
Bram Reps
**Communication avoiding strategies for linear solvers in a
space-filling curve framework**

Department of Mathematics and Computer Science
University of Antwerp
Middelheimlaan 1
B-2020 Antwerp
Belgium
Intel ExaScience Lab
Kapeldreef 75
B-3001 Leuven
Belgium
`bram.reps@ua.ac.be`
Bart Verleye
Tobias Weinzierl
Dirk Roose
Pieter Ghysels
Wim Vanroose

Iterative methods such as Krylov subspace and multigrid methods are typically used to solve large linear systems that arise from the discretization of a partial differential equation (PDE). As these equations mainly connect neighboring points in space, resulting in very sparse systems, a data structure that is based on space-filling curves can be exploited to develop numerical solvers, since the associated data mappings carry specific locality properties [1].

The cost of an iterative method is largely determined by the efficient use of two basic calculations: sparse matrix-vector multiplications (SpMV) for stencil operations, and vector multiplications (dot products) for computing projections and norms. Both of these common steps in a numerical algorithm require time-consuming data communication. The result of a SpMV is a vector and communication is locally restricted to a few neighboring points in space, as given by the sparsity pattern of the matrix, whereas a dot product results in a scalar that depends on all points in space.

We study a framework where a multidimensional numerical grid is traversed in a particular order defined by a one-dimensional space-filling curve. In each cell along this curve only local grid data, belonging to the cell and its adjacent vertices, is manipulated as specified by the iterative method, before moving on to the next cell in line. This approach is particularly interesting for the sequentialization of adaptive grids, since the curve can recursively be extended on-the-fly in refining regions and the underlying data is typically stored in a

spacetree. In addition, the parallelization of a solver is straightforward if the space-filling curve is simply cut and divided over different processors.

Although the data dependency for the SpMV is usually relatively local in space, an exchange of information between cells is unavoidable in between computations. This can easily be invoked by simply running extra grid traversals dedicated to data communication through mutual vertices; however, it dramatically increases the runtime of a programming code, especially when communicating cells are handled by different parallel processors. The result of a dot product is stored in a state variable, accessible from all cells, and needs to be communicated back and forth between parallel processors at the end of a traversal.

It fits better to the state-of-the-art in high performance computing if the two phases of computation and communication are combined as much as possible. This means that from the point of view of space-filling curves, the current tendency of communication avoiding and hiding programming [4, 2] naturally translates into the basic concept of minimizing the number of grid traversals by postponing and bundling the communication of SpMV and dot product results.

To elucidate these strategies, we present a parallel implementation of both the Conjugant Gradient (CG) method [3] and an additive multigrid method that respects the idea of minimal grid traversals and illustrate them in the space-filling curve framework Peano [5, 6]. The necessary communication of dot product and SpMV results, is aggregated in few instances of the algorithm's main loop, requiring e.g. only one grid traversal per CG-iteration.

Bibliography

- [1] M. Bader. *Space-Filling Curves*. Springer-Verlag Berlin Heidelberg, 2013.
- [2] P. Ghysels, T. J. Ashby, K. Meerbergen, and W. Vanroose. Hiding global communication latency in the GMRES algorithm on massively parallel machines. *SIAM Journal on Scientific Computing*, 2012.
- [3] M. Hestenes and E. Stiefel. Methods of Conjugate Gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6), December 1952.
- [4] M. Hoemmen. *Communication-avoiding Krylov subspace methods*. PhD thesis, University of California, 2010.
- [5] T. Weinzierl. *A Framework for Parallel PDE Solvers on Multiscale Adaptive Cartesian Grids*. PhD thesis, Technische Universitt Mnchen, 2009.
- [6] T. Weinzierl et al. Peano – a Framework for PDE Solvers on Spacetree Grids, 2012. <http://www.peano-framework.org>.