Restarting for the Tensor Infinite Arnoldi

Giampaolo Mele

KTH Royal Institute of technology
Dept. Math, Numerical analysis group

11 July 2016

Joint work with Elias Jarlebring

ILAS 2016, Leuven

The nonlinear eigenvalue problem

Find $\lambda \in \mathbb{C}$, $v \neq 0$ such that

$$M(\lambda)v = 0$$

where M analytic in a disk $\Omega \subset \mathbb{C}$.

Selection of interesting works

[Ruhe '73], [Mehrmann, Voss '04], [Lancaster '02], [Tisseur, et al. '01], [Voss '05], [Unger '50], [Mackey, et al. '09], [Kressner '09], [Bai, et al. '05], [Meerbergen '09], [Breda, et al. '06], [Betcke, et al. '04, '10], [Asakura, et a. '10], [Beyn '12], [Szyld, Xue '13], [Hochstenbach, et al. '08], [Neumaier '85], [Gohberg, et al. '82], [Effenberger '13], [Van Beeumen, et al '15] . . .

Outline

- Tensor Infinite Arnoldi framework
- Restarting techniques for Tensor Infinite Arnoldi
- Numerical experiments

Properties / features of infinite Arnoldi method

- Equivalent to Arnoldi's method on a companion matrix, for any truncation parameter N with $N > k$
- Equivalent to Arnoldi's method on an operator $\mathcal{B}$
- Convergence theory (?)
- Requires adaption of computation of y_0.

$$y_0 = M(\hat{\lambda})^{-1}(M'(\hat{\lambda})x_1 + \cdots + M^{(k)}(\hat{\lambda})x_k)$$

- Complexity of orthogonalization at step k: $O(k^2 n)$

Described in, e.g. : [Jarlebring, et al. '12, '14, '15]

By Taylor expansion

$$M(\lambda) = \sum_{i=0}^{\infty} \frac{M_i}{i!} \lambda^i$$

Associated operator

$$\mathcal{B}(\psi)(\theta) := \int_0^{\theta} \psi(\hat{\theta}) d\hat{\theta} + C(\psi)$$

where

$$C(\psi) := -M_0^{-1} \sum_{i=1}^{\infty} \frac{M_i}{i!} \psi^{(i)}(0)$$

Theorem (operator equivalence) [Jarlebring, et al. '12]

$$M(\lambda)v = 0 \iff \lambda \mathcal{B}\psi = \psi$$

where $\psi(\theta) = ve^{\lambda\theta}$

Approximate the eigenfunctions in

$$K_{m+1}(\mathcal{B}, \psi) = \text{span}\left(\psi, \mathcal{B}\psi, \ldots, \mathcal{B}^m\psi\right)$$

Algorithm 1 Arnoldi method for $\mathcal{B}$

Require: $<\psi_1, \psi_1>= 1$

for $k = 1$ to m **do**

  $\phi = \mathcal{B}\psi_k$

  for $i = 1$ to k **do**

    $h_{i,k} =<\phi, \psi_i>$, $\phi = \phi - h_{i,k}\psi_i$

  end for

  $h_{k+1,k} = \sqrt{<\phi, \phi>}$, $\psi_{k+1} = \phi / h_{k+1,k}$

end for

Arnoldi factorization

$$\mathcal{B}\Psi_m = \Psi_{m+1} H_{m+1,m}$$

with $\Psi_{m+1} := (\psi_1, \ldots, \psi_{m+1})$

Structured functions

$$\psi_j(\theta) = \sum_{i=0}^{d-1} v_i^j \theta^i + Y \exp_{d-1}(\theta S) c_j, \qquad \exp_{d-1}(\theta S) := \sum_{i=d}^{\infty} \theta^i S^i$$

Tensor Structured functions

$$\psi_j(\theta) = P_{d-1}(\theta) \left(\sum_{\ell=1}^{r} a_{:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} b_{:,\ell} \otimes w_\ell \right) + Y \exp_{d-1}(\theta S) c_j$$

with $P_d(\theta) := (1, \theta, \ldots, \theta^d) \otimes I_n$

- $z_1, \ldots, z_r, w_1, \ldots, w_p$ are orthonormal,
- $\mathrm{span}(w_1, \ldots, w_p) = \mathrm{span}(Y)$,
- $< \sum_{i=0}^{\infty} \alpha_i \theta^i, \sum_{i=0}^{\infty} \beta_i \theta^i >= \sum_{i=0}^{\infty} \alpha_i \beta_i$.

