M\cr NA
Manchester Numerical Analysis

Stefan Güttel

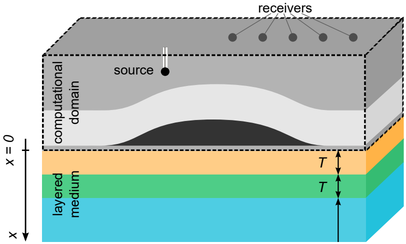
Compressing variable-coefficient exterior Helmholtz problems via RKFIT

Motivation

Consider acoustic waves in the Earth modelled by the Helmholtz equation

$$\Delta u + [k_\infty^2 + c(x)]u = f$$

on an unbounded domain $\Omega = [0, +\infty] \times [0, 1]^2$ with compactly supported offset function $c(x)$ for the wave number k_∞, & boundary conditions.

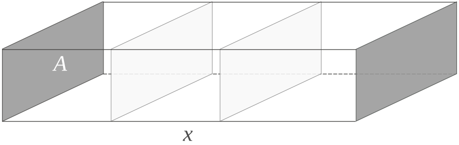


We want to discretize the layered medium in $x \geq 0$ using a "coarsest possible" finite difference grid in the x-direction.

A simple boundary value problem

For Hermitian $A \in \mathbb{C}^{N\times N}$ and $\{\mathbf{b}, \mathbf{u}(x)\} \subset \mathbb{C}^N$, $x \in [0, +\infty)$, consider the following **constant-coefficient BVP:**

$$\frac{\partial^2}{\partial x^2}\mathbf{u} = A\mathbf{u}, \quad \mathbf{u}(0) = \mathbf{u}_0, \quad \mathbf{u}(x) \text{ bounded as } x \to \infty.$$



Exact solution: $\mathbf{u}(x) = \exp(-xA^{1/2})\mathbf{u}_0$.

In particular: At $x = 0$ we simply have $\mathbf{u}'(0) = -A^{1/2}\mathbf{u}_0 = f(A)\mathbf{u}_0$.

Hence $f(A) = -A^{1/2}$ is the **Dirichlet-to-Neumann map.**

If A is discretization of differential operator on spatial domain $\Omega \subseteq \mathbb{R}^d$, BVP is a semidiscretization of a $(d+1)$-dimensional PDE on $[0, +\infty) \times \Omega$.

In the constant-coefficient indefinite Helmholtz case,

$$A \approx -\frac{\partial^2}{\partial^2 y_1} - \cdots - \frac{\partial^2}{\partial^2 y_d} - k_\infty^2.$$

The DtN operator $A^{1/2}$ "stores" all relevant information about the PDE solution on the unbounded domain $[0, +\infty) \times \Omega$.

If A is discretization of differential operator on spatial domain $\Omega \subseteq \mathbb{R}^d$, BVP is a semidiscretization of a $(d+1)$-dimensional PDE on $[0, +\infty) \times \Omega$.

In the constant-coefficient indefinite Helmholtz case,

$$A \approx -\frac{\partial^2}{\partial^2 y_1} - \cdots - \frac{\partial^2}{\partial^2 y_d} - k_\infty^2.$$

The DtN operator $A^{1/2}$ "stores" all relevant information about the PDE solution on the unbounded domain $[0, +\infty) \times \Omega$.

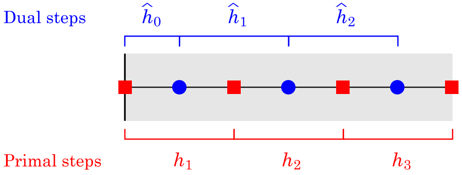
Using a rational approximation $R_n(A) \approx A^{1/2}$, one can convert R_n into a finite difference grid that approximates $A^{1/2}\mathbf{u}_0 \approx R_n(A)\mathbf{u}_0$.

Rational approximation $\iff$ finite difference grids

Consider the finite difference relations

$$\begin{aligned} \frac{1}{\widehat{h}_0}\left(\frac{\mathbf{u}_1 - \mathbf{u}_0}{h_1} - \mathbf{b}\right) &= A\mathbf{u}_0, \\ \frac{1}{\widehat{h}_j}\left(\frac{\mathbf{u}_{j+1} - \mathbf{u}_j}{h_{j+1}} - \frac{\mathbf{u}_j - \mathbf{u}_{j-1}}{h_j}\right) &= A\mathbf{u}_j, \quad j = 1, \ldots, n-1, \end{aligned}$$

with the convention that $\mathbf{u}_n = \mathbf{0}$.



Rational approximation $\iff$ finite difference grids

Can show via back-substitution that $\mathbf{b} = R_n(A)\mathbf{u}_0$ with

$$R_n(z) = \widehat{h}_0 z + \cfrac{1}{h_1 + \cfrac{1}{\widehat{h}_1 z + \cdots + \cfrac{1}{h_{n-1} + \cfrac{1}{\widehat{h}_{n-1} z + \cfrac{1}{h_n}}}}}.$$

Loosely speaking, the continued fraction $\widetilde{R}_n$ is equivalent to a finite difference grid implementation of a DtN map.

$\Longrightarrow$ **complex coordinate stretching/perfectly matched layers.**

[Engquist/Majda '74; Bérenger '94; Chew/Weedon '94; Guddati/Tassoulas '00; Ingerman/Druskin/Knizhnerman '00; Lisitsa 08; Appelö/Hagstrom '09; ...]

Our approach: Rational least squares approximation

We aim to solve

$$\|F\mathbf{b} - R_n(A)\mathbf{b}\|_2^2 \rightsquigarrow \min, \quad F = f(A),$$

with the minimum taken over all rational functions $R_n(z) = \frac{p_n(z)}{q_n(z)}$.

This is a nonlinear weighted rational least squares problem.

Observation

If $q_n(z) = \prod_{j=1}^{n}(z - \xi_j)$ was known, the LS problem became linear:

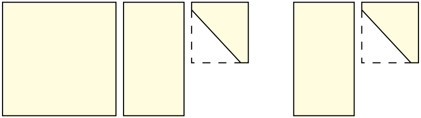
Find vector $R_n(A)\mathbf{b}$ via orthogonal projection onto *rational Krylov space*

$$\mathcal{Q}_{n+1}(A, \mathbf{b}, q_n) := q_n(A)^{-1} \underbrace{\operatorname{span}\{\mathbf{b}, A\mathbf{b}, \ldots, A^n\mathbf{b}\}}_{\mathcal{K}_{n+1}(A,\mathbf{b})}.$$

Rational Arnoldi decompositions

The rational Arnoldi algorithm [Ruhe 1994] is used to compute a rational Arnoldi decomposition of the form

$$A \quad V_{n+1} \; \underline{K_n} \; = \; V_{n+1} \; \underline{H_n}$$



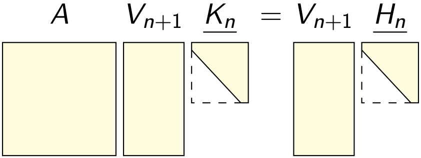
where

- the columns of V_{n+1} are an orthonormal basis of $\mathcal{Q}_{n+1}(A, \mathbf{b}, q_n)$,
- first column of V_{n+1} is $\mathbf{v}_1 = \mathbf{b}/\|\mathbf{b}\|_2$,
- $(\underline{H_n}, \underline{K_n})$ is unreduced upper-Hessenberg $(n+1) \times n$ pencil,
- the quotients $\{h_{j+1,j}/k_{j+1,j}\}_{j=1}^{n}$ are roots of $q_n(z) = \prod_{j=1}^{n}(z - \xi_j)$.

What can we say about the uniqueness of such decompositions?

