Low-rank techniques for integrating large Sylvester-like equations with Parareal

Benjamin Carrel
Joint work with Martin J. Gander and Bart Vandereycken

UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES

Section de Mathématiques

PinT 2022, Marseille

Schedule

Introduction: Setup

Consider the Sylvester differential equation

$$\begin{aligned}\dot{X}(t) &= AX(t) + X(t)B + C, \quad t \in [0,T] \\ X(0) &= X_0 \in \mathbb{R}^{m\times n}\end{aligned}$$

for symmetric matrices $A \in \mathbb{R}^{m\times m}$, $B \in \mathbb{R}^{n\times n}$ and a low-rank matrix $C \in \mathbb{R}^{m\times n}$.

Example: if $A = B$ is 1D Laplacian, it is a model for 2D heat propagation.

We'll focus on stiff problems.

Suppose the solution admits a good low-rank approximation,

$$X(t) \approx Y(t) \in \mathcal{M}_r = \{A \in \mathbb{R}^{m\times n} \mid \operatorname{rank}(A) = r\}, \quad \text{for all } t \in [0,T].$$

Goal: for every t, find $Y(t) \in \mathcal{M}_r$ such that $\|X(t) - Y(t)\| \overset{!}{=} min$.

Introduction: Setup

Consider the Sylvester differential equation

$$\begin{aligned}\dot{X}(t) &= AX(t) + X(t)B + C, \quad t \in [0, T] \\ X(0) &= X_0 \in \mathbb{R}^{m \times n}\end{aligned}$$

for symmetric matrices $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{n \times n}$ and a low-rank matrix $C \in \mathbb{R}^{m \times n}$.

Example: if $A = B$ is 1D Laplacian, it is a model for 2D heat propagation.

We'll focus on stiff problems.

Suppose the solution admits a good low-rank approximation,

$$X(t) \approx Y(t) \in \mathcal{M}_r = \{A \in \mathbb{R}^{m \times n} \mid \text{rank}(A) = r\}, \quad \text{for all } t \in [0, T].$$

Goal: for every t, find $Y(t) \in \mathcal{M}_r$ such that $\|X(t) - Y(t)\| \overset{!}{=} min$.

Introduction: Setup

Consider the Sylvester differential equation

$$\begin{aligned} \dot{X}(t) &= AX(t) + X(t)B + C, \quad t \in [0, T] \\ X(0) &= X_0 \in \mathbb{R}^{m \times n} \end{aligned}$$

for symmetric matrices $A \in \mathbb{R}^{m \times m}$, $B \in \mathbb{R}^{n \times n}$ and a low-rank matrix $C \in \mathbb{R}^{m \times n}$.

Example: if $A = B$ is 1D Laplacian, it is a model for 2D heat propagation.

We'll focus on stiff problems.

Suppose the solution admits a good low-rank approximation,

$$X(t) \approx Y(t) \in \mathcal{M}_r = \{A \in \mathbb{R}^{m \times n} \mid \operatorname{rank}(A) = r\}, \quad \text{for all } t \in [0, T].$$

Goal: for every t, find $Y(t) \in \mathcal{M}_r$ such that $\|X(t) - Y(t)\| \overset{!}{=} min$.

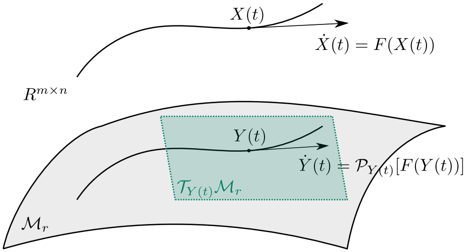
Introduction: DLRA

Original problem

$$\dot{X}(t) = F(X(t)), \quad t \in [0,T]$$
$$X(0) = X_0 \in \mathbb{R}^{m\times n}.$$

Dynamical low-rank approximation ([Koch and Lubich, 2007])

$$\dot{Y}(t) = \mathcal{P}_{Y(t)}\left[F(Y(t))\right], \quad t \in [0,T]$$
$$Y(0) = Y_0 \in \mathcal{M}_r.$$



Question: Can we solve efficiently the DLRA in parallel in time?

Introduction: Dynamical low-rank approximation

Standard DLRA assumptions:

- F is Lipschitz: $\|F(X) - F(Y)\| \leq L \|X - Y\|$.
- F is one-sided Lipschitz[1]: $\langle X - Y, F(X) - F(Y) \rangle \leq \ell \|X - Y\|^2$.
- F maps to a tangent bundle of $\mathcal{M}_r$: $\|F(Y) - \mathcal{P}_Y F(Y)\| \leq \varepsilon$.

Theorem (Accuracy of DLRA [Koch and Lubich, 2007])

Under the three assumptions above, the error made by DLRA verifies

$$\left\| \psi_r^h(Y_0) - \phi^h(X_0) \right\| \leq \underbrace{e^{lt} \|Y_0 - X_0\|}_{\text{initial error}} + \underbrace{\varepsilon \int_0^t e^{\ell s} ds}_{\text{modeling error}},$$

where ψ_r^h is the flow of the DLRA, and ϕ^h is the flow of the original problem.