Structured functions

$$\Psi_k(\theta) = \sum_{i=0}^{d-1} V_i \theta^i + Y \exp_{d-1}(\theta S) C, \qquad \exp_{d-1}(\theta S) := \sum_{i=d}^{\infty} \theta^i S^i$$

Tensor Structured functions

$$\Psi_k(\theta) = P_{d-1}(\theta) \left(\sum_{\ell=1}^{r} a_{:,:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell} \otimes w_\ell \right) + Y \exp_{d-1}(\theta S) C$$

with $P_d(\theta) := (1, \theta, \dots, \theta^d) \otimes I_n$

- $z_1, \dots, z_r, w_1, \dots, w_p$ are orthonormal,
- $\mathrm{span}(w_1, \dots, w_p) = \mathrm{span}(Y)$,
- $< \sum_{i=0}^{\infty} \alpha_i \theta^i, \sum_{i=0}^{\infty} \beta_i \theta^i >= \sum_{i=0}^{\infty} \alpha_i \beta_i$.

Action of $\mathcal{B}$ on tensor structured functions

$$\psi_k(\theta) = P_{d-1}(\theta)\left(\sum_{\ell=1}^{r} a_{:,\ell}\otimes z_\ell + \sum_{\ell=1}^{p} b_{:,\ell}\otimes w_\ell\right) + Y\exp_{d-1}(\theta S)c_j$$

$$\mathcal{B}\psi_k(\theta) = P_{d}(\theta)\left(\sum_{\ell=1}^{r+1} \tilde{a}_{:,\ell}\otimes z_\ell + \sum_{\ell=1}^{p} \tilde{b}_{:,\ell}\otimes w_\ell\right) + Y\exp_{d}(\theta S)\tilde{c}$$

$$\tilde{a}_{i,r+1} := 0, \qquad \tilde{a}_{i+1,\ell} := a_{i,\ell}/i, \qquad \tilde{b}_{i+1,\bar{\ell}} := b_{i,\bar{\ell}}/i,$$

$$i = 1,\dots,d, \qquad \ell = 1,\dots,r, \qquad \bar{\ell} = 1,\dots,p,$$

$$\tilde{z} := -M_0^{-1}\left[\sum_{i=d+1}^{\infty}\frac{M_i Y S^i}{i!}\tilde{c} - \sum_{i=1}^{d} M_i\left(\sum_{i=1}^{r}\frac{a_{i,\ell}}{i}z_\ell + \sum_{\ell=1}^{p}\frac{b_{i,\ell}}{i}w_\ell\right)\right],$$

$$\tilde{z} = \sum_{i=1}^{r+1}\tilde{a}_{1,\ell}z_i + \sum_{i=1}^{p}\tilde{b}_{1,\ell}w_i \quad \text{(G–S orth.)}, \qquad \tilde{c} = S^{-1}c.$$

Action of $\mathcal{B}$ on tensor structured functions

$$\psi_k(\theta) = P_{d-1}(\theta)\left(\sum_{\ell=1}^{r} a_{:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} b_{:,\ell} \otimes w_\ell\right) + Y \exp_{d-1}(\theta S) c_j$$

$$\mathcal{B}\psi_k(\theta) = P_{d}(\theta)\left(\sum_{\ell=1}^{r+1} \tilde{a}_{:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} \tilde{b}_{:,\ell} \otimes w_\ell\right) + Y \exp_{d}(\theta S) \tilde{c}$$

$$\tilde{a}_{i,r+1} := 0, \qquad \tilde{a}_{i+1,\ell} := a_{i,\ell}/i, \qquad \tilde{b}_{i+1,\bar{\ell}} := b_{i,\bar{\ell}}/i,$$

$$i = 1,\ldots,d, \qquad \ell = 1,\ldots,r, \qquad \bar{\ell} = 1,\ldots,p,$$

$$\tilde{z} := -M_0^{-1}\left[\sum_{i=d+1}^{\infty} \frac{M_i Y S^i}{i!}\tilde{c} - \sum_{i=1}^{d} M_i \left(\sum_{i=1}^{r} \frac{a_{i,\ell}}{i} z_\ell + \sum_{\ell=1}^{p} \frac{b_{i,\ell}}{i} w_\ell\right)\right],$$

$$\tilde{z} = \sum_{i=1}^{r+1} \tilde{a}_{1,\ell} z_i + \sum_{i=1}^{p} \tilde{b}_{1,\ell} w_i \quad \text{(G–S orth.)}, \qquad \tilde{c} = S^{-1} c.$$

