MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

Generalizing Reduced Rank Extrapolation (RRE) to Low-Rank Matrix Sequences

Xiaobo Liu[3]

[3]MPI Magdeburg, Germany

ILAS 2025,
National Sun Yat-sen University, Kaohsiung, June 26, 2025

Joint work with **Pascal den Boef**[1] **Patrick Kürschner**[2] **Jos Maubach**[1] **Jens Saak**[3]
Wil Schilders[1] **Jonas Schulze**[3] **Nathan van de Wouw**[1]
[1]TU Eindhoven, Netherlands, [2]HWTK Leipzig, Germany

1. Reduced Rank Extrapolation
2. Methodological Extensions
3. Numerical Experiments
4. Conclusions

Original sequence: x_1

Improved sequence:

Non-Cycling Mode

for $i \leftarrow 1, 2, \ldots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if $i \geq n$ **then**
    $\hat{x}_i \leftarrow \text{extr.}(x_{i-n+1}, \ldots, x_i, x_{i+1})$
    if $\hat{x}_i$ converged **then** break
  end
end

Cycling Mode

for $i \leftarrow 1, 2, \ldots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if should start new cycle **then**
    $x_{i+1} \leftarrow \text{extr.}(x_{i-n+1}, \ldots, x_i, x_{i+1})$
  end
  if x_{i+1} converged **then** break
end

- Non-cycling mode generates new extrapolated sequence $\widehat{x}_n, \widehat{x}_{n+1}, \ldots$ (via window size n).
- Cycling mode restarts using the extrapolated iterate, to generate, e.g., $x_1, x_2, x_3, x_4, \ldots$.

Original sequence: x_1 x_2

Improved sequence:

Non-Cycling Mode

for $i \leftarrow 1, 2, \ldots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if $i \geq n$ **then**
    $\hat{x}_i \leftarrow$ extr.$(x_{i-n+1}, \ldots, x_i, x_{i+1})$
    if $\hat{x}_i$ converged **then** break
  end
end

Cycling Mode

for $i \leftarrow 1, 2, \ldots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if should start new cycle **then**
    $x_{i+1} \leftarrow$ extr.$(x_{i-n+1}, \ldots, x_i, x_{i+1})$
  end
  if x_{i+1} converged **then** break
end

- Non-cycling mode generates new extrapolated sequence $\widehat{x}_n, \widehat{x}_{n+1}, \ldots$ (via window size n).
- Cycling mode restarts using the extrapolated iterate, to generate, e.g., $x_1, x_2, x_3, x_4, \ldots$.

Original sequence: x_1 x_2 x_3

Improved sequence:

Non-Cycling Mode

for $i \leftarrow 1, 2, \ldots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if $i \geq n$ **then**
    $\hat{x}_i \leftarrow \text{extr.}(x_{i-n+1}, \ldots, x_i, x_{i+1})$
    if $\hat{x}_i$ converged **then** break
  end
end

Cycling Mode

for $i \leftarrow 1, 2, \ldots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if should start new cycle **then**
    $x_{i+1} \leftarrow \text{extr.}(x_{i-n+1}, \ldots, x_i, x_{i+1})$
  end
  if x_{i+1} converged **then** break
end

- Non-cycling mode generates new extrapolated sequence $\hat{x}_n, \hat{x}_{n+1}, \ldots$ (via window size n).
- Cycling mode restarts using the extrapolated iterate, to generate, e.g., $x_1, x_2, x_3, x_4, \ldots$.

Original sequence: $x_1 \quad x_2 \quad x_3$

Improved sequence: $\hat{x}_3$

Non-Cycling Mode

for $i \leftarrow 1, 2, \dots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if $i \geq n$ **then**
    $\hat{x}_i \leftarrow$ extr.$(x_{i-n+1}, \dots, x_i, x_{i+1})$
    if $\hat{x}_i$ converged **then** break
  end
end

Cycling Mode

for $i \leftarrow 1, 2, \dots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if should start new cycle **then**
    $x_{i+1} \leftarrow$ extr.$(x_{i-n+1}, \dots, x_i, x_{i+1})$
  end
  if x_{i+1} converged **then** break
end

- Non-cycling mode generates new extrapolated sequence $\hat{x}_n, \hat{x}_{n+1}, \dots$ (via window size n).
- Cycling mode restarts using the extrapolated iterate, to generate, e.g., $x_1, x_2, x_3, x_4, \dots$.

Original sequence: x_1 x_2 x_3 x_4

Improved sequence: $\hat{x}_3$ $\hat{x}_4$

Non-Cycling Mode

for $i \leftarrow 1, 2, \ldots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if $i \geq n$ **then**
    $\hat{x}_i \leftarrow \text{extr.}(x_{i-n+1}, \ldots, x_i, x_{i+1})$
    if $\hat{x}_i$ converged **then** break
  end
end

Cycling Mode

for $i \leftarrow 1, 2, \ldots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if should start new cycle **then**
    $x_{i+1} \leftarrow \text{extr.}(x_{i-n+1}, \ldots, x_i, x_{i+1})$
  end
  if x_{i+1} converged **then** break
end

- Non-cycling mode generates new extrapolated sequence $\widehat{x}_n, \widehat{x}_{n+1}, \ldots$ (via window size n).
- Cycling mode restarts using the extrapolated iterate, to generate, e.g., $x_1, x_2, x_3, x_4, \ldots$.

