On the Convergence and the Stability of the Parareal Algorithm to Solve Partial Differential Equations

Guillaume Bal

Columbia University, APAM (`http://www.columbia.edu/~gb2030/`)

Summary. After stating an abstract convergence result for the parareal algorithm used in the parallelization in time of general partial differential equations, we analyze the stability and convergence properties of the algorithm for equations with constant coefficients. We show that suitably damping coarse schemes ensure unconditional stability of the parareal algorithm and analyze how the regularity of the initial condition influences convergence in the absence of sufficient damping.

1 Introduction

The parareal algorithm pioneered in Lions et al. [2000] and slightly modified in Bal and Maday [2002] allows us to speed up the numerical simulation of solutions to time dependent equations provided that we have enough processors. We refer to Baffico et al. [2002], Bal [2003], Farhat and Chandesris [2003], Maday and Turinici [2002] for additional detailed presentations of the method and applications; see also Section 2 below. Natural questions then arise related to the stability and convergence of the method. Here are some elements of answers to these questions. In Section 2 an abstract result in a general setting shows convergence of the algorithm provided that regularity conditions are satisfied. In the simplified setting of linear partial differential equations with constant coefficients, more refined estimates are provided for the convergence and stability of the parareal algorithm in Section 3. A typical result we can show is as follows: whereas the parareal algorithm is unconditionally stable for most discretizations of parabolic equations, it is not for hyperbolic equations.

2 An abstract convergence result

Let us consider a possibly non-linear partial differential equation of the form

$$\frac{du}{dt} = A(t, u), \qquad t > 0$$
$$u(0) = u_0, \tag{1}$$

where $u(t)$ takes values in a Banach space B and $A(t, u)$ is a possibly time-dependent partial differential operator. Let us assume that we want to simulate this equation on an interval of time $(0, T)$ and define the discretization $0 = T^0 < T^1 < \cdots < T^N = T$. We assume that the equation (1) admits a solution operator $g(T^n, v)$, which to $v \in B$ associates $u(T^{n+1}) \in B$ solution of

$$\frac{du}{dt} = A(t, u), \qquad T^n < t < T^{n+1}$$
$$u(T^n) = v. \tag{2}$$

Let us now assume that we have at our disposal an approximate solution operator $g_\Delta(T^n, v)$. We then define the approximate sequence

$$u_1^{n+1} = g_\Delta(T^n, u_1^n), \qquad n \in I = \{0 \le n \le N - 1\}, \tag{3}$$

and $u_1^0 = u_0$. Let ΔT be the maximal lag between successive time steps T^n and assume that $T^{n+1} - T^n \ge \eta_0 \Delta T$ for some positive constant η_0 for all $n \in I$. The convergence of u^N to $u(T^n)$ as ΔT converges to 0 is obtained under the following hypotheses. Let us assume that g_Δ is Lipschitz in a Banach space B_0 and is an approximation of order m of g in the sense that

$$\sup_{n \in I} \|g_\Delta(T^n, u) - g_\Delta(T^n, v)\|_{B_0} \le (1 + C\Delta T)\|u - v\|_{B_0}, \tag{4}$$

$$\sup_{n \in I} \|\delta g(T^n, u)\|_{B_0} \le C(\Delta T)^{m+1}\|u\|_{B_1}, \tag{5}$$

where $\delta g(T^n, u) = g(T^n, u) - g_\Delta(T^n, u)$, C is a constant independent of ΔT, $\|\cdot\|_B$ denotes a norm on the Banach space B and B_1 is another Banach space (usually a strict subset of B_0).