Extend the polynomial part

$$\begin{aligned}
\Psi_k(\theta) &= P_{d-1}(\theta)\left(\sum_{\ell=1}^{r+1} a_{:,:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell} \otimes w_\ell\right) + Y \exp_{d-1}(\theta S) C \\
&= P_{d}(\theta)\left(\sum_{\ell=1}^{r+1} a_{:,:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell} \otimes w_\ell\right) + Y \exp_{d}(\theta S) C
\end{aligned}$$

$$E := \frac{W^H Y S^d C}{d!}$$

$$\begin{aligned}
a_{d,j,\ell} &:= 0 & \ell &= 1, \dots, r+1 \\
b_{d,j,\ell} &:= e_{\ell,j} & \ell &= 1, \dots, p \\
& & j &= 1, \dots, k
\end{aligned}$$

Extend the polynomial part

$$\begin{aligned}\Psi_k(\theta) &= P_{d-1}(\theta)\left(\sum_{\ell=1}^{r+1} a_{:,:,\ell}\otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell}\otimes w_\ell\right) + Y\exp_{d-1}(\theta S)C \\ &= P_{d}(\theta)\left(\sum_{\ell=1}^{r+1} a_{:,:,\ell}\otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell}\otimes w_\ell\right) + Y\exp_{d}(\theta S)C\end{aligned}$$

$$E := \frac{W^H Y S^d C}{d!}$$

$$\begin{aligned} a_{d,j,\ell} &:= 0 & \ell &= 1,\dots,r+1 \\ b_{d,j,\ell} &:= e_{\ell,j} & \ell &= 1,\dots,p \\ & & j &= 1,\dots,k \end{aligned}$$

Orthogonalization in the Arnoldi process

$$\Psi_k(\theta) = P_d(\theta)\left(\sum_{\ell=1}^{r+1} a_{:,:,\ell}\otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell}\otimes w_\ell\right) + Y\exp_d(\theta S)C$$

$$\mathcal{B}\psi_k(\theta)^\perp = P_d(\theta)\left(\sum_{\ell=1}^{r+1} a^\perp_{:,\ell}\otimes z_\ell + \sum_{\ell=1}^{p} b^\perp_{:,\ell}\otimes w_\ell\right) + Y\exp_d(\theta S)c^\perp$$

$$h = \sum_{\ell=1}^{r}(a_{:,:,\ell})^H\bar{a}_{:,\ell} + \sum_{\ell=1}^{p}(b_{:,:,\ell})^H\bar{b}_{:,\ell} + \sum_{i=d}^{\infty} C^H\frac{(S^i)^H Y^H Y S^i}{(i!)^2}\bar{c},$$

$$a^\perp_{:,\ell} = \bar{a}_{:,\ell} - a_{:,:,\ell}h, \qquad b^\perp_{:,\bar{\ell}} = \bar{b}_{:,\bar{\ell}} - b_{:,:,\bar{\ell}}h, \qquad c^\perp = \bar{c} - Ch,$$

$$\ell = 1,\ldots,r+1, \qquad \bar{\ell} = 1,\ldots,p,$$

$$\beta := \sqrt{\|b^\perp\|_F^2 + \|a^\perp\|_F^2 + \sum_{i=d}^{\infty}\frac{(c^\perp)^H (S^i)^H Y^H Y S^i c^\perp}{(i!)^2}}$$

Orthogonalization in the Arnoldi process

$$\Psi_k(\theta) = P_d(\theta)\left(\sum_{\ell=1}^{r+1} a_{:,:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell} \otimes w_\ell\right) + Y \exp_d(\theta S) C$$

$$\mathcal{B}\psi_k(\theta)^\perp = P_d(\theta)\left(\sum_{\ell=1}^{r+1} a^\perp_{:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} b^\perp_{:,\ell} \otimes w_\ell\right) + Y \exp_d(\theta S) c^\perp$$

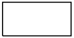
$$h = \sum_{\ell=1}^{r} (a_{:,:,\ell})^H \bar{a}_{:,\ell} + \sum_{\ell=1}^{p} (b_{:,:,\ell})^H \bar{b}_{:,\ell} + \sum_{i=d}^{\infty} C^H \frac{(S^i)^H Y^H Y S^i}{(i!)^2} \bar{c},$$