Rational implicit-Q thm [Mach/Pranic/Vandebril 2014; Berljafa/G. 2015]

Let $A \in \mathbb{C}^{N \times N}$ satisfy an orthonormal rational Arnoldi decomposition $AV_{n+1}\underline{K_n} = V_{n+1}\underline{H_n}$ with poles $\xi_j = h_{j+1,j}/k_{j+1,j}$.
Then the orthonormal matrix V_{n+1} and the pencil $\left(\underline{H_n}, \underline{K_n}\right)$ are essentially uniquely determined by the first column of V_{n+1} and the poles $\xi_1, \ldots, \xi_n$.

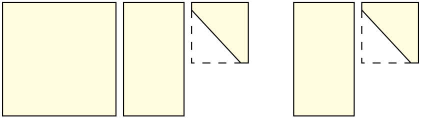


Rational implicit-Q thm [Mach/Pranic/Vandebril 2014; Berljafa/G. 2015]

Let $A \in \mathbb{C}^{N\times N}$ satisfy an orthonormal rational Arnoldi decomposition $AV_{n+1}\underline{K_n} = V_{n+1}\underline{H_n}$ with poles $\xi_j = h_{j+1,j}/k_{j+1,j}$.
Then the orthonormal matrix V_{n+1} and the pencil $(\underline{H_n}, \underline{K_n})$ are essentially uniquely determined by the first column of V_{n+1} and the poles $\xi_1, \dots, \xi_n$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{n+1}:

$$A \quad V_{n+1} \quad \underline{K_n} \quad = \quad V_{n+1} \quad \underline{H_n}$$



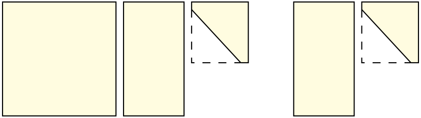
Rational implicit-Q thm [Mach/Pranic/Vandebril 2014; Berljafa/G. 2015]

Let $A \in \mathbb{C}^{N\times N}$ satisfy an orthonormal rational Arnoldi decomposition $AV_{n+1}\underline{K_n} = V_{n+1}\underline{H_n}$ with poles $\xi_j = h_{j+1,j}/k_{j+1,j}$.
Then the orthonormal matrix V_{n+1} and the pencil $(\underline{H_n}, \underline{K_n})$ are essentially uniquely determined by the first column of V_{n+1} and the poles $\xi_1, \ldots, \xi_n$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{n+1}:

insert unitary

$$A \quad V_{n+1} \overbrace{\qquad}^{P_{n+1}P^*_{n+1}} \underline{K_n} = V_{n+1} \overbrace{\qquad}^{P_{n+1}P^*_{n+1}} \underline{H_n}$$



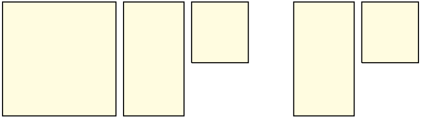
Rational implicit-Q thm [Mach/Pranic/Vandebril 2014; Berljafa/G. 2015]

Let $A \in \mathbb{C}^{N\times N}$ satisfy an orthonormal rational Arnoldi decomposition $AV_{n+1}\underline{K_n} = V_{n+1}\underline{H_n}$ with poles $\xi_j = h_{j+1,j}/k_{j+1,j}$.
Then the orthonormal matrix V_{n+1} and the pencil $\left(\underline{H_n}, \underline{K_n}\right)$ are essentially uniquely determined by the first column of V_{n+1} and the poles $\xi_1, \ldots, \xi_n$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{n+1}:

change basis to $\widetilde{V}_{n+1} = V_{n+1}P_{n+1}$

$$A \quad \widetilde{V}_{n+1} \; \underline{\widetilde{K}_n} \; = \; \widetilde{V}_{n+1} \; \underline{\widetilde{H}_n}$$



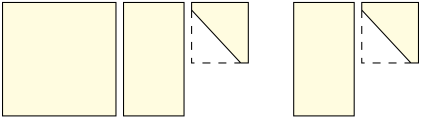
Rational implicit-Q thm [Mach/Pranic/Vandebril 2014; Berljafa/G. 2015]

Let $A \in \mathbb{C}^{N\times N}$ satisfy an orthonormal rational Arnoldi decomposition $AV_{n+1}\underline{K_n} = V_{n+1}\underline{H_n}$ with poles $\xi_j = h_{j+1,j}/k_{j+1,j}$.
Then the orthonormal matrix V_{n+1} and the pencil $(\underline{H_n}, \underline{K_n})$ are essentially uniquely determined by the first column of V_{n+1} and the poles $\xi_1, \ldots, \xi_n$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{n+1}:

QZ transform on lower $n \times n$ part of pencil $(\underline{\widetilde{H}_n}, \underline{\widetilde{K}_n})$

$$A \quad \widehat{V}_{n+1} \; \underline{\widehat{K}_n} \; = \; \widehat{V}_{n+1} \; \underline{\widehat{H}_n}$$

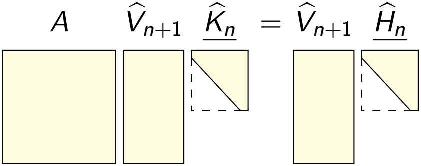


Rational implicit-Q thm [Mach/Pranic/Vandebril 2014; Berljafa/G. 2015]

Let $A \in \mathbb{C}^{N \times N}$ satisfy an orthonormal rational Arnoldi decomposition $AV_{n+1}\underline{K_n} = V_{n+1}\underline{H_n}$ with poles $\xi_j = h_{j+1,j}/k_{j+1,j}$.
Then the orthonormal matrix V_{n+1} and the pencil $\left(\underline{H_n}, \underline{K_n}\right)$ are essentially uniquely determined by the first column of V_{n+1} and the poles $\xi_1, \ldots, \xi_n$.

$\Longrightarrow$ Allows us to move poles ξ_j by changing first column of V_{n+1}:

QZ transform on lower $n \times n$ part of pencil $(\underline{\widetilde{H}_n}, \underline{\widetilde{K}_n})$



Read off new poles $\widehat{\xi}_j := \widehat{h}_{j+1,j}/\widehat{k}_{j+1,j}$

Rational Krylov fitting $\|F\mathbf{b} - R_n(A)\mathbf{b}\|_2 \rightsquigarrow \min$

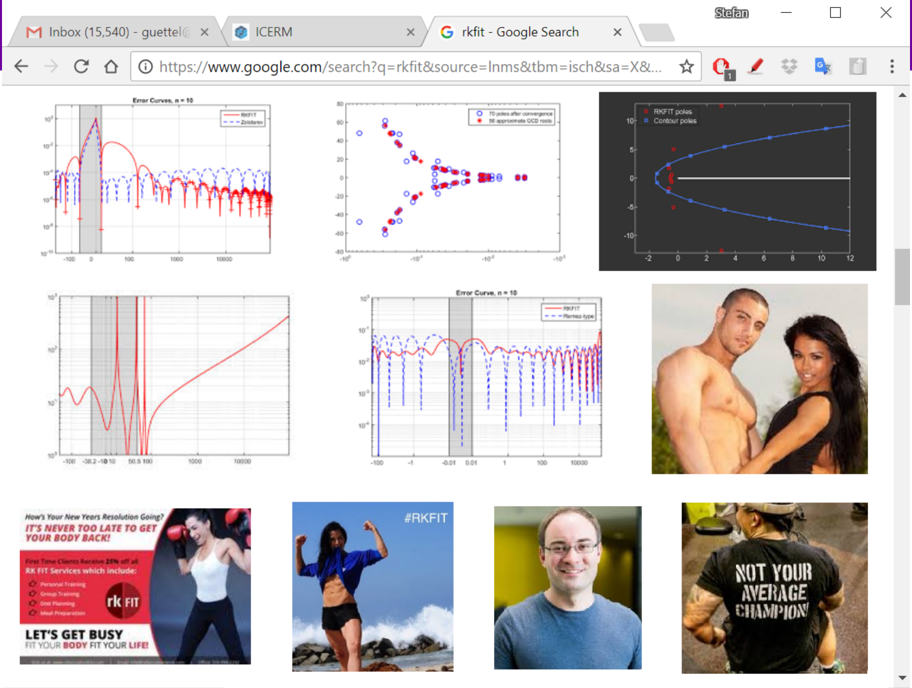
Take initial guess $q_n(z) = \prod_{j=1}^{n}(z - \xi_j)$ and iterate:

1. Compute orthonormal basis V_{n+1} for $\mathscr{Q}_{n+1}(A, \mathbf{b}, q_n)$.
2. Solve the following linear problem:

> Find $\widehat{\mathbf{v}} \in \mathscr{Q}_{n+1}(A, \mathbf{b}, q_n)$ such that $F\widehat{\mathbf{v}}$ is best approximated by an element of $\mathscr{Q}_{n+1}(A, \mathbf{b}, q_n)$, i.e.,
>
> $$\widehat{\mathbf{v}} = \underset{\substack{\mathbf{y}=V_{n+1}\mathbf{c} \\ \|\mathbf{y}\|_2=1}}{\operatorname{argmin}} \|(I - V_{n+1}V_{n+1}^*)F\mathbf{y}\|_2.$$

3. Move $\widehat{\mathbf{v}}$ to the first column of $\widehat{V}_{n+1}$ and find new poles $\widehat{\xi}_1, \ldots, \widehat{\xi}_n$.

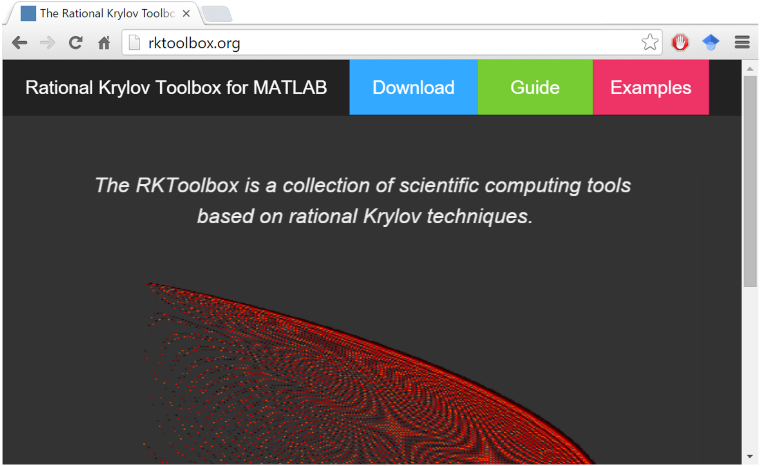
We call this algorithm **RKFIT** [Berljafa & G., SISC 2017].


Inbox (15,540) - guettel@
ICERM
rkfit - Google Search
Stefan
https://www.google.com/search?q=rkfit&source=lnms&tbm=isch&sa=X&...
Error Curves, n = 10
RKFIT
Zolotarev
70 poles after convergence
56 approximate GCD roots
RKFIT poles
Contour poles
Error Curve, n = 10
RKFIT
Remez-type
How's Your New Years Resolution Going?
IT'S NEVER TOO LATE TO GET YOUR BODY BACK!
First Time Clients Receive 25% off all
RK FIT Services which include:
Personal Training
Group Training
Diet Planning
Meal Preparation
rk FIT
LET'S GET BUSY
FIT YOUR BODY FIT YOUR LIFE!
#RKFIT
NOT YOUR AVERAGE CHAMPION

Rational Krylov Toolbox

RKFIT is part of our MATLAB Rational Krylov Toolbox:

`www.rktoolbox.org`



Convergence result: Exactness

Theorem

Assume that $\mathscr{K}_{2n+1}(A, \mathbf{b}) = \mathrm{span}\{\mathbf{b}, A\mathbf{b}, \ldots, A^{2n}\mathbf{b}\}$ is not A-invariant and that $F = p_n(A)q_n^\star(A)^{-1}$ for some $p_n, q_n^\star \in \mathscr{P}_n$. Then, in exact arithmetic, RKFIT will find $q_n^\star$ in a single iteration independent of the initial guess q_n.

Convergence result: Exactness

Theorem

Assume that $\mathscr{K}_{2n+1}(A, \mathbf{b}) = \mathrm{span}\{\mathbf{b}, A\mathbf{b}, \ldots, A^{2n}\mathbf{b}\}$ is not A-invariant and that $F = p_n(A) q_n^\star(A)^{-1}$ for some $p_n, q_n^\star \in \mathscr{P}_n$. Then, in exact arithmetic, RKFIT will find $q_n^\star$ in a single iteration independent of the initial guess q_n.

- A similar result has been shown in [Lefteriu/Antoulas 2013] for the vector fitting (VFIT) algorithm by [Gustavsen/Semlyen 1999].
- VFIT is based on a representation of p_n and q_n in barycentric form, with an implicit pole reallocation by changing weights.
- RKFIT differs from VFIT by its basis representation (partial fractions vs orthogonal rational functions) and the normalization of p_n/q_n.

Degree revelation

Theorem

Assume that $\mathscr{K}_{2n+1}(A, \mathbf{b})$ is not A-invariant and that $F = p_n(A)q_n^\star(A)^{-1}$ for some $p_n, q_n^\star \in \mathscr{P}_n$. Let V_{n+1} be an orthonormal basis of $\mathscr{Q}_{n+1}(A, \mathbf{b}, q_n)$ Let $M := (I - V_{n+1}V_{n+1}^*)FV_{n+1}$. Then $d = \dim(\operatorname{null} M) - 1$ is the largest integer such that F is of type $(n-d, n-d)$.

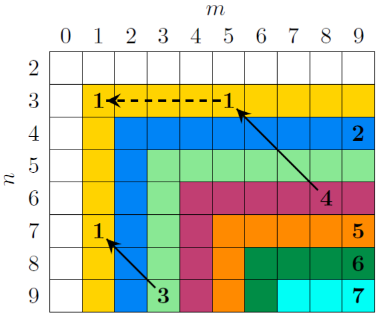
Degree revelation

Theorem

Assume that $\mathscr{K}_{2n+1}(A,\mathbf{b})$ is not A-invariant and that $F = p_n(A)q_n^\star(A)^{-1}$ for some $p_n, q_n^\star \in \mathscr{P}_n$. Let V_{n+1} be an orthonormal basis of $\mathscr{Q}_{n+1}(A,\mathbf{b},q_n)$

Let $M := (I - V_{n+1}V_{n+1}^*)FV_{n+1}$. Then $d = \dim(\mathrm{null}\, M) - 1$ is the largest integer such that F is of type $(n-d, n-d)$.

- We have implemented automatic degree reduction based on the numerical rank of M.
- Related ideas in numerical GCD and robust interpolation; e.g. [Gonnet/Pachón/Trefethen '14] [Beckermann/Labahn/Matos '17].



RKFUN format

When RKFIT stops with $R_n(A)\mathbf{b} \approx F\mathbf{b}$, it returns the approximant R_n represented in a rational Krylov basis:

$$R_n(z) = \sum_{j=0}^{n} c_j \cdot r_j(z) \equiv \underbrace{(\underline{H_n}, \underline{K_n}, \mathbf{c})}_{\text{RKFUN}}.$$

Here $\mathbf{c} = [c_0, c_1, \ldots, c_n]$ and the basis functions r_j are encoded in

$$z[r_0(z), r_1(z), \ldots, r_n(z)]\underline{K_n} = [r_0(z), r_1(z), \ldots, r_n(z)]\underline{H_n}, \;\; r_0 \equiv 1.$$

RKFUN methods currently implemented in RKToolbox:

`basis`, `coeffs`, `contfrac`, `diff`, `disp`, `double`, `ezplot`, `feval`, `hess`, `isreal`, `minus`, `mp`, `mrdivide`, `mtimes`, `plus`, `poles`, `poly`, `residue`, `roots`, `size`, `subsref`, `type`, `uminus`, `uplus`, `vpa`, ...

