Rational Krylov methods and approximation of $f(A)\,b$

B. Beckermann[‡], M. Eiermann[†], O. G. Ernst[†], Stefan Güttel[†], and R. Vandebril[‡]

[†]TU Bergakademie Freiberg, Germany
Institut für Numerische Mathematik und Optimierung

[‡]Université Lille 1, France
Equipe d'Analyse Numérique et d'Optimisation

Matrix Functions

Given a square matrix A of size $N \times N$, a vector $\boldsymbol{b}$ of length N and a scalar function $f(z)$,

$$f(A)\boldsymbol{b} := p(A)\boldsymbol{b},$$

where $p(\cdot)$ is a polynomial of degree $< N$ that Hermite-interpolates f in the eigenvalues of A. In typical applications the matrix A is large and sparse.

Some Applications

- $f(z) = (z - i\omega)^{-1}$: model reduction in the frequency domain,
- $f(z) = \exp(-tz)$: time-integration of linear ODE's, exponential integrators, e.g., in geophysics or chemistry,
- $f(z) = \sqrt{tz}$: simulation of Brownian motion of molecules or sampling from Gaussian Markov random fields,
- $f(z) = \operatorname{sign}(z)$: Quantum Chromodynamics.

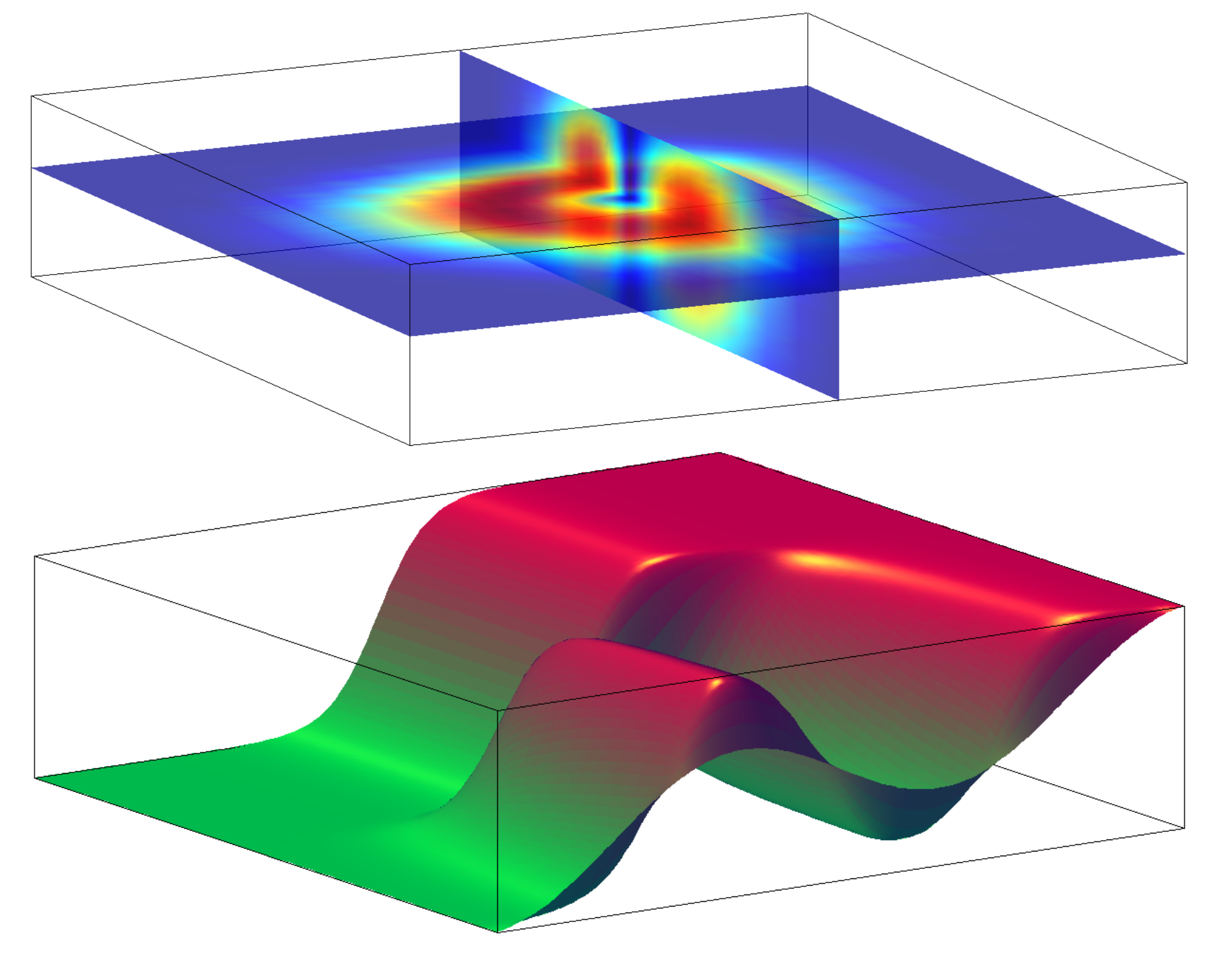

Figure 1: Modelling an electric field in a geophysical application [1]. Time-evolution of a reaction-diffusion process [4].

Rational Krylov Spaces

Definition: Given a sequence of polynomials

$$q_{m-1}(z) = \prod_{\substack{j=1 \\ \xi_j \neq \infty}}^{m-1} (z - \xi_j), \quad m = 1, 2, \ldots,$$

where $\xi_j \in \overline{\mathbb{C}} \setminus \Lambda(A)$. Then the associated rational Krylov spaces of order m are defined as

$$\mathcal{Q}_m(A, \boldsymbol{b}) := q_{m-1}(A)^{-1} \mathcal{K}_m(A, \boldsymbol{b}),$$

