Low-rank Parareal: a low-rank parallel-in-time integrator

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

UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES

Section de Mathématiques

IMPDE 2023, Paris

Schedule

Matrix differential equations

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

Examples

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

Sylvester: $\dot{X}(t) = AX(t) + X(t)B^T + C,$

Riccati: $\dot{X}(t) = AX(t) + X(t)A^T - X(t)BX(t) + C$

Suppose the solution has 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 $\|X(t) - Y(t)\| \overset{!}{=} min.$

Matrix differential equations

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

Examples

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

Sylvester: $\dot{X}(t) = AX(t) + X(t)B^T + C,$

Riccati: $\dot{X}(t) = AX(t) + X(t)A^T - X(t)BX(t) + C$

Suppose the solution has 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 $\|X(t) - Y(t)\| \overset{!}{=} min.$

Low-rank approximation

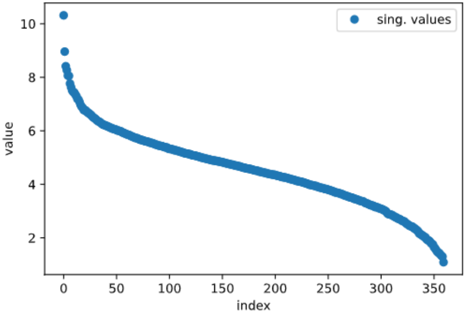
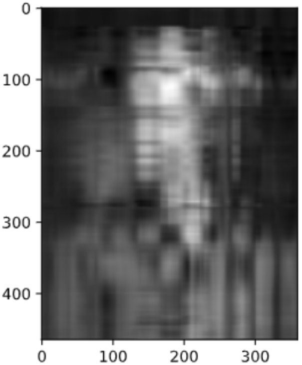
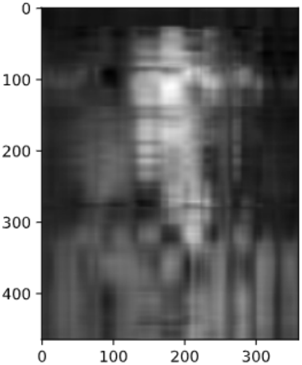


Figure: Leonhard Euler and its singular values.

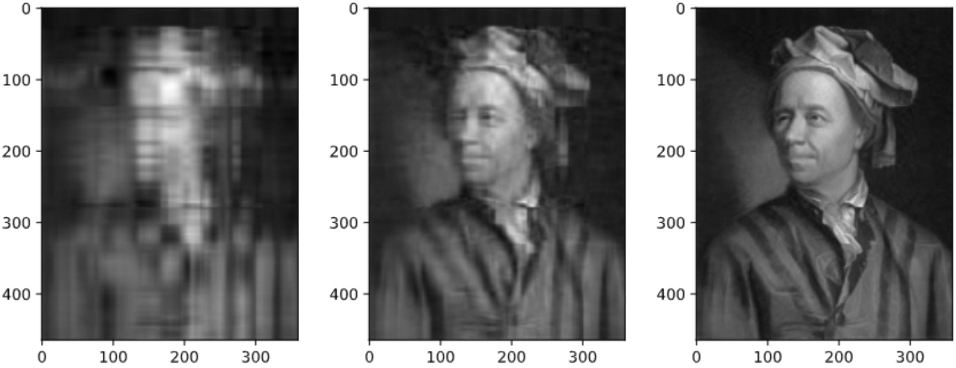


(a) Rank 5 approximation.
Relative error: 23.5%
Compression : 72× smaller.



(a) Rank 5 approximation.
Relative error: 23.5%
Compression : 72× smaller.