Example: RKFUN $\Longrightarrow$ continued fraction (`contfrac`)

1. Start with RKFUN $R_n \equiv (\underline{H_n}, \underline{K_n}, \mathbf{c})$ and associated Arnoldi relation

$$z[r_0, r_1, r_2, r_3, r_4] \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ & * & * & * \\ & & * & * \\ & & & * \end{bmatrix} = [r_0, r_1, r_2, r_3, r_4] \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ & * & * & * \\ & & * & * \\ & & & * \end{bmatrix}.$$

2. Left-multiply pencil by invertible matrix such that

$$z[{\color{red}R_n}, {\color{red}r_0}, r_1, r_2, r_3] \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} = [{\color{red}R_n}, {\color{red}r_0}, r_1, r_2, r_3] \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}.$$

Example: RKFUN $\Longrightarrow$ continued fraction (contfrac)

3. Apply sequence of left- and right-transformations to obtain:

$$z[R_n, r_0, \tilde{r}_1, \tilde{r}_2, \tilde{r}_3] \begin{bmatrix} 0 & & & \\ 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} = [R_n, r_0, \tilde{r}_1, \tilde{r}_2, \tilde{r}_3] \begin{bmatrix} 1 & & & \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}.$$

4. Run the two-sided Lanczos algorithm on the lower $n \times n$-part of $\underline{H_n}$ with $\mathbf{e}_1$ as left and right starting vector:

$$z[R_n, r_0, \hat{r}_1, \hat{r}_2, \hat{r}_3] \begin{bmatrix} 0 & & & \\ 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} = [R_n, r_0, \hat{r}_1, \hat{r}_2, \hat{r}_3] \begin{bmatrix} 1 & & & \\ * & * & & \\ * & * & * & \\ & * & * & * \\ & & * & * \end{bmatrix}.$$

Example: RKFUN $\Longrightarrow$ continued fraction (`contfrac`)

Using a classical connection (e.g., Gragg, Gutknecht, Brezinski, Bultheel) the tridiagonal matrix generated by two-sided Lanczos is related to continued fraction parameters of the Krylov basis functions. In particular,

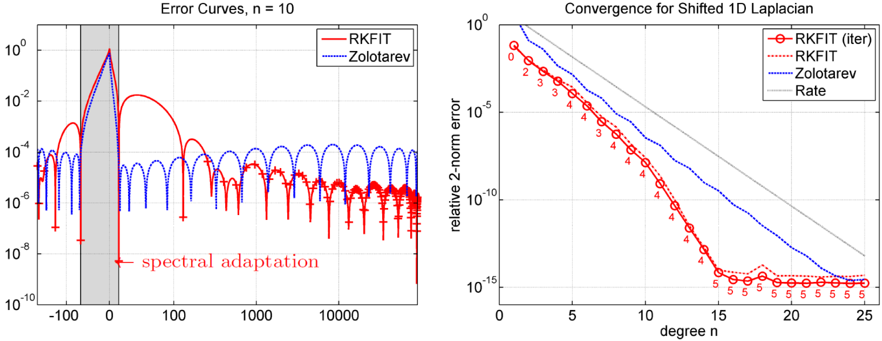
$$R_n(z) = \widehat{h}_0 z + \cfrac{1}{h_1 + \cfrac{1}{\widehat{h}_1 z + \cdots + \cfrac{1}{h_{n-1} + \cfrac{1}{\widehat{h}_{n-1} z + \cfrac{1}{h_n}}}}}$$

with the numbers $h_j, \widehat{h}_{j+1}$ obtained from the entries of $\underline{H_n}$.

Disclaimer: Conversion is potentially ill-conditioned (use `mp`!) and Lanczos may break down unluckily (never seen in practice; use another random starting vector $\mathbf{b}$ for RKFIT).

Recall:
$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2^2 = \sum_{j=1}^{N} |\lambda_j^{1/2} - R_n(\lambda_j)|^2 |w_j|^2$$
$$\leq C \max_{z\in[a_1,b_1]\cup[a_2,b_2]} |1 - z^{-1/2} R_n(z)|^2,$$

so RKFIT approximant expected to converge exponentially with at least the same rate as Zolotarev approximant. Advantage, spectral adaptation:



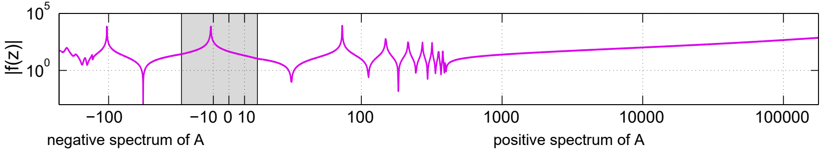
Variable-coefficient DtN maps

Consider BVP for $x \in [0, +\infty)$,

$$\frac{\partial^2}{\partial x^2}\mathbf{u} = [A + c(x)I]\mathbf{u}, \quad \mathbf{u}(0) = -\mathbf{b}, \quad \mathbf{u}(x) \text{ bounded as } x \to \infty$$

with $c(x)$ compactly supported. Consider DtN map $\mathbf{u}'(0) = f(A)\mathbf{b}$, then

- $f(z)$ may have many singularities near $\Lambda(A)$:



- no "explicit" rational approximant, but RKFIT still applicable since
  1.) only action of $f(A)$ onto vectors required,
  2.) RKFIT approximant adapts to $\Lambda(A)$ — this is crucial here!

Two-layer example

Consider
$$\begin{aligned} u''(x) &= (\lambda + c)u(x), \quad && x \in [0, T), \\ u''(x) &= \lambda u(x), && x \in [T, \infty). \end{aligned}$$

Then the DtN function satisfying $u'(0) = f(\lambda)u_0$ is given as

$$f(\lambda) = -\frac{\sqrt{\lambda + c} \cdot \sinh(T\sqrt{\lambda + c}) + \sqrt{\lambda} \cdot \cosh\left(T\sqrt{\lambda + c}\right)}{\sqrt{\lambda + c} \cdot \cosh(T\sqrt{\lambda + c}) + \sqrt{\lambda} \cdot \sinh\left(T\sqrt{\lambda + c}\right)} \cdot \sqrt{\lambda + c}.$$

Can show:

- If $c > 0$, then f has no real poles.
- If $c < 0$, then f has $m = \left\lfloor \frac{T\sqrt{c}}{\pi} \right\rfloor + q$ real poles ($q \in \{0, 1\}$)

$\Longrightarrow$ **2 real poles per wavelength** $\ell = 2\pi/\sqrt{-c}$

$\Longrightarrow$ Degree $\geq m$ required for uniform rational approximation,

but not necessarily for discrete RKFIT approximation!