where $\mathcal{K}_m(A, \boldsymbol{b}) = \operatorname{span}\{A^0\boldsymbol{b}, A^1\boldsymbol{b}, \ldots, A^{m-1}\boldsymbol{b}\}$.

Properties: Let M be the invariance index of $\mathcal{K}_m(A, \boldsymbol{b})$. Then

- $\mathcal{Q}_m \simeq \mathcal{P}_{m-1}/q_{m-1}$ for $m \leq M$,
- $\operatorname{span}\{\boldsymbol{b}\} = \mathcal{Q}_1 \subset \mathcal{Q}_2 \subset \cdots \subset \mathcal{Q}_M = \mathcal{K}_M(A, \boldsymbol{b})$,
- $f(A)\boldsymbol{b} \in \mathcal{Q}_M(A, \boldsymbol{b})$.

The main aims are always

1. choose q_{m-1} such that $\operatorname{dist}(f(A)\boldsymbol{b}, \mathcal{Q}_m)$ is small (not considered on this poster, see e.g., [3]),
2. extract near-best approximation $\boldsymbol{f}_m \in \mathcal{Q}_m$.

Rayleigh-Ritz extraction

1. Compute orthonormal basis Q_m of $\mathcal{Q}_m$,
2. Define Rayleigh quotient $A_m := Q_m^* A Q_m$,
3. Define $\boldsymbol{f}_m := Q_m f(A_m) Q_m^* \boldsymbol{b}$.

Typically, the *rational Arnoldi method* [6] is used to compute the basis Q_m iteratively. It yields *rational Arnoldi decompositions*

$$A Q_{m+1} \underline{K_m} = Q_{m+1} \underline{H_m}, \quad \text{where} \tag{1}$$

Q_{m+1}	collects orthonormal basis vectors of $\mathcal{Q}_{m+1}$,
$\underline{K_m}$	is an $m+1 \times m$ upper Hessenberg matrix,
$\underline{H_m}$	is an $m+1 \times m$ unreduced upper Hessenberg matrix.

The Rayleigh quotient A_m can be obtained without explicit projection. The eigenvalues $\Lambda(A_m)$ are called *rational Ritz values*.

Interpolation property: There holds

$$\boldsymbol{f}_m = q_{m-1}(A)^{-1} p_{m-1}(A)\boldsymbol{b},$$

where p_{m-1} Hermite-interpolates $(q_{m-1} \cdot f)$ at $\Lambda(A_m)$.