Original sequence: $x_1 \quad x_2 \quad x_3 \quad x_4 \quad \dots$

Improved sequence: $\hat{x}_3 \quad \hat{x}_4 \quad \dots$

Non-Cycling Mode

for $i \leftarrow 1, 2, \dots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if $i \geq n$ **then**
    $\hat{x}_i \leftarrow \text{extr.}(x_{i-n+1}, \dots, x_i, x_{i+1})$
    if $\hat{x}_i$ converged **then** break
  end
end

Cycling Mode

for $i \leftarrow 1, 2, \dots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if should start new cycle **then**
    $x_{i+1} \leftarrow \text{extr.}(x_{i-n+1}, \dots, x_i, x_{i+1})$
  end
  if x_{i+1} converged **then** break
end

- Non-cycling mode generates new extrapolated sequence $\widehat{x}_n, \widehat{x}_{n+1}, \dots$ (via window size n).
- Cycling mode restarts using the extrapolated iterate, to generate, e.g., $x_1, x_2, x_3, x_4, \dots$.

Reduced Rank Extrapolation

Extrapolation Methods for Fixed-Point Processes

Original sequence: $x_1 \quad x_2 \quad x_3 \quad x_4 \quad \dots$

Improved sequence: $\hat{x}_3 \quad \hat{x}_4 \quad \dots$

Non-Cycling Mode

for $i \leftarrow 1, 2, \dots$ **do**
 $x_{i+1} \leftarrow f(x_i)$
 if $i \geq n$ **then**
  $\hat{x}_i \leftarrow \text{extr.}(x_{i-n+1}, \dots, x_i, x_{i+1})$
  if $\hat{x}_i$ converged **then** break
 end
end

Cycling Mode

for $i \leftarrow 1, 2, \dots$ **do**
 $x_{i+1} \leftarrow f(x_i)$
 if should start new cycle **then**
  $x_{i+1} \leftarrow \text{extr.}(x_{i-n+1}, \dots, x_i, x_{i+1})$
 end
 if x_{i+1} converged **then** break
end

- Non-cycling mode generates new extrapolated sequence $\hat{x}_n, \hat{x}_{n+1}, \dots$ (via window size n).
- Cycling mode restarts using the extrapolated iterate, to generate, e.g., $x_1, x_2, x_3, x_4, \dots$.

Original sequence: $x_1 \quad x_2 \quad x_3 \quad x_4 \quad \dots$

Improved sequence: $\hat{x}_3 \quad \hat{x}_4 \quad \dots$

Non-Cycling Mode

for $i \leftarrow 1, 2, \dots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if $i \geq n$ **then**
    $\hat{x}_i \leftarrow \text{extr.}(x_{i-n+1}, \dots, x_i, x_{i+1})$
    if $\hat{x}_i$ converged **then** break
  end
end

Cycling Mode

for $i \leftarrow 1, 2, \dots$ **do**
  $x_{i+1} \leftarrow f(x_i)$
  if should start new cycle **then**
    $x_{i+1} \leftarrow \text{extr.}(x_{i-n+1}, \dots, x_i, x_{i+1})$
  end
  if x_{i+1} converged **then** break
end

- Non-cycling mode generates new extrapolated sequence $\widehat{x}_n, \widehat{x}_{n+1}, \dots$ (via window size n).
- Cycling mode restarts using the extrapolated iterate, to generate, e.g., $x_1, x_2, x_3, x_4, \dots$.

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = I - (I - A)$,

$$x_{i+1} := f(x_i) := (I - A)x_i + b.$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|f(\hat{x}) - \hat{x}\|,$$

where $\gamma_1 + \cdots + \gamma_n = 1$.
- Key observations:

$$f(x) - x = b - Ax$$
$$f(\hat{x}) - \hat{x} = \sum_{i=1}^{n} \gamma_i (f(x_i) - x_i) \quad \text{(linearity)}$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = I - (I - A)$,

$$x_{i+1} := f(x_i) := (I - A)x_i + b.$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|f(\hat{x}) - \hat{x}\|,$$

where $\gamma_1 + \cdots + \gamma_n = 1$.

- Key observations:

$$f(x) - x = b - Ax$$
$$f(\hat{x}) - \hat{x} = \sum_{i=1}^{n} \gamma_i (f(x_i) - x_i) \quad \text{(linearity)}$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = I - (I - A)$,

$$x_{i+1} := f(x_i) := (I - A)x_i + b.$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|f(\hat{x}) - \hat{x}\|,$$

where $\gamma_1 + \cdots + \gamma_n = 1$.
- Key observations:

$$f(x) - x = b - Ax$$
$$f(\hat{x}) - \hat{x} = \sum_{i=1}^{n} \gamma_i \big(f(x_i) - x_i\big) \quad \text{(linearity)}$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = I - (I - A)$,

$$x_{i+1} := f(x_i) := (I - A)x_i + b.$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|f(\hat{x}) - \hat{x}\|,$$

where $\gamma_1 + \cdots + \gamma_n = 1$.
- Key observations:

$$f(x) - x = b - Ax$$
$$f(\hat{x}) - \hat{x} = \sum_{i=1}^{n} \gamma_i \big(f(x_i) - x_i\big) \quad \text{(linearity)}$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = M - N$,

$$x_{i+1} := f(x_i) := M^{-1}(Nx_i + b).$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|f(\hat{x}) - \hat{x}\|,$$

where $\gamma_1 + \cdots + \gamma_n = 1$.
- Key observations:

$$f(x) - x = M^{-1}(b - Ax)$$
$$f(\hat{x}) - \hat{x} = \sum_{i=1}^{n} \gamma_i \big(f(x_i) - x_i\big) \quad \text{(linearity)}$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = M - N$,

$$x_{i+1} := f(x_i) := M^{-1}(Nx_i + b).$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|f(\hat{x}) - \hat{x}\|,$$

  where $\gamma_1 + \cdots + \gamma_n = 1$.
- Key observations:

$$f(x) - x = M^{-1}(b - Ax)$$
$$f(\hat{x}) - \hat{x} = \sum_{i=1}^{n} \gamma_i \big(f(x_i) - x_i\big) \quad \text{(linearity)}$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = M - N$,

$$x_{i+1} := f(x_i) := M^{-1}(Nx_i + b).$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|f(\hat{x}) - \hat{x}\|,$$

where $\gamma_1 + \cdots + \gamma_n = 1$.

- Key observations:

$$f(x) - x = M^{-1}(b - Ax)$$
$$f(\hat{x}) - \hat{x} = \sum_{i=1}^{n} \gamma_i \big(f(x_i) - x_i\big) \quad \text{(linearity)}$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

Most derivations of reduced rank extrapolation (RRE) read off the iteration scheme from an underlying equation where both notions of the residual coincide:

$x = Ax + b$	[Mešina, '77], [Eddy, '79](not spelled out), [Sidi, '88], [Sidi, '91]
$Ax = b$	[Sidi & Shapira, '98]

or only consider the residual of the iteration:

$Ax = b$	[Kaniel & Stein, '74]
$x = f(x)$	[Sidi, '20]

Theorem (Sidi, '88), (Sidi, '17)

Let $\{x_m\}$ be the sequence generated by the fixed-point iteration

$$x_{m+1} = Fx_m + b, \quad m = 0, 1, 2, \ldots,$$

with initial guess x_0. The extrapolated solutions x_k^{RRE} (with window size $k+1$) obtained by applying non-cycling RRE to $\{x_m\}$ ***coincide*** *with the iterates x_k^{GMRES} obtained by applying GMRES (after k steps) to the linear system $(I - F)x = b$ with the same initial guess.*

Sketch proof: $x_{m+1} - x_m = b - (I - F)x_m = r_m \implies$ both RRE and GMRES minimize the residual over $\mathcal{K}_k(A, r_0) = \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$, where $A = I - F$.

- No known mathematical equivalence between cycling RRE and restarted GMRES, though they are similar in *spirit* and *practical behavior*.

Theorem (Sidi, '88), (Sidi, '17)

Let $\{x_m\}$ be the sequence generated by the fixed-point iteration

$$x_{m+1} = Fx_m + b, \quad m = 0, 1, 2, \dots,$$

with initial guess x_0. The extrapolated solutions x_k^{RRE} (with window size $k+1$) obtained by applying non-cycling RRE to $\{x_m\}$ ***coincide*** *with the iterates x_k^{GMRES} obtained by applying GMRES (after k steps) to the linear system $(I-F)x = b$ with the same initial guess.*

Sketch proof: $x_{m+1} - x_m = b - (I-F)x_m = r_m \implies$ both RRE and GMRES minimize the residual over $\mathcal{K}_k(A, r_0) = \mathrm{span}\{r_0, Ar_0, \dots, A^{k-1}r_0\}$, where $A = I - F$.

- No known mathematical equivalence between cycling RRE and restarted GMRES, though they are similar in *spirit* and *practical behavior*.

Theorem (Sidi, '88), (Sidi, '17)

Let $\{x_m\}$ be the sequence generated by the fixed-point iteration

$$x_{m+1} = Fx_m + b, \quad m = 0, 1, 2, \ldots,$$

with initial guess x_0. The extrapolated solutions x_k^{RRE} (with window size $k+1$) obtained by applying non-cycling RRE to $\{x_m\}$ ***coincide*** *with the iterates x_k^{GMRES} obtained by applying GMRES (after k steps) to the linear system $(I - F)x = b$ with the same initial guess.*

Sketch proof: $x_{m+1} - x_m = b - (I - F)x_m = r_m \implies$ both RRE and GMRES minimize the residual over $\mathcal{K}_k(A, r_0) = \mathrm{span}\{r_0, Ar_0, \ldots, A^{k-1}r_0\}$, where $A = I - F$.

- No known mathematical equivalence between cycling RRE and restarted GMRES, though they are similar in *spirit* and *practical behavior*.

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = M - N$,

$$x_{i+1} := f(x_i) := M^{-1}(Nx_i + b).$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|f(\hat{x}) - \hat{x}\|,$$

where $\gamma_1 + \cdots + \gamma_n = 1$.
- Key observations:

$$f(x) - x = M^{-1}(b - Ax)$$
$$f(\hat{x}) - \hat{x} = \sum_{i=1}^{n} \gamma_i \big(f(x_i) - x_i\big) \quad \text{(linearity)}$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

- Objective: solve $Ax = b$.
- Iterative scheme: splitting $A = M_i - N_i$,

$$x_{i+1} := f_i(x_i) := M_i^{-1}(N_i x_i + b).$$

- RRE Ansatz: find $\hat{x} := \sum_{i=1}^{n} \gamma_i x_i$ minimizing the **residual of the fix-point iteration**

$$\|A\hat{x} - b\|,$$

where $\gamma_1 + \cdots + \gamma_n = 1$.
- Key observations:

$$f_i(x) - x = M_i^{-1}(b - Ax)$$
$$b - A\hat{x} = \sum_{i=1}^{n} \gamma_i \left(b - Ax_i\right)$$

