Five Decades of Time Parallel Time Integration, and a Note on the Degradation of the Performance of the Parareal Algorithm as a Function of the Reynolds Number

Martin J. Gander

Time parallel time integration has received renewed attention over the past decade, because, finally, in 2004, the visionary prediction of Nievergelt from 1964 became true [1]:

> "For the last 20 years, one has tried to speed up numerical computation mainly by providing ever faster computers. **Today, as it appears that one is getting closer to the maximal speed of electronic components**, emphasis is put on allowing operations to be performed in parallel. In the near future, much of numerical analysis will have to be recast in a more 'parallel' form."

There are nowadays so many processors on supercomputers that space parallelization often saturates before the available number of processors is employed, and thus speedup is not possible any more. This happens especially often for evolution problems, whose solution is usually computed by time-stepping, a process which is classically considered completely sequential.

Over the past five decades since Nievergelt, there have been many attempts to parallelize the classical time stepping approach, and the ideas employed can be classified into four different categories [2]:

Shooting Type Methods: These methods are based on a decomposition of the time interval into subintervals, and then a multiple shooting method is used, together with a Newton solver for the non-linear system of equations that needs to be solved for the shooting parameters. Nievergelt's method is a precursor of this, the complete algorithm was first proposed at the continuous level in [3]. Renewed interest was sparked by the presentation of the parareal algorithm [4], which is a multiple shooting method with an approximate Jacobian calculated on a coarse grid. A complete convergence analysis of parareal can be found in [5] for linear problems, and in [6] for non-linear problems. The parareal algorithm converges superlinearly, and works quite well for diffusive type problems, but is not effective for hyperbolic problems.

Waveform Relaxation: these methods go back to the existence proofs by Picard and Lindelöf for solutions of ordinary differential equations, and were adapted to become computational methods for circuits in [7]: the idea is to decompose the circuit into subcircuits, and then to solve iteratively only subcircuits, relaxing signals (waveforms) along wires that had to be cut in the decomposition to the previous iteration, which gave the method its name. Modern variants of waveform relaxation for PDEs use a domain decomposition of the spatial domain into subdomains, and then optimized transmission conditions, which leads to optimized Schwarz waveform relaxation methods [8]. These methods are well understood for parabolic problems [9, 10, 11], and also work for hyperbolic problems [12, 13]. Optimized transmission conditions are the key ingredient of many recently developed solvers for wave propagation problems, e.g. the sweeping preconditioner, source transfer and the method of polarized traces,

for an overview, see [14, 15]. There are also waveform relaxation methods based on other domain decomposition methods, like Dirichlet-Neumann and Neumann-Neumann waveform relaxation methods [16, 18, 17].

Space-Time Multigrid Methods: The first space-time multigrid method was the parabolic multigrid method proposed by Hackbusch in [19]. The idea is to apply the smoother not only on one time level, but over several time levels, like advancing in time before having converged at the present time level. The parabolic multigrid method has multigrid performance, provided one is not coarsening in time. A multigrid waveform relaxation algorithm was proposed in [20], using waveform relaxation at the continuous level as a smoother. The first full space-time multigrid method was proposed and analyzed in [21], using special restriction and extension operators in time. A new space-time multigrid method for parabolic problems using optimized block Jacobi smoothers can be found in [22]. A complete convergence analysis exists for this method, and it has excellent weak and even strong scalability properties.

Direct time parallel methods: These methods solve the full space-time problem without iteration. The first such methods were special predictor corrector methods, where the prediction and correction step can be performed in parallel [23]. Such methods are ideal for small scale parallelism on multicore architectures, and the most modern variants are currently the RIDC methods [24]. These methods can solve for example a system of ordinary differential equations to 8th order accuracy at the cost of an Euler method using 8 cores; this factor can thus multiply the number of processors already used for the space parallelization. A very recent time parallel direct method is the ParaExp algorithm, which is based on a completely overlapping time domain decomposition and the rapid propagation of homogeneous problems using rational Krylov approximations [25]. This method is currently restricted to linear problems, but works very well also for hyperbolic problems.

Many more details about the historical development of time parallel time integration methods can be found in [2].