If in addition, (1) is stable in B_1, in the sense that $u(t) \in B_1$ uniformly in t provided that $u_0 \in B_1$, then we have the classical result

$$\|u(T^N) - u_1^N\|_{B_0} \le C(\Delta T)^m \|u_0\|_{B_1}, \tag{6}$$

based on the above regularity assumptions and the decomposition

$$u(T^{n+1}) - u_1^{n+1} = \delta g(T^n, u(T^n)) + g_\Delta(T^n, u(T^n)) - g_\Delta(T^n, u_1^n).$$

We now consider the *parareal algorithm*, which allows us to speed up the calculation of $u(t)$ provided that we have access to a sufficiently large number of processors. The parareal algorithm is given by

$$u_{k+1}^{n+1} = g_\Delta(T^n, u_{k+1}^n) + \delta g(T^n, u_k^n), \qquad n \in I, \quad k \ge 1 \tag{7}$$

with initial condition $u_{k+1}^0 = u_0$. The idea of the algorithm is to add to the prediction term $g_\Delta(T^n, u_{k+1}^n)$ a correction involving the previous iteration

u_k^n and a "fine" calculation that can be done in parallel on every time step (T^n, T^{n+1}) of the coarse discretization since all the terms u_k^n are known when the calculation of u_{k+1}^n starts. This requires to have N processors available. In practice, we cannot simulate $\delta g(T^n, u)$ exactly, but rather an approximation of the form $g_\delta(T^n, u) - g_\Delta(T^n, u)$, where $g_\delta(T^n, u)$ is a sufficiently accurate approximation of $g(T^n, u)$ so that in all the results that follow we can safely replace $g(T^n, u)$ by $g_\delta(T^n, u)$. See the bibliographical references for additional details.

We now show that the above algorithm ideally replaces the accuracy of order m of the non-parallel algorithm by an accuracy of order km to solve (1). Such an accuracy cannot be obtained in general solely under the above assumptions for g and g_Δ. It rather requires much stronger regularity constraints. Let us define a scale of Banach spaces B_0, B_1, ..., B_k, where in practice $B_k \subset B_{k-1} \subset \cdots B_1 \subset B_0$. We have then the following result:

(H1) The equation (1) is stable in all spaces B_j for $0 \leq j \leq k$, in the sense that $\|u(t)\|_{B_j} \leq C\|u_0\|_{B_j}$ where C is independent of u_0 and $t \in (0, T)$.

(H2) The operator g_Δ is Lipschitz in the sense that

$$\sup_{n \in I} \|g_\Delta(T^n, u) - g_\Delta(T^n, v)\|_{B_j} \leq (1 + C\Delta T)\|u - v\|_{B_j}, \quad 0 \leq j \leq k-1,$$

where C is independent of ΔT, u, and v.

(H3) The operator δg is an approximation of order m in the sense that

$$\sum_{n \in I} \|\delta g(T^n, u)\|_{B_j} \leq C(\Delta T)^{m+1}\|u\|_{B_{j+1}}, \quad 0 \leq j \leq k-1. \tag{8}$$

Theorem 1. *Under hypotheses (H1)-(H3), the order of accuracy of the parallel algorithm (7) is mk. More precisely, for $u_0 \in \cap_{0 \leq j \leq k} B_j$, we have*

$$\|u(T^N) - u_k^N\|_{B_0} \leq C(\Delta T)^{mk}\|u_0\|_{B_k},$$

where C is a constant independent of ΔT and u_0.

Proof. The result is obtained by induction. We know it to hold when $k = 1$ thanks to (H1)-(H3) and assume that it holds for k given. We then apply the result with the sequence of Banach spaces $B_1, \ldots, B_{k+1}$, so that

$$\|u(T^N) - u_k^N\|_{B_1} \leq C(\Delta T)^{mk}\|u_0\|_{B_{k+1}}.$$

By definition (7), we have

$$u(T^{n+1}) - u_{k+1}^{n+1} = g_\Delta(T^n, u(T^n)) - g_\Delta(T^n, u_{k+1}^n) + \delta g(T^n, u(T^n)) - \delta g(T^n, u_k^n).$$

