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

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

UNIVERSITÉ DE GENEVE
FACULTÉ DES SCIENCES
Section de Mathématiques
IMPDE 2023, Paris
Schedule

1. Setup and motivation
2. Dynamical low-rank approximation
   - Definition and theory
   - Numerical integration
3. Low-rank Parareal
   - Description of the algorithm
   - Theoretical analysis
   - Numerical experiments
4. Conclusion
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 \( \| X(t) - Y(t) \| \overset{1}{=} \text{min} \).
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.

(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.
(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])**

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

$$\left\| \psi^h_r(Y_0) - \phi^h(X_0) \right\| \leq e^{\ell t} \|Y_0 - X_0\| + \varepsilon \int_0^t e^{\ell s} \, ds,$$

where $\psi^h_r$ is the flow of the DLRA, and $\phi^h$ is the flow of the original problem.

\(^1\)If $F$ is linear, $\ell$ 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

\[
\left\| \psi_r^h(Y_0) - \phi^h(X_0) \right\| \leq e^{\ell t} \| Y_0 - X_0 \| + \varepsilon \int_0^t e^{\ell s} ds ,
\]

where \( \psi_r^h \) is the flow of the DLRA, and \( \phi^h \) is the flow of the original problem.

\(^1\)If \( F \) is linear, \( \ell \) is the largest eigenvalue, potentially negative.
\textbf{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**

\[
\dot{X}(t) = AX(t) + X(t)B^T + \mathcal{G}(X(t)) = L(X(t)) + \mathcal{G}(X(t)), \quad t \in [0, T]
\]

\[
X(0) = X_0 \in \mathbb{R}^{m \times n},
\]

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 \( h\varphi_1(h\mathcal{L}) = \mathcal{L}^{-1}(e^{h\mathcal{L}} - \text{Id}) \). 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:

\[ X_1 = e^{h\mathcal{L}}(X_0) + h\phi_1(h\mathcal{L})(\mathcal{G}(X_0)). \]

where \( h\phi_1(h\mathcal{L}) = \mathcal{L}^{-1}(e^{h\mathcal{L}} - Id) \). 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

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

**Question:** Can we apply the scheme efficiently?

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

Let us write the inner term differently,

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

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

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

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

Let us write the inner term differently,

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

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

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

**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:

\[ Y_1 = \mathcal{T}_r \left( e^{h\mathcal{L}}(Y_0) + h\varphi_1(h\mathcal{L})\mathcal{P}_Y_0 [\mathcal{G}(Y_0)] \right) \]

Let us write the inner term differently,

\[ Z(t) = e^{t\mathcal{L}}(Y_0) + t\varphi_1(t\mathcal{L})\mathcal{P}_Y_0 [\mathcal{G}(Y_0)] \iff \begin{cases} \dot{Z}(t) = AZ(t) + Z(t)B + \mathcal{P}_Y_0 [\mathcal{G}(Y_0)] \\ Z(0) = Y_0 \end{cases} \]

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

\( Y_0 \) and \( \mathcal{P}_Y_0 [\mathcal{G}(Y_0)] \) are low-rank \( Y_0 = U_1 \Sigma V_1^T, \quad \mathcal{P}_Y_0 [\mathcal{G}(Y_0)] = [U_1, U_2] \tilde{\Sigma} [V_1, V_2]^T \)

**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):

\[
\begin{align*}
\dot{Z}_k(t) &= U_k^T A U_k Z_k(t) + Z_k(t) V_k^T B V_k + U_k^T P Y_0 [G(Y_0)] V_k \\
Z_k(0) &= U_k^T Y_0 V_k
\end{align*}
\]

Final solution:

\[ Z(t) \approx U_k Z_k(t) V_k^T \quad \text{where} \quad Z_k(t) \in \mathbb{R}^{2kr \times 2kr} \]
**DLRA: Lucky Krylov approximation**

**Left rational Krylov space:**

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

**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):**

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

**Final solution:**

\[
Z(t) \approx U_k Z_k(t) V_k^T
\]

where \(Z_k(t) \in \mathbb{R}^{2kr \times 2kr}\)
Application to Riccati $\dot{X} = AX + XA + C^T C - XBB^T X$

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

(Initial value) \( X_0^k = X_0 \),

(Initial approximation) \( X_{n+1}^0 = G^h(X_n^0) \),

(Iteration) \( X_{n+1}^{k+1} = F^h(X_n^k) + G^h(X_n^{k+1}) - G^h(X_n^k) \),

where \( F^h \) and \( G^h \) 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

\[
Y_0^k = Y_0, \\
Y_{n+1}^0 = \psi_q^h \circ T_q(Y_n^0) + \mathcal{E}_n, \\
Y_{n+1}^{k+1} = \psi_r^h \circ T_r(Y_n^k) + \psi_q^h \circ T_q(Y_n^{k+1}) - \psi_q^h \circ T_q(Y_n^k),
\]