Rational Ritz values

… are the eigenvalues of the Rayleigh quotient $A_m = Q_m^* A Q_m$, denoted by $\Theta = \{\theta_1, \ldots, \theta_m\}$.

Let A be Hermitian. Then the θ_k's lie in the spectral interval of A and *interlace* the eigenvalues $\Lambda(A) = \{\lambda_1, \ldots, \lambda_N\}$:

(*) In any interval $(\theta_\kappa, \theta_{\kappa+1})$ there is at least one eigenvalue λ_k of A.

Moreover, the rational Ritz values are zeros of orthogonal rational functions and may be characterized as

(**) The θ_k's are the zeros of the minimizer of $\| p(A) q_{m-1}(A)^{-1} \boldsymbol{b} \|$ among all monic $p \in \mathcal{P}_m^\infty$.

Logarithmic potential theory can explain the *asymptotic* distribution of the rational Ritz values. Therefore we consider

- a sequence of Hermitian matrices (A_N), each of size $N \times N$, whose eigenvalue counting measures converge to a Borel probability measure σ in the weak-* sense,
- a sequence of vectors $(\boldsymbol{b}_N)$, each of length N,
- a ray sequence of integers (m_N) such that
$$m_N / N \to t \in (0,1) \quad \text{as } N \to +\infty,$$
- a sequence of polynomials (q_N), each of degree $m_N - 1$, whose zero counting measures converge to a Borel measure ν, $\|\nu\| = t$,
- the sequence (Θ_N) of rational Ritz values of order m_N.

Tools from Potential Theory Associated with a (signed) Borel measure μ_1 is the logarithmic potential

$$U^{\mu_1}(z) := \int \frac{1}{\log|x - z|} \, \mathrm{d}\mu_1(x).$$

The mutual logarithmic energy of the measures μ_1 and μ_2 is

$$I(\mu_1, \mu_2) := \int U^{\mu_1}(z) \, \mathrm{d}\mu_2(z), \quad I(\mu_1) := I(\mu_1, \mu_1).$$

We also define the set of σ-constrained measures of mass t,

$$\mathcal{M}_t^\sigma := \{\mu_1 \text{ Borel measure} : \mu_1 \leq \sigma, \ \|\mu_1\| = t\}.$$

With this notation, (*) and (**) can be translated into a constrained weighted minimal energy problem from potential theory:

(*') Among all $\mu \in \mathcal{M}_t^\sigma$
(**') minimize $I(\mu - \nu) = I(\mu) - 2I(\mu, \nu) + I(\nu) \geq 0$.

Under moderate assumptions on $(A_N, \boldsymbol{b}_N, q_N)$, we have

Theorem (see [2])**:** The counting measures of (Θ_N) converge to a positive Borel measure μ being the unique minimizer of

$$I(\mu) - 2I(\mu, \nu) \quad \text{among } \mu \in \mathcal{M}_t^\sigma.$$

Let F be the maximum of $U^{\mu-\nu}$ in the complex plane and $\Sigma_t^* := \{z \in \mathbb{C} : U^{\mu-\nu}(z) = F\}$. In a closed interval $J \subset \mathbb{R} \setminus \Sigma_t^*$ all eigenvalue sequences $J \ni \lambda_{k(N)} \to \lambda$ satisfy

$$\lim_{N \to \infty} \operatorname{dist}(\lambda_{k(N)}, \Theta_N)^{1/N} = e^{2(U^{\mu-\nu}(\lambda) - F)}.$$