Using (H1)-(H3), this implies that

$$\|u(T^{n+1}) - u_{k+1}^{n+1}\|_{B_0} \leq (1 + C\Delta T)\|u(T^n) - u_{k+1}^n\|_{B_0} +$$
$$C(\Delta T)^{m+1}\|u(T^n) - u_k^n\|_{B_1}$$
$$\leq (1 + C\Delta T)\|u(T^n) - u_{k+1}^n\|_{B_0} +$$
$$C(\Delta T)^{m(k+1)+1}\|u_0\|_{B_{k+1}}.$$

Since $u(T^0) = u_{k+1}^0 = u_0$, it is then a routine calculation to obtain (8). $\square$

3 Stability and convergence for linear operators

The above abstract result shows that the parareal algorithm allows us to replace a scheme of order m by a scheme of order km *provided that* regularity constraints are met. In practice we need to choose B_{k+1} as a strict subset of B_k in (H3); see below. These constraints may not be optimal as they do not account for possible dissipative effects of the coarse scheme.

To address this issue, we consider a pseudo-differential operator $P(D)$ with symbol $P(\xi)$ such that $\Re(P(\xi)) \geq 0$ (otherwise consider $P + \alpha$ with α sufficiently large and solve for $v = e^{-\alpha t}u$), and define $A(t, u) = P(D)u$. To simplify notation, we assume that $u(t)$ is a function on $\mathbb{R}$. In the Fourier domain, the evolution of $\hat{u}(t, \xi) = \int_{\mathbb{R}} e^{-ix\xi} u(t, x) dx$ is thus given by

$$\frac{\partial \hat{u}}{\partial t}(t, \xi) + P(\xi)\hat{u}(t, \xi) = 0 \qquad \xi \in \mathbb{R}, \, t > 0$$
$$\hat{u}(0, \xi) = \hat{u}_0(\xi), \qquad \xi \in \mathbb{R}. \tag{9}$$

The evolution operator is independent of time T^n $(= n\Delta T$ from now on to simplify) and is in the frequency domain

$$g(T^n, \hat{u}) = e^{-P(\xi)\Delta T}\hat{u}. \tag{10}$$

We define $\delta(\xi) = P(\xi)\Delta T$ and using the same notation g define the propagator

$$g(\delta(\xi)) = e^{-\delta(\xi)}. \tag{11}$$

We now want to define approximate solutions to the above equation. Let us assume that the symbol $P(\xi)$ is approximated by $P_H(\xi)$ and that the time propagator $g(\delta)$ is approximated by $g_\Delta(\delta_H)$, where $\delta_H(\xi) = P_H(\xi)\Delta T$. For instance $g_\Delta(\delta) = (1 + \delta)^{-1}$ for implicit first-order Euler. We then define the parareal scheme as

$$\hat{u}_{k+1}^{n+1}(\xi) = g_\Delta(\delta_H(\xi))\hat{u}_{k+1}^n(\xi) + (g(\delta(\xi)) - g_\Delta(\delta_H(\xi)))\hat{u}_k^n(\xi) \tag{12}$$

for $n \in I$ and $k \geq 0$. The boundary conditions are $\hat{u}_{k+1}^0(\xi) = \hat{u}_0(\xi)$ and $\hat{u}_0^n(\xi) \equiv 0$. In the above equations, ξ is a parameter so stability and error of convergence can be analyzed for each frequency separately. We verify that the exact solution $\hat{u}^n(\xi)$ also satisfies (12) with different initial conditions so that the error term $\varepsilon_k^n(\xi) = \hat{u}^n(\xi) - \hat{u}_k^n(\xi)$ satisfies the following equation

$$\varepsilon_{k+1}^{n+1}(\xi) = g_\Delta(\delta_H(\xi))\varepsilon_{k+1}^n(\xi) + (g(\delta(\xi)) - g_\Delta(\delta_H(\xi)))\varepsilon_k^n(\xi) \tag{13}$$

with boundary conditions $\varepsilon_{k+1}^0(\xi) = 0$ and $\varepsilon_0^n(\xi) = \hat{u}^n(\xi)$. We verify that $\varepsilon_1^n(\xi) = (g^n(\delta(\xi)) - g^n(\delta_H(\xi))\hat{u}_0(\xi)$. Upon defining