[1] If F is linear, ℓ is the largest eigenvalue, potentially negative.

Introduction: Dynamical low-rank approximation

Standard DLRA assumptions:

- F is Lipschitz: $\|F(X) - F(Y)\| \leq L \|X - Y\|$.
- F is one-sided Lipschitz[1]: $\langle X - Y, F(X) - F(Y) \rangle \leq \ell \|X - Y\|^2$.
- F maps to a tangent bundle of $\mathcal{M}_r$: $\|F(Y) - \mathcal{P}_Y F(Y)\| \leq \varepsilon$.

Theorem (Accuracy of DLRA [Koch and Lubich, 2007])

Under the three assumptions above, the error made by DLRA verifies

$$\left\| \psi_r^h(Y_0) - \phi^h(X_0) \right\| \leq \underbrace{e^{lt} \|Y_0 - X_0\|}_{\textit{initial error}} + \underbrace{\varepsilon \int_0^t e^{\ell s} ds}_{\textit{modeling error}} ,$$

where ψ_r^h is the flow of the DLRA, and ϕ^h is the flow of the original problem.

[1] If F is linear, ℓ is the largest eigenvalue, potentially negative.

Part 1: Parallel-in-time integration of DLRA

Low-rank Parareal: Definition

Definition (Low-rank Parareal [Carrel et al.,])

Choose a *coarse rank* q and a *fine rank* r such that $q < r$. The low-rank Parareal algorithm iterates

$$
\begin{aligned}
&\text{(Initial value)} && Y_0^k = Y_0, \\
&\text{(Initial approximation)} && Y_{n+1}^0 = \psi_q^h \circ \mathcal{T}_q(Y_n^0), \\
&\text{(Iteration)} && Y_{n+1}^{k+1} = \psi_r^h \circ \mathcal{T}_r(Y_n^k) + \psi_q^h \circ \mathcal{T}_q(Y_n^{k+1}) - \psi_q^h \circ \mathcal{T}_q(Y_n^k),
\end{aligned}
$$

where ψ_r^h is the solution of the DLRA of rank r at time h, and $\mathcal{T}_r$ is the orthogonal projection onto $\mathcal{M}_r$. The notations ψ_q^h and $\mathcal{T}_q$ are similar but apply to rank q.

Remark: The rank of each iteration is at most $r + 2q$.

Low-rank Parareal: Analysis

The error verifies the double recursion

$$\left\|E_{n+1}^{k+1}\right\| \leq \alpha \left\|E_n^k\right\| + \beta \left\|E_n^{k+1}\right\| + \kappa, \quad \left\|E_n^0\right\| \leq \gamma, \quad n, k \geq 0,$$

with the positive constants

- $\alpha = e^{\ell h} C_{r,q}$ and $\beta = e^{\ell h} C_q$ where C_q (resp. $C_{r,q}$) stands for the Lipschitz constant of $\mathcal{T}_q$ (resp. $\mathcal{T}_r - \mathcal{T}_q$). By [Feppon and Lermusiaux, 2018],

$$C_q \lessapprox \frac{1}{1 - \frac{\sigma_{q+1}}{\sigma_q}} \approx \frac{1}{1 - e^{-c}} \text{ when, for some } c > 0, \sigma_k \approx e^{-ck}.$$

- $\gamma = \max_{n \geq 0} \left\|E_n^0\right\|$,
- $\kappa = e^{\ell h} \max_{n \geq 0} \|X_n - \mathcal{T}_r(X_n)\| + (2\varepsilon_q + \varepsilon_r) \int_0^h e^{\ell(h-s)} ds.$

Low-rank Parareal: Analysis

The error verifies the double recursion

$$\left\|E_{n+1}^{k+1}\right\| \leq \alpha \left\|E_n^k\right\| + \beta \left\|E_n^{k+1}\right\| + \kappa, \quad \left\|E_n^0\right\| \leq \gamma, \quad n, k \geq 0,$$

with the positive constants

- $\alpha = e^{\ell h} C_{r,q}$ and $\beta = e^{\ell h} C_q$ where C_q (resp. $C_{r,q}$) stands for the Lipschitz constant of $\mathcal{T}_q$ (resp. $\mathcal{T}_r - \mathcal{T}_q$). By [Feppon and Lermusiaux, 2018],

$$C_q \lessapprox \frac{1}{1 - \frac{\sigma_{q+1}}{\sigma_q}} \approx \frac{1}{1 - e^{-c}} \text{ when, for some } c > 0, \sigma_k \approx e^{-ck}.$$

- $\gamma = \max_{n \geq 0} \left\|E_n^0\right\|$,
- $\kappa = e^{\ell h} \max_{n \geq 0} \|X_n - \mathcal{T}_r(X_n)\| + (2\varepsilon_q + \varepsilon_r) \int_0^h e^{\ell(h-s)} ds.$