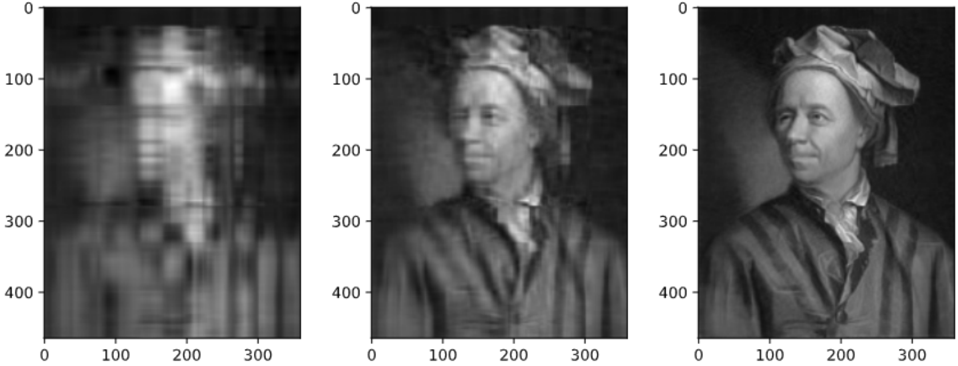
(b) Rank 15 approximation.
Relative error: 14.5%
Compression : 24× smaller.



(a) Rank 5 approximation.
Relative error: 23.5%
Compression : 72× smaller.

(b) Rank 15 approximation.
Relative error: 14.5%
Compression : 24× smaller.

(c) Rank 50 approximation.
Relative error: 7.8%
Compression : 7.2× smaller.



(a) Rank 5 approximation.
Relative error: 23.5%
Compression : 72× smaller.

(b) Rank 15 approximation.
Relative error: 14.5%
Compression : 24× smaller.

(c) Rank 50 approximation.
Relative error: 7.8%
Compression : 7.2× smaller.

→ Euler has a good low-rank approximation!!

Part 1: Dynamical low-rank approximation

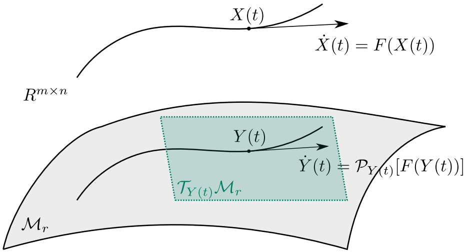
DLRA: Definition

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])

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



DLRA: Theory

Standard DLRA assumptions:

1. F is Lipschitz: $\|F(X) - F(Y)\| \leq L \|X - Y\|$.
2. F is one-sided Lipschitz[1]: $\langle X - Y, F(X) - F(Y) \rangle \leq \ell \|X - Y\|^2$.
3. 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

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

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.

DLRA: Theory

Standard DLRA assumptions:

1. F is Lipschitz: $\|F(X) - F(Y)\| \leq L \|X - Y\|$.
2. F is one-sided Lipschitz[1]: $\langle X - Y, F(X) - F(Y) \rangle \leq \ell \|X - Y\|^2$.
3. 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

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

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.

DLRA: Exponential Euler

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

Typically,

- $\mathcal{L}$ is linear and stiff,
- $\mathcal{G}$ is non-linear and non-stiff.

Closed form solution: $X(t) = e^{t\mathcal{L}}(X_0) + \int_0^t e^{(t-s)\mathcal{L}}(\mathcal{G}(X(s)))ds,$

Exponential Euler: $X_1 = e^{h\mathcal{L}}(X_0) + h\varphi_1(h\mathcal{L})(\mathcal{G}(X_0)).$

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

DLRA: Exponential Euler

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

Typically,

- $\mathcal{L}$ is linear and stiff,
- $\mathcal{G}$ is non-linear and non-stiff.

Closed form solution: $X(t) = e^{t\mathcal{L}}(X_0) + \int_0^t e^{(t-s)\mathcal{L}}(\mathcal{G}(X(s)))ds,$

Exponential Euler: $X_1 = e^{h\mathcal{L}}(X_0) + h\varphi_1(h\mathcal{L})(\mathcal{G}(X_0)).$

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

DLRA: Exponential Euler

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

Typically,

- $\mathcal{L}$ is linear and stiff,
- $\mathcal{G}$ is non-linear and non-stiff.