$$\theta_k^n = \frac{1}{g_\Delta^{n-k+1}(\delta_H)} \frac{\varepsilon_k^n}{(g(\delta) - g_\Delta(\delta_H))^{k-1}}, \qquad \theta_1^n = \frac{(g^n(\delta) - g_\Delta^n(\delta_H))}{g_\Delta^n(\delta_H)}\hat{u}_0(\xi),$$

we find that $\theta_{k+1}^{n+1} = \theta_{k+1}^n + \theta_k^n$ and $\theta_{k+1}^0 = 0$. The constraint $\Re(P(\xi)) \geq 0$ implies that $g^n(\delta)$ is uniformly bounded and we assume that $g_\Delta^n(\delta_H)$ is uniformly bounded for $n \in I$ (this is nothing but stability of the coarse scheme). We also assume that $|(g(\delta)/g_\Delta(\delta_H))(\xi)|$ is uniformly bounded. This allows us to obtain the following bound

$$|\theta_1^n(\xi)| \leq C(n|g(\delta) - g_\Delta(\delta_H)| \wedge 1)|\hat{u}_0(\xi)|. \tag{14}$$

Here $a \wedge b = \min(a, b)$. This implies then the following bound on the error

$$|\varepsilon_{k+1}^{n+1}|(\xi) \leq C\left(n|g(\delta) - g_\Delta(\delta_H)| \wedge 1\right)|g_\Delta(\delta_H)|^{n-k}|g(\delta) - g_\Delta(\delta_H)|^k \binom{n}{k}, \tag{15}$$

for $n \in I$ and $k \geq 1$, where C is proportional to $|\hat{u}_0(\xi)|$ only. This estimate gives us optimal bounds to prove convergence and stability of the parareal scheme. For $k + 1 = 1$ in the above formula we recover that the coarse scheme is stable. We can also obtain the maximal order of convergence of the scheme after k parareal iterations. Assuming that $\delta = \delta_H$ and that g_Δ is of order m so that $|(g - g_\Delta)(\delta)| \leq C(\xi)(\Delta T)^{m+1}$, we deduce that $|\varepsilon_k^N|$ is of order $N(\Delta T)^{m+1}[(\Delta T)^{m+1}]^{k-1}N^{k-1} = (\Delta T)^{km}$. We recover that the parareal algorithm replaces an algorithm of order m by an algorithm of order km.

A central difficulty with the iterative scheme is that for $k \geq 2$, the error term $\varepsilon_k^n(\xi)$, hence the solution $\hat{u}_k^n$, may blow up for large frequencies. We now consider this stability analysis.

The above parallel algorithm requires to solve a fine scale problem $k - 1$ times to obtain an accuracy of order km on the coarse time grid (i.e. at the times T^n, $n \in N$). The best use of the available processors is thus obtained for $k = 2$; see Bal [2003]. In any case the algorithm is useful when the value of k is small. We therefore assume from now on that $1 \sim k \ll n$. This implies that $\binom{n}{k} \sim n^k$.

Let us assume here that $\delta_H = \delta$, i.e., the spatial discretization is the same for the coarse and fine schemes. Stability at all frequencies is thus ensured provided that

$$R_{k+1,n}(\delta) = \left(|g_\Delta|^{n-k}|g - g_\Delta|^k\right)(\delta)n^k, \tag{16}$$

remains bounded for all values of $\delta(\xi) = P(\xi)\Delta T$, $\xi \in \mathbb{R}$. The first term $|g_\Delta|^{n-k}$ is clearly bounded since the coarse algorithm is stable. The second term $|g - g_\Delta|^k$ is also bounded. However it may not be small for values of δ of order 1. The relation (16) indicates how the parareal algorithm blows up. Unless high frequencies (δ of order $O(1)$ or higher) are damped by the coarse scheme $|g_\Delta|$, an instability of size n^k will appear at the iteration step $k + 1$.