After the presentation, Dörte Jando asked how the performance of the parareal algorithm degrades when one passes smoothly from the parabolic to a hyperbolic problem. I show here a numerical experiment to illustrate this, solving

$$u_t = \nu u_{xx} + u_x \quad \text{in } (0,2) \times (0,T), \quad T = 4, \tag{1}$$

with periodic boundary conditions[1] and the initial condition $u(0,x) = e^{-20(x-1)^2}$. I discretize the problem using centered finite differences with mesh size $h = 1/12$ and backward Euler in time with time step $\Delta t = 1/60$, which leads to the approximate solution shown on the left in Figure 1 for $\nu = 1/1024$, where ν represents the inverse of the Reynolds number. I then apply the parareal algorithm with 8 coarse time intervals, using one backward Euler step as the coarse solver. In Figure 1 on the right, I plot the decay of the maximum of the L^2 error in space over all coarse time points, as the parareal iteration progresses for a decreasing sequence of ν. One can clearly observe that when ν becomes small, the algorithm converges more and more slowly, and for very small ν, the

[1] otherwise the inflow boundary conditions would dominate the solution later in time, see [26]

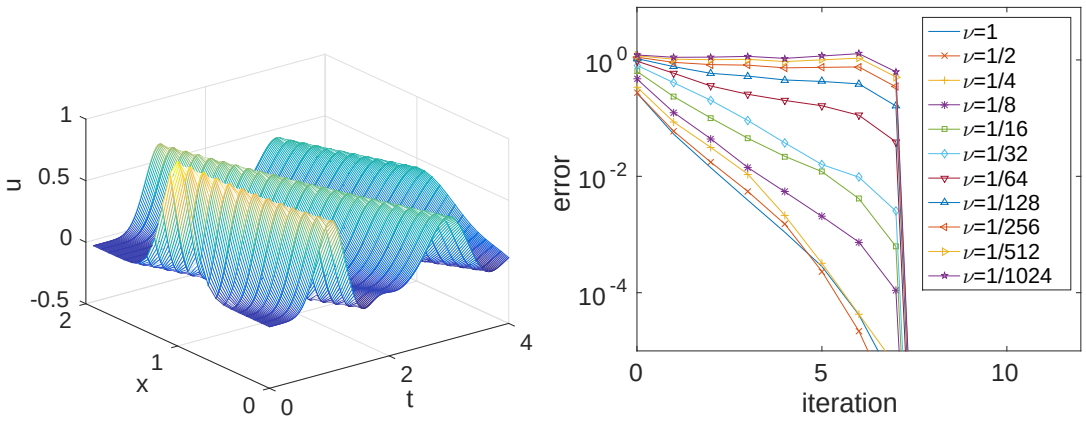

FIGURE 1. Solution for $\nu = 1/1024$ on the left, and degradation of the convergence of the parareal algorithm for decreasing ν (increasing Reynolds number) on the right.

only convergence mechanism left for the parareal algorithm is the convergence in a finite number of steps (here the 8th step), which however requires a number of iterations corresponding to the number of processors, and thus makes the method useless for parallelization, see also [5].

REFERENCES

[1] J. NIEVERGELT, *Parallel methods for integrating ordinary differential equations*, Comm. ACM, 7 (1964), pp. 731–733.

[2] M.J. GANDER, *50 Years of Time Parallel Time Integration*, in 'Multiple Shooting and Time Domain Decomposition', T. Carraro, M. Geiger, S. Körkel, R. Rannacher, editors, Springer Verlag (2015), pp. 69–114,

[3] P. CHARTIER AND B. PHILIPPE, *A parallel shooting technique for solving dissipative ODEs*, Computing, 51 (1993), pp. 209–236.

[4] J.-L. LIONS, Y. MADAY, AND G. TURINICI, *A parareal in time discretization of PDEs*, C.R. Acad. Sci. Paris, Serie I, 332 (2001), pp. 661–668.

[5] M. J. GANDER AND S. VANDEWALLE, *Analysis of the parareal time-parallel time-integration method*, SIAM J. Sci.Comput., 29 (2007), pp. 556–578.

[6] M. J. GANDER AND E. HAIRER, *Nonlinear convergence analysis for the parareal algorithm*, in Domain Decomposition Methods in Science and Engineering XVII, O. B. Widlund and D. E. Keyes, eds., vol. 60 of LNCS, Springer, 2008, pp. 45–56.