Closed form solution: $X(t) = e^{t\mathcal{L}}(X_0) + \int_0^t e^{(t-s)\mathcal{L}}(\mathcal{G}(X(s)))ds,$

Exponential Euler: $X_1 = e^{h\mathcal{L}}(X_0) + h\varphi_1(h\mathcal{L})(\mathcal{G}(X_0)).$

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

DLRA: Projected exponential Euler (PERK1)

Now considering the DLRA problem,

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

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

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

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

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

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

DLRA: Projected exponential Euler (PERK1)

Now considering the DLRA problem,

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

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

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

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

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

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

DLRA: Projected exponential Euler (PERK1)

Now considering the DLRA problem,

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

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

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

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

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

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

DLRA: Lucky Krylov approximation

Question: Can we apply the scheme efficiently?

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

Let us write the inner term differently,

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

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.

DLRA: Lucky Krylov approximation

Question: Can we apply the scheme efficiently?

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

Let us write the inner term differently,

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

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.

DLRA: Lucky Krylov approximation

Question: Can we apply the scheme efficiently?

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

Let us write the inner term differently,

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

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.

DLRA: Lucky Krylov approximation

Left rational Krylov space:

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

Right rational Krylov space:

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

Reduced differential equation (via Galerkin projection):

$$\begin{cases} \dot{Z}_k(t) = U_k^T A U_k Z_k(t) + Z_k(t) V_k^T B V_k + U_k^T \mathcal{P}_{Y_0} [\mathcal{G}(Y_0)] V_k \\ Z_k(0) = U_k^T Y_0 V_k \end{cases}$$

Final solution: $Z(t) \approx U_k Z_k(t) V_k^T$ where $Z_k(t) \in \mathbb{R}^{2kr \times 2kr}$

DLRA: Lucky Krylov approximation

Left rational Krylov space:

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

Right rational Krylov space:

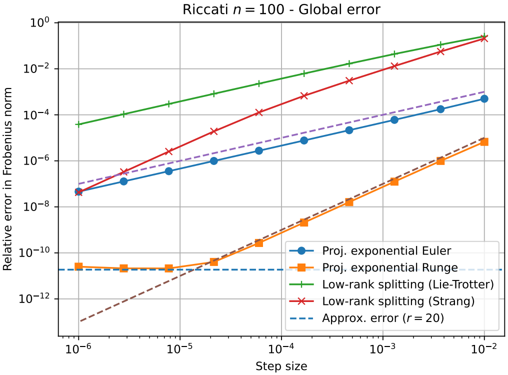
$$\mathcal{RK}_k(B, V = [V_1, V_2]) = \text{span}\left\{V, (B - \xi_2 I)^{-1} BV, \dots, \prod_{i=2}^{k} (B - \xi_i I)^{-1} B^{k-1} V\right\}$$

Reduced differential equation (via Galerkin projection):

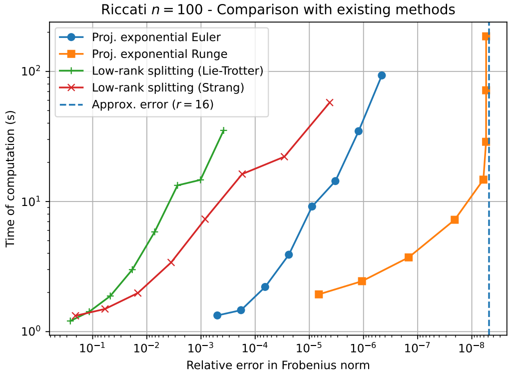
$$\begin{cases} \dot{Z}_k(t) = U_k^T A U_k Z_k(t) + Z_k(t) V_k^T B V_k + U_k^T \mathcal{P}_{Y_0}\left[\mathcal{G}(Y_0)\right] V_k \\ Z_k(0) = U_k^T Y_0 V_k \end{cases}$$

Final solution: $\quad Z(t) \approx U_k Z_k(t) V_k^T \quad$ where $Z_k(t) \in \mathbb{R}^{2kr \times 2kr}$

Application to Riccati $\dot{X} = AX + XA + C^T C - XBB^T X$



(a) Global errors



(b) Performances

Figure: New methods applied to the Riccati equation, compared to [?]