Low-rank Parareal: Analysis

The error verifies the double recursion

$$\left\|E_{n+1}^{k+1}\right\| \leq \alpha \left\|E_n^k\right\| + \beta \left\|E_n^{k+1}\right\| + \kappa, \quad \left\|E_n^0\right\| \leq \gamma, \quad n, k \geq 0,$$

with the positive constants

- $\alpha = e^{\ell h} C_{r,q}$ and $\beta = e^{\ell h} C_q$ where C_q (resp. $C_{r,q}$) stands for the Lipschitz constant of $\mathcal{T}_q$ (resp. $\mathcal{T}_r - \mathcal{T}_q$). By [Feppon and Lermusiaux, 2018],

$$C_q \lessapprox \frac{1}{1 - \frac{\sigma_{q+1}}{\sigma_q}} \approx \frac{1}{1 - e^{-c}} \text{ when, for some } c > 0, \sigma_k \approx e^{-ck}.$$

- $\gamma = \max_{n \geq 0} \left\|E_n^0\right\|$,
- $\kappa = e^{\ell h} \max_{n \geq 0} \|X_n - \mathcal{T}_r(X_n)\| + (2\varepsilon_q + \varepsilon_r) \int_0^h e^{\ell(h-s)} ds.$

Low-rank Parareal: Analysis

Theorem (Convergence of low-rank Parareal [Carrel et al.,])

Under the standard DLRA assumptions, and if $\alpha + \beta < 1$, the error satisfies

1st linear bound:
$$\max_{n\geq 0} \left\| E_n^k \right\| \leq \left(\frac{\alpha}{1-\beta} \right)^k \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta},$$

2nd linear bound:
$$\left\| E_n^k \right\| \leq \alpha^k (1+\beta)^{n-1} \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta},$$

Superlinear bound:
$$\left\| E_n^k \right\| \leq \frac{\alpha^k}{(k-1)!} \frac{\prod_{j=2}^{k}(n-j)}{1-\beta} \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta}.$$

Low-rank Parareal: Analysis

Theorem (Convergence of low-rank Parareal [Carrel et al.,])

Under the standard DLRA assumptions, and if $\alpha + \beta < 1$, the error satisfies

1st linear bound:
$$\max_{n\geq 0} \left\| E_n^k \right\| \leq \left(\frac{\alpha}{1-\beta} \right)^k \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta},$$

2nd linear bound:
$$\left\| E_n^k \right\| \leq \alpha^k (1+\beta)^{n-1} \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta},$$

Superlinear bound:
$$\left\| E_n^k \right\| \leq \frac{\alpha^k}{(k-1)!} \frac{\prod_{j=2}^{k}(n-j)}{1-\beta} \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta}.$$

Low-rank Parareal: Analysis

Theorem (Convergence of low-rank Parareal [Carrel et al.,])

Under the standard DLRA assumptions, and if $\alpha + \beta < 1$*, the error satisfies*

1st linear bound:
$$\max_{n\geq 0} \left\| E_n^k \right\| \leq \left(\frac{\alpha}{1-\beta} \right)^k \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta},$$

2nd linear bound:
$$\left\| E_n^k \right\| \leq \alpha^k (1+\beta)^{n-1} \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta},$$

Superlinear bound:
$$\left\| E_n^k \right\| \leq \frac{\alpha^k}{(k-1)!} \frac{\prod_{j=2}^{k}(n-j)}{1-\beta} \max_{n\geq 0} \left\| E_n^0 \right\| + \frac{\kappa}{1-\alpha-\beta}.$$

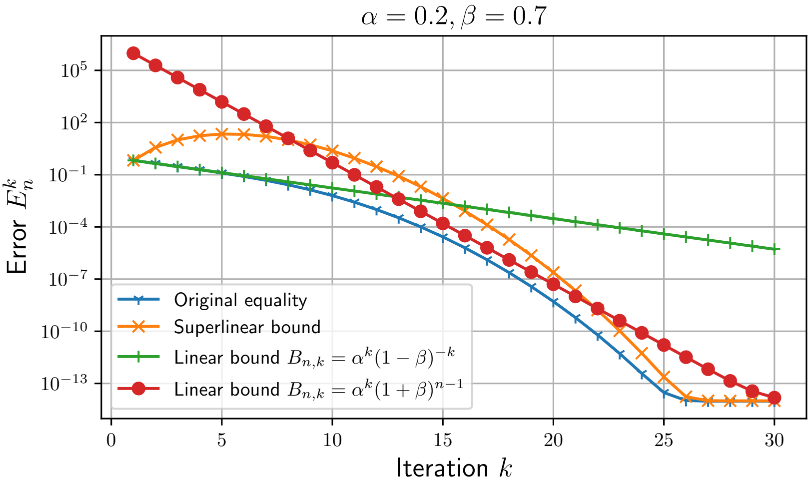


Figure: Comparison of the bounds

Low-rank Parareal: numerical results

Lyapunov equation

$$\dot{X}(t) = AX(t) + X(t)A^T + CC^T,$$

where

- $t \in [0, T]$
- $X(0) = X_0$ is low-rank.
- $A \in \mathbb{R}^{n \times n}$ is sparse.
- $C \in \mathbb{R}^{n \times \ell}$ is a tall matrix.

