Qeios

Peer Review

Review of: "Evaluation of Bfloat16, Posit, and Takum Arithmetics in Sparse Linear Solvers"

Yves Durand[1]

1. Atomic Energy and Alternative Energies Commission, Gif-sur-Yvette, France

This study entitled "Evaluation of Bfloat16, Posit, and Takum Arithmetics in Sparse Linear Solvers" is a valuable contribution to the SOA regarding floating point formats.
It provides experimental results on the forward error obtained with these different formats with direct methods (LU, QR), iterative refinement and GMRES.
Globally, this work is clear, well documented and its conclusions are well and fairly expressed.
Only issue, this work is based on a set of example matrices whose characterics are not documented precisely.

Comparing forward errors is useful for direct methods. It also makes sense to compare low-precision formats in that case. For larger, ill-conditionned datasets, direct methods saturate the memory and iterative methods are used with larger formats.

The study details the methodology, but does not really mention the sizes of the matrices: the reader is aware of 295 matrices but the size distributions (and NNZ) are not known, just mentionning that they exceed 10^4 in NNZ, which is not really significative (it is achieved for some small 2Kx2K matrices).
The condition number is useful, but since incomplete LU is used, the actual matrix being solved (after preconditionning) is considerably easier, it does not really represent the problem. It is worth mentionning that iLU works very well with rather easy matrices, but better preconditionners are necessary for larger and more difficult cases.

 doi.org/10.32388/7TL64C

The question of using Posit and Takum for Krylov-based solvers, for which these formats seem unappropriate at first sight, has not been fully addressed yet (only with CG with small matrices in [6] [7]).

Thus, the experimental results provided with GMRES are welcome. However, the raw cumulated iteration count doers not answer the real question.

The problem of tapered precision with Krylov is that at some point, the correction to the residual steps outside of the format "golden zone" and the solver diverges. Since this case is difficult to predict in theory (impossible to predict except for CG), there are discussions whether this happens in practice or not.

In brief, the question with Krlov methods is not so much to compare the number of iteration (which are quite close, as predictable), but to know if Posit & Takum is stable enough for large matrices.

Thus I kindly ask to document a little more precisely the choice of matrices

Declarations

Potential competing interests: No potential competing interests to declare.