DLRA: Conclusion

- Faster computations
- Low memory footprint
- Nice theoretical tools
- Active topic of research

Can we solve it in parallel in time?

Discretized DLRA

- Projector-splitting [Lubich and Oseledets, 2014] and [Kieri et al., 2016]
- Projection methods [Kieri and Vandereycken, 2019]
- An unconventional algorithm [Ceruti and Lubich, 2022]
- A robust-to-stiffness low-rank splitting [Ostermann et al., 2019]

Continuous DLRA

- Application to the Vlasov–poisson equation [Einkemmer and Lubich, 2018]
- Stability (parabolic problems) [Kazashi et al., 2021]
- Stability (hyperbolic problems) [Kusch et al., 2023]

Rank-adaptive DLRA

- Rank-adaptive unconventional [Ceruti et al., 2022]
- Rank-adaptive DORK [Charous and Lermusiaux, 2022]
- Rank-adaptive for second-order MDEs [Hochbruck et al., 2023]

Part 2: Low-rank Parareal

Low-rank Parareal: Motivation

Definition (Parareal [Lions et al., 2001])

The Parareal algorithm iterates

$$
\begin{aligned}
&\text{(Initial value)} && X_0^k = X_0,\\
&\text{(Initial approximation)} && X_{n+1}^0 = \mathcal{G}^h(X_n^0),\\
&\text{(Iteration)} && X_{n+1}^{k+1} = \mathcal{F}^h(X_n^k) + \mathcal{G}^h(X_n^{k+1}) - \mathcal{G}^h(X_n^k),
\end{aligned}
$$

where $\mathcal{F}^h$ and $\mathcal{G}^h$ are the numerical integrators of the fine and coarse problems, respectively.

- Analysis [Gander and Vandewalle, 2007] and [Gander and Hairer, 2008].
- A link with multigrid [Gander et al., 2018].
- Task scheduling [Aubanel, 2011].
- A unified framework [Gander et al., 2022].

Low-rank Parareal: Definition

Definition (Low-rank Parareal [Carrel et al., 2023])

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

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

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: Definition

Definition (Low-rank Parareal [Carrel et al., 2023])

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

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

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

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

with the positive constants

- Let C_q (resp. $C_{r,q}$) be the Lipschitz constant of $\mathcal{T}_q$ (resp. $\mathcal{T}_r - \mathcal{T}_q$). Then,

$$\alpha = e^{\ell h} C_{r,q} \quad \text{and} \quad \beta = e^{\ell h} C_q.$$

  By [Feppon and Lermusiaux, 2018],

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

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

Low-rank Parareal: Analysis

The error verifies the double recursion

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

with the positive constants

- Let C_q (resp. $C_{r,q}$) be the Lipschitz constant of $\mathcal{T}_q$ (resp. $\mathcal{T}_r - \mathcal{T}_q$). Then,

$$\alpha = e^{\ell h} C_{r,q} \quad \text{and} \quad \beta = e^{\ell h} C_q.$$

  By [Feppon and Lermusiaux, 2018],

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

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

Low-rank Parareal: Analysis

Theorem (Convergence of low-rank Parareal [Carrel et al., 2023])

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., 2023])

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., 2023])