Target properties

1. Nonstationary f
2. Low-rank x_i and x_{i+1}

Alert

1. Two notions of residual
2. One not suitable for nonstationary f

So far:

$$\hat{x} := \sum_{i=1}^{n} \gamma_i x_i, \quad \text{where}$$

Increment-Based RRE **(old)**

$$\gamma = \arg\min_{g \in \mathbb{R}^n} \left\| \sum_{i=1}^{n} g_i (x_{i+1} - x_i) \right\|,$$

$$\text{s.t.} \quad \sum_{i=1}^{n} g_i = 1.$$

Residual-Based RRE **(new)**

$$\gamma = \arg\min_{g \in \mathbb{R}^n} \left\| \sum_{i=1}^{n} g_i (A x_i - b) \right\|,$$

$$\text{s.t.} \quad \sum_{i=1}^{n} g_i = 1.$$

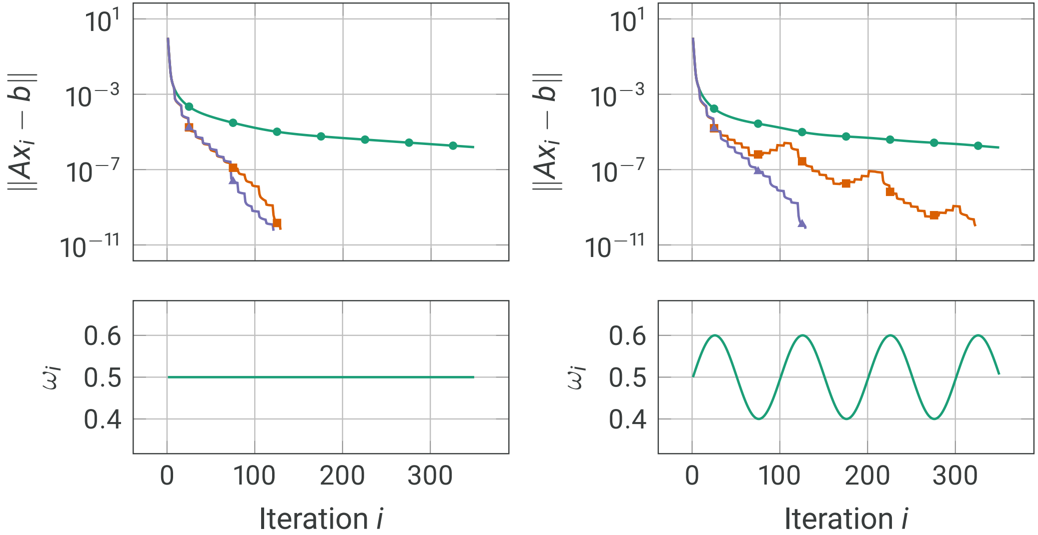


Figure: Comparison of increment-based (—■—) and residual-based (—▲—) RRE formulations applied to successive over-relaxation schemes (—●—).

- RRE for vector sequences $x_i \in \mathbb{R}^d$: via method of Lagrange multipliers

$$\gamma = \arg\min_{g \in \mathbb{R}^n} \left\| \sum_{i=1}^{n} g_i (Ax_i - b) \right\| \quad \text{s.t.} \quad \sum_{i=1}^{n} g_i = 1 \quad \Longleftrightarrow \quad U^{\mathsf{T}} U \alpha = 1 \in \mathbb{R}^n \text{ and } \gamma := \alpha / \|\alpha\|$$

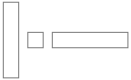
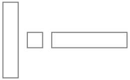
$$\boxed{U} := \left[Ax_1 - b \middle| \dots \middle| Ax_n - b \right] \in \mathbb{R}^{d \times n}$$

- Low-rank matrix sequences $X_i =$ with residuals $\mathcal{R}(X_i) =$
- Assembling $X_i \in \mathbb{R}^{d \times d}$ is prohibitively expensive, despite inner factor having size r
  $U \in \mathbb{R}^{d^2 \times n}$ is even more expensive, with $\mathrm{vec}(\cdot)$ applied to X_i

- RRE for vector sequences $x_i \in \mathbb{R}^d$: via method of Lagrange multipliers

$$\gamma = \underset{g \in \mathbb{R}^n}{\arg\min} \left\| \sum_{i=1}^{n} g_i (A x_i - b) \right\| \quad \text{s.t.} \quad \sum_{i=1}^{n} g_i = 1 \qquad \Longleftrightarrow \qquad U^\mathsf{T} U \alpha = 1 \in \mathbb{R}^n \text{ and } \gamma := \alpha / \|\alpha\|$$

$$U := \left[A x_1 - b \middle| \dots \middle| A x_n - b \right] \in \mathbb{R}^{d \times n}$$

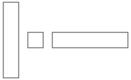
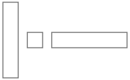
- Low-rank matrix sequences $X_i =$  with residuals $\mathcal{R}(X_i) =$ 

- Assembling $X_i \in \mathbb{R}^{d \times d}$ is prohibitively expensive, despite inner factor having size r
  $U \in \mathbb{R}^{d^2 \times n}$ is even more expensive, with $\mathrm{vec}(\cdot)$ applied to X_i