If A is the 1D discrete Laplacian, this equation models a 2D heat problem.

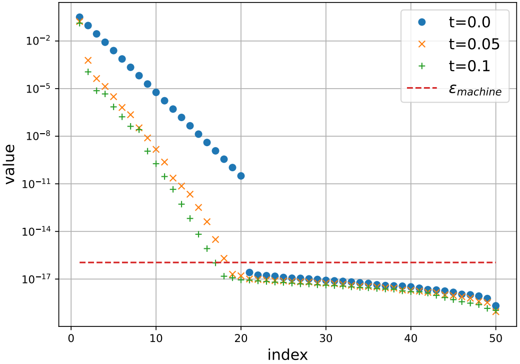
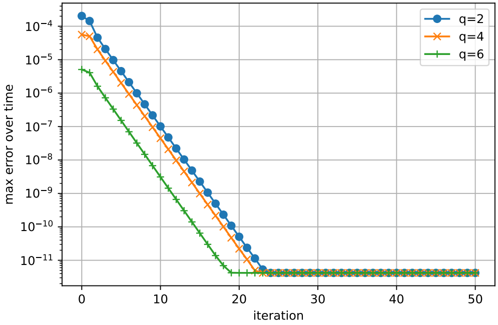
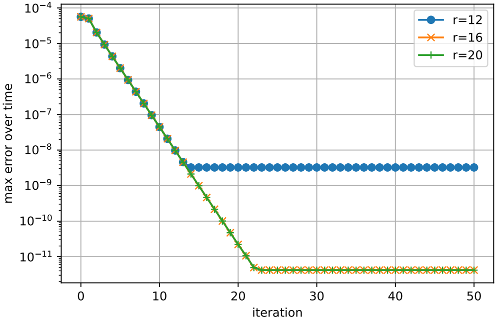


Figure: Singular values of the reference solution.

Low-rank Parareal: numerical results



(a) Several coarse ranks q with fine rank $r = 16$.



(b) Several fine ranks r with coarse rank $q = 4$.

Figure: Convergence of the error of low-rank Parareal for the Lyapunov ODE with $n = 100$ and $T = 2.0$. Influence of the coarse and fine ranks.

Low-rank Parareal: conclusion and challenges

Conclusion:

- Low-rank Parareal is the first parallel-in-time integrator for DLRA
- Linear and superlinear bounds
- Good behavior on the heat equation

Links:

- Preprint on arXiv: https://doi.org/10.48550/arXiv.2203.08455
- Code on GitHub: https://github.com/BenjaminCarrel/Low-rank-Parareal

Question:

How DLRA is solved?
Current DLRA solvers are not robust to stiffness!

Part 2: A new robust-to-stiffness DLRA integrator

Exponential integrators

Consider a Sylvester-like differential equation

$$\dot{X}(t) = \mathcal{L}_S(X(t)) + \mathcal{G}(X(t)), \quad t \in [0, T]$$
$$X(0) = X_0 \in \mathbb{R}^{m \times n},$$

where $\mathcal{L}_S(X) = AX + XB$ and $\mathcal{G}$ is a non-linear, non-stiff operator.
The closed form solution is

$$X(t) = e^{t\mathcal{L}_S}(X_0) + \int_0^t e^{(t-s)\mathcal{L}_S}(\mathcal{G}(X(s)))ds.$$

The exponential Euler scheme is

$$X_1 = e^{h\mathcal{L}_S}(X_0) + h\varphi_1(h\mathcal{L}_S)(\mathcal{G}(X_0)),$$

where $h\varphi_1(h\mathcal{L}_S) = \mathcal{L}_S^{-1}(e^{h\mathcal{L}_S} - Id)$. See [Hochbruck and Ostermann, 2010].

Exponential integrators

Consider a Sylvester-like differential equation

$$\dot{X}(t) = \mathcal{L}_S(X(t)) + \mathcal{G}(X(t)), \quad t \in [0, T]$$
$$X(0) = X_0 \in \mathbb{R}^{m \times n},$$

where $\mathcal{L}_S(X) = AX + XB$ and $\mathcal{G}$ is a non-linear, non-stiff operator.
The closed form solution is

$$X(t) = e^{t\mathcal{L}_S}(X_0) + \int_0^t e^{(t-s)\mathcal{L}_S}(\mathcal{G}(X(s)))ds.$$

The exponential Euler scheme is

$$X_1 = e^{h\mathcal{L}_S}(X_0) + h\varphi_1(h\mathcal{L}_S)(\mathcal{G}(X_0)),$$

where $h\varphi_1(h\mathcal{L}_S) = \mathcal{L}_S^{-1}(e^{h\mathcal{L}_S} - Id)$. See [Hochbruck and Ostermann, 2010].