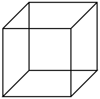
$$a^\perp_{:,\ell} = \bar{a}_{:,\ell} - a_{:,:,\ell} h, \qquad b^\perp_{:,\bar{\ell}} = \bar{b}_{:,\bar{\ell}} - b_{:,:,\bar{\ell}} h, \qquad c^\perp = \bar{c} - Ch,$$

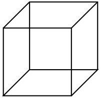
$$\ell = 1, \ldots, r+1, \qquad \bar{\ell} = 1, \ldots, p,$$

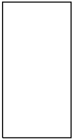
$$\beta := \sqrt{\|b^\perp\|_F^2 + \|a^\perp\|_F^2 + \sum_{i=d}^{\infty} \frac{(c^\perp)^H (S^i)^H Y^H Y S^i c^\perp}{(i!)^2}}$$

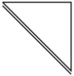
TIAR factorization

$C =$ 

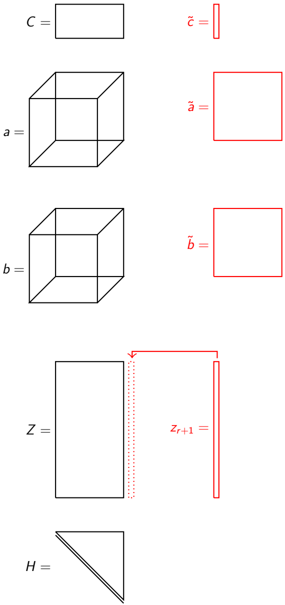
$a =$ 

$b =$ 

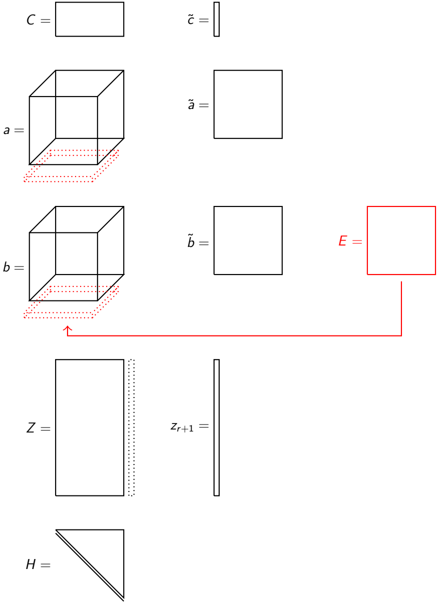
$Z =$ 

$H =$ 

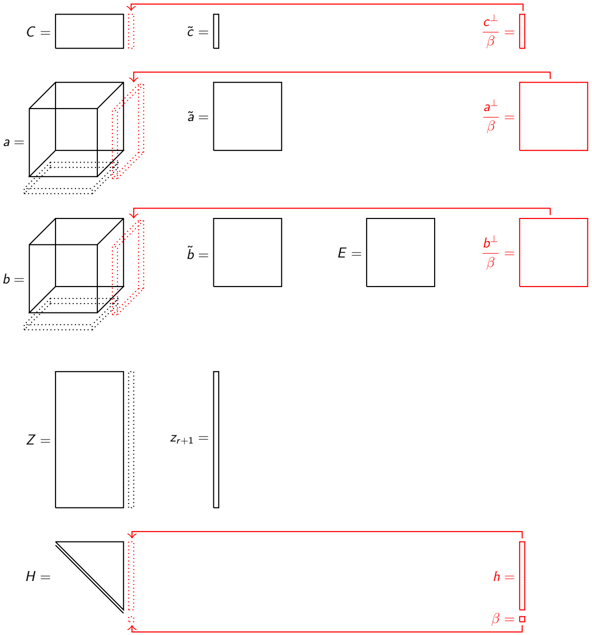
Action of $\mathcal{B}$



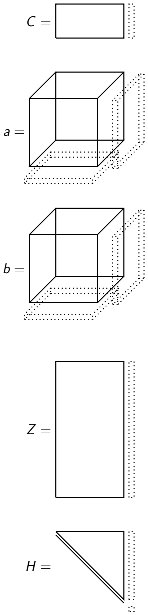
Expand polynomial part



Orthogonalization



Expanded TIAR factorization



Theorem: Krylov–Schur restart

Let

$$\mathcal{B}\Psi_m = \Psi_{m+1} H_{m+1,m}$$

be a TIAR factorization with Ritz values

$$\underbrace{\theta_1, \dots, \theta_p}_{\text{wanted}}, \underbrace{\theta_{p+1}, \dots, \theta_m}_{\text{unwanted}}$$

It exists Q and such that $\tilde{\Psi}_{p+1} := \Psi_{m+1} Q$ gives an TIAR factorization

$$\mathcal{B}\tilde{\Psi}_p = \tilde{\Psi}_{p+1} \tilde{H}_{p+1,p}$$

with Ritz values $\theta_1, \dots, \theta_p$.

