

A Parareal Algorithm without Coarse Propagator?

Martin J. Gander^[0000-0001-8450-9223], Mario Ohlberger^[0000-0002-6260-3574], and
Stephan Rave^[0000-0003-0439-7212]

1 Introduction

The Parareal algorithm was introduced in [9] as a non-intrusive way to parallelize time stepping methods in the time direction for solving partial differential equations. Its convergence is now well understood: for linear problems of parabolic type, Parareal converges superlinearly on bounded time intervals, and satisfies a linear convergence bound on unbounded time intervals [8], which means it is a contraction even if computations are performed on arbitrary long time intervals. It was shown in the same reference also that for hyperbolic problems, the linear convergence estimate over long time does not predict contraction, and the superlinear estimate only indicates contraction when there are already too many iterations performed and one is approaching the finite step convergence property of Parareal. A non-linear convergence analysis for Parareal can be found in [6], and also the performance of Parareal for Hamiltonian problems is well understood, see [7], where also a derivative variant of Parareal was proposed.

An essential ingredient in the Parareal algorithm is the coarse propagator, and it is its accuracy that has a decisive influence on the convergence of Parareal, as one can see in the estimates from [8, 6]. We consider here a Parareal algorithm in the standard form for solving an evolution problem $\partial_t u = \mathcal{F}(u, t)$, and time partition $0 = T_0 < T_1 < \dots < T_N = T$,

$$U_{n+1}^{k+1} = F(U_n^k, T_n, T_{n+1}) + G(U_n^{k+1}, T_n, T_{n+1}) - G(U_n^k, T_n, T_{n+1}), \quad (1)$$

Martin J. Gander
Université de Genève, e-mail: martin.gander@unige.ch

Mario Ohlberger
University of Münster, e-mail: mario.ohlberger@uni-muenster.de

Stephan Rave
University of Münster, e-mail: stephan.rave@uni-muenster.de

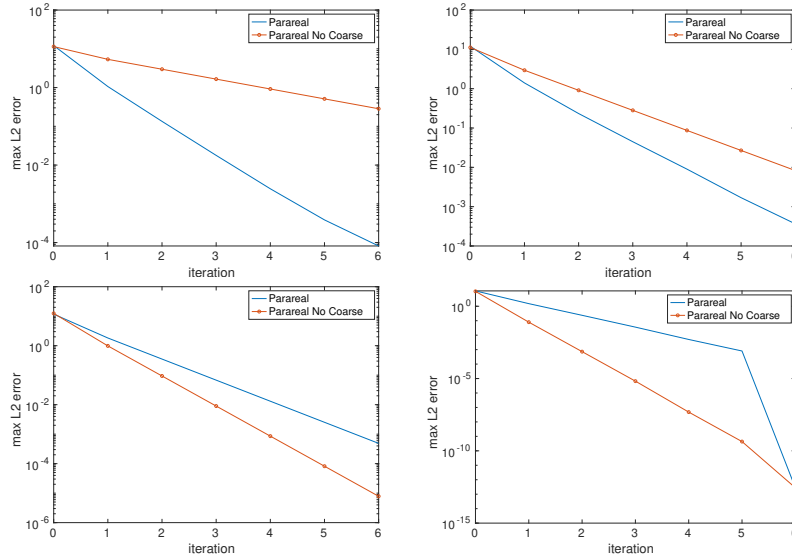


Fig. 1 Running a classical Parareal algorithm with and without coarse propagator. From top left to bottom right: 48, 24, 12, and 6 coarse time intervals with $T = 3$.

where the fine solver $F(U_n^k, T_n, T_{n+1})$ and the coarse solver $G(U_n^k, T_n, T_{n+1})$ solve the underlying evolution problem on time intervals $\Delta T := T_{n+1} - T_n$,

$$\partial_t u = \mathcal{F}(u, t), \quad u(T_n) = U_n^k, \quad (2)$$

with different accuracy. We are interested in understanding what happens when we remove the coarse solver in the Parareal algorithm (1), i.e. we run instead the iteration

$$U_{n+1}^k = F(U_n^k, T_n, T_{n+1}). \quad (3)$$

Note that this is quite different from the Identity Parareal algorithm [10], where the coarse propagator was replaced by the identity, a very coarse approximation, while here we remove the coarse propagator altogether.