Exponential integrators

Consider a Sylvester-like differential equation

$$\dot{X}(t) = \mathcal{L}_S(X(t)) + \mathcal{G}(X(t)), \quad t \in [0, T]$$
$$X(0) = X_0 \in \mathbb{R}^{m \times n},$$

where $\mathcal{L}_S(X) = AX + XB$ and $\mathcal{G}$ is a non-linear, non-stiff operator.
The closed form solution is

$$X(t) = e^{t\mathcal{L}_S}(X_0) + \int_0^t e^{(t-s)\mathcal{L}_S}(\mathcal{G}(X(s)))ds.$$

The exponential Euler scheme is

$$X_1 = e^{h\mathcal{L}_S}(X_0) + h\varphi_1(h\mathcal{L}_S)(\mathcal{G}(X_0)),$$

where $h\varphi_1(h\mathcal{L}_S) = \mathcal{L}_S^{-1}(e^{h\mathcal{L}_S} - Id)$. See [Hochbruck and Ostermann, 2010].

Projected exponential Euler scheme

By definition, the DLRA formulation of the problem is

$$\dot{Y}(t) = \mathcal{P}_{Y(t)}\left[\mathcal{L}_S(Y(t)) + \mathcal{G}(Y(t))\right], \quad t \in [0,T]$$
$$Y(0) = Y_0 \in \mathcal{M}_r.$$

Since $AY + YB \in \mathcal{T}_Y\mathcal{M}_r$ for any matrices A, B, it is equivalent to

$$\dot{Y}(t) = \mathcal{L}_S(Y(t)) + \mathcal{P}_{Y(t)}\left[\mathcal{G}(Y(t))\right], \quad t \in [0,T]$$
$$Y(0) = Y_0 \in \mathcal{M}_r.$$

Definition (Projected exponential Euler [$C., Vandereycken$])

For a given stepsize h, the projected exponential Euler scheme is defined by

$$Y_1 = \mathcal{T}_r\left(e^{h\mathcal{L}_S}(Y_0) + h\varphi_1(h\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right]\right).$$

Projected exponential Euler scheme

By definition, the DLRA formulation of the problem is

$$\dot{Y}(t) = \mathcal{P}_{Y(t)}\left[\mathcal{L}_S(Y(t)) + \mathcal{G}(Y(t))\right], \quad t \in [0,T]$$
$$Y(0) = Y_0 \in \mathcal{M}_r.$$

Since $AY + YB \in \mathcal{T}_Y\mathcal{M}_r$ for any matrices A, B, it is equivalent to

$$\dot{Y}(t) = \mathcal{L}_S(Y(t)) + \mathcal{P}_{Y(t)}\left[\mathcal{G}(Y(t))\right], \quad t \in [0,T]$$
$$Y(0) = Y_0 \in \mathcal{M}_r.$$

Definition (Projected exponential Euler [$C., Vandereycken$])

For a given stepsize h, the projected exponential Euler scheme is defined by

$$Y_1 = \mathcal{T}_r\left(e^{h\mathcal{L}_S}(Y_0) + h\varphi_1(h\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right]\right).$$

Projected exponential Euler scheme

By definition, the DLRA formulation of the problem is

$$\begin{aligned} \dot{Y}(t) &= \mathcal{P}_{Y(t)}\left[\mathcal{L}_S(Y(t)) + \mathcal{G}(Y(t))\right], \quad t \in [0,T] \\ Y(0) &= Y_0 \in \mathcal{M}_r. \end{aligned}$$

Since $AY + YB \in \mathcal{T}_Y\mathcal{M}_r$ for any matrices A, B, it is equivalent to

$$\begin{aligned} \dot{Y}(t) &= \mathcal{L}_S(Y(t)) + \mathcal{P}_{Y(t)}\left[\mathcal{G}(Y(t))\right], \quad t \in [0,T] \\ Y(0) &= Y_0 \in \mathcal{M}_r. \end{aligned}$$

Definition (Projected exponential Euler [$C., Vandereycken$])

For a given stepsize h, the projected exponential Euler scheme is defined by

$$Y_1 = \mathcal{T}_r\left(e^{h\mathcal{L}_S}(Y_0) + h\varphi_1(h\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right]\right).$$

Proj. exp. Euler: Lucky Krylov approximation

Question: Can we apply the scheme efficiently?

Projected exponential Euler: $$Y_1 = \mathcal{T}_r\left(e^{h\mathcal{L}_S}(Y_0) + h\varphi_1(h\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right]\right)$$

Let us write the inner term differently,

$$\tilde{Y}(t) = e^{t\mathcal{L}_S}(Y_0) + t\varphi_1(t\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] \iff \begin{cases}\dot{\tilde{Y}}(t) = A\tilde{Y}(t) + \tilde{Y}(t)B + \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right]\\ \tilde{Y}(0) = Y_0\end{cases}$$

We are back to a Sylvester differential equation. Interesting because ...

Y_0 and $\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right]$ are low-rank $\implies Y_0 = U_1\Sigma V_1{}^T, \quad \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] = [U_1, U_2]\tilde{\Sigma}[V_1, V_2]^T$

New idea: Use two rational Krylov[2] spaces to approximate the solution.

[2] See [Güttel, 2013] for overview.