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

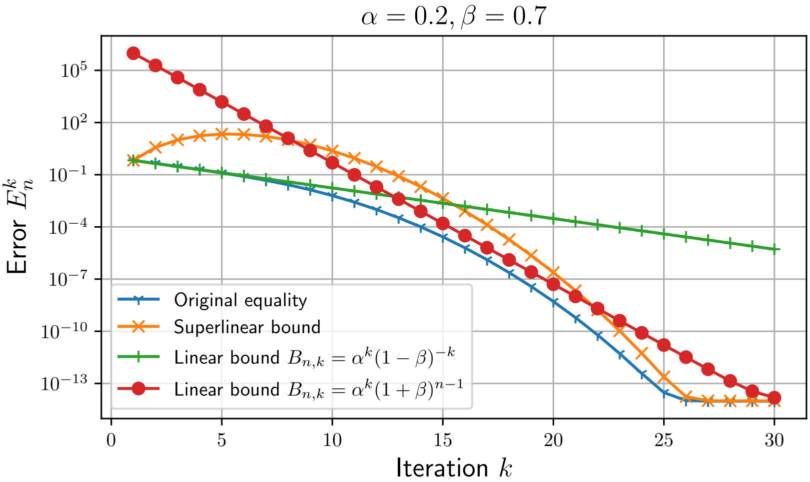


Figure: Comparison of the bounds

Low-rank Parareal: numerical results

Lyapunov equation

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

where

- $t \in [0, 2]$
- $X(0) = X_0$ is low-rank.
- $A \in \mathbb{R}^{100\times 100}$ is sparse.
- $C \in \mathbb{R}^{100\times 5}$ is a tall matrix.

$\rightarrow$ Model for 2D heat diffusion (stiff).

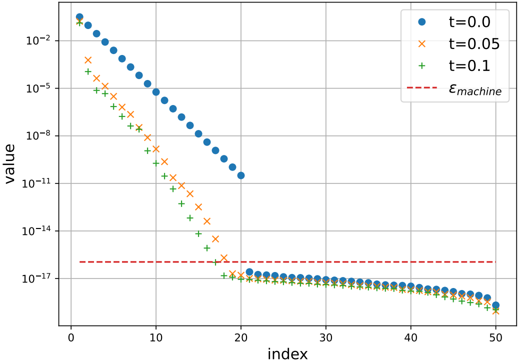
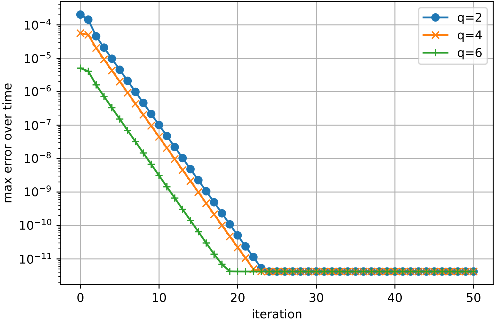


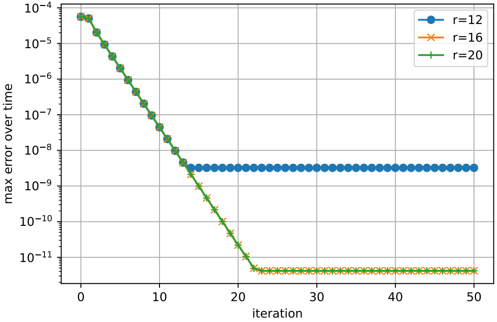
Figure: Singular values of the reference solution.

Low-rank Parareal: numerical results

$$\text{Lyapunov:} \quad \dot{X}(t) = AX(t) + X(t)A^T + C$$



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

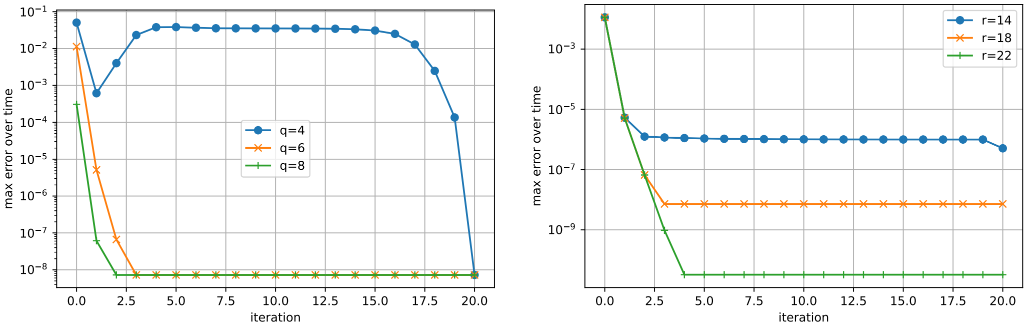


(b) Several fine ranks r with coarse rank $q = 4$.

Figure: Low-rank Parareal applied to Lyapunov ODE with $n = 100$ and $T = 2.0$.