- RRE for vector sequences $x_i \in \mathbb{R}^d$: via method of Lagrange multipliers

$$\gamma = \underset{g \in \mathbb{R}^n}{\arg\min} \left\| \sum_{i=1}^{n} g_i (Ax_i - b) \right\| \quad \text{s.t.} \quad \sum_{i=1}^{n} g_i = 1 \qquad \Longleftrightarrow \qquad U^\mathsf{T} U \alpha = 1 \in \mathbb{R}^n \text{ and } \gamma := \alpha / \|\alpha\|$$

$$U := \left[Ax_1 - b \middle| \dots \middle| Ax_n - b \right] \in \mathbb{R}^{d \times n}$$

- Low-rank matrix sequences $X_i =$  with residuals $\mathcal{R}(X_i) =$ 
- Assembling $X_i \in \mathbb{R}^{d \times d}$ is prohibitively expensive, despite inner factor having size r
  $U \in \mathbb{R}^{d^2 \times n}$ is even more expensive, with $\mathrm{vec}(\cdot)$ applied to X_i

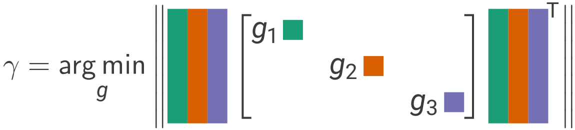
Main idea:

1. Move arithmetic onto inner factors of dimension $r \ll d$
2. Reduce dimension via QR decomposition of outer factors

Benefits:

Solution of $U^{\mathsf{T}} U \alpha = 1 \in \mathbb{R}^n$ now requires $U \in \mathbb{R}^{(nr)^2 \times n}$ where $(nr)^2 \ll d^2$.
Recall that $\gamma := \alpha / \|\alpha\|$, and that ■/■/■ may represent increments or residuals.

Main idea:

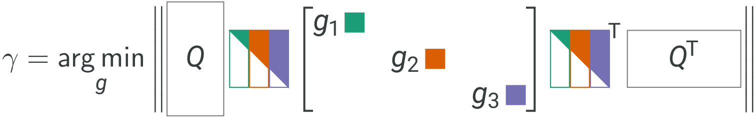


1. Move arithmetic onto inner factors of dimension $r \ll d$
2. Reduce dimension via QR decomposition of outer factors

Benefits:

Solution of $U^{\mathsf{T}}U\alpha = 1 \in \mathbb{R}^n$ now requires $U \in \mathbb{R}^{(nr)^2 \times n}$ where $(nr)^2 \ll d^2$.
Recall that $\gamma := \alpha / \|\alpha\|$, and that ■/■/■ may represent increments or residuals.

Main idea:



1. Move arithmetic onto inner factors of dimension $r \ll d$
2. Reduce dimension via QR decomposition of outer factors

Benefits:

Solution of $U^{\mathsf{T}} U \alpha = 1 \in \mathbb{R}^n$ now requires $U \in \mathbb{R}^{(nr)^2 \times n}$ where $(nr)^2 \ll d^2$.

Recall that $\gamma := \alpha / \|\alpha\|$, and that ■/■/■ may represent increments or residuals.

Main idea:

1. Move arithmetic onto inner factors of dimension $r \ll d$
2. Reduce dimension via QR decomposition of outer factors

Benefits:

Solution of $U^\top U\alpha = 1 \in \mathbb{R}^n$ now requires $U \in \mathbb{R}^{(nr)^2 \times n}$ where $(nr)^2 \ll d^2$.
Recall that $\gamma := \alpha / \|\alpha\|$, and that ■/■/■ may represent increments or residuals.

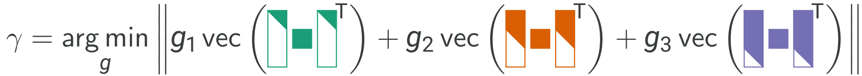
Main idea:

1. Move arithmetic onto inner factors of dimension $r \ll d$
2. Reduce dimension via QR decomposition of outer factors

Benefits:

Solution of $U^{\mathsf{T}} U \alpha = 1 \in \mathbb{R}^n$ now requires $U \in \mathbb{R}^{(nr)^2 \times n}$ where $(nr)^2 \ll d^2$.
Recall that $\gamma := \alpha / \|\alpha\|$, and that / / may represent increments or residuals.

Main idea:



1. Move arithmetic onto inner factors of dimension $r \ll d$
2. Reduce dimension via QR decomposition of outer factors

Benefits:

Solution of $U^\mathsf{T} U \alpha = 1 \in \mathbb{R}^n$ now requires $U \in \mathbb{R}^{(nr)^2 \times n}$ where $(nr)^2 \ll d^2$.
Recall that $\gamma := \alpha / \|\alpha\|$, and that ■/■/■ may represent increments or residuals.

Xiaobo Liu, xliu@mpi-magdeburg.mpg.de

Figure: Increment-based RRE for low-rank matrix sequences, where $\tau_i := \gamma_i + \cdots + \gamma_n$.