Proj. exp. Euler: Lucky Krylov approximation

Question: Can we apply the scheme efficiently?

Projected exponential Euler: $$Y_1 = \mathcal{T}_r \left(e^{h\mathcal{L}_S}(Y_0) + h\varphi_1(h\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] \right)$$

Let us write the inner term differently,

$$\tilde{Y}(t) = e^{t\mathcal{L}_S}(Y_0) + t\varphi_1(t\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] \iff \begin{cases} \dot{\tilde{Y}}(t) = A\tilde{Y}(t) + \tilde{Y}(t)B + \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] \\ \tilde{Y}(0) = Y_0 \end{cases}$$

We are back to a Sylvester differential equation. Interesting because ...

Y_0 and $\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right]$ are low-rank $\implies Y_0 = U_1 \Sigma {V_1}^T, \quad \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] = [U_1, U_2]\tilde{\Sigma}[V_1, V_2]^T$

New idea: Use two rational Krylov[2] spaces to approximate the solution.

[2] See [Güttel, 2013] for overview.

Proj. exp. Euler: Lucky Krylov approximation

Question: Can we apply the scheme efficiently?

$$\text{Projected exponential Euler:} \qquad Y_1 = \mathcal{T}_r \left(e^{h\mathcal{L}_S}(Y_0) + h\varphi_1(h\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] \right)$$

Let us write the inner term differently,

$$\tilde{Y}(t) = e^{t\mathcal{L}_S}(Y_0) + t\varphi_1(t\mathcal{L}_S)\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] \iff \begin{cases} \dot{\tilde{Y}}(t) = A\tilde{Y}(t) + \tilde{Y}(t)B + \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] \\ \tilde{Y}(0) = Y_0 \end{cases}$$

We are back to a Sylvester differential equation. Interesting because ...

Y_0 and $\mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right]$ are low-rank $\implies Y_0 = U_1 \Sigma V_1{}^T, \quad \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] = [U_1, U_2]\tilde{\Sigma}[V_1, V_2]^T$

New idea: Use two rational Krylov[2] spaces to approximate the solution.

[2] See [Güttel, 2013] for overview.

Proj. exp. Euler: Lucky Krylov approximation

Left rational Krylov space:

$$\mathcal{RK}_k(A, U = [U_1, U_2]) = \operatorname{span}\left\{U, (A - \eta_2 I)^{-1} U, \ldots, \prod_{i=2}^{k} (A - \eta_i I)^{-1} U\right\}$$

Right rational Krylov space:

$$\mathcal{RK}_k(B, V = [V_1, V_2]) = \operatorname{span}\left\{V, (B - \xi_2 I)^{-1} V, \ldots, \prod_{i=2}^{k} (B - \xi_i I)^{-1} V\right\}$$

Reduced differential equation (via Galerkin projection):

$$\begin{cases} \dot{\tilde{Y}}_k(t) = V_k^T A V_k \tilde{Y}_k(t) + \tilde{Y}_k(t) W_k^T B W_k + V_k^T \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] W_k \\ \tilde{Y}_k(0) = V_k^T Y_0 W_k \end{cases}$$

Final solution: $\tilde{Y}(t) \approx V_k \tilde{Y}_k(t) W_k^T$ where $\tilde{Y}_k(t) \in \mathbb{R}^{2kr \times 2kr}$

Proj. exp. Euler: Lucky Krylov approximation

Left rational Krylov space:

$$\mathcal{RK}_k(A, U = [U_1, U_2]) = \operatorname{span}\left\{ U, (A - \eta_2 I)^{-1} U, \ldots, \prod_{i=2}^{k} (A - \eta_i I)^{-1} U \right\}$$

Right rational Krylov space:

$$\mathcal{RK}_k(B, V = [V_1, V_2]) = \operatorname{span}\left\{ V, (B - \xi_2 I)^{-1} V, \ldots, \prod_{i=2}^{k} (B - \xi_i I)^{-1} V \right\}$$

Reduced differential equation (via Galerkin projection):

$$\begin{cases} \tilde{Y}_k(t) = V_k^T A V_k \tilde{Y}_k(t) + \tilde{Y}_k(t) W_k^T B W_k + V_k^T \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] W_k \\ \tilde{Y}_k(0) = V_k^T Y_0 W_k \end{cases}$$

Final solution: $\tilde{Y}(t) \approx V_k \tilde{Y}_k(t) W_k^T$ where $\tilde{Y}_k(t) \in \mathbb{R}^{2kr \times 2kr}$

Application to the Riccati differential equation

Riccati differential equation: $\dot{X} = AX + XA^T + C^TC - XSX, \quad X(0) = X_0.$

Our setup[3]:

- $S = I \in \mathbb{R}^{100\times 100}$
- $A \in \mathbb{R}^{100\times 100}$ is the spatial discretization of the diffusion operator

$$\mathcal{D} = \partial_x(\alpha(x)\partial_x(\cdot)) - \lambda I, \quad \Omega = [0,1], \quad \alpha(x) = 2 + 2\cos(2\pi x), \quad \lambda = 1.$$

- Tall matrix $C \in \mathbb{R}^{100\times 9}$ obtained from 9 independent vectors.
- Time interval: $[0,T] = [0, 0.1]$.
- Rank of approximation: $r = 20$.