[7] E. LELARASMEE, A. E. RUEHLI, AND A. L. SANGIOVANNI-VINCENTELLI, *The waveform relaxation method for time-domain analysis of large scale integrated circuits*, IEEE Trans. on CAD of IC and Syst., 1 (1982), pp. 131–145.

[8] M. J. GANDER, L. HALPERN AND F. NATAF, *Optimal Convergence for Overlapping and Non-Overlapping Schwarz Waveform Relaxation*, in Proceedings of the 11th International Conference on Domain Decomposition, ddm.org, (1999), pp. 27–36.

[9] M. J. GANDER, AND L. HALPERN, *Optimized Schwarz Waveform Relaxation for Advection Reaction Diffusion Problems*, SIAM J. on Num. Anal., 45(2) (2007), pp. 666–697.

[10] D. BENNEQUIN, M. J. GANDER, AND L. HALPERN, *A homographic best approximation problem with application to optimized Schwarz waveform relaxation*, Math. of Comp., 78 (2009), pp. 185–223.

[11] D. BENNEQUIN, M. J. GANDER, L. GOUARIN, AND L. HALPERN, *Optimized Schwarz Waveform Relaxation for Advection Reaction Diffusion Equations in Two Dimensions*, 134(3), (2016), pp. 513–567.

[12] M. J. GANDER, L. HALPERN, AND F. NATAF, *Optimal Schwarz waveform relaxation for the one dimensional wave equation*, SIAM J. Num. Anal., 41 (2003), pp. 1643–1681.

[13] M. J. Gander and L. Halpern, *Absorbing boundary conditions for the wave equation and parallel computing*, Math. of Comp., 74 (2004), pp. 153–176.
[14] M.J. Gander, H. Zhang, *Iterative Solvers for the Helmholtz Equation: Factorizations, Sweeping Preconditioners, Source Transfer, Single Layer Potentials, Polarized Traces, and Optimized Schwarz Methods*, submitted to SIAM Review (2016).
[15] M.J. Gander, *On the limitations of sweeping type preconditioners for wave propagation*, Oberwolfach Report for the Mini-Workshop on Fast Solvers for Highly Oscillatory Problems (2016).
[16] F. Kwok, *Neumann-Neumann waveform relaxation for the time-dependent heat equation*, in Domain decomposition methods in science and engineering XVII, LNCS Springer (2014), pp. 189–198.
[17] M.J. Gander, F. Kwok and B. Mandal, *Dirichlet-Neumann and Neumann-Neumann waveform relaxation for the wave equation*, in Domain decomposition methods in science and engineering XVII, LNCS Springer (2014), pp. 501–509.
[18] M.J. Gander, F. Kwok and B. Mandal, *Dirichlet-Neumann and Neumann-Neumann Waveform Relaxation Algorithms for Parabolic Problems*, ETNA, 45 (2016), pp. 424–456.
[19] W. Hackbusch, *Parabolic multi-grid methods*, in Computing Methods in Applied Sciences and Engineering, VI, R. Glowinski and J.-L. Lions, eds., North-Holland, 1984, pp. 189–197.
[20] C. Lubich and A. Ostermann, *Multi-grid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216–234.
[21] S. Vandewalle and E. Van de Velde, *Space-time concurrent multigrid waveform relaxation*, Annals of Numer. Math, 1 (1994), pp. 347–363.
[22] M. J. Gander and M. Neumüller, *Analysis of a new space-time parallel multigrid algorithm for parabolic problems*, SIAM J. Sci. Comp., 38(4) (2016), pp. A2173–A2208.
[23] W. L. Miranker and W. Liniger, *Parallel methods for the numerical integration of ordinary differential equations*, Math. Comp., 91 (1967), pp. 303–320.
[24] A. J. Christlieb, C. B. Macdonald, and B. W. Ong, *Parallel high-order integrators*, SIAM Journal on Scientific Computing, 32 (2010), pp. 818–835.
[25] M. J. Gander and S. Güttel, *ParaExp: A parallel integrator for linear initial-value problems*, SIAM Journal on Scientific Computing, 35 (2013), pp. C123–C142.
[26] M. J. Gander, *Analysis of the Parareal Algorithm Applied to Hyperbolic Problems Using Characteristics*, Bol. Soc. Esp. Mat. Apl., 42 (2008), pp. 21–35.