Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., author comment AC2
https://doi.org/10.5194/gmd-2021-216-AC2, 2021

Reply on RC2

Rafael Lago et al.

Author comment on "MagIC v5.10: a two-dimensional message-passing interface (MPI) distribution for pseudo-spectral magnetohydrodynamics simulations in spherical geometry" by Rafael Lago et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-216-AC2, 2021

We'd like to thank the referee for her/his very thorough and constructive review, especially concerning other architectures. In the following we individually address all points raised by the Referee #2:

> Considering that a large portion of todays supercomputers include
> accelerators, possibly making up most of their raw computational power, does
> the 2D strategy bring any advantages with regards to using accelerators?

This is a very valid point. Unfortunately, porting a large production code like MagIC to GPUs (or more generally speaking to a discrete "accelerator") is a significant challenge which is way beyond the scope of this work. Currently, and at least for the next couple of years, the community using MagIC has access to large-scale CPU (x86_64 CPUs from Intel or AMD) resources, e.g. in Germany or France, justifying the development efforts for "modernizing" the CPU-only version of MagIC.

That said, we do fully agree with the referee that porting to GPUs eventually might even become unavoidable in the light of the current hardware trends and technological developments in HPC. Realistically, a GPU port of MagIC would start out from the OpenMP parallelization which is already there for the 1D-MPI parallelization (which is envisaged also for the 2D-MPI parallelization). In that sense, the newly developed 2D MPI-strategy does not bring immediate advantages with regards to using accelerators but it would add flexibility to a future GPU version of MagIC in the same way as it now does for the CPU-only version. We have added a short paragraph at the end of Section 5, Conclusions and Future Work (lines 601-607 of the revised manuscript) which addresses this point and sketches a conceivable GPU-porting strategy for MagIC.

> The discussion of the transposition in section 3.3 is confusing. The paragraph
> starting at line 335 discusses the importance of the size of the queue, but
> the MPI algorithms (l. 325) specify a MPI communication "per scalar field".
> The discussion in Section 4.1 also seems to imply a single communication call
> per "queue". Can you clarify this?

Indeed, the description in lines 325 was wrong. The algorithm performs one communication for all queued fields. The text has been updated accordingly.

> The performance benchmarks provide insight into the behaviour of the 2D
> parallelization implementation but are executed on a single cluster. A more

> general discussion on what kind of performance to expect on another HPC
> cluster depending on its technical characteristics would be helpful for the
> wider community.

Indeed such a discussion was missing from the main manuscript. We added a paragraph in in Section 4.5 of the manuscript (lines 523-529), elaborating on how these performances should translate into other clusters.

> The strong scaling experiment with the 1D hybrid strategy increases the
> number of threads for runs with higher number of cores. This is counter
> intuitive as it leads to even smaller computational load per thread which
> I would expect to affect the scaling negatively. Can you explain this
> behaviour?

We performed tests with the 1d-hybrid code with 10 and 20 threads, but always with the same total number of threads. For instance, the test with 1,000 cores was executed using 100 MPI ranks and 10 threads per rank, or 50 MPI ranks and 20 threads per rank. In both scenarios, the workload per thread is the same. What changes is the amount of data being handled by each MPI rank. The transition from 10 threads to 20 threads shows that, at some point it is better for MPI to handle larger chunks of data, which is expected.

> At high resolution, the memory footprint of a pseudo-spectral code can become
> important. Is there a benefit, at the memory level, to use a 2D distribution
> in MagIC?

This is indeed another benefit of the 2d-MPI implementation. Since it allows the use of more resources, the user may have access to more (distributed) memory as well. We added a comment in the revised version of the text, in Section 4.5, lines 577 and 578.

It may be noted that the proposed 2D MPI distribution requires additional memory for storing the queued fields during the θ-transposition. However, in section 3.3 we clarify that the queue size (and thus, extra memory requirement) can be controlled by the user as well. The benefit of having more distributed memory available can easily overcome the disadvantage of having to store a controllable extra number of scalar fields.

> l. 132: Plm should be the "associated Legendre polynomials" and not the
> "Legendre polynomials".

This has been corrected as requested.

> Table 3: Adding a κ Tr column would make it easier to follow the discussion
> at the end of Section 4.6

This has been updated as suggested.