Semi–explicit restart

Given $\mathcal{B}\tilde{\Psi}_p = \tilde{\Psi}_{p+1}\tilde{H}_{p+1,p}$, if $\theta_1, \dots, \theta_{p_\ell}$ converged,

$$\mathcal{B}\tilde{\Psi}_p = \tilde{\Psi}_{p+1}\begin{pmatrix} \Lambda & * \\ & \underline{\hat{H}} \end{pmatrix} \qquad \Lambda = \begin{pmatrix} \theta_1 & * & * \\ & \ddots & * \\ & & \theta_{p_\ell} \end{pmatrix}$$

Observation: $\mathcal{B}\tilde{\Psi}_{p_\ell} = \tilde{\Psi}_{p_\ell}\Lambda$ is an invariant pair.

Theorem: invariant pairs [Jarlebring, et al. '14]

If $\mathcal{B}\tilde{\Psi}_{p_\ell} = \tilde{\Psi}_{p_\ell}\Lambda$ then $\tilde{\Psi}_{p_\ell}(\theta) = \hat{Y}\exp(\theta\Lambda^{-1})$

Imposing the structure

$$\hat{Y} := \tilde{\Psi}_{p_\ell}(0), S := \begin{pmatrix} \Lambda & * \\ & \hat{H} \end{pmatrix}^{-1}, \Psi_{p_\ell+1} := \hat{Y}\exp(\hat{S}\theta)\begin{pmatrix} I_{p_\ell+1} \\ 0 \end{pmatrix}$$

$$\mathcal{B}\Psi_{p_\ell} = \Psi_{p_\ell+1}\begin{pmatrix} \Lambda \\ 0 \end{pmatrix}$$

Naive implicit restart

$$\boxed{\mathcal{B}\Psi_m = \Psi_{m+1} H_{m+1,m}} \longrightarrow \mathcal{B}\tilde{\Psi}_p = \tilde{\Psi}_{p+1} H_{p+1,p}$$

New TIAR factorization

$$\Psi_{m+1}(\theta) Q = \left[P_{d-1}(\theta) \left(\sum_{\ell=1}^{r} a_{:,:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell} \otimes w_\ell \right) + Y \exp_{d-1}(\theta S) C \right] Q$$

Naive implicit restart

$$\mathcal{B}\Psi_m = \Psi_{m+1} H_{m+1,m} \longrightarrow \mathcal{B}\tilde{\Psi}_p = \tilde{\Psi}_{p+1} H_{p+1,p}$$

New TIAR factorization

$$\Psi_{m+1}(\theta)Q = P_{d-1}(\theta)\left(\sum_{\ell=1}^{r} a_{:,:,\ell} Q \otimes z_\ell + \sum_{\ell=1}^{p} b_{:,:,\ell} Q \otimes w_\ell\right) + Y \exp_{d-1}(\theta S) C Q$$

Naive implicit restart

$$\mathcal{B}\Psi_m = \Psi_{m+1} H_{m+1,m} \longrightarrow \boxed{\mathcal{B}\tilde{\Psi}_p = \tilde{\Psi}_{p+1} H_{p+1,p}}$$

New TIAR factorization

$$\tilde{\Psi}_{p+1}(\theta) = P_{d-1}(\theta)\left(\sum_{\ell=1}^{r} \tilde{a}_{:,:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} \tilde{b}_{:,:,\ell} \otimes w_\ell\right) + Y \exp_{d-1}(\theta S)\tilde{C}$$

Naive implicit restart

$$\mathcal{B}\Psi_m = \Psi_{m+1} H_{m+1,m} \longrightarrow \boxed{\mathcal{B}\tilde{\Psi}_p = \tilde{\Psi}_{p+1} H_{p+1,p}}$$

New TIAR factorization

$$\tilde{\Psi}_{p+1}(\theta) = P_{d-1}(\theta)\left(\sum_{\ell=1}^{r} \tilde{a}_{:,:,\ell} \otimes z_\ell + \sum_{\ell=1}^{p} \tilde{b}_{:,:,\ell} \otimes w_\ell\right) + Y \exp_{d-1}(\theta S)\tilde{C}$$

Observation: the number of vectors $z_1, \ldots, z_r$ increase independently on the restarts.