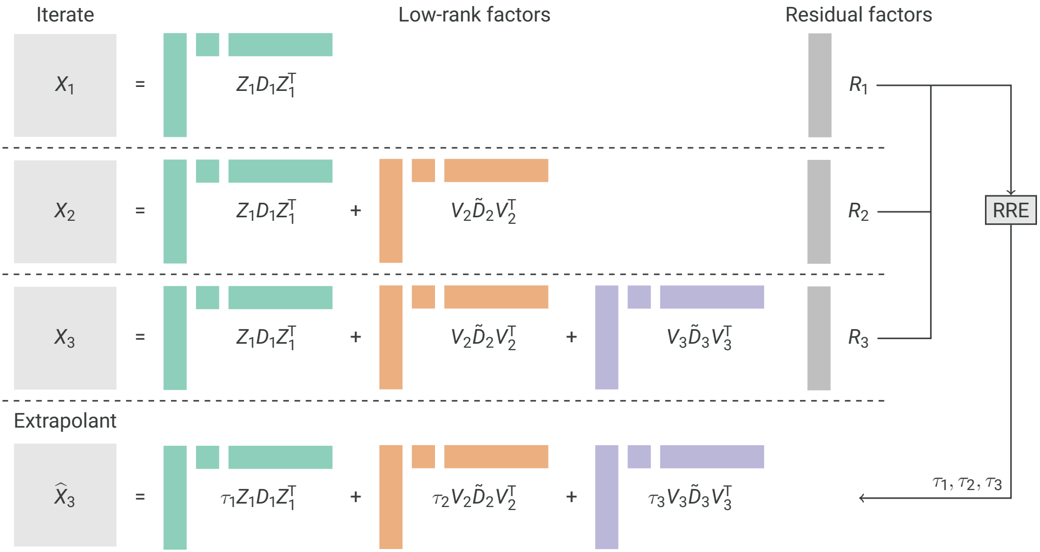


Figure: Residual-based RRE for low-rank matrix sequences, where $\tau_i := \gamma_i + \cdots + \gamma_n$.

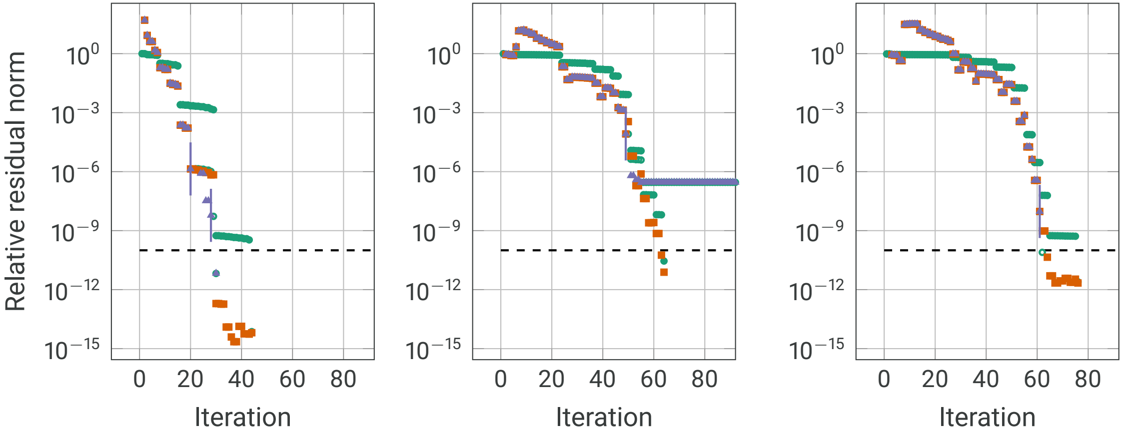
1. Reduced Rank Extrapolation
2. Methodological Extensions
3. Numerical Experiments
4. Conclusions

Numerical Experiments

- Underlying equation: algebraic Riccati equation (ARE)

$$C^{\mathsf{T}} C + A^{\mathsf{T}} X E + E^{\mathsf{T}} X A - E^{\mathsf{T}} X B H^{-1} B^{\mathsf{T}} X E = 0 \in \mathbb{R}^{d\times d}$$

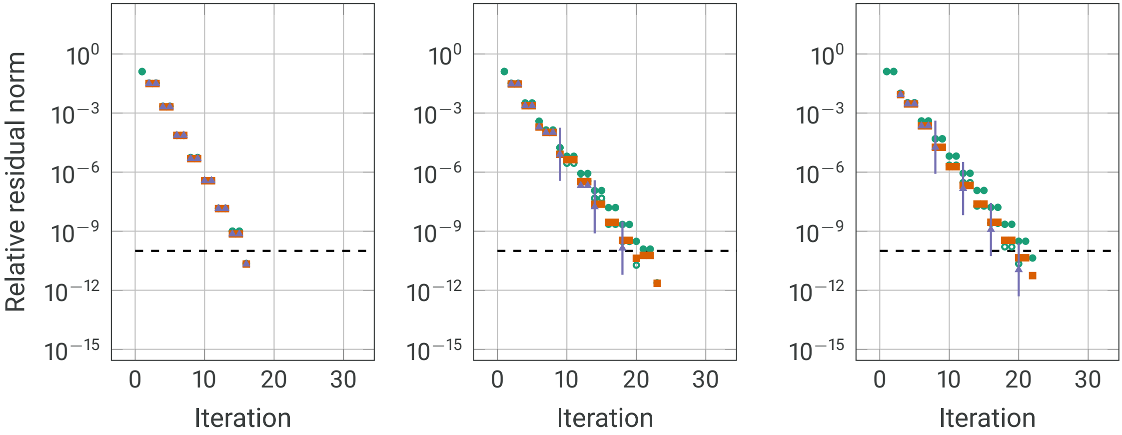
- $X \approx Z\, D\, Z^{\mathsf{T}}$ and $\mathcal{R}(X) \approx R\, T\, R^{\mathsf{T}}$ **[Benner & Bujanović, ’16]**
- Iterative process: RIccati Alternating Directions Implicit (RADI) **[Benner et al., ’18]**
- Implementation derived from M-M.E.S.S (The Matrix Equation Sparse Solvers library) **[Saak, Köhler, & Benner, ’25]**



(a) $q = 1$. (b) $q = 20$. (c) $q = 40$.