Low-rank Parareal: numerical results

Riccati: $\dot{X}(t) = AX(t) + X(t)A^T - X(t)BX(t) + C$



(a) Several coarse ranks with fine rank $r = 18$.

(b) Several fine ranks with coarse rank $q = 6$.

Figure: Low-rank Parareal applied to Riccati ODE with $n = 200$ and $T = 0.1$.

Outlook

Conclusion:

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

Links:

- Published in BIT Numerical Mathematics: https://link.springer.com/article/10.1007/s10543-023-00953-3
- Code on GitHub: https://github.com/BenjaminCarrel/Low-rank-Parareal

Merci pour votre attention! Questions?

References I

Aubanel, E. (2011).
Scheduling of tasks in the parareal algorithm.
Parallel Computing, 37(3):172–182.

Carrel, B., Gander, M. J., and Vandereycken, B. (2023).
Low-rank parareal: a low-rank parallel-in-time integrator.
BIT Numerical Mathematics, 63(1):13.

Ceruti, G., Kusch, J., and Lubich, C. (2022).
A rank-adaptive robust integrator for dynamical low-rank approximation.
BIT Numerical Mathematics, pages 1–26.

Ceruti, G. and Lubich, C. (2022).
An unconventional robust integrator for dynamical low-rank approximation.
BIT Numerical Mathematics, 62(1):23–44.

Charous, A. and Lermusiaux, P. F. (2022).
Stable rank-adaptive dynamically orthogonal runge-kutta schemes.
arXiv preprint arXiv:2211.07852.

Einkemmer, L. and Lubich, C. (2018).
A low-rank projector-splitting integrator for the vlasov–poisson equation.
SIAM Journal on Scientific Computing, 40(5):B1330–B1360.

References II

Feppon, F. and Lermusiaux, P. F. (2018).
A geometric approach to dynamical model order reduction.
SIAM Journal on Matrix Analysis and Applications, 39(1):510–538.

Gander, M. J. and Hairer, E. (2008).
Nonlinear convergence analysis for the parareal algorithm.
Springer.

Gander, M. J., Kwok, F., and Zhang, H. (2018).
Multigrid interpretations of the parareal algorithm leading to an overlapping variant and mgrit.
Computing and Visualization in Science, 19:59–74.

Gander, M. J., Lunet, T., Ruprecht, D., and Speck, R. (2022).
A unified analysis framework for iterative parallel-in-time algorithms.
arXiv preprint arXiv:2203.16069.

Gander, M. J. and Vandewalle, S. (2007).
Analysis of the parareal time-parallel time-integration method.
SIAM Journal on Scientific Computing, 29(2):556–578.

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

References III

Hochbruck, M., Neher, M., and Schrammer, S. (2023).
Rank-adaptive dynamical low-rank integrators for first-order and second-order matrix differential equations.
BIT Numerical Mathematics, 63(1):9.

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

Kazashi, Y., Nobile, F., and Vidličková, E. (2021).
Stability properties of a projector-splitting scheme for dynamical low rank approximation of random parabolic equations.
Numerische Mathematik, 149:973–1024.

Kieri, E., Lubich, C., and Walach, H. (2016).
Discretized dynamical low-rank approximation in the presence of small singular values.
SIAM Journal on Numerical Analysis, 54(2):1020–1038.

Kieri, E. and Vandereycken, B. (2019).
Projection methods for dynamical low-rank approximation of high-dimensional problems.
Computational Methods in Applied Mathematics, 19(1):73–92.

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

References IV

Kusch, J., Einkemmer, L., and Ceruti, G. (2023).
On the stability of robust dynamical low-rank approximations for hyperbolic problems.
SIAM Journal on Scientific Computing, 45(1):A1–A24.

Lions, J.-L., Maday, Y., and Turinici, G. (2001).
Résolution d'edp par un schéma en temps «pararéel».
Comptes Rendus de l'Académie des Sciences-Series I-Mathematics, 332(7):661–668.

Lubich, C. and Oseledets, I. V. (2014).
A projector-splitting integrator for dynamical low-rank approximation.
BIT Numerical Mathematics, 54(1):171–188.

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.