Properties of TIAR factorization

Theorem

Let $\mathcal{B}\Psi_k = \Psi_{k+1}\underline{H}_k$ be a TIAR factorization, if $Y = W = 0$

$$\Psi_k(\theta) = P_{d-1}(\theta) \left(\sum_{\ell=1}^{r} a_{:,:,\ell} \otimes z_\ell \right)$$

- Fast decay :

$$\|a_{i,:,:}\| \leq \frac{C}{(i-1)!}$$

- Singular values decay :
  Let $A = [A_1, \ldots, A_d]$ such that $A_i := (a_{i,:,:})^T$ then

$$\sigma_{R+1} \leq C \frac{d-R-k+2}{(R-k+1)!}$$

Compression of TIAR factorization

Theorem

Let $\mathcal{B}\Psi_k = \Psi_{k+1}\underline{H}_k$ be a TIAR factorization, then

$$\Psi_k(\theta) = P_{d-1}(\theta)\left(\sum_{\ell=1}^{r} a_{:,:,\ell} \otimes z_\ell\right)$$

can be approximated as

$$\tilde{\Psi}_k(\theta) = P_{d-1}(\theta)\left(\sum_{\ell=1}^{\tilde{r}} \tilde{a}_{:,:,\ell} \otimes z_\ell\right)$$

such that $\tilde{r} << r$ and

$$\|\Psi_{k+1} - \tilde{\Psi}_{k+1}\| \leq \sqrt{(d+1)(k+1)}\sigma_{\tilde{r}+1}$$
$$\|\mathcal{B}\tilde{\Psi}_k - \tilde{\Psi}_{k+1}\underline{H}_k\| \leq \sqrt{k}C\sigma_{\tilde{r}+1}$$

Idea: replace the unfolding of "a" with a low rank approximation.

Reduce the degree

Theorem

Let $\mathcal{B}\Psi_k = \Psi_{k+1}\underline{H}_k$ be a TIAR factorization, then

$$\Psi_k(\theta) = P_{d-1}(\theta) \left(\sum_{\ell=1}^{r} a_{:,:,\ell} \otimes z_\ell \right)$$

can be approximated as

$$\tilde{\Psi}_k(\theta) = P_{\tilde{d}-1}(\theta) \left(\sum_{\ell=1}^{r} \tilde{a}_{:,:,\ell} \otimes z_\ell \right)$$

such that $\tilde{d} << d$ and

$$\|\tilde{\Psi}_{k+1} - \Psi_{k+1}\| \leq C_1 \sqrt{k+1} \frac{(d-\tilde{d})}{\tilde{d}!}$$

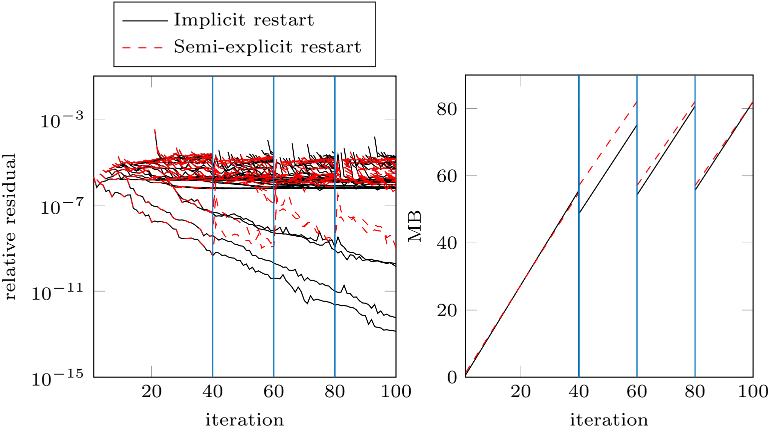
$$\|\mathcal{B}\tilde{\Psi}_k - \tilde{\Psi}_{k+1}\underline{H}_k\| \leq C_2 \sqrt{k+1} \frac{d-\tilde{d}}{(\tilde{d}+1)!}$$

Idea: Neglect the high terms in the polynomial, i.e., truncate of the tensor "a".

