Convergence Bounds: Differences and Similarities

Ausra Pogozelskyte

University of Geneva

PinT 2022, 11–15 July 2022

Setting and the Parareal algorithm

Consider the ODE (coming from the method-of-lines),

$$\frac{\partial u(t)}{\partial t} = L\,u(t) + f(t), \quad t \in (0, T], \quad u(0) = u_0 \,.$$

After time-discretization, we obtain the time-stepping procedure

$$u_{n+1} = \Phi\, u_n + f_n,\; n = 0, \ldots, n_t \,.$$

For example, for Forward Euler,

$$\Phi = 1 + \Delta t\, L \,.$$

Setting and the Parareal algorithm

Consider the ODE (coming from the method-of-lines),

$$\frac{\partial u(t)}{\partial t} = L\, u(t) + f(t), \quad t \in (0, T], \quad u(0) = u_0 \ .$$

After time-discretization, we obtain the time-stepping procedure

$$u_{n+1} = \Phi\, u_n + f_n,\ n = 0, \ldots, n_t \ .$$

The Parareal algorithm is given by the iterations

$$\begin{cases} U_0^k = u_0 & k = 0, \ldots, K \\ U_{n+1}^0 = G\, U_n^0 + f_n & n = 0, \ldots, n_t \\ U_{n+1}^{k+1} = F\, U_n^k + G\, U_n^{k+1} - G\, U_n^k + f_n & k = 0, \ldots, K, n = 0, \ldots, n_t \end{cases}$$

where the coarse operator G is a cheap approximation of the fine operator $F = \Phi^m$.

Parareal error

From the iteration

$$U_{n+1}^{k+1} = F\,U_n^k + G\,U_n^{k+1} - G\,U_n^k + f_n, \quad n = 1, \ldots, n_t \ ,$$

the error, $\epsilon_n^k := u_n - U_n^k$, can be computed as

$$\epsilon_{n+1}^{k+1} = F\,\epsilon_n^k + G\,\epsilon_n^{k+1} - G\,\epsilon_n^k \ .$$

Parareal error

From the iteration

$$U_{n+1}^{k+1} = F\, U_n^k + G\, U_n^{k+1} - G\, U_n^k + f_n, \quad n = 1, \ldots, n_t \ ,$$

the error, $\epsilon_n^k := u_n - U_n^k$, can be computed as

$$\epsilon_{n+1}^{k+1} = F\, \epsilon_n^k + G\, \epsilon_n^{k+1} - G\, \epsilon_n^k \ .$$

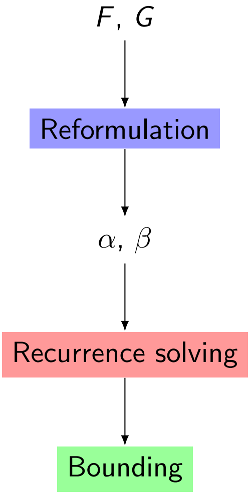
In turn, it can be bounded as

$$\|\epsilon_{n+1}^{k+1}\| =: e_{n+1}^{k+1} \leq \underbrace{\|F - G\|}_{=\alpha} e_n^k + \underbrace{\|G\|}_{=\beta} e_n^{k+1} \ .$$

We thus only need to solve the iteration

$$e_{n+1}^{k+1} = \alpha\, e_n^k + \beta\, e_n^{k+1} \ .$$

General structure



$$\epsilon_{n+1}^{k+1} = (F - G)\,\epsilon_n^k + G\,\epsilon_n^{k+1}$$

Bounding

$$e_{n+1}^{k+1} = \alpha e_n^k + \beta e_n^{k+1}$$

The iteration $e_{n+1}^k = \alpha\, e_n^{k-1} + \beta\, e_n^k$ can be written in matrix form as

$$\begin{pmatrix} I & & & \\ -\beta & I & & \\ & \ddots & \ddots & \\ & & -\beta & I \end{pmatrix} \begin{pmatrix} e_0^k \\ e_1^k \\ \vdots \\ e_{n_t}^k \end{pmatrix} = \begin{pmatrix} 0 & & & \\ \alpha & 0 & & \\ & \ddots & \ddots & \\ & & \alpha & 0 \end{pmatrix} \begin{pmatrix} e_0^{k-1} \\ e_1^{k-1} \\ \vdots \\ e_{n_t}^{k-1} \end{pmatrix} .$$