Two-layer example: spectral adaptation

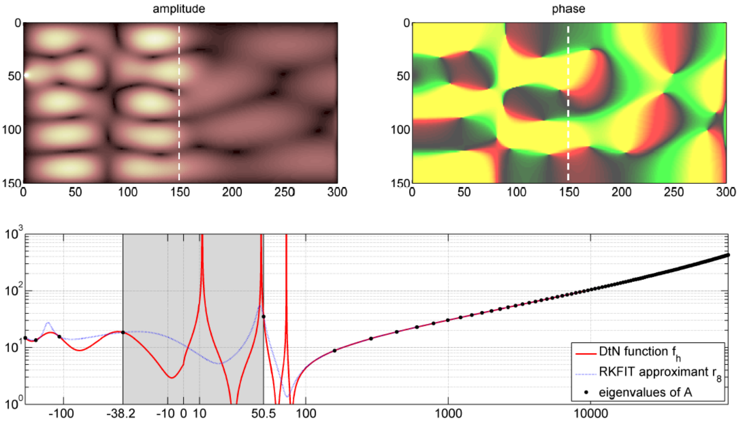
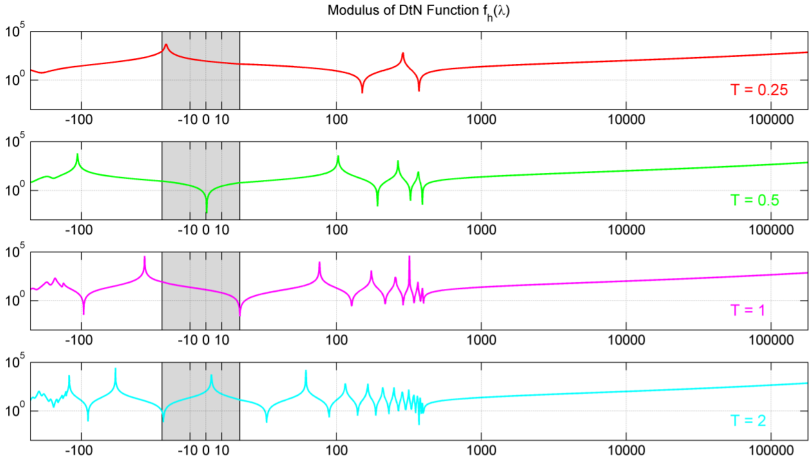


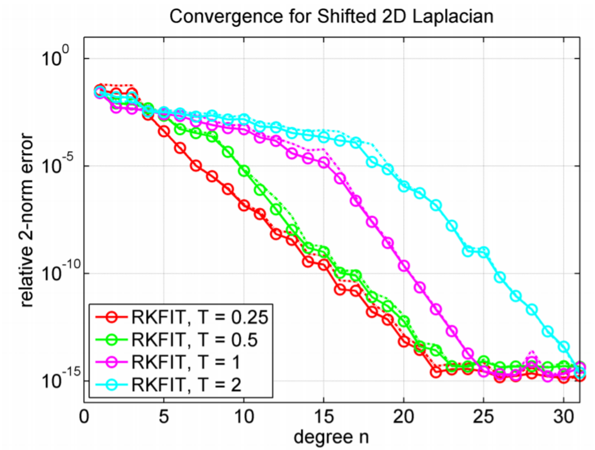
FIGURE 1.2. *A waveguide with varying coefficient (wave number) in the x-direction (piecewise constant over the first 150 grid points and the remaining grid points until infinity). The top row shows the amplitude and phase of the solution, with the position of the coefficient jump highlighted by vertical dashed line. The bottom shows a plot of the exact DtN function f_h (solid red line) over the spectral interval of the indefinite matrix A. The plot is doubly logarithmic on both axis, with the x-axis showing a negative and positive part of the real axis, glued together by the gray linear part in between. The RKFIT approximant of degree $n = 8$ (dotted blue curve) exhibits spectral adaptation to some of the eigenvalues of A (black dots).*

Three-layer example

3D Helmholtz problem on $[0, +\infty] \times [0, 1]^2$, finite difference discr.
Three staggered x-layers with wavenumber offsets $c(x) \in \{-400, 125, 0\}$.
The thickness of each layer is varied, resulting in DtN functions:



Three-layer example ctd.



	$T = 0.25$	$T = 0.5$	$T = 1$	$T = 2$
Nyquist minimum N	8.75	17.5	35	70
SEM minium $\frac{\pi}{2}$N	13.7	27.5	55.0	110.0
RKFIT-FD	8	10	16	19

Conclusions

- Have presented new approach to DtN approximation via RKFIT.
- Rational approximants $\Leftrightarrow$ continued fractions $\Leftrightarrow$ finite difference grids.
- RKFIT solely based on unitary rational Krylov transformations.
- RKFIT exploits "discreteness" of spectrum $\Lambda(A)$ and weights $\mathbf{b}$.
- RKFIT outperforms analytic best approximants for spectral intervals.
- RKFIT works for variable-coefficient problems, and we observe exponential convergence. (No guarantee though!)
- Rational Krylov Toolbox: `www.rktoolbox.org`

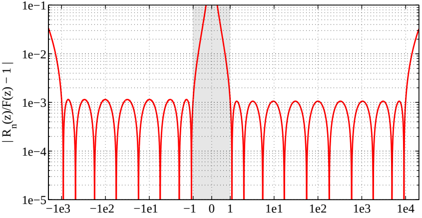
M. Berljafa and S. Güttel, *Generalized rational Krylov decompositions with an application to rational approximation*, SIAM J. Matrix Anal. Appl., 2015.

V. Druskin, S. Güttel, and L. Knizhnerman, *Near-optimal perfectly matched layers for indefinite Helmholtz problems*, SIAM Review, 2016.

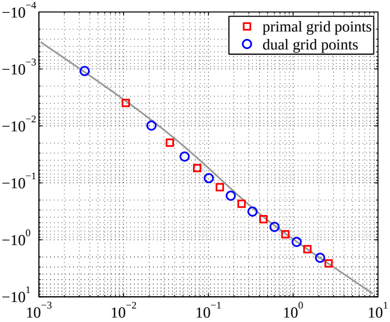
M. Berljafa and S. Güttel, *The RKFIT algorithm for nonlinear rational appproximation*, SIAM Journal on Scientific Computing, 2017.

DtN can hence be implemented into existing FD scheme by changing n trailing primal and dual grid steps to continued fraction coefficients of R_n.

In the indefinite Helmholtz case the grid steps will be complex:



$\Longrightarrow$

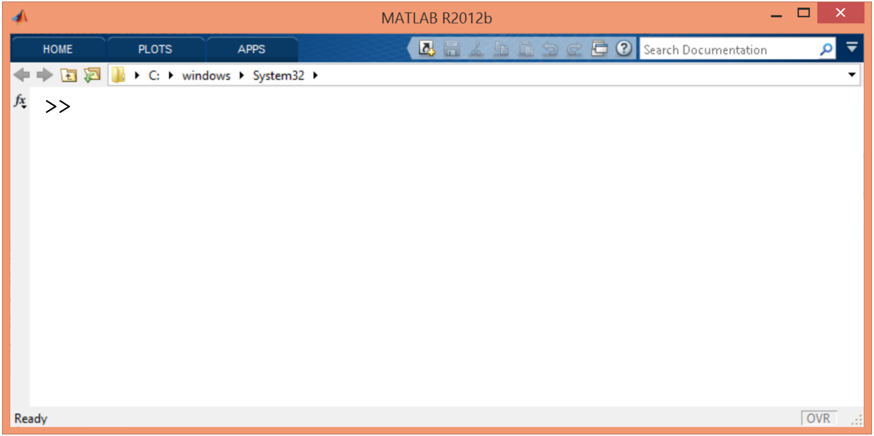


RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1,2,-1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.



RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1, 2, -1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1, 2, -1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1,2,-1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1, 2, -1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

```
>> unzip('http://guettel.com/rktoolbox/rktoolbox.zip');
>> cd('rktoolbox'); addpath(fullfile(cd));
>> N = 100; A = gallery('tridiag',N) - .5*speye(N);
>> F = @(X) sqrtm(full(A))*X;
>>
```

RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1, 2, -1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

MATLAB R2012b

HOME PLOTS APPS

Search Documentation

C: ▸ windows ▸ System32 ▸

```
>> unzip('http://guettel.com/rktoolbox/rktoolbox.zip');
>> cd('rktoolbox'); addpath(fullfile(cd));
>> N = 100; A = gallery('tridiag',N) - .5*speye(N);
>> F = @(X) sqrtm(full(A))*X;
>> b = randn(N,1) xi = inf(1,4);
>>
```

Ready

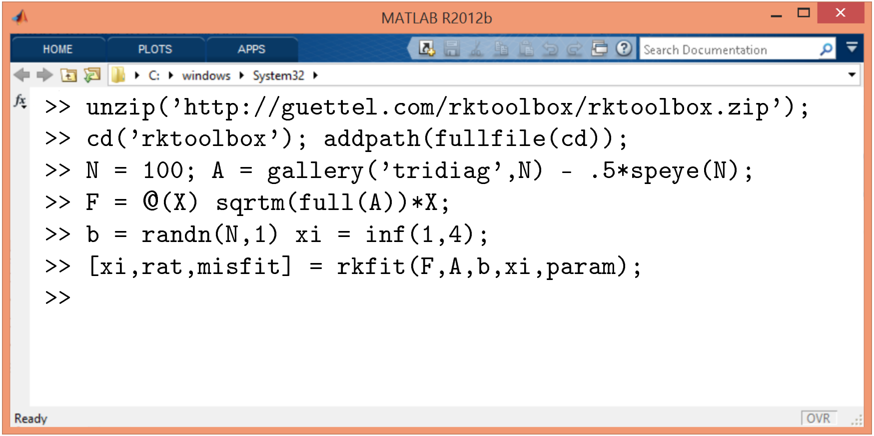
OVR

RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1,2,-1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.



RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small},$$

where $A = \mathrm{tridiag}(-1,2,-1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

```
>> unzip('http://guettel.com/rktoolbox/rktoolbox.zip');
>> cd('rktoolbox'); addpath(fullfile(cd));
>> N = 100; A = gallery('tridiag',N) - .5*speye(N);
>> F = @(X) sqrtm(full(A))*X;
>> b = randn(N,1) xi = inf(1,4);
>> [xi,rat,misfit] = rkfit(F,A,b,xi,param);
>> norm(rat(A,b) - F(b))/norm(F(b))
```

RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1,2,-1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

MATLAB R2012b

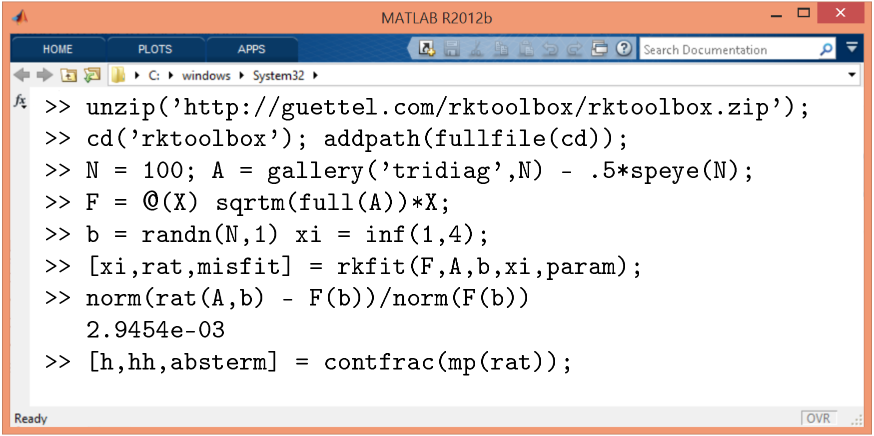
```
>> unzip('http://guettel.com/rktoolbox/rktoolbox.zip');
>> cd('rktoolbox'); addpath(fullfile(cd));
>> N = 100; A = gallery('tridiag',N) - .5*speye(N);
>> F = @(X) sqrtm(full(A))*X;
>> b = randn(N,1) xi = inf(1,4);
>> [xi,rat,misfit] = rkfit(F,A,b,xi,param);
>> norm(rat(A,b) - F(b))/norm(F(b))
   2.9454e-03
>>
```

RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1,2,-1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.



RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1,2,-1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

MATLAB R2012b

HOME PLOTS APPS Search Documentation

C: ▸ windows ▸ System32 ▸

```
>> unzip('http://guettel.com/rktoolbox/rktoolbox.zip');
>> cd('rktoolbox'); addpath(fullfile(cd));
>> N = 100; A = gallery('tridiag',N) - .5*speye(N);
>> F = @(X) sqrtm(full(A))*X;
>> b = randn(N,1) xi = inf(1,4);
>> [xi,rat,misfit] = rkfit(F,A,b,xi,param);
>> norm(rat(A,b) - F(b))/norm(F(b))
   2.9454e-03
>> [h,hh,absterm] = contfrac(mp(rat));
>>
```

Ready OVR

RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1, 2, -1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.

MATLAB R2012b

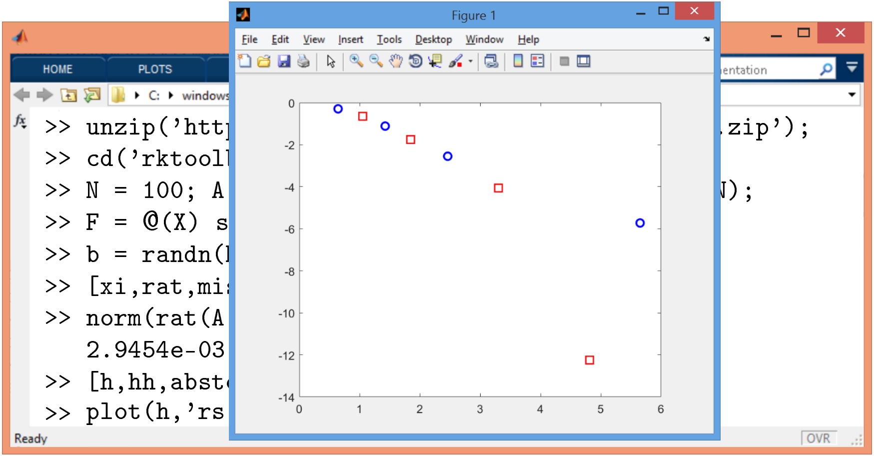
```
>> unzip('http://guettel.com/rktoolbox/rktoolbox.zip');
>> cd('rktoolbox'); addpath(fullfile(cd));
>> N = 100; A = gallery('tridiag',N) - .5*speye(N);
>> F = @(X) sqrtm(full(A))*X;
>> b = randn(N,1) xi = inf(1,4);
>> [xi,rat,misfit] = rkfit(F,A,b,xi,param);
>> norm(rat(A,b) - F(b))/norm(F(b))
   2.9454e-03
>> [h,hh,absterm] = contfrac(mp(rat));
>> plot(h,'rs'); hold on; plot(hh,'bo');
```

RKToolbox Demo (toy example)

Find rational function $R_n(z)$ of type $(4,4)$ such that

$$\|A^{1/2}\mathbf{b} - R_n(A)\mathbf{b}\|_2 \quad \text{is small,}$$

where $A = \mathrm{tridiag}(-1,2,-1) - 0.5I$ and $\mathbf{b} = \texttt{randn}$ of size $N = 100$.



PML test: Waveguide problem

- Consider the constant-coefficient Helmholtz equation

$$\Delta u(x, y) + k_\infty^2 u(x, y) = f(x, y)$$

on rectangular domain $\Omega = [0, L] \times [0, \pi]$ of length $L \in \{\pi, 2\pi\}$.

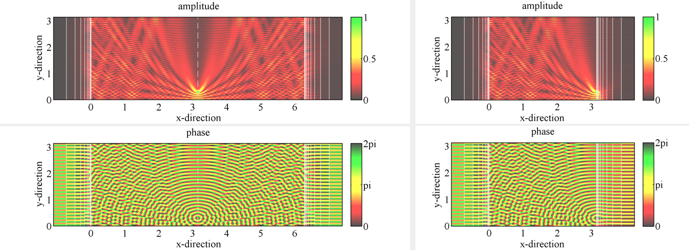
- Source term

$$f(x, y) = 10 \cdot \delta(x - 511\pi/512) \cdot \delta(y - 50\pi/512).$$

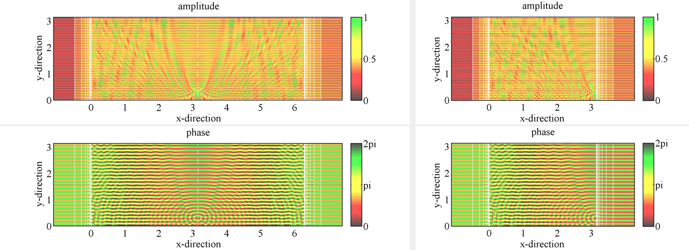
- Homogeneous Dirichlet conditions at upper and lower boundaries in y.
- Uniform finite difference discretization with $h = \pi/512$.
- Append n RKFIT grid points to the left of $x = 0$ and right of $x = L$.
- We use a random vector $\mathbf{b}$ as input to RKFIT.
- Solution should behave like on an infinite strip $(-\infty, \infty) \times [0, \pi]$.