Figure: RADI (●○), non-cycling RRE (■), and cycling RRE (▲) applied to nonlinear *Toeplitz* example **[Benner et al., '20]**, $d = 100\,000$, for varying numbers of outputs; $C \in \mathbb{R}^{q \times d}$.

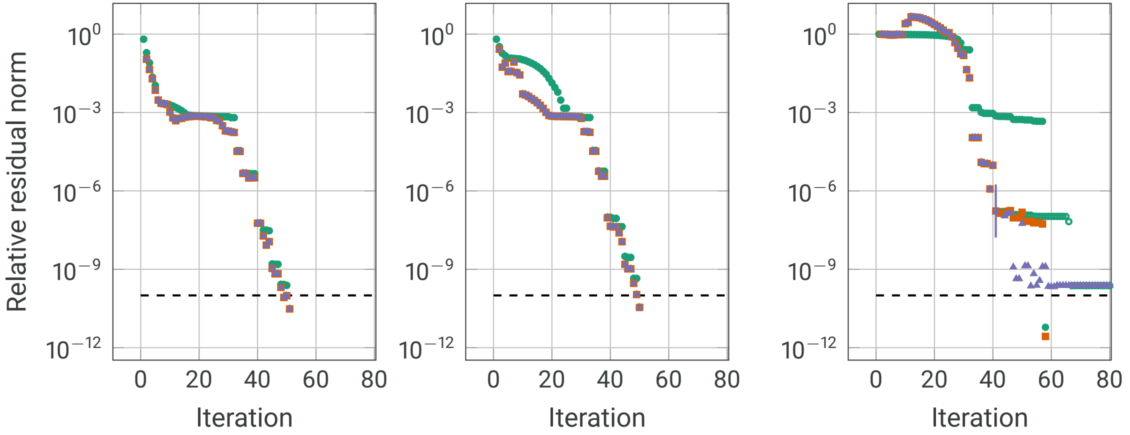
- Stagnation of cycling mode is likely due to non-PSD res. of the extrapolant (as initializer).



(a) $q = 1$. (b) $q = 20$. (c) $q = 40$.

Figure: ADI (●○), non-cycling RRE (■), and cycling RRE (▲) applied to *Toeplitz* example **[Benner et al., '20]**, $d = 10^5$ and **B=0** ($\Rightarrow$ Algebraic Lyapunov Equation (ALE)), for varying outputs q; $C \in \mathbb{R}^{q \times d}$.

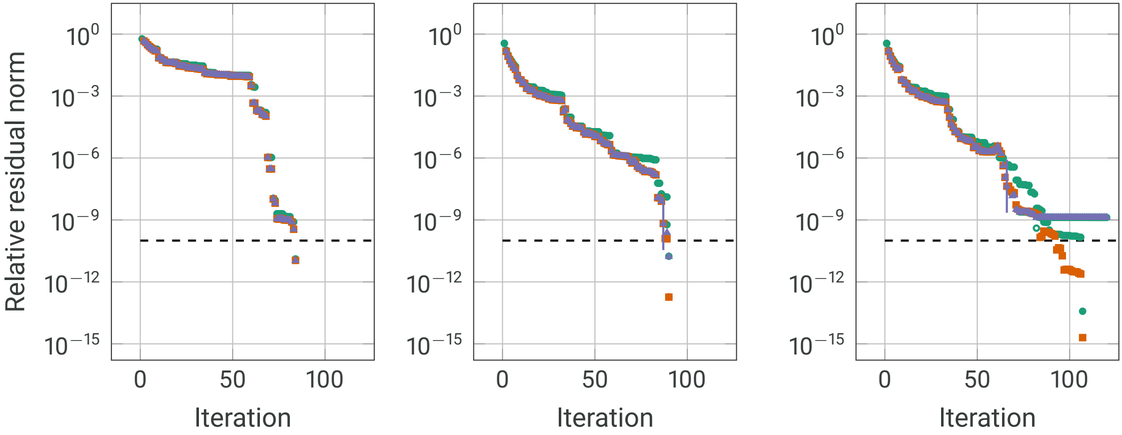
- No stagnation observed in this linear case. More benefits of RRE with larger q.



(a) $\lambda = 10^0$. (b) $\lambda = 10^{-2}$. (c) $\lambda = 10^{-4}$.

Figure: RADI (●○), non-cycling RRE (■), and cycling RRE (▲) applied to *Chip* example **[Moosmann et al., '04]**, $d = 20\,082$, for varying contributions of the quadratic factor matrix $H^{-1} = \lambda^{-1} I_p$.

