Time-parallel methods for massively parallel solution of PDEs

D. Ruprecht†, R. Speck*†, R. Krause†,
†Institute of Computational Science, Università della Svizzera italiana, Switzerland.
*Jülich Supercomputing Centre, Forschungszentrum Jülich, Germany.

I. Overview

Todays fastest supercomputers already feature more than a million cores and this number is expected to rise beyond 100 million over the next decade. Because at the same time frequencies of individual processors remain constant or even decrease for reasons of efficiency, developers are increasingly confronted with the fact that accelerating numerical (but also other) codes *necessarily requires to exploit concurrency*. This, in turn, means that concurrency is more and more becoming a critical property of numerical algorithms. To this effect, methods for solving initial value problems that provide concurrency in the temporal direction have been shown to be an effective way to increase the degree of parallelism in the solution of time-dependent PDEs. Several such "time-parallel" methods exist, ranging from the very early interpolation-based scheme in [1] over the parabolic multi-grid method [2] to, more recently, the Parareal algorithm [3] and the "parallel full approximation scheme in space and time" (PFASST) [4]. The talk will focus on both Parareal and PFASST, discuss some of their critical mathematical properties and present benchmarks of the performance of these methods for large-scale parallel solutions of time-dependent PDEs. A very brief summary of the key aspects including multiple references is given below.

II. Parareal

Parareal is at present probably the most popular time-parallel method, as it allows to use basically arbitrary one-step methods within the Parareal iteration. Denoting an accurate but computationally expensive method by $\mathcal{F}$ and a coarse but computationally cheap method by $\mathcal{G}$, Parareal replaces the straightforward serial time-stepping procedure $y_{n+1} = \mathcal{F}(y_n)$, $n = 0, \dots N-1$ by an iteration

$$y_{n+1}^{k+1} = \mathcal{G}(y_n^{k+1}) + \mathcal{F}(y_n^k) - \mathcal{G}(y_n^k), \;\; k = 0, \dots, N_{\text{it}}. \qquad (1)$$

The key here is that once the values y_n^k from the previous iteration are known, the computationally expensive computation of $\mathcal{F}(y_n^k)$ for $n = 0, \dots, N-1$ can be done in parallel on N processors, followed by a serial but cheap correction in which $\mathcal{G}(y_n^{k+1})$ is evaluated and y_{n+1}^{k+1} computed. A detailed discussion of the algorithm plus many additional references can be found e.g. in [5], a detailed mathematical analysis is conducted in [6]. Its speedup using N processors and performing N_{it} iterations can by design not be optimal and is restricted by two competing bounds

$$s(N) \leq \min \left\{ \frac{N}{N_{\text{it}}}, \frac{\text{Runtime } \mathcal{G}}{\text{Runtime } \mathcal{F}} \right\}. \qquad (2)$$

Nevertheless, Parareal can provide additional speedup for the solution of time-dependent problems after spatial parallelization is saturated.

III. PFASST

The PFASST method has been introduced in [4]. It is based on "spectral deferred correction" (SDC) methods [7], an iterative approach for computing collocation solutions. PFASST employs a hierarchy of space-time levels on which iterations of SDC (so-called "sweeps") are performed. These levels are coupled, as in nonlinear multi-grid methods, by an FAS-correction that allows the solution on the coarser levels to converge up to an accuracy determined by the discretization on the finest level. PFASST can also be interpreted as a time-parallel version of a multi-level spectral deferred correction method (MLSDC) [8]. By not solving the fine-level problem to full accuracy but only performing SDC sweeps, PFASST has a significantly improved speedup bound compared to Parareal. Ideal speedup, however, is also not obtainable. The capability of PFASST to be used in parallel simulations on $\mathcal{O}(100k)$ cores has been demonstrated for different scenarios. In [9], it is shown that PFASST can accelerate a particle-based Navier-Stokes solver beyond the saturation point of the underlying spatial parallelization of a Barnes-Hut tree-code [10]. Scaling of PFASST combined with a mesh-based discretization and a parallel multi-grid (PMG) as space-parallel solver for implicit time-stepping is studied in [11] and the impact of using spatial coarsening strategies in large-scale parallel simulations is discussed. Finally, the study is extended in [12], where scaling results of PFASST+PMG on up to all 448K cores of the IBM Blue Gene/Q JUQUEEN at Jülich Supercomputing Centre are reported.