Visually no difference to solution with Zolotarev PML:

wave number $k_\infty = 50$

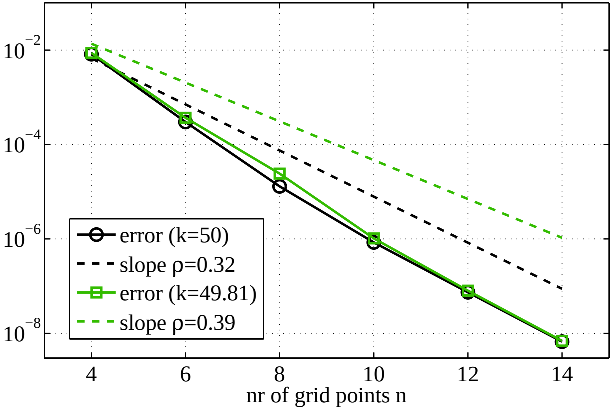


wave number $k_\infty = 49.81$



How accurate is the RKFIT PML?

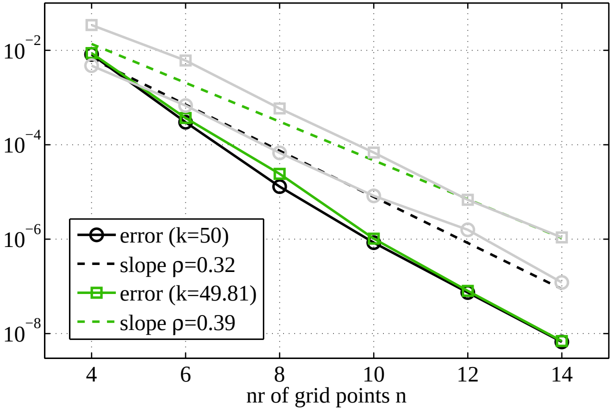
$$\mathrm{err} = \max_{0 \le x,y \le \pi} |u_1(x,y) - u_2(x,y)| \Big/ \max_{0 \le x,y \le \pi} |u_1(x,y)|.$$



RKFIT-PML is finite difference scheme with exponential convergence.

How accurate is the RKFIT PML?

$$\mathrm{err} = \max_{0\le x,y\le\pi} |u_1(x,y) - u_2(x,y)| \Big/ \max_{0\le x,y\le\pi} |u_1(x,y)|.$$



RKFIT-PML more accurate and less affected by resonance than Zolotarev!

Numerical example: Tensorized PML

The PML construction can be extended to higher dimensions, and it allows for varying coefficients in the interior and tangential to the boundary.

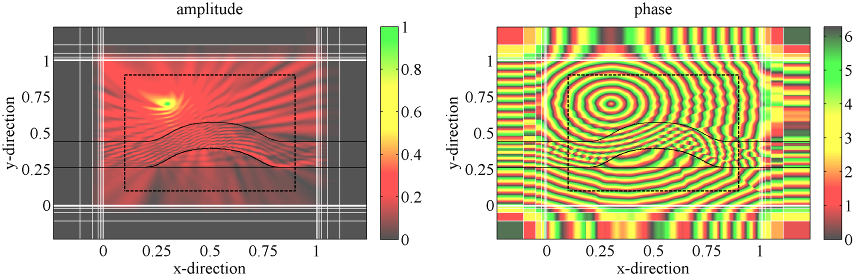
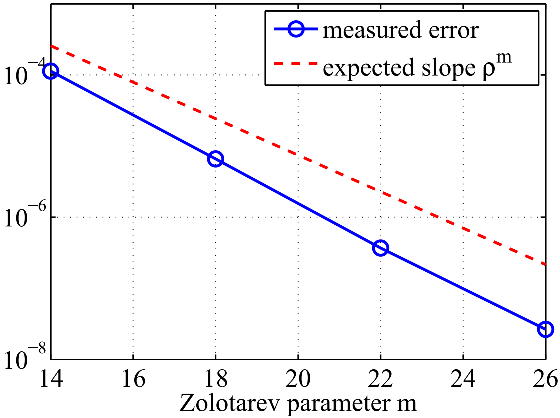


Figure: Amplitude (left) and phase (right) of the solution to a Helmholtz problem on $\Omega_1 = [0,1]^2$ appended with absorbing boundary layers at all boundary edges. There are $n = 7$ grid points appended to all boundaries. The step size in the interior is $h = 1/400$. Comparison with $\Omega_2 = [0.1, 0.9]^2$ for verification.

How accurate is the PML?

$$\mathrm{err} = \max_{0.1 \le x,y \le 0.9} |u_1(x,y) - u_2(x,y)| \Big/ \max_{0.1 \le x,y \le 0.9} |u_1(x,y)|.$$



Again: The PML is a finite difference scheme with exponential convergence.

More precisely, the interpolation points of $R_n(z)$ are roots of

$$H_m(s) = Z_{m_1}^{[\sqrt{-b_1},\sqrt{-a_1}]}(-is) \cdot Z_{m_2}^{[\sqrt{a_2},\sqrt{b_2}]}(s),$$

where $m = m_1 + m_2 = 2n$.

Here, $Z_j^{[c,d]}$ is a real monic polynomial of degree j minimizing

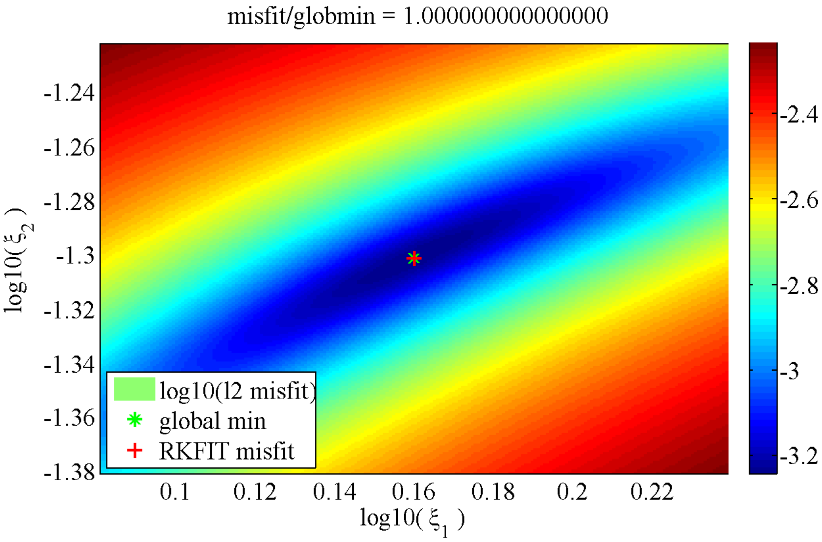
$$\max_{s\in[c,d]} \left| \frac{Z(s)}{Z(-s)} \right|.$$

