Efficiency and Scaling of the Multigrid Beta Filter for Modeling Background Error Covariance

Miodrag Rancic, R. James Purser and Manuel Pondeca
IMSG at NOAA/NCEP/EMC
Email: miodrag.rancic@noaa.gov

1. Multigrid beta filter

A new technique for modeling of background error covariance, the Multigrid Beta Filter (MGBF), is under development at the Environmental Modeling Center - EMC (Purser et al. 2022). The new technique approximates a Gaussian using the Beta function with a finite support but achieves large spatial spread through the application of a multigrid technique. This combination results in a very efficient and easily parallelizable algorithm, which is expected to perform successfully on the large numbers of parallel processors of contemporary high-performance machines. In addition, the filter response can produce a large range of shapes of covariances, including those with negative side lobes, and even produce cross-correlations within the variational method, which was until now exclusive to ensemble-based methods.

The MGBF comes in two main flavors: a radial filter and a line filter. While the radial filter spreads the filter response within an ellipse of influence around the initial delta impulse, the line filter, like the recursive filters used in data assimilation at EMC until now, spreads this response along the selected lines. In the 2D case, three stages of filtering are needed; in the 3D case, seven stages are needed, with only six active filters at any generic grid point. The number of stages corresponds to the number of degrees of freedom in a general symmetric "aspect" tensor which controls the shape of the filter. In the isotropic case, where the *x* and *y* components of the 2D aspect tensor are equal, the line filters need only be applied in two dimensions, which provides a fast, isotropic version of the filter.

2. Tests with a standalone version of MGBF

A standalone version of the MGBF was developed to handle the same number of variables (six 3D and four 2D) as the recursive filter used in the GSI (EMC's operational data assimilation system). It operates at the same resolution as the GSI: 1792 x 1956 horizontal grid intervals and 50 vertical levels, but outside of the framework of the whole data assimilation system. It can therefore more speedily facilitate new developments, and the testing of different flavors of the filter.

Fig.1 shows the "speedup" achieved in a 2D+1D option of the radial and the line versions of the MGB filter, in a series of tests running on various constellations of processors (PEs), starting with 64 all the way up to 11264. Here, 2D+1D is a shorthand to describe the method where the 3D variables are filtered through the application of a horizontal 2D filter and a vertical 1D filter. Speedup is defined as inverse time – thus, on this graph, a higher value is better.

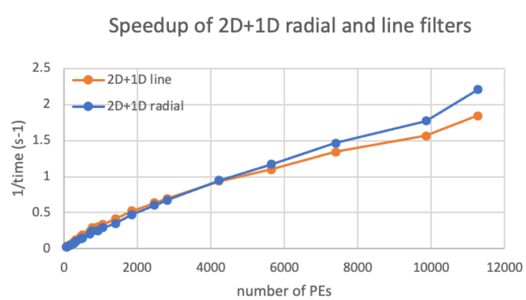

Figure 1 Speedup of the 2D+1D radial and line filters in the standalone version of the MGBF.

In this case, the line filter performs better up to about 2800 PEs, then the radial filter becomes faster thanks to the less frequent need to call the halo-exchange "side sending" subroutines. Development is already in progress for a new method which will group line filters into planar combinations to reduce the number of side-send halo exchanges and address this issue.

In the case of fully anisotropic 3D MGB filtering, the line version is systematically superior to the radial (Fig. 2), but its speedup with the increasing number of PEs still cannot match that of the isotropic 2D+1D version.

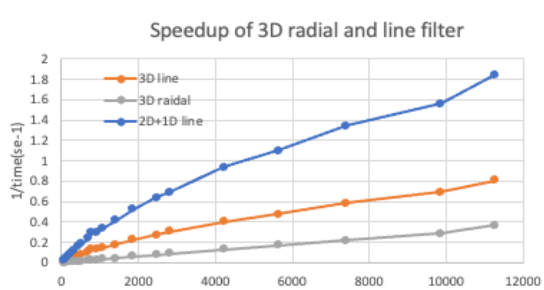

Figure 2 Speedup of the 3D radial and line filters in the standalone version of the MGBF.

Finally, we tested the "fast" isotropic version of the line filter (1D+1D+1D), which requires only two passes of the line filter horizontally (Fig. 3) and, as expected, it clearly has the best speedup rate.

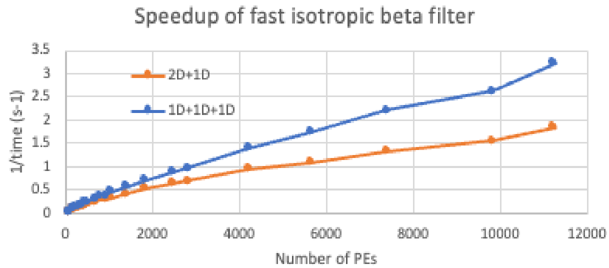

Figure 3 Speedup of the 1D+1D+1D line filter compared with the 2D+1D radial filter in the standalone version of the MGBF.

Note that the performance of all tested versions of the filter keeps improving all the way to 11264 PEs. A further increase in the number of PEs would require a substantial modification of the code to include halos over two rows of processors, which has not yet been done.

3. Tests with MGBF in GSI

We tested the isotropic version of the MGB line filter against the recursive filter (RF) in the GSI, using two different resolutions for the filter grid for the multigrid generation. The first configuration used the resolution of the analysis grid, and the second had about 10% lower horizontal resolution. Results are summarized in Figs. 4 and 5.

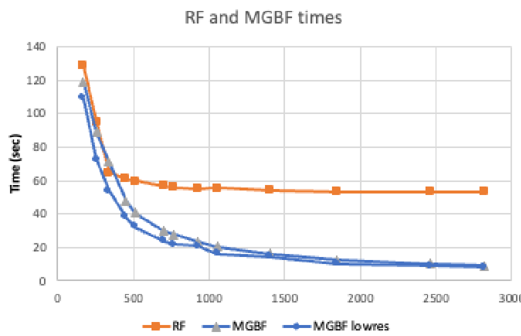

Figure 4 Times spent on filtering in runs of the GSI over various constellations of PEs. The isotropic version of the line MGBF is run at two different resolutions.

The isotropic version of the MGB line filter is about three times faster than the RF and scales much better (Fig. 4).

Figure 5 Total times spent on filtering in runs of the GSI over various constellations of PEs. The isotropic version of line MGBF is used, with two different resolutions.

Yet, though the total time spent in the GSI on filtering (Fig. 5) is systematically at least 20-25% lower with the MGBF than with the RF, the MGBF alone is not sufficient to push down the overall execution time of the GSI for processor numbers over ~600 PEs, presumably due to too many all-to-all communications and direct exchanges of data with disks. Therefore, some re-engineering of the GSI may be needed if we are to take full advantage of the new MGBF.

References

Purser, R. J., M. Rančić, M.S.F.V. De Pondeca, 2022: The Multigrid Beta function approach for modeling of background error covariance in the Real Time Mesoscale Analysis (RTMA), *Mon. Wea. Rev.,* **150,** (in press). **https://doi.org/10.1175/MWR-D-20-0405.1**