Well crafted and relevant paper by the authors, it was very nice to see how well the BBM rheology compares to the satellite data for the LKFs. The PDF analysis is another good results that shows the BBM does a good job in describing deformations.

We thank the reviewer for reviewing our paper and raising interesting questions, and also for suggesting the inclusion of the idealized “cyclone” test-case. We have implemented this test-case and results are now shown and discussed in the new version of the manuscript.

Before moving to a point-by-point reply format, here is some important information regarding the new simulations performed for the new version of the manuscript.

Based on the concerns expressed by both the anonymous reviewer #1 and Dr Plante on our initialization/spinup strategy, we have decided to use a simpler approach. As such, our simulations have been performed again: initialized with the new strategy (described in section 3.2 of the new manuscript), and consequently, all diagnostics, figures, and table values have been processed and generated again, based on these new simulations.

Also based on concerns raised by the reviewers, on what has been wrongly interpreted as unfairness towards the aEVP rheology, we now make it very clear, all over the paper, that what we use as the reference SI3 simulation is the default workhorse setup of SI3 as provided in the current version of NEMO. As such, for instance, experiment “SI3-aEVP” has been renamed to “SI3-default”, and it is clearly stated in the paper that a better-tuned “aEVP” would likely perform better (L513-516).

Moreover, in the first version of the manuscript, our developments and simulations were based on the version 4.2.1 of NEMO. At the time, it was the current stable release of NEMO. But during the course of the review process of this paper, a bug related to the ocean-ice drag parameterization has been identified and fixed by the NEMO team. This bug has been judged sufficiently severe by the NEMO team to justify the release of a new stable version: the 4.2.2. Link to release note: https://forge.nemo-ocean.eu/nemo/nemo/-/releases/4.2.2
Consequently we have switched to version 4.2.2 of NEMO for all the simulations presented in the new version of the manuscript.
The bug fix has significantly impacted the values of the simulated deformations and has therefore required a new tuning of the air-ice drag coefficients.

However, I am not convinced that the argument has been conducted in a way that is entirely fair for aEVP for these reasons:

1. If the default number of iterations for aEVP is 100 why has that been changed to 180? That is 80% more. You mention it is for "fairness" but is it fair to make the one you are comparing against slower? As far as I understand it, with 180 iterations your framework is 60% slower than aEVP, does that mean that the comparing to the default number of iterations your framework would be around twice as slow?

Based on your remark and that of RW1, these diagnostics have been performed again with the default value of N_{EVP}=100. Also note that in all our new BBM simulations (including those on which the deformation analysis is performed) we now use a number of sub-time-steps of N_S=100 for the time-splitting, and not N_S=180 as before; because based on sensitivity experiments that we have performed during the course of the review we have come to the conclusion that N_S=100 in our BBM implementation is sufficient to give very similar sea ice deformations, both qualitatively and quantitatively. So now, both aEVP and BBM simulations perform 100 subcycles under one advective time-step, in all the results presented. In that regard, we have also added two new figures in the new Appendix C (figures C1 & C2) that compare the cyclone test-case results for aEVP and BBM using more sub-cycles, N_{EVP}=1000 and N_S=200, respectively.

2. It is also not clear why when coupling with ocean the difference in time decreases. Can you clarify?

Because in regular ocean/sea-ice coupled mode, NEMO couples the ocean and sea-ice modules in a sequential, and not in a parallel, way. This means that the cost of SI3 simply adds up to that of OCE.
We have clarified this part in the discussion section 4.1 (607):
"This lower value is explained by the fact that by default, the coupling between OCE and SI3 is done sequentially. As such, the cost of SI3 simply adds up to that of OCE, and the cost of OCE is expected to be independent of the mode used (in our case: 113 and 114 cpu h for SI3-default and SI3-BBM, respectively)."
We also now provide more numbers in the new Table 4, these should help the reader figure out the smaller values obtained when using the coupled setup.

3. You make a comment saying that the resolution does not change with your approach, but you are considering more than double the number of degrees of freedom. When discussing accuracy, I think you should have also compared BBM to a case where aEVP has the same degrees of freedom as BBM.

Yes, and it turns out that it is this higher degree of freedom that is actually responsible for causing our problem with the use of the E-grid, namely the separation of the solutions between the two grids, which necessitates the use of our cross-nudging treatment to prevent it. And this cross-nudging by smoothing and forcing the two solutions to remain consistent with one another does nothing else than "taming" the consequences of this increased degree

of freedom. So in short, this higher degree of freedom has some clear disadvantages and is sort of “destroyed” by the use of the cross-nudging. We now clearly state in our manuscript, that the benefit of not having to interpolate some fields between corner and center points of the grid (using a 4-point-averaging, as done in the aEVP implementation of SI3) comes with the disadvantage of having to smooth them with the cross-nudging (e.g. in the conclusion L640-646).
We think that it would indeed be very interesting to port aEVP on the E-grid in a way similar to what we have done in our BBM implementation, as we expect that this would improve the aEVP results. This, however, is out of scope of this paper that primarily aim is to focus on the BBM rheology, its implementation and the evaluation of its performance compared to observations in a Pan-Arctic realistic simulation.

4. It is great to use satellite data as comparison and the results you show are definitely very good for BBM. I think however that you should also consider a test that has been widely used by other modelers so that your results with BBM can be put into a common context and it would be easier to compare against existing approaches. The test case I am referring to is the one moving anti-cyclone that can be found for instance here: *Simulating Linear Kinematic Features in Viscous-Plastic Sea Ice Models on Quadrilateral and Triangular Grids With Different Variable Staggering*, but has been also used for instance here: *Simulating Sea-Ice Deformation in Viscous-Plastic Sea-Ice Models With CD-Grids* and here: *CD-type discretization for sea ice dynamics in FESOM version2.*

Thank you for this recommendation which enhances the validation of our BBM implementation and helped put BBM into a known context. You will see that this idealized test-case, along with the appropriate references, has been added and is discussed in the new version of the manuscript (new section 3.1, and Figures 5, C1 & C2). It is also used to show the effect of using a higher number of time subcycles in both SI3-default (known as SI3-aEVP in the previous version of our manuscript) and SI3-BBM (Figures C2 & C3).

I recommend the authors address the issues I raised here before I can advise for publication.