References

[1] J. Nievergelt, "Parallel methods for integrating ordinary differential equations," *Commun. ACM*, vol. 7, no. 12, pp. 731–733, 1964. [Online]. Available: http://dx.doi.org/10.1145/355588.365137

[2] W. Hackbusch, "Parabolic multi-grid methods," *Computing Methods in Applied Sciences and Engineering, VI*, pp. 189–197, 1984. [Online]. Available: http://dl.acm.org/citation.cfm?id=4673.4714

[3] J.-L. Lions, Y. Maday, and G. Turinici, "A "parareal" in time discretization of PDE's," *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, vol. 332, pp. 661–668, 2001. [Online]. Available: http://dx.doi.org/10.1016/S0764-4442(00)01793-6

[4] M. Emmett and M. L. Minion, "Toward an efficient parallel in time method for partial differential equations," *Communications in Applied Mathematics and Computational Science*, vol. 7, pp. 105–132, 2012. [Online]. Available: http://dx.doi.org/10.2140/camcos.2012.7.105

[5] D. Ruprecht and R. Krause, "Explicit parallel-in-time integration of a linear acoustic-advection system," *Computers & Fluids*, vol. 59, no. 0, pp. 72 – 83, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.compfluid.2012.02.015

[6] M. J. Gander and S. Vandewalle, "Analysis of the parareal time-parallel time-integration method," *SIAM Journal on Scientific Computing*, vol. 29, no. 2, pp. 556–578, 2007. [Online]. Available: http://dx.doi.org/10.1137/05064607X

[7] A. Dutt, L. Greengard, and V. Rokhlin, "Spectral deferred correction methods for ordinary differential equations," *BIT Numerical Mathematics*, vol. 40, no. 2, pp. 241–266, 2000. [Online]. Available: http://dx.doi.org/10.1023/A:1022338906936

[8] R. Speck, D. Ruprecht, M. Emmett, M. Minion, M. Bolten, and R. Krause, "A multi-level spectral deferred correction method," 2013, arXiv:1307.1312 [math.NA]. [Online]. Available: http://arxiv.org/abs/1307.1312

[9] R. Speck, D. Ruprecht, R. Krause, M. Emmett, M. Minion, M. Winkel, and P. Gibbon, "A massively space-time parallel N-body solver," in *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis*, ser. SC '12. Los Alamitos, CA, USA: IEEE Computer Society Press, 2012, pp. 92:1–92:11. [Online]. Available: http://dx.doi.org/10.1109/SC.2012.6

[10] M. Winkel, R. Speck, H. Hübner, L. Arnold, R. Krause, and P. Gibbon, "A massively parallel, multi-disciplinary Barnes-Hut tree code for extreme-scale N-body simulations," *Computer Physics Communications*, vol. 183, no. 4, pp. 880 – 889, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.cpc.2011.12.013

[11] R. Speck, D. Ruprecht, M. Emmett, M. Bolten, and R. Krause, "A space-time parallel solver for the three-dimensional heat equation," 2013, arXiv:1307.7867 [cs.NA]. [Online]. Available: http://arxiv.org/abs/1307.7867

[12] D. Ruprecht, R. Speck, M. Emmett, M. Bolten, and R. Krause, "Extreme-scale space-time parallelism," 2013, poster, accepted for SC'13.

Workshop "Efficient solution of large systems of non-linear PDEs in science"
Lyon, France, October 7–9, 2013.