Lemma (Recurrence solving)

Assuming that α and β are scalars, the error at step k is given by

$$\boldsymbol{e}^k = M(\beta)\,(I_{n_t} \otimes \alpha)\,\boldsymbol{e}^{k-1} = \ldots = M(\beta)^k\,(I_{n_t} \otimes \alpha^k)\,\boldsymbol{e}^0 \ ,$$

where the Toeplitz matrix M is defined as

$$M(\beta) = \begin{pmatrix} 0 & & & & \\ I & 0 & & & \\ \beta & I & 0 & & \\ \vdots & \ddots & \ddots & \ddots & \\ \beta^{nt-1} & \cdots & \beta & I & 0 \end{pmatrix} .$$

General structure

F, G — $$\epsilon_{n+1}^{k+1} = (F - G)\,\epsilon_n^k + G\,\epsilon_n^{k+1}$$

↓

Reformulation — Bounding

↓

α, β — $$e_{n+1}^{k+1} = \alpha e_n^k + \beta e_n^{k+1}$$

↓

Recurrence solving — $$\boldsymbol{e}^k = M(\beta)^k\,(I_{n_t} \otimes \alpha^k)\,\boldsymbol{e}^0$$

↓

Bounding

Bounding

We then want to bound

$$\boldsymbol{e}^k = M(\beta)^k \left(I_{n_t} \otimes \alpha^k\right) \boldsymbol{e}^0 \ .$$

Linear Bound [Lemma 4.4, Gander, Vandewalle, 2007]

$$\|M(\beta)^k\|_\infty \leq \|M(\beta)\|_\infty^k = \left(\frac{1 - |\beta|^{n_t}}{1 - |\beta|}\right)^k \ .$$

Superlinear Bound [Lemma 4.3, Gander, Vandewalle, 2007]

$$\begin{aligned}\|M(\beta)^k\|_\infty &= \sum_{i=0}^{n_t - k} \binom{i + k - 1}{k - 1} |\beta|^i \\ &= \frac{1}{(k-1)!} \sum_{i=0}^{n_t - k} \left[\prod_{l=1}^{k-1} (i + l)\right] |\beta|^i\end{aligned}$$

Bounding

We then want to bound

$$\boldsymbol{e}^k = M(\beta)^k \left(I_{n_t} \otimes \alpha^k \right) \boldsymbol{e}^0 \ .$$

Linear Bound [Lemma 4.4, Gander, Vandewalle, 2007]

$$\|M(\beta)^k\|_\infty \leq \|M(\beta)\|_\infty^k = \left(\frac{1 - |\beta|^{n_t}}{1 - |\beta|} \right)^k \ .$$

Explicit Superlinear Bound [Lemma 4.4, Gander, Vandewalle, 2007]

If $|\beta| < 1$, then

$$\|M(\beta)^k\|_\infty \leq \binom{n_t}{k} = \frac{1}{k!} \prod_{l=0}^{k-1} (n_t - l) \ .$$

M. J. Gander, S. Vandewalle. "Analysis of the Parareal Time-Parallel Time-Integration Method", 2007.

General structure

F, G — $\epsilon_{n+1}^{k+1} = (F - G)\,\epsilon_n^k + G\,\epsilon_n^{k+1}$

↓

Reformulation — Bounding

↓

α, β — $e_{n+1}^{k+1} = \alpha e_n^k + \beta e_n^{k+1}$

↓

Recurrence solving — $\boldsymbol{e}^k = M(\beta)^k\,(I_{n_t} \otimes \alpha^k)\,\boldsymbol{e}^0$

↓

Bounding — Linear, (explicit) superlinear bounds for $\|M(\beta)^k\|$

Formulation for eigenvalues

We said that $\alpha = \|F - G\|$ and $\beta = \|G\|$, but they do not have to be.