Let us compare our method to an other DLRA integrator that is also robust-to-stiffness, proposed in [Ostermann et al., 2019].

[3] Code available on https://github.com/BenjaminCarrel/Low-rank-Parareal

Application to the Riccati differential equation

Riccati differential equation: $\dot{X} = AX + XA^T + C^T C - XSX, \quad X(0) = X_0.$

Our setup[3]:

- $S = I \in \mathbb{R}^{100 \times 100}$
- $A \in \mathbb{R}^{100 \times 100}$ is the spatial discretization of the diffusion operator

$$\mathcal{D} = \partial_x(\alpha(x)\partial_x(\cdot)) - \lambda I, \quad \Omega = [0, 1], \quad \alpha(x) = 2 + 2\cos(2\pi x), \quad \lambda = 1.$$

- Tall matrix $C \in \mathbb{R}^{100 \times 9}$ obtained from 9 independent vectors.
- Time interval: $[0, T] = [0, 0.1]$.
- Rank of approximation: $r = 20$.

Let us compare our method to an other DLRA integrator that is also robust-to-stiffness, proposed in [Ostermann et al., 2019].

[3] Code available on https://github.com/BenjaminCarrel/Low-rank-Parareal

Application to the Riccati differential equation

Riccati differential equation: $\dot{X} = AX + XA^T + C^T C - XSX, \quad X(0) = X_0.$

Our setup[3]:

- $S = I \in \mathbb{R}^{100\times 100}$
- $A \in \mathbb{R}^{100\times 100}$ is the spatial discretization of the diffusion operator

$$\mathcal{D} = \partial_x(\alpha(x)\partial_x(\cdot)) - \lambda I, \quad \Omega = [0,1], \quad \alpha(x) = 2 + 2\cos(2\pi x), \quad \lambda = 1.$$

- Tall matrix $C \in \mathbb{R}^{100\times 9}$ obtained from 9 independent vectors.
- Time interval: $[0,T] = [0, 0.1]$.
- Rank of approximation: $r = 20$.

Let us compare our method to an other DLRA integrator that is also robust-to-stiffness, proposed in [Ostermann et al., 2019].

[3] Code available on https://github.com/BenjaminCarrel/Low-rank-Parareal

Application to the Riccati differential equation

Riccati differential equation: $\dot{X} = AX + XA^T + C^TC - XSX, \quad X(0) = X_0.$

Our setup[3]:

- $S = I \in \mathbb{R}^{100\times100}$
- $A \in \mathbb{R}^{100\times100}$ is the spatial discretization of the diffusion operator

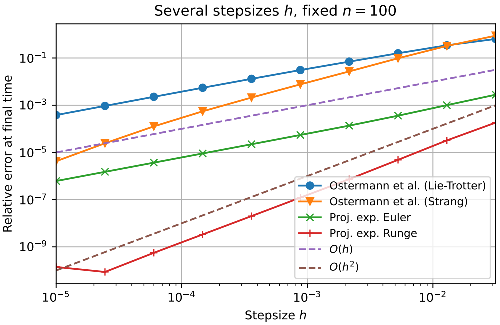
$$\mathcal{D} = \partial_x(\alpha(x)\partial_x(\cdot)) - \lambda I, \quad \Omega = [0,1], \quad \alpha(x) = 2 + 2\cos(2\pi x), \quad \lambda = 1.$$

- Tall matrix $C \in \mathbb{R}^{100\times9}$ obtained from 9 independent vectors.
- Time interval: $[0,T] = [0,0.1]$.
- Rank of approximation: $r = 20$.

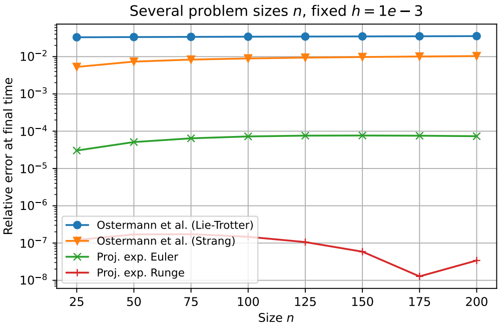
Let us compare our method to an other DLRA integrator that is also robust-to-stiffness, proposed in [Ostermann et al., 2019].

[3] Code available on https://github.com/BenjaminCarrel/Low-rank-Parareal

Projected exponential methods: Numerical results



(a) Global error comparison



(b) Robust-to-stiffness property

Figure: PERK 1 and 2 applied to the Riccati equation and comparison with Ostermann et al.

Final conclusion and outlook

Conclusion:

- DLRA $\rightarrow$ recent topic for large-scale problems integration.
- Low-rank Parareal $\rightarrow$ a new parallel-in-time integrator of DLRA using a coarse rank q and a fine rank r.
- Projected exponential methods $\rightarrow$ a new DLRA integrator robust-to-stiffness.