Consider a real-valued $P(\xi) > 0$ for $|\xi| > 0$ and the centered scheme

$$g_\Delta(\delta) = \frac{1 - \delta/2}{1 + \delta/2} = e^{-\delta} + O(\delta^3). \tag{17}$$

As $\delta \to \infty$, $|g_\Delta(\delta)| \to 1$ and $g(\delta) \to 0$. We thus observe that high frequencies will grow like n^{k-1} in (16) and the parareal scheme is unstable as soon as

$k \geq 2$ although the scheme is unconditionally stable for $k = 1$. However we can still apply the general theory. For instance for $P(\xi) \leq |\xi|^M$ and a coarse scheme of order m we can verify that all the hypotheses of Theorem 1 are verified provided that $B_k = H^{(m+1)Mk}(\mathbb{R})$. The instability of the parareal scheme can thus be overcome by assuming sufficient regularity of the initial conditions.

The growth n^{k-1} may be compensated when the coarse scheme is dissipative. A result that covers many classical examples is the following:

Theorem 2. *Let us assume that the coarse scheme is an approximation of order m of the exact propagator and that it is dissipative in the sense that there exist three constants $C, \gamma > 0$ and $1 \leq \beta \leq m+1$ such that for all $\xi \in \mathbb{R}$,*

$$|g(\delta(\xi)) - g_\Delta(\delta(\xi))| \leq C(\delta(\xi)^{m+1} \wedge 1) \qquad (18)$$

$$|g_\Delta(\delta(\xi))| \leq (1 + C\Delta T)e^{-\gamma(|\delta(\xi)|^\beta \wedge 1)}. \qquad (19)$$

Then the parallel algorithm is stable in the sense that $R_{k,n}(\delta)$ is bounded uniformly in $k = O(1)$, $n \in N$, and $\delta = \delta(\xi)$ for $\xi \in \mathbb{R}$.

Proof. Consider the case $|\delta| \geq 1$ first. We observe that $|R_{k+1,n}(\delta)|$ is bounded by $e^{-\gamma n} n^k$ which is clearly bounded independent of $k = O(1)$, n, and $|\delta| \geq 1$.

For $|\delta| \leq 1$ we obtain that $|R_{k+1,n}(\delta)|$ is bounded by $|\delta|^{(m+1)k} n^k e^{-n\gamma|\delta|^\beta}$. Upon differentiating the above majorizing function with respect to $|\delta|$ we obtain that the maximum is reached for $|\delta_0|^\beta = (k(m+1))/(\gamma\beta n)$, so that a bound for $|R_{k+1,n}(\delta)|$ is given by

$$|R_{k+1,n}(\delta)| \leq Ce^{-k(m+1)/\beta}\left(n^{1-(m+1)/\beta}\right)^k.$$

The latter power of n does not grow as $n \to \infty$ provided that $\beta \leq m + 1$. □

The above result shows that sufficient exponential damping of the large frequencies is sufficient to ensure stability. Notice that the centered scheme defined in (17) does not verify the hypotheses of the theorem since large values of δ are not damped at all by the coarse scheme. For real valued non-negative symbols $P(\xi)$, we can use Theorem 2 to deduce that the θ scheme

$$g_\Delta(\delta) = \frac{1 - (1 - \theta)\delta}{1 + \theta\delta}, \qquad (20)$$

makes the parareal algorithm stable as soon as $\theta > 1/2$. Indeed we have then $|\delta(\xi)| = \delta(\xi)$. Since $g'_\Delta(0) = -1$, we verify that $\beta = 1$ and γ sufficiently small (all the more that $\theta \to 1/2+$) does the job. This covers then all parabolic equations (such as the Laplancian $P(\xi) = \xi^2$) and many spatial discretizations (such as the centered finite difference scheme $P(\xi) = 2h^{-2}[1 - \cos(h\xi)]$).