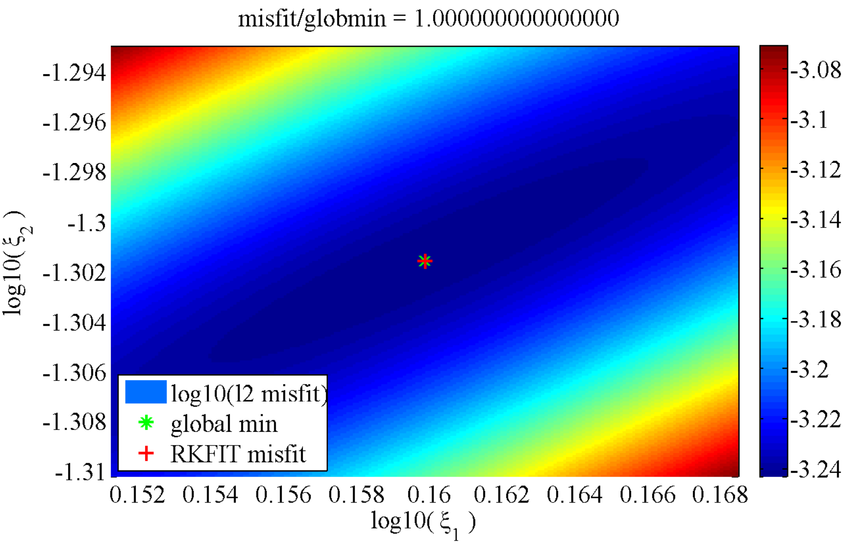
Zolotarev's classical work (1877):

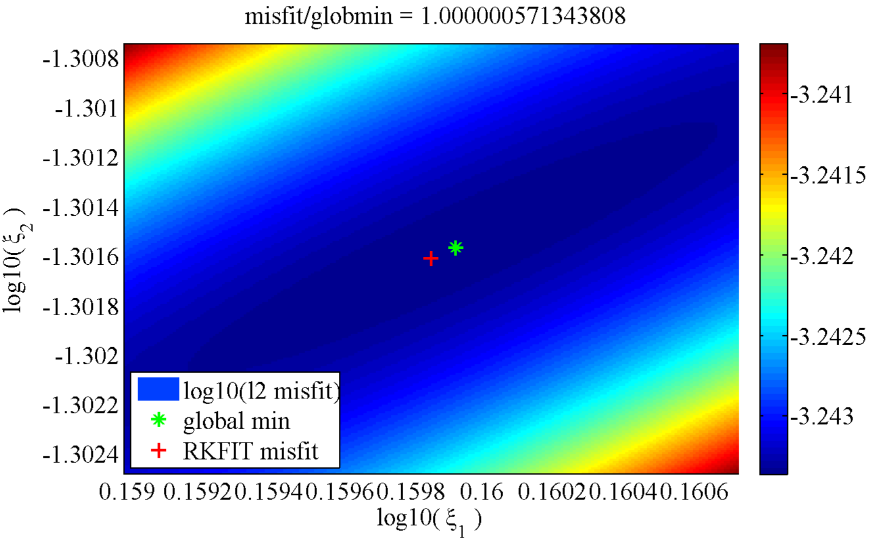
- $Z_j^{[c,d]}$ is uniquely defined.
- Its roots are explicitly known in terms of elliptic functions.
- All its roots are contained in $[c, d]$.

We have no convergence proof for RKFIT but performed some numerical tests, comparing its output to a brute-force search of the global minimum.

For the example in the talk, i.e.,
$N = 10$, $F = A^{-1/2}$, $A = \mathrm{tridiag}(-1, 2, -1)$, with two poles ξ_1, ξ_2 we obtained the following.

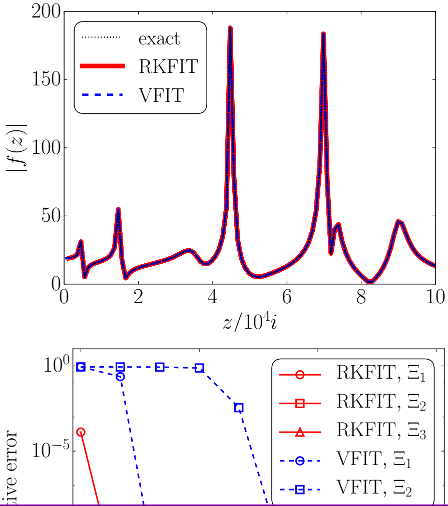

misfit/globmin = 1.000000000000000
-1.24
-1.26
-1.28
-1.3
-1.32
-1.34
-1.36
-1.38
log10(ξ2)
0.1
0.12
0.14
0.16
0.18
0.2
0.22
log10(ξ1)
-2.4
-2.6
-2.8
-3
-3.2
log10(l2 misfit)
global min
RKFIT misfit


misfit/globmin = 1.000000000000000
-1.294
-1.296
-1.298
-1.3
-1.302
-1.304
-1.306
-1.308
-1.31
log10(ξ2)
0.152
0.154
0.156
0.158
0.16
0.162
0.164
0.166
0.168
log10(ξ1)
log10(l2 misfit)
global min
RKFIT misfit
-3.08
-3.1
-3.12
-3.14
-3.16
-3.18
-3.2
-3.22
-3.24

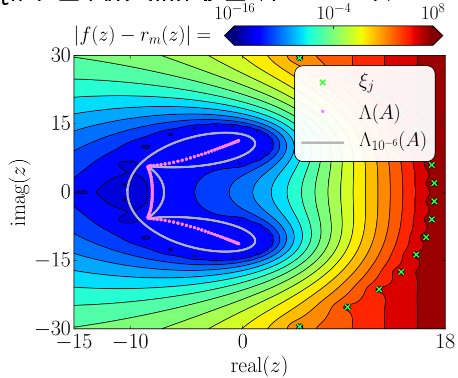

misfit/globmin = 1.000000571343808
-1.3008
-1.301
-1.3012
-1.3014
-1.3016
-1.3018
-1.302
-1.3022
-1.3024
log10(ξ2)
0.159
0.1592
0.1594
0.1596
0.1598
0.16
0.1602
0.1604
0.1606
log10(ξ1)
log10(l2 misfit)
global min
RKFIT misfit
-3.241
-3.2415
-3.242
-3.2425
-3.243

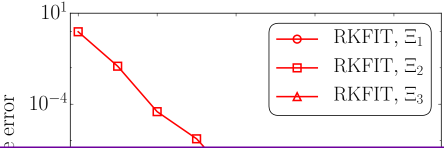
Fitting an artificial frequency response

- f is a $(19, 18)$ rational function, $f(\overline{z}) = \overline{f(z)}$
- $N = 200$



- $A = -5\,\mathrm{grcar}(100, 3)$
- $F = f(A)$, with $f = \exp$, and $\mathbf{v} = [1 \quad 1]^T$





Vector functions with identical poles

Given: $\{A, F_1, \ldots, F_k\} \subset \mathbb{C}^{N\times N}$, and a unit 2-norm vector $\mathbf{v} \in \mathbb{C}^N$.

Find rational functions $r_m^{[\ell]} = \dfrac{p_m^{[\ell]}}{q_m}$ with common denominator s.t.

$$\sum_{\ell=1}^{k} \|F_\ell \mathbf{v} - r_m^{[\ell]}(A)\mathbf{v}\|_2^2 \to \min.$$

In step 2 of RKFIT consider the SVD of

$$\begin{bmatrix} F_1 V_{m+1} - V_{m+1} \left(V_{m+1}^* F_1 V_{m+1}\right) \\ F_2 V_{m+1} - V_{m+1} \left(V_{m+1}^* F_2 V_{m+1}\right) \\ \vdots \\ F_k V_{m+1} - V_{m+1} \left(V_{m+1}^* F_k V_{m+1}\right) \end{bmatrix}.$$

Vector functions with identical poles, an example

- Fitting all elements of the admittance matrix of a six-terminal system (power system distribution network).
- $N = 300 + 300, \quad m = 25 + 25$
- $f^{[\ell]}(\overline{z}) = \overline{f^{[\ell]}(z)}, \quad \ell = 1, \dots, 21$