm also slightly surprised that the matching scale variation is quite smaller for CASCADE3 than for PYTHIA8, as the two tools have NLO+PS accuracy. Given the amount of approximations underlying the method I don't think this is necessarily an indication of enhanced robustness, rather of uncertainty underestimate.
ANSWER: We thank the referee for this comment, which we used as an opportunity to clarify the manuscript (regarding the matching scale interpretation) and amend the manuscript (regarding statements about robustness).

As further clarification: In an angular-ordered shower, the highest-pT emission does not necessarily arise in the first branching. Thus, limiting the influence of the matching on the first branching alone is not sufficient. The main point that we were making is that in an angular-ordered shower, the matching scale thus acts as a "veto scale" for not only first, but also for all subsequent branchings. This is at odds with a transverse-momentum-ordered shower. In the latter case (and as the referee correctly describes), the matching scale acts as a veto scale for the first branching alone, since all subsequent branchings would be limited by the first one. We found it relevant to highlight this qualitative difference between matching to different showers in the manuscript, since it is (non-trivially) related to the spread of predictions obtained by varying the matching scale.

We have updated the text in the paper draft accordingly.

7) It seems the proposed strategy to measure factorisation-breaking effects only uses the high pt(leading) region. How would the measurement at low pt(leading) be used then? Could the authors give some more information on this?
Moreover, for this measurement to be effective, the equality between DPhi distributions at high pt(leading) in Z+jet and dijet should be stable against perturbative corrections, which would be established using more accurate resummation tools and higher fixed orders. Could the authors explain if/why this feature is expected to hold at higher orders as well?
ANSWER: We propose to use the ratio of measurement over prediction of \Delta \phi at low and at high pt(leading). A difference between prediction and measurement can be a hint for a factorization breaking effect. Since factorization breaking happens because of interactions between the colored initial and final state partons, we expect a difference in size of the effect in dijets and Zjet (since a different number of colored final state partons are involved). However, also the initial state (gluons or

quarks) can influence the decorrelation. In order to minimize the effect coming from different initial state configuration, we focus on high pt(leading), since there the qq channel becomes important in both dijets and Z+jets.
We have added a sentence clarifying this: ”The number of colored partons involved in Z+jet and multijet events is different, and deviations from factorization will depend on the structure of the colored initial and final state. In order to minimize the effect of different initial state configurations, a measurement at high pleading, could hint more clearly possible factorization - breaking effects. “

In addition we have also included a reference in the introduction to a new paper on Zjet factorization breaking: arXiv 2205.05104