Assume that F and G are simultaneously diagonalizable, then there exists a matrix U such that $F = U \Lambda U^{-1}$ and $G = UHU^{-1}$,
where $\Lambda = \mathrm{diag}(\lambda_1^m, \dots, \lambda_{n_x}^m)$ and $H = \mathrm{diag}(\mu_1, \dots, \mu_{n_x})$.

$$\epsilon_{n+1}^{k+1} = (F - G)\, \epsilon_n^k + G\, \epsilon_n^{k+1}$$

becomes the decoupled system,

$$\tilde{\epsilon}_{n+1}^{k+1} = (\Lambda - H)\, \tilde{\epsilon}_n^k + H\, \tilde{\epsilon}_n^{k+1} \; .$$

We are left with bounding equations

$$\tilde{\epsilon}_{n+1}^{k+1}(\omega) = \underbrace{(\lambda_\omega^m - \mu_\omega)}_{=\alpha}\, \tilde{\epsilon}_n^k(\omega) + \underbrace{\mu_\omega}_{=\beta}\, \tilde{\epsilon}_n^{k+1}(\omega) \; .$$

General structure

F, G — $\epsilon_{n+1}^{k+1} = (F - G)\,\epsilon_n^k + G\,\epsilon_n^{k+1}$

↓

Reformulation — Bounding, Simultaneous diagonalization

↓

α, β — $e_{n+1}^{k+1} = \alpha e_n^k + \beta e_n^{k+1}$

↓

Recurrence solving — $\boldsymbol{e}^k = M(\beta)^k\,(I_{n_t} \otimes \alpha^k)\,\boldsymbol{e}^0$

↓

Bounding — Linear, (explicit) superlinear bounds for $\|M(\beta)^k\|$

Parareal and MGRIT bounds

Recall: $\boldsymbol{e}^k = M(\beta)^k \left(I_{n_t} \otimes \alpha^k\right) \boldsymbol{e}^0$.

$\rightarrow$ Do these bounds correspond to some we already know?

- If $\alpha = \|F - G\|$ and $\beta = \|G\|$, ("norm bounds")
  - LINEAR: Gander, Vandewalle (2007).
  - EXPLICIT SUPERLINEAR: Gander, Vandewalle (2007), Gander, Hairer (2008).
  - SUPERLINEAR: Gander, Lunet, Ruprecht, de Steck (2022).
- If $\alpha = |\lambda_\omega^m - \mu_\omega|$ and $\beta = |\mu_\omega|$, ("eigenvalue bounds")
  - LINEAR: Dobrev, Kolev, Petersson, Schroder (2017).
  - (EXPLICIT) SUPERLINEAR: do not exist yet.

Questions:

$\rightarrow$ How do those bounds compare to each other?

$\rightarrow$ What are the advantages of each of these bounds?

Example: Dahlquist equation $\partial_t u = (-1 + i)\, u$

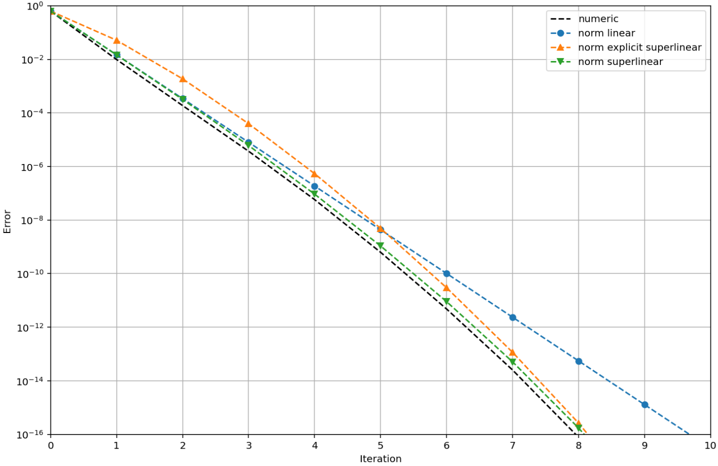


Figure: Norm bounds for the Dahlquist equation for $T = 5$.