Future works:

- Extend low-rank Parareal to the tensor setting.[4]
- Prove further convergence bounds for the projected exponential methods.
- Apply Schwarz waveform relaxation to DLRA.

Thank to the organizers and
Thank you for your attention!

[4] Project with G. Ceruti (EPFL).

Final conclusion and outlook

Conclusion:

- DLRA $\rightarrow$ recent topic for large-scale problems integration.
- Low-rank Parareal $\rightarrow$ a new parallel-in-time integrator of DLRA using a coarse rank q and a fine rank r.
- Projected exponential methods $\rightarrow$ a new DLRA integrator robust-to-stiffness.

Future works:

- Extend low-rank Parareal to the tensor setting.[4]
- Prove further convergence bounds for the projected exponential methods.
- Apply Schwarz waveform relaxation to DLRA.

Thank to the organizers and
Thank you for your attention!

[4] Project with G. Ceruti (EPFL).

Final conclusion and outlook

Conclusion:

- DLRA → recent topic for large-scale problems integration.
- Low-rank Parareal → a new parallel-in-time integrator of DLRA using a coarse rank q and a fine rank r.
- Projected exponential methods → a new DLRA integrator robust-to-stiffness.

Future works:

- Extend low-rank Parareal to the tensor setting.[4]
- Prove further convergence bounds for the projected exponential methods.
- Apply Schwarz waveform relaxation to DLRA.

Thank to the organizers and
Thank you for your attention!

[4] Project with G. Ceruti (EPFL).

References

Carrel, B., Gander, M. J., and Vandereycken, B.
Low-rank Parareal: a low-rank parallel-in-time integrator.

Feppon, F. and Lermusiaux, P. F. J. (2018).
A Geometric Approach to Dynamical Model Order Reduction.
SIAM Journal on Matrix Analysis and Applications, 39(1):510–538.

Güttel, S. (2013).
Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection.
GAMM-Mitteilungen, 36(1):8–31.

Hochbruck, M. and Ostermann, A. (2010).
Exponential integrators.
Acta Numerica, 19:209–286.
Publisher: Cambridge University Press.

Koch, O. and Lubich, C. (2007).
Dynamical Low-Rank Approximation.
SIAM Journal on Matrix Analysis and Applications, 29(2):434–454.

Ostermann, A., Piazzola, C., and Walach, H. (2019).
Convergence of a Low-Rank Lie–Trotter Splitting for Stiff Matrix Differential Equations.
SIAM Journal on Numerical Analysis, 57(4):1947–1966.

Appendix: Rational Krylov error bound

Lemma (Approximation of the exponential by rational Krylov [Güttel, 2013])

Let V_k *be the orthogonal matrix given by the basis of the rational Krylov space* $\mathcal{RK}_k(A,b)$. *Then,* $\forall M > 0, \exists C_M, \forall k > M$,

$$\left\| e^{hA} b - V_k e^{hH_k} V_k^T b \right\| \leq C_M \left(\frac{1}{R}\right)^k \|b\|, \quad R = \exp\left(\frac{\pi^2}{4\log(2/\delta)}\right),$$

where $\delta = \sqrt{\frac{\lambda_{\min}(A)}{\lambda_{\max}(A)}}$.

Projected exponential Runge–Kutta (PERK)

Definition (PERK $[C., Vandereycken]$)

The *projected exponential Runge–Kutta* method with s stages is defined by

$$G_{nj} = \mathcal{P}_{Y_{nj}}\left[\mathcal{G}(Y_{nj})\right],$$

$$Y_{ni} = \mathcal{T}_r\left(e^{c_i h\mathcal{L}_S}(Y_n) + h\sum_{j=1}^{s} a_{ij}(h\mathcal{L}_S)(G_{nj})\right),$$

$$Y_{n+1} = \mathcal{T}_r\left(e^{h\mathcal{L}_S}(Y_n) + h\sum_{i=1}^{s} b_i(h\mathcal{L}_S)(G_{ni})\right),$$

where the operators $a_{ij}(h\mathcal{L}_S)$, $b_i(h\mathcal{L}_S)$, and the coefficients c_i are the same as for exponential methods.

Remark: The most expensive computations come from $a_{ij}(h\mathcal{L}_S)$ and $b_i(h\mathcal{L}_S)$.

Appendix: Left and right Krylov spaces

For the Sylvester operator $\mathcal{L}_S = A \otimes I + I \otimes B$, we have shown that the Krylov space

$$\mathcal{K}_k(\mathcal{L}_S, \text{vec}([X_0, \mathcal{P}_{X_0}[\mathcal{G}(X_0)]]))$$

is a subset of

$$\mathcal{K}_k(A \otimes I, \text{vec}([X_0, \mathcal{P}_{X_0}[\mathcal{G}(X_0)]])) + \mathcal{K}_k(I \otimes B, \text{vec}([X_0, \mathcal{P}_{X_0}[\mathcal{G}(X_0)]])),$$

which correspond to the left and the right Krylov spaces.