The result also applies to more general equations with complex-valued symbol. Consider the same θ scheme given in (20). We define $\delta = \delta_r + i\delta_i$. The assumption on $P(\xi)$ implies that $\delta_r \geq 0$. We now find that

$$|g_\Delta(\delta)| = \sqrt{\frac{(1 + (1 - \theta)^2|\delta|^2 - 2(1 - \theta)\delta_r)}{(1 + \theta^2|\delta|^2 + 2\theta\delta_r)}} \leq \sqrt{\frac{(1 + (1 - \theta)^2|\delta|^2)}{1 + \theta^2|\delta|^2}}. \qquad (21)$$

Asymptotically as $\delta \to 0$ we find for the θ scheme that $|g_\Delta(\delta)| \leq 1 - \frac{\theta^2 - (1-\theta)^2}{2}|\delta|^2 + O(|\delta|^3)$. We see that we have to choose here $\beta = 2$ to verify the assumptions of Theorem 2 (instead of $\beta = 1$ when δ is real-valued). Since $m = 1$, $\beta = 2$ is the only value allowed. It is then easy to find a value of γ such that the hypotheses of Theorem 2 are satisfied. The theorem thus addresses the case of the transport equation $P(\xi) = ia\xi$ and of the advection diffusion equation $P(\xi) = 1 + ia\xi + b\xi^2$, $b > 0$.

In the latter case we can actually do better. Indeed consider $P(\xi) = \alpha_0 + \sum_{k=1}^{M-1} \alpha_k \xi^k + |\xi|^M$ with $\alpha_k \in \mathbb{C}$ arbitrary and $\alpha_0 > 0$ such that $\Re(P(\xi)) > 0$. Then there exists a constant $\rho > 0$ such that $|\delta_i| \leq \rho\delta_r$, whence $|\delta| \leq \sqrt{1 + \rho^2}\delta_r$. We then deduce from (21) that, on $(0, \delta_0)$ for $\delta_0 = 2((1 - \theta)\sqrt{1 + \rho^2})^{-1}$, we have $|g_\Delta(\delta)| \leq (1 + \frac{\theta}{\sqrt{1+\rho^2}}|\delta|)^{-1}$, and still from (21), that $|g_\Delta(\delta)| < 1 - \varepsilon$ for $\varepsilon > 0$ on (δ_0, ∞) as $\theta > 1/2$. So we can choose $\beta = 1$ (this is important in Theorem 3 below) and γ small enough in Theorem 2.

Let us now turn to convergence. We have seen that the error term at final time T is bounded by $|R_{k,N}(\xi)| \leq C|\delta|^{(m+1)k}N^k e^{-N\gamma|\delta|^\beta}$ for $|\delta| < 1$. Let us assume that $\beta = 1$. We then deduce from the above analysis that $|R_{k,N}(\xi)| \leq CN^{-km}$. So all frequencies are uniformly bounded by CN^{-km} and the algorithm has an accuracy of order $(\Delta T)^{km}$ in $\mathcal{L}(H^\alpha(\mathbb{R}))$ for all $\alpha \in \mathbb{R}$. The case $\beta > 1$ is much less favorable. Let us assume that $|P(\xi)| \leq |\xi|^M$ for $M > 0$. The bound for R is then $|R_{k,N}(\xi)| \leq C(\Delta T)^{km}[|\xi|^{M(m+1)k}e^{-\gamma|\xi|^{M\beta}(\Delta T)^{\beta-1}}]$. This implies that there is no damping for all frequencies of order up to $|\xi| \sim (\Delta T)^{(1-\beta)/(M\beta)} \gg 1$. So the L^2 norm (for instance) of the error term is bounded by

$$\int_{|\xi| \leq (\Delta T)^{(1-\beta)/(M\beta)}} |\hat{u}_0(\xi)|^2 |\xi|^{2M(m+1)k}d\xi \leq C\|u_0\|_{M(m+1)k}^2,$$

where $\|\cdot\|_\alpha$ is the norm in the Hilbert space $H^\alpha(\mathbb{R})$. We have thus proved the following result:

Theorem 3. *Under the assumptions of Theorem 2 we have the following convergence result for all $\alpha \in \mathbb{R}$. When $\beta = 1$, we have*

$$\|u^N - u_k^N\|_\alpha \leq C(\Delta T)^{km}\|u_0\|_\alpha.$$

When $\beta > 1$, we have for $0 \leq \tau \leq 1$,

$$\|u^N - u_k^N\|_\alpha \leq C(\Delta T)^{km\tau}\|u_0\|_{\alpha + \tau M(m+1)k}.$$

The latter estimate follows from stability by interpolation. At $\tau = 1$, this is nothing but Theorem 1. However the result for $\tau < 1$ requires stability.

For parabolic equations ($P(\xi) > 0$ real-valued), we thus obtain that the θ scheme for $\theta > 1/2$ has a very strong convergence property as the error is of order $(\Delta T)^{km}$ in the space where u_0 is defined. This generalizes the results obtained in Bal and Maday [2002] for $\theta = 1$. The same result holds for symbols of the form $P(\xi) = |\xi|^M + lower\ order\ terms$ with $\Re(P(\xi)) > 0$ since we can then choose $\beta = 1$. However for the transport equation $P(\xi) = ia\xi$ (or symbols with purely imaginary leading term) we see that convergence of implicit Euler ($m = M = 1$, $\beta = 2$) is of order $(\Delta T)^{k\tau}\|u_0\|_{2k\tau}$ in $L^2(\mathbb{R})$.

Let us conclude by a remark. In the above analysis we have assumed that $\delta = \delta_H$, i.e. the spatial discretization is the same for the coarse and the fine steps. This need not be so. For lack of space, we postpone the general analysis to future work and only mention the result where $P(\xi) = \xi^2$, $P_H(\xi) = 2(1 - \cos(H\xi))/H^2$, and the implicit Euler scheme $g_\Delta(\delta_H) = (1 + \delta_H)^{-1}$. We can then show that $\|u^N - u_2^N\|_\alpha \leq C(\Delta T)^2\|u_0\|_{\alpha+4}$ so that the optimal accuracy $(\Delta T)^2$ is attained for $k = 2$ for a coarse spatial discretization $H = (\Delta T)^{1/2}$. The loss of "4" derivatives comes from the fact that the coarse scheme damps frequencies up to H^{-1} only. This is an intermediate result between Theorems 1 and 3. It should be compared with the case $\delta = \delta_H$ where the spatial discretization need be chosen as $h = \Delta T$. So $H \gg h$ for the same final accuracy (but it requires more regularity of the initial condition).

References

L. Baffico, S. Bernard, T. Maday, G. Turinici, and G. Zérah. Parallel-in-time molecular-dynamics simulations. *Phys. Rev. E*, 66:057701, 2002.

G. Bal. Parallelization in time of (stochastic) ordinary differential equations. *Preprint;* www.columbia.edu/~gb2030/PAPERS/ParTimeSODE.ps, 2003.

G. Bal and Y. Maday. A "parareal" time discretization for non-linear PDE's with application to the pricing of an american put. *Recent developments in domain decomposition methods (Zürich, 2001), Lect. Notes Comput. Sci. Eng., Springer, Berlin*, 23:189–202, 2002.

C. Farhat and M. Chandesris. Time-decomposed parallel time-integrators: theory and feasibility studies for fluid, structure, and fluid-structure applications. *Int. J. Numer. Meth. Engng.*, 58(9):1397–1434, 2003.

J.-L. Lions, Y. Maday, and G. Turinici. Résolution d'EDP par un schéma en temps "pararéel". *C.R.A.S. Sér. I Math.*, 332(7):661–668, 2000.

Y. Maday and G. Turinici. A parareal in time procedure for the control of partial differential equations. *C.R.A.S. Sér. I Math.*, 335:387–391, 2002.