Example: Dahlquist equation $\partial_t u = (-1 + i)\, u$

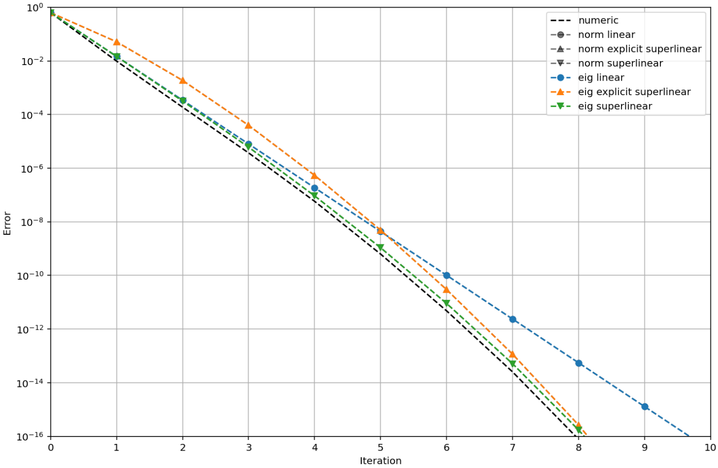
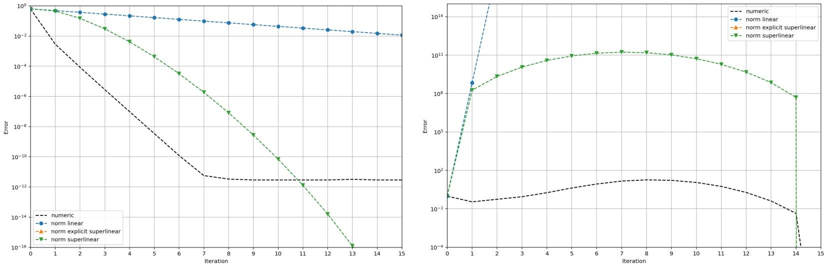


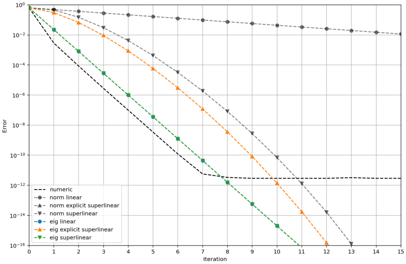
Figure: Eigenvalue bounds for the Dahlquist equation for $T = 5$.

Example: heat and advection equations

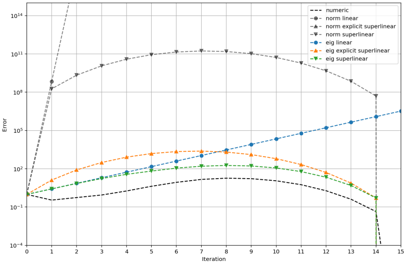


(a) Heat (b) Advection

Example: heat and advection equations



(a) Heat



(b) Advection

Generalization to Multigrid Reduction in Time

For Multigrid Reduction in Time (MGRIT) with *FCF*-relaxation, the error propagator is given by

$$E = B_{\Delta}^{-1}(B_{\Delta} - A_{\Delta})(I - A_{\Delta}) \, ,$$

that is,

$$E = \Big(\sum_{j=0}^{n_t-1} U_{-j-1} \otimes \Psi^j\Big) \, \big(I \otimes (\Phi^m - \Psi)\big) \, \big(U_{-1} \otimes \Phi^m\big) \, .$$

