This paper presents a method that extends previously published work on Blocked-HGFDR to achieve better lossy compression – both in terms of the distribution of residuals as well as compression ratios. This new method is called Adaptive-HGFDR. The paper includes results from compression experiments to justify the claims about the method.

Although the manuscript has improved since the last revision, many of the previously pointed-out issues remain, and I also have some concerns not previously raised.

Previously pointed-out issues that were not convincingly addressed yet, in my opinion:

1. References still seem off: e.g. Anon, 2011; Of and Acm, 2000; Anon, 2013; None, 1970; Text contains (Springer, 2011) but not the References section; Diffenderfer (2019 a) vs Diffenderfer (2019 b) in the text when there is only one Diffenderfer et al. in the References section
2. An exhaustive list of all outstanding language issues would be too big to list out here. I have included a subset of minor language-related corrections in minor issues, but perhaps the authors should use an "autocorrect tool" to list out all the issues (e.g. I use Grammarly for this purpose).
3. Line 54: "For the file-based compression method, it is difficult to arbitrarily adjust the compression parameter according to the given compression error." This is related to something previously pointed out as not true - I still think this is not true - see e.g. https://github.com/LLNL/H5Z-ZFP
4. Line 73: "...ZFP (Diffenderfer et al., 2019b) are typical methods that use the[sic] feature prediction to achieve lossy compression." Also pointed out in previous reviews that it is not clear why this is being called feature prediction. I looked up the cited paper and it doesn't mention the word "feature". So where is this insight from?
5. The captions for all figures should be expanded so that the reader can answer the questions of "What is going on in this figure?", "What does that mean?", "Why should I care?" are answered right there in the caption, i.e. without having to read the full text.
6. Repeating a previous reviewer's comment: "Not all ESMD data is high-precision. In fact, it is typical that calculations are done in double precision, but that data is output in single precision (e.g., for CESM). What is the precision of the data that you are compressing?" While this was answered in the rebuttal document, but the manuscript was not updated correspondingly (I did a quick Ctrl+F for the word "precision")
7. The above is one example where the authors responded to a comment in the rebuttal but the manuscript was not updated correspondingly, but there are many more. I would advise the authors to go through all the comments from the previous round and make sure the comments are addressed in the manuscript text and not just the rebuttal document.

Issues not previously pointed out but would be nice to fix regardless:

1. The motivation for why a uniform distribution of compression errors is something to strive for is not convincing for me. I understand that lines 36-39 are attempting to do this but I honestly can't follow the line of reasoning. Since this is the core "Why should I care?" of this paper, I think the authors would do well to explain this better.
2. The paper would be easier to read if it had a clear "Contributions" section. As I understand it, the delta of the method described here, as compared to Blocked-HGFDR is that a) each block can have its own rank b) a proposed method for

calculating the optimal rank per block. I think the readers would appreciate having this spelled out towards the beginning of the paper.

3. The “fast search algorithm” described in Definition 4 (line 194) appears to be stated as an original contribution in this paper. To me, this appears to be a rephrasing of “Binary Search”, a method that is commonly used in this space. In fact, Grasedyk et al, cited here, also uses this algorithm. The authors would do the reader a favour by making this clearer – either by clearly stating that this is a description of Binary Search, or by clarifying the difference between the stated method and Binary Search.
4. Assuming that I’m not mistaken in the above comment, I’m not convinced that Rouillier et. al. is the best article to cite for the log(n) complexity of Binary Search.
5. Line 55: “For the error truncation-based[sic] compression, the distribution of floating-point precision of ESMD is not uniform, which could lead to the unevenly[sic] distribution of compression errors.” Either I misunderstood the statement or this is not true. That rounding/truncation errors are approximately uniformly distributed is a well-known result and used in fields from signal processing to machine learning (e.g. https://arxiv.org/pdf/1802.01436.pdf - this paper refers to it as quantization error)
6. Line 60: “To summarize, it is hard to achieve flexible control of the compression ratio and errors for the description-based lossy compression methods.” Perhaps as a result of the other concerns I raise elsewhere, but there doesn’t seem to be enough justification provided to make this claim.
7. Line 83: “None of these methods considers ESMD as the[sic] high dimensional data with the heterogeneous correlation between different dimensions.” I’m not sure about the other methods but ZFP and SZ surely consider higher-dimensional data. In fact, Tao et. al (2017), cited here, talks specifically about multidimensional prediction. Can the authors please clarify why this does not meet their definition of considering ESMD as high dimensional data?
8. Equation 3 is from Yuan et al (2015) and should be cited as such
9. As I understand it, the hardware used here (HP Compaq Elite 8380 219 MT with Intel Core i7-3770 3.4 GHz processors and 8 GB of RAM) is really small compared to what would be used in realistic runs of this problem. I think this should be pointed out since the compression times are an essential result being reported.
10. It's not clear how the data for Figure 3 was obtained. e.g. text says "… zfp algorithm are affected by the tolerance parameter, which is set to 0.5 ". Does that mean all the data points in Fig 3 for zfp were obtained by the same setting? Then what was varied to get a variation of compression ratios?
11. When I attempted to answer the above question myself by looking at the provided code, I realised that the code does not include ZFP anywhere. Please correct me if I’m mistaken in this conclusion. I think the authors should provide the code to reproduce all of the plots in the paper.

Minor issues:
Line 40: “The main idea of ESMD lossy compressions is to eliminate unnecessary information in data to reduce the data size.” This is true of any compression, not just ESMD lossy compression. Also compressions->compression.

develope -> develop

exiting -> existing

unbalance -> imbalance
prominent components -> principal components?