2 A Numerical Experiment

We start with a numerical experiment for the one dimensional heat equation on the spatial domain $\Omega := (0, 1)$ and the time interval $(0, T]$,

$$\partial_t u = \partial_{xx} u + f, \quad u(x, 0) = u_0(x), \quad u(0, t) = u(1, t) = 0. \quad (4)$$

We show in Figure 1 the results obtained with zero initial conditions, $u_0(x) = 0$, and

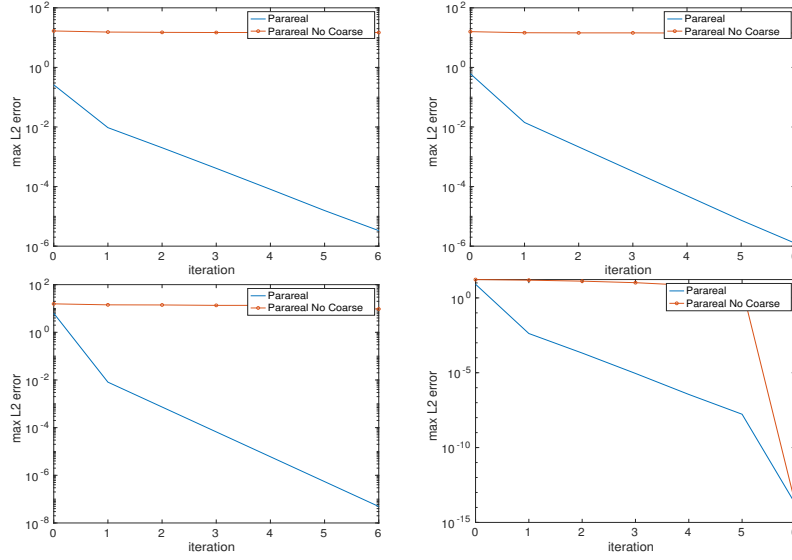


Fig. 2 Classical Parareal with and without coarse propagator with zero Neumann boundary conditions. From top left to bottom right: 48, 24, 12, and 6 coarse time intervals and $T = 3$.

the source term

$$f(x, t) = 10e^{-100(x-0.5)^2} \sum_{j=1}^4 e^{-100(t-t_j)^2}, \quad t_j = 0.1, 0.6, 1.35, 1.85, \quad (5)$$

like a heating device which is turned on and off at specific time instances. We discretized the heat equation (4) using a centered finite difference discretization in space with mesh size $\Delta x = \frac{1}{128}$, and Backward Euler in time with time step $\Delta t = \frac{1}{96}$, and run the Parareal algorithm up to $T = 3$ with various numbers N of coarse time intervals, where the coarse propagator G just does one big Backward Euler step. We see two really interesting results: first, the Parareal algorithm without coarse propagator is actually also converging when applied to the heat equation, both methods are a contraction. And second, if there are not many coarse time intervals, i.e. if the coarse time interval length ΔT is becoming large, the Parareal algorithm without coarse correction is converging even faster than with coarse correction!

In order to test this surprising property further, we repeat the above experiment, but use now homogeneous Neumann boundary conditions in (4), $\partial_x u(0, t) = \partial_x u(1, t) = 0$, instead of the Dirichlet conditions. We show the corresponding results in Figure 2. We see that while the Parareal algorithm with coarse propagator is converging as before with Dirichlet conditions, the Parareal algorithm without coarse propagator is not converging any more, except in the last case when only 6 coarse time intervals are used, and here it is the finite step convergence property of the Parareal algorithm

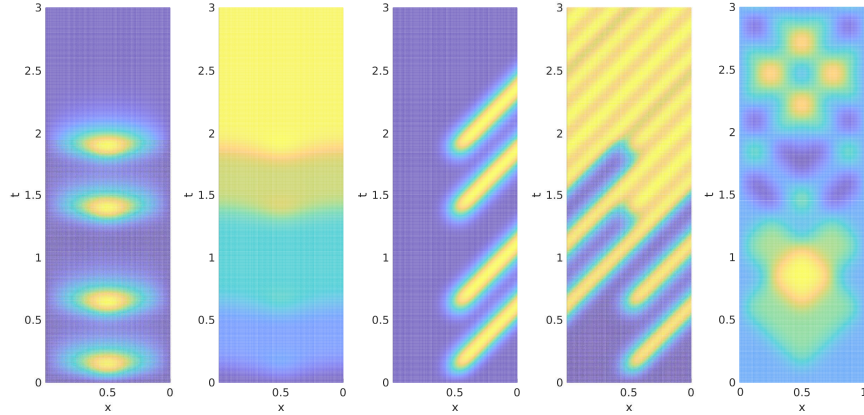


Fig. 3 From left to right: solution of the heat equation model problem with zero Dirichlet and zero Neumann boundary conditions. Solution of the corresponding advection equation with Dirichlet boundary condition on the left and periodic boundary condition, and solution of the second order wave equation with zero Dirichlet boundary conditions.

when iterating 6 times that leads to convergence, there is no contraction without the coarse propagator with Neumann boundary conditions.