In our notation,

$$E = M(G)(I \otimes (F - G))(U_{-1} \otimes F) \, .$$

Generalization to Multigrid Reduction in Time

For Multigrid Reduction in Time (MGRIT) with $F(CF)^{\nu}$-relaxation, the error propagator is given by

$$E = B_{\Delta}^{-1}(B_{\Delta} - A_{\Delta})(I - A_{\Delta})^{\nu} ,$$

that is,

$$E = (\sum_{j=0}^{n_t-1} U_{-j-1} \otimes \Psi^{j}) \; (I \otimes (\Phi^{m} - \Psi)) \; (U_{-\nu} \otimes \Phi^{m\nu}) \; .$$

In our notation,

$$E = M(G)(I \otimes (F - G))(U_{-\nu} \otimes F^{\nu}) \; .$$

Conclusion

Summary:

- General structure for Parareal and MGRIT bounds.
- The new superlinear bound captures convergence the best.
- Noticed that:
  - Eigenvalue bounds are more accurate.
  - Norm bounds are general.

Future Work:

- Accurate convergence bounds for spatial coarsening.
- Generalization of this bounding strategy to other smoothers.
- Convergence bounds for variable time-step sizes.

References

V. A. Dobrev et al. "Two-Level Convergence Theory for Multigrid Reduction in Time (MGRIT)". In: *SIAM J. Sci. Comput.* 39.5 (Jan. 2017). Publisher: Society for Industrial and Applied Mathematics, S501–S527.

M. J. Gander and E. Hairer. "Nonlinear convergence analysis for the parareal algorithm". In: *Domain decomposition methods in science and engineering XVII*. Springer, 2008, pp. 45–56.

M. J. Gander and S. Vandewalle. "Analysis of the Parareal Time-Parallel Time-Integration Method". In: *SIAM J. Sci. Comput.* 29.2 (Jan. 2007), pp. 556–578.

M. J. Gander et al. *A unified analysis framework for iterative parallel-in-time algorithms*. en. Number: arXiv:2203.16069 arXiv:2203.16069 [cs, math]. Mar. 2022.

Why eigenvalue bounds are better than norm bounds?

If we express norm bound in terms of eigenvalues, the linear bound is given by

$$\|E_\omega\|_\infty \leq \sup_\omega \frac{1-|\mu_\omega|^{n_t}}{1-|\mu_\omega|} \sup_\omega |\lambda_\omega^m - \mu_\omega| \ ,$$

whereas the linear eigenvalue bound is given by

$$\|E_\omega\|_\infty \leq \sup_\omega \left\{ \frac{1-|\mu_\omega|^{n_t}}{1-|\mu_\omega|} |\lambda_\omega^m - \mu_\omega| \right\} \ .$$

Bounds with coarse initialization

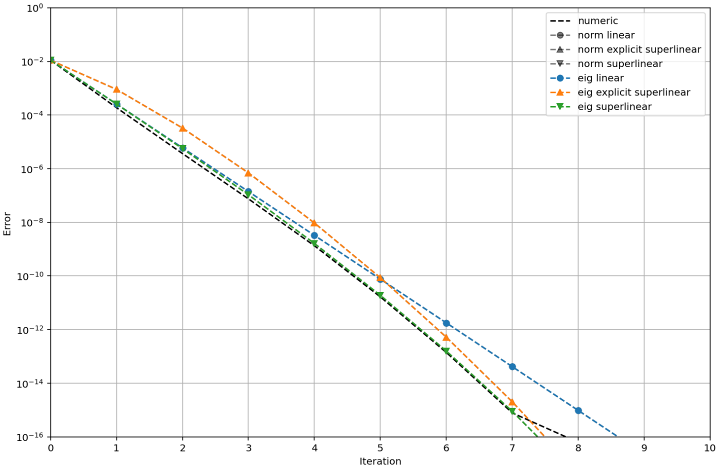


Figure: Bounds with coarse initialization for the Dahlquist equation $\partial_t u = (-1 + i)\, u$ with the same parameters as on slide 12.