Comments to Reviewers

L. Minah Yang, Alyson Fox, Geoffrey Sanders

May 28, 2020

We would like to thank both reviewers for your time and effort. First, we would like to make a few comments on the major revisions:

- **Column partitioned Householder QR factorization algorithm (HQR) has been added to reflect on mixed precision settings of GPU tensor cores units.**
  We added the analysis of the level-3 BLAS variant of HQR since it is the standard HQR implementation in many libraries and can be effortlessly adapted to utilize the block Fused Multiply-Add operations (bFMAs) of NVIDIA TensorCore units. The WY representation of [1] is used instead of the compact, storage-efficient version of [4] since the former is discussed in both [2, 3], which we often refer to in the text.
- **We have added the mixed precision setting of NVIDIA's TensorCore bFMAs.** This setting is more relevant and practical since these hardware units are already in use, and result in fewer low precision errors than the inner product mixed precision setting we had introduced. While we do not discuss the speed-up advantages in depth, we do refer to the speed benchmarks for these operations that already exist and may be of interest ot the readers.
- **Section 5 (Applications) has been removed from this manuscript.**
  Given the length of the first submission and the additional materials introduced as explained above, we have removed the applications section from this text. However, we plan on presenting our work on using mixed precision arithmetic in graph clustering in a future work, where we will include other graph problems.

Next, we would like to address concerns voiced by both referees:

- **Missing references to relevant and prior works**
- **Inconsistencies in notation and adhering to standard notation:**
  - **Hyphens have been removed from "low-", "mixed-", and "high-precision".**
  - **We changed the notation for γ for k accumulated rounding errors in precision type q from $\gamma_q^{(k)}$ to $\gamma_k^{(q)}$.**
- **line 44-45: "QR factorization is known to provide a backward stable solution to the linear least squares problem ...' Using what algorithm/under what conditions? / P2L44 "QR factorization is known to..."**

Requests from Referee #1

1. **The primary problem that I see is that the authors have not convinced me that there is any novelty in what they call a "new framework" for doing mixed-precision floating point error analysis.**
   asdf

2. **The paper also suffers a complete lack of acknowledgement of prior work in this area and of the current state-of-the-art.**

3. **line 39: What is meant by "exact products"?**

4. **lines 157-159: citation missing at the end of this sentence.**

5. **lines 163-165: Be more specific about what rounding error analysis framework was established in the textbook [13]. As far as I know, the textbook does not establish any new error analysis framework, nor is it limited to analyses using a single precision.**

6. **Section 2.2: Prior work should be discussed and cited here.**

7. **Pages 2-6 contain standard introductory textbook material and can be significantly shortened.**

Requests from Referee #2

1. Choice of mixed precision assumptions
   - Relation with GPU tensor core units
   - Assumptions on storage precision types
   - Distinction between Lemma 2.4 and Corollary 2.5
2. Rounding error analysis framework
3. Conclusions from the mixed precision HQR analysis (section 3)
4. Conclusion from HQR vs TSQR comparison (section 4)
5. Mislabel of forward and backward errors
6. P1L15: "standard algorithms may no longer be numerically stable when using half precision"
7. P2L49: fp16 should be removed, bfloat should be bfloat16.
8. P5L125: Rewording is needed to clarify "k represents the number of FLOPs"
9. Title suggestion:
10. P1L20: what does "weight" refer to in this context?
11. P2L57 "can successfully" $\rightarrow$ "can be succcessfully".
12. P5L134: $\gamma_p^{(d+2)}$ has not been defined yet.
13. P12L329: the middle term should be $(1+\delta_w)(x_1-\sigma-\Delta\sigma)$, rather than $(1+\delta_w)(\sigma+\Delta\sigma)$. Moreover, the last equality is only true because no cancellation can happen, since x_1 and σ have the same sign: this should be commented on.
14. Equations (4.6) and (4.7): isn't the $\sqrt{n}$ factor on the wrong equation?

15. P18L527: “for the a meaningful”.
16. P18L531: as mentioned above, the $(L+1)/2^L$ factor is reversed
17. P20L599: I find it very strange that the backward error depends on the condition number of the matrix! Is it rather the forward error that is being plotted?
18. Section 5: given the relatively theoretical nature of this article, section 5 felt slightly out of place to me. Given that the article is quite long, perhaps the authors could consider including section 4 in another piece of work?

References

[1] C. Bischof and C. Van Loan, *The WY Representation for Products of Householder Matrices*, SIAM Journal on Scientific and Statistical Computing, 8 (1987), pp. s2–s13, `https://doi.org/10.1137/0908009`.

[2] G. H. Golub and C. F. Van Loan, *Matrix computations*, JHU press, 4 ed., 2013.

[3] N. J. Higham, *Accuracy and Stability of Numerical Methods*, 2002, `https://doi.org/10.2307/2669725`.

[4] R. Schreiber and C. Van Loan, *A Storage-Efficient WY Representation for Products of Householder Transformations*, SIAM Journal on Scientific and Statistical Computing, 10 (1989), pp. 53–57, `https://doi.org/10.1137/0910005`.