It is very helpful at this point to take a closer look at the solutions we were computing in these two experiments for zero Dirichlet and zero Neumann conditions, shown in Figure 3 in the two leftmost panels. We see that the solution with Dirichlet boundary conditions is very much local in time: in order to compute the solution around $t = 2$, we only need to know if the right hand side function was on just before, but not at the much earlier times $t \ll 2$, since the solution does not contain this information any more. This is very different in the case of Neumann conditions, when using the same source term: now the solution at time $t = 2$ strongly depends on the fact that the right hand side function was on at earlier times, since the heat injected into the system is kept, as a constant value in space. In analogy, it is very easy to predict the temperature in the room you are currently in a year from now, you only need to know if the heater or air-conditioner is on and the windows and doors are open or closed then, but not what happened with the room over the entire year before.

3 Convergence Analysis

In order to study our observation mathematically, it is easiest to consider a Parareal algorithm which uses a spectral method in space and is performing exact integrations in time for our heat equation model problem (4). The fine propagator $F(U_n^k, T_n, T_{n+1})$ is thus using a highly accurate spectral method in space with m_F basis functions to solve (4), and the coarse propagator $G(U_n^k, T_n, T_{n+1})$ solves the same problem using

a very cheap spectral method in space using only $m_G \ll m_F$ basis functions. Note that we allow explicitly that $m_G = 0$, which means the coarse solver is not present in the Parareal algorithm (1)!

A spectral representation of the solution of the heat equation (4) can be obtained using separation of variables, and considering a spatial domain $\Omega := (0, \pi)$ to avoid to have to carry the constant π in all the computations (see also Remark 3 later), we get

$$u(x, t) := \sum_{m=1}^{\infty} \hat{u}_m(t) \sin(mx). \quad (6)$$

Expanding also the right hand side in a sine series,

$$f(x, t) := \sum_{m=1}^{\infty} \hat{f}_m(t) \sin(mx), \quad (7)$$

and the initial condition,

$$u_0(x) := \sum_{m=1}^{\infty} \hat{u}_{0,m}(t) \sin(mx), \quad (8)$$

we find that the Fourier coefficients in the solution satisfy

$$\partial_t \hat{u}_m = -m^2 \hat{u}_m + \hat{f}_m, \quad \hat{u}_m(0) = \hat{u}_{0,m}. \quad (9)$$

The solution of this equation can be written in closed form using an integrating factor,

$$\hat{u}_m(t) = \hat{u}_{0,m} e^{-m^2 t} + \int_0^t \hat{f}_m(\tau) e^{-m^2(t-\tau)} d\tau. \quad (10)$$

If we use in the Parareal algorithm (1) a spectral approximation of the solution for the fine propagator using m_F modes, we obtain

$$F(U_n^k, T_n, T_{n+1}) = \sum_{m=1}^{m_F} \left(\hat{U}_{m,n}^k e^{-m^2 \Delta T} + \int_{T_n}^{T_{n+1}} \hat{f}_m(\tau) e^{-m^2(t-\tau)} d\tau \right) \sin(mx), \quad (11)$$

and similarly for the coarse approximation when using m_G modes

$$G(U_n^k, T_n, T_{n+1}) = \sum_{m=1}^{m_G} \left(\hat{U}_{m,n}^k e^{-m^2 \Delta T} + \int_{T_n}^{T_{n+1}} \hat{f}_m(\tau) e^{-m^2(t-\tau)} d\tau \right) \sin(mx). \quad (12)$$

Because of the orthogonality of the sine functions, the Parareal algorithm (1) is diagonalized in this representation, which is the main reason that we chose a spectral method, to simplify the present analysis for this short note. For Fourier modes with $m \leq m_G$, the first and the third term in (1) coincide, as we see from (11) and (12), and thus cancel, and hence the Parareal update formula simplifies to

$$\hat{U}_{m,n+1}^{k+1} = \hat{U}_{m,n}^{k+1} e^{-m^2 \Delta T} + \int_{T_n}^{T_{n+1}} \hat{f}_m(\tau) e^{-m^2(t-\tau)} d\tau, \quad \text{for } m = 1, 2, \dots, m_G. \quad (13)$$