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

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

(Initial value) \[ Y_0^k = Y_0, \]

(Initial approximation) \[ Y_{n+1}^0 = \psi_q^h \circ T_q(Y_n^0) + \mathcal{E}_n, \]

(Iteration) \[ Y_{n+1}^{k+1} = \psi_r^h \circ T_r(Y_n^k) + \psi_q^h \circ T_q(Y_{n+1}^{k+1}) - \psi_q^h \circ T_q(Y_n^k), \]

where $\psi_r^h$ is the solution of the DLRA of rank $r$ at time $h$, and $T_r$ is the orthogonal projection onto $\mathcal{M}_r$. The notations $\psi_q^h$ and $T_q$ 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

\[ \|E_{n+1}^{k+1}\| \leq \alpha \|E_{n}^{k}\| + \beta \|E_{n}^{k+1}\| + \kappa, \quad n, k \geq 0, \]

with the positive constants

- Let \( C_q \) (resp. \( C_{r,q} \)) be the Lipschitz constant of \( T_q \) (resp. \( T_r - T_q \)). Then,

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

By [Feppon and Lermusiaux, 2018],

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

- \( \kappa = e^{\ell h} \max_{n \geq 0} \|X_n - T_r(X_n)\| + (2\varepsilon_q + \varepsilon_r) \int_{0}^{h} e^{\ell(h-s)} \, ds. \)
Low-rank Parareal: Analysis

The error verifies the double recursion

\[ \|E_{n+1}^{k+1}\| \leq \alpha \|E_n^k\| + \beta \|E_n^{k+1}\| + \kappa, \quad n, k \geq 0, \]

with the positive constants

- Let \( C_q \) (resp. \( C_{r,q} \)) be the Lipschitz constant of \( T_q \) (resp. \( T_r - T_q \)). Then,

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

By [Feppon and Lermusiaux, 2018],

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

- \( \kappa = e^{\ell h} \max_{n \geq 0} \|X_n - T_r(X_n)\| + (2\varepsilon_q + \varepsilon_r) \int_0^h e^{\ell(h-s)} ds. \)
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: \[
\max_{n \geq 0} \| E_n^k \| \leq \left( \frac{\alpha}{1 - \beta} \right)^k \max_{n \geq 0} \| E_n^0 \| + \frac{\kappa}{1 - \alpha - \beta},
\]

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

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

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

Superlinear bound: \[
\| E_n^k \| \leq \frac{\alpha^k}{(k - 1)!} \frac{\prod_{j=2}^{k} (n - j)}{1 - \beta} \max_{n \geq 0} \| E_n^0 \| + \frac{\kappa}{1 - \alpha - \beta}.
\]
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:  \[
\max_{n \geq 0} \| E_n^k \| \leq \left( \frac{\alpha}{1 - \beta} \right)^k \max_{n \geq 0} \| E_0^n \| + \frac{\kappa}{1 - \alpha - \beta},
\]

2nd linear bound:  \[
\| E_n^k \| \leq \alpha^k (1 + \beta)^{n-1} \max_{n \geq 0} \| E_0^n \| + \frac{\kappa}{1 - \alpha - \beta},
\]

Superlinear bound:  \[
\| E_n^k \| \leq \frac{\alpha^k}{(k-1)!} \frac{\prod_{j=2}^{k} (n-j)}{1 - \beta} \max_{n \geq 0} \| E_0^n \| + \frac{\kappa}{1 - \alpha - \beta}.
\]
\( \alpha = 0.2, \beta = 0.7 \)

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.

→ Model for 2D heat diffusion (stiff).

Figure: Singular values of the reference solution.
Low-rank Parareal: numerical results

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

(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 II

A geometric approach to dynamical model order reduction.

*Nonlinear convergence analysis for the parareal algorithm.*
Springer.

Multigrid interpretations of the parareal algorithm leading to an overlapping variant and mgrit.

A unified analysis framework for iterative parallel-in-time algorithms.

Analysis of the parareal time-parallel time-integration method.

Rational krylov approximation of matrix functions: Numerical methods and optimal pole selection.
References III

Rank-adaptive dynamical low-rank integrators for first-order and second-order matrix differential equations.

Exponential integrators.

Stability properties of a projector-splitting scheme for dynamical low rank approximation of random parabolic equations.
*Numerische Mathematik*, 149:973–1024.

Discretized dynamical low-rank approximation in the presence of small singular values.

Projection methods for dynamical low-rank approximation of high-dimensional problems.

Dynamical low-rank approximation.
On the stability of robust dynamical low-rank approximations for hyperbolic problems.

Résolution d’edp par un schéma en temps «pararéel».

Lubich, C. and Oseledets, I. V. (2014).
A projector-splitting integrator for dynamical low-rank approximation.

Convergence of a low-rank lie–trotter splitting for stiff matrix differential equations.