Comparison: implicit and semi–explicit

	Implicit	Semi–explicit
Singularities	✓	✗
p small	✗	✓
Slow convergence	✓	✗
Memory	✗	✓
Complexity	✓	✗

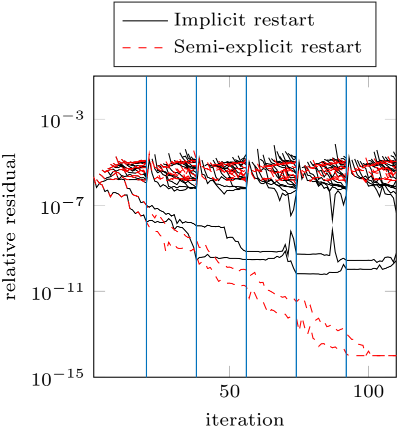
Waveguide eigenvalue problem



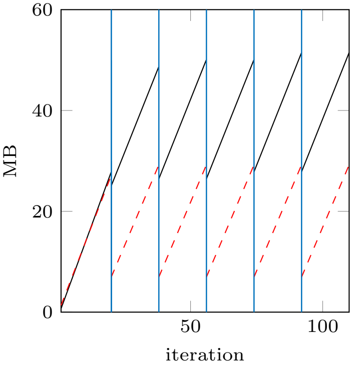
(a) Convergence (b) Memory

Size $n = 91203$ with $m = 40$, $p = 20$ and restart=4

Waveguide eigenvalue problem



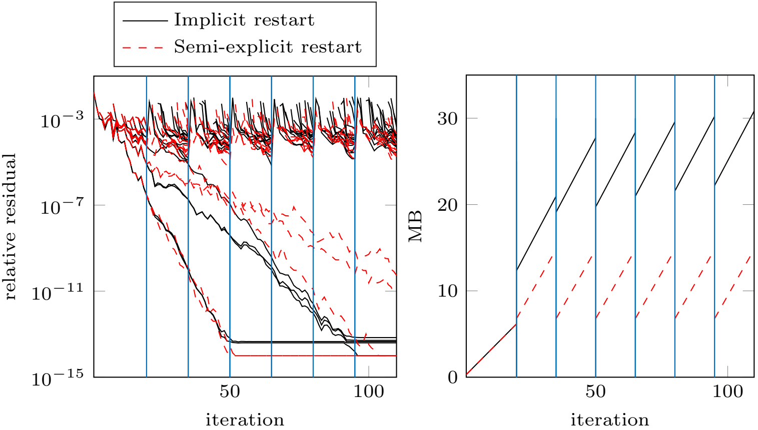
(a) Convergence



(b) Memory

Size $n = 91203$ with $m = 20$, $p = 4$ and restart=6

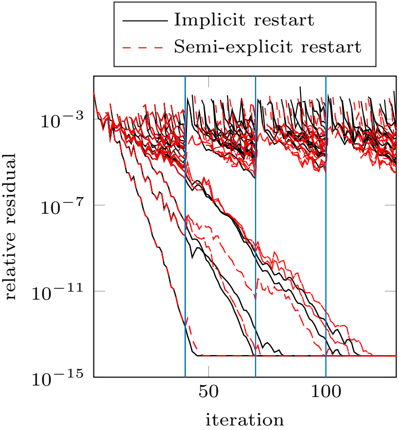
Delay eigenvalue problem



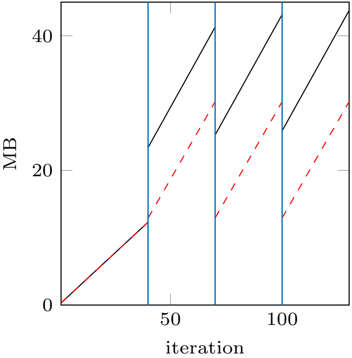
(a) Convergence (b) Memory

Size $n = 40401$ with $m = 20$, $p = 5$ and restart=7

Delay eigenvalue problem



(a) Convergence



(b) Memory

Size $n = 40401$ with $m = 40$, $p = 10$ and restart=4

Conclusion

Scientific contributions:

- ✐ extension of TIAR for tensor structured functions,
- ✐ implicit and semi–explicit restarts,
- ✐ bounds on the approximations.

Online material:

- Preprint:
  `http://arxiv.org/abs/1606.08595`