For $m > m_G$, the coarse correction term is not present in the Parareal update formula (1), and thus only the contribution from the fine propagator remains,

$$\hat{U}_{m,n+1}^{k+1} = \hat{U}_{m,n}^k e^{-m^2 \Delta T} + \int_{T_n}^{T_{n+1}} \hat{f}_m(\tau) e^{-m^2(t-\tau)} d\tau, \quad \text{for } m = m_G + 1, m_G + 2, \dots, m_F. \quad (14)$$

The only difference is the iteration index k in the first term on the right, but this term makes all the difference: for $m \leq m_G$, the update formula (13) represent the exact integration of the corresponding mode sequentially going through the entire time domain $(0, T)$. For $m > m_G$, the update formula (14) represents simply a block Jacobi iteration solving all time subintervals (T_n, T_{n+1}) in parallel starting from the current approximation at iteration k . Using this insight, we obtain the following convergence result.

Theorem 1 (Parareal convergence even without coarse propagator) *The Parareal algorithm (1) for the heat equation (4) using the spectral coarse propagator (12) and the spectral fine propagator (11) satisfies for any initial guess U_n^0 the convergence estimate*

$$\sup_n \|U_n^k(\cdot) - u(\cdot, T_n)\|_2 \leq e^{-(m_g+1)^2 k \Delta T} \sup_n \|U_n^0(\cdot) - u(\cdot, T_n)\|_2, \quad (15)$$

and this estimate also holds if the coarse propagator does not contain any modes, $m_G = 0$, which means it is not present. The Parareal algorithm therefore converges also without coarse propagator.

Proof. We introduce the error in Fourier space at the interfaces, $\hat{E}_{m,n}^k := \hat{U}_{m,n} - \hat{U}_{m,n}^k$, where the converged solution $\hat{U}_{m,n}$ satisfies

$$\hat{U}_{m,n+1} = \hat{U}_{m,n} e^{-m^2 \Delta T} + \int_{T_n}^{T_{n+1}} \hat{f}_m(\tau) e^{-m^2(t-\tau)} d\tau, \quad \text{for all } m = 1, 2, \dots \quad (16)$$

Taking the difference with the Parareal update formula for small m in (13) and using the fact that $U_{m,0}^k = U_{m,0} = u^0$, we thus obtain a vanishing error in the first Fourier coefficients after having performed one Parareal iteration,

$$\hat{E}_{m,n}^k = 0, \quad m = 1, 2, \dots, m_G, \quad k = 1, 2, \dots$$

For $m > m_G$, taking the difference between (16) and the update formula for larger m in (14), the part with the source term f cancels, and we are left with

$$\hat{E}_{m,n}^{k+1} = \hat{E}_{m,n}^k e^{-m^2 \Delta T}, \quad m = m_G + 1, m_G + 2, \dots$$

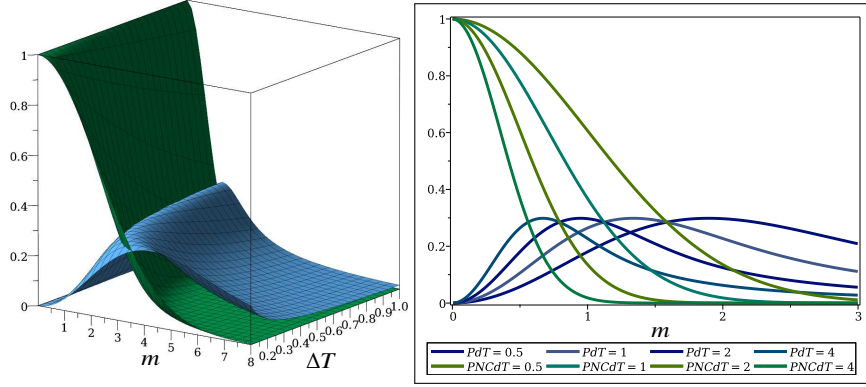


Fig. 4 Contraction for each Fourier mode m as function of the Parareal time interval length ΔT with (blue, $PdT=\Delta T$) and without (green, $PNCdT=\Delta T$) coarse propagator.

Now using the Parseval-Plancherel identity, and taking the largest convergence factor $e^{-(m_G+1)^2\Delta T}$ out of the sum, we obtain the convergence estimate (15) taking the sup over all time intervals n . \square

We show in Figure 4 a graphical comparison of the convergence factor of each Fourier mode without coarse propagator from (15), i.e. $e^{-m^2\Delta T}$, and with coarse propagator $\frac{|e^{-m^2\Delta T} - R_G(-m^2\Delta T)|}{1 - |R_G(-m^2\Delta T)|}$ from [8] for one Backward Euler step, $R_G(z) = \frac{1}{1-z}$, as we used for Figures 1 and 2. We see that indeed there are situations where the coarse correction with one Backward Euler step is detrimental, as observed in Figures 1 and 2, and there is also quantitative agreement.