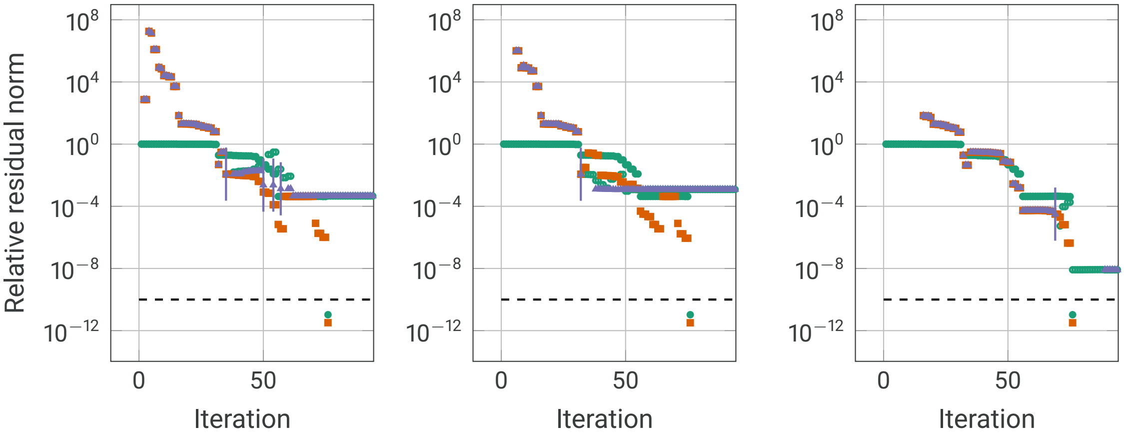
- ARE more nonlinear with smaller λ, greater benefits of RRE observed.



(a) ADI for Controllability ALE (b) ADI for Observability ALE (c) RADI for ARE

Figure: (R)ADI (●○), non-cycling RRE (■), and cycling RRE (▲) applied to *Steel Profile* example **[Benner & Saak, '05]**, $d = 317\,377$.

- Little acceleration of RRE for ALEs (the ADI iterates near optimal?); clear benefits for ARE.



(a) $n = 3$.

(b) $n = 6$.

(c) $n = 12$.

Figure: RADI (●○), non-cycling RRE (■), and cycling RRE (▲) applied to *Triple Chain* mass-spring-damper example [Truhar & Veselić, '09], $d = 60\,002$, with varying RRE window size $n \in \mathbb{N}^+$.

- Residual of non-cycling mode RRE decreases faster for larger n.

Conclusions

- Proposed two extensions to RRE, applicable to any sequence of low-rank matrices:
  1. Nonstationary fixed-point iterations
  2. Low-rank matrix sequences
- Demonstrated feasibility of RADI + non-cycling RRE for ARE
  under a simple condition on RRE coefficients γ
  for various system behaviors, parameters, and equation types
- Problem: non-PSD residuals
  which leads to stagnation for RADI + cycling RRE for nonlinear equations

Outlook:

- Convergence analysis
- Other equation types: multi-term Sylvester, Lyapunov-plus-positive, etc.

https://arxiv.org/abs/2502.09165

Peter Benner and Zvonimir Bujanović.
On the solution of large-scale algebraic riccati equations by using low-dimensional invariant subspaces.
Linear Algebra Appl., 488:430–459, 2016.

Peter Benner, Zvonimir Bujanović, Patrick Kürschner, and Jens Saak.
RADI: a low-rank ADI-type algorithm for large scale algebraic Riccati equations.
Numer. Math., 138:301–330, 2018.

Peter Benner, Zvonimir Bujanović, Patrick Kürschner, and Jens Saak.
A numerical comparison of different solvers for large-scale, continuous-time algebraic Riccati equations and LQR problems.
SIAM J. Sci. Comput., 42(2):A957–A996, 2020.

Peter Benner and Jens Saak.
Linear-quadratic regulator design for optimal cooling of steel profiles.
Technical Report SFB393/05-05, Sonderforschungsbereich 393 *Parallele Numerische Simulation für Physik und Kontinuumsmechanik*, TU Chemnitz, D-09107 Chemnitz (Germany), 2005.

Robert P. Eddy.
Extrapolating to the limit of a vector sequence.
In *Information linkage between applied mathematics and industry*, pages 387–396. Elsevier, 1979.

Bibliographies II

S. Kaniel and J. Stein.
Least-square acceleration of iterative methods for linear equations.
J. Opt. Th. Appl., 14:431–437, 1974.

Marián Mešina.
Convergence acceleration for the iterative solution of the equations $X = AX + f$.
Comp. Meth. Appl. Mech. Eng., 10(2):165–173, 1977.

Christian Moosmann, Evgenii B. Rudnyi, Andreas Greiner, and Jan G. Korvink.
Model order reduction for linear convective thermal flow.
In *Proceedings of 10th International Workshops on THERMal INvestigations of ICs and Systems*, 2004.

Jens Saak and Martin Köhler and Peter Benner
M-M.E.S.S. – The Matrix Equations Sparse Solvers library.

Avram Sidi.
Extrapolation vs. projection methods for linear systems of equations.
J. Comput. Appl. Math., 22(1):71–88, 1988.

Avram Sidi.
Efficient implementation of minimal polynomial and reduced rank extrapolation methods.
J. Comput. Appl. Math., 36(3):305–337, 1991.

Bibliographies III

Avram Sidi.
Minimal polynomial and reduced rank extrapolation methods are related.
Adv. Comput. Math., 43:151–170, 2017.

Avram Sidi.
A convergence study for reduced rank extrapolation on nonlinear systems.
Numer. Algorithms, 84:957–982, 2020.

Avram Sidi and Yair Shapira.
Upper bounds for convergence rates of acceleration methods with initial iterations.
Numer. Algorithms, 18(2):113–132, 1998.

Ninoslav Truhar and Krešimir Veselić.
An efficient method for estimating the optimal dampers' viscosity for linear vibrating systems using Lyapunov equation.
SIAM J. Matrix Anal. Appl., 31(1):18–39, 2009.