Remark 1 Theorem 1 shows that the Parareal algorithm (1) for the heat equation is scalable, provided the time interval length ΔT is held constant: no matter how many such time intervals are solved simultaneously, the convergence rate remains the same, and this even in the case without coarse propagator! This is like the scalability of one level Schwarz methods discovered in [1] for molecular simulations in solvent models, and rigorously proved in [2, 3, 4] using three different techniques: Fourier analysis, maximum principle, and projection arguments in Hilbert spaces.

Remark 2 If we had zero Neumann boundary conditions, a similar result could also be derived using a cosine expansion, but then no coarse correction would not lead to a convergent method, since the zero mode would have the contraction factor one. One must therefore in this case have at least the constant mode in the coarse correction for the Parareal algorithm to converge. Otherwise however all remains the same.

Remark 3 A precise dependence on the domain size can be obtained by introducing the spatial domain of length L , $\Omega = (0, L)$. In this case the convergence factor simply becomes $e^{-\left(\frac{(m_G+1)\pi}{L}\right)^2\Delta T}$. The same estimate can also be obtained for a more general parabolic equation, the convergence factor then becomes $e^{-\lambda_{m_G+1}\Delta T}$, where λ_{m_G+1} is the m_G plus first eigenvalue of the corresponding spatial operator.

4 Conclusions

We have shown that Parareal without a coarse propagator can converge very well for the prototype parabolic problem of the heat equation, because such problems “forget” all fine information over time, and only very coarse information, if at all, remains, depending on the boundary conditions. This is very different for hyperbolic problems: while for the advection equation $u_t + au_x = f$ with Dirichlet boundary condition, whose solution is shown in Figure 3 (middle) the information is also “forgotten” by leaving the domain on the right and Parareal converges very well [5], also without coarse propagator, for the advection equation with periodic boundary conditions in the fourth plot of Figure 3 this is not the case any more: nothing is forgotten, and the coarse propagator needs to have fine accuracy quality for the Parareal algorithm not to fail. The situation is similar for the second order wave equation $u_{tt} = u_{xx}$, whose solution is shown on the right in Figure 3.

Acknowledgment: Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy EXC 2044 –390685587, Mathematics Münster: Dynamics–Geometry–Structure and the Swiss National Science Foundation.

References

1. Eric Cances, Yvon Maday, and Benjamin Stamm. Domain decomposition for implicit solvation models. *The Journal of chemical physics*, 139(5), 2013.
2. Gabriele Ciaramella and Martin J. Gander. Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part I. *SIAM Journal on Numerical Analysis*, 55(3):1330–1356, 2017.
3. Gabriele Ciaramella and Martin J. Gander. Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part II. *SIAM Journal on Numerical Analysis*, 56(3):1498–1524, 2018.
4. Gabriele Ciaramella and Martin J. Gander. Analysis of the parallel Schwarz method for growing chains of fixed-sized subdomains: Part III. *Electron. Trans. Numer. Anal.*, 49:201–243, 2018.
5. Martin J. Gander. Analysis of the parareal algorithm applied to hyperbolic problems using characteristics. *SeMA Journal: Boletín de la Sociedad Española de Matemática Aplicada*, (42):21–36, 2008.
6. Martin J. Gander and Ernst Hairer. Nonlinear convergence analysis for the parareal algorithm. In *Domain decomposition methods in science and engineering XVII*, pages 45–56. Springer, 2008.
7. Martin J. Gander and Ernst Hairer. Analysis for parareal algorithms applied to Hamiltonian differential equations. *Journal of Computational and Applied Mathematics*, 259:2–13, 2014.
8. Martin J. Gander and Stefan Vandewalle. Analysis of the parareal time-parallel time-integration method. *SIAM Journal on Scientific Computing*, 29(2):556–578, 2007.
9. Jacques-Louis Lions, Yvon Maday, and Gabriel Turinici. Résolution d’edp par un schéma en temps pararéel. *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics*, 332(7):661–668, 2001.
10. Toshiya Takami and Daiki Fukudome. An identity parareal method for temporal parallel computations. In *Parallel Processing and Applied Mathematics: 10th International Conference, PPAM 2013, Warsaw, Poland, September 8-11, 2013, Revised Selected Papers, Part I 10*, pages 67–75. Springer, 2014.