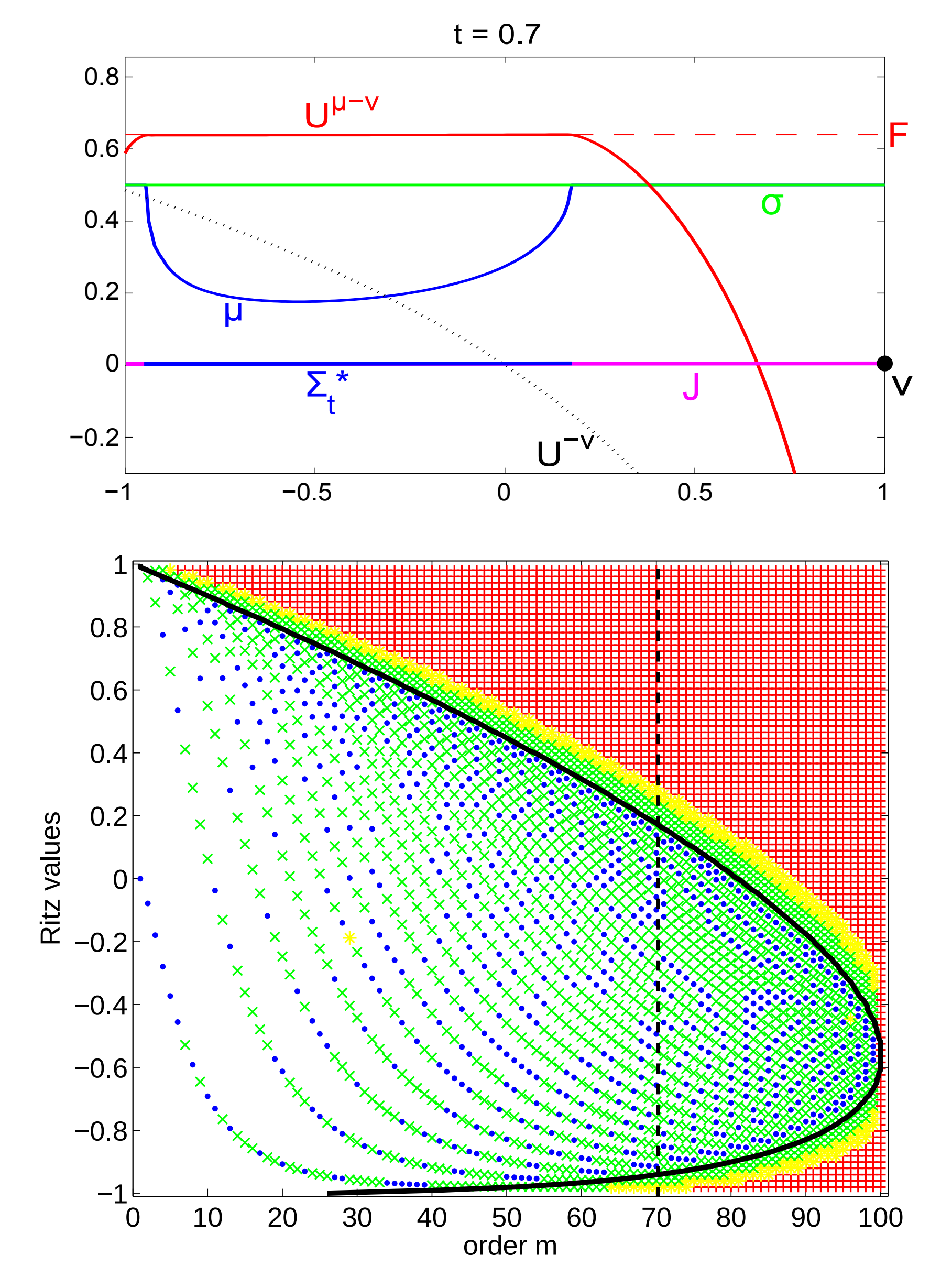

Figure 2: Top: Energy problem, where the constraint σ stems from a matrix A with equispaced ev's in $[-1, 1]$, $\nu = t\delta_0$. Bottom: Rational Ritz values. The colors blue, green, yellow, red indicate the distance to the ev's in decreasing order. The black curve is the predicted region of converged Ritz values.

Inexact solves & error estimators

In each iteration of the rational Arnoldi method a linear system of the form $(A - \xi_j I)\boldsymbol{x}_j = \boldsymbol{q}_j$ is solved. If the residuals are collected in a matrix R_m, then (1) becomes

$$A Q_{m+1} \underline{K_m} = Q_{m+1} \underline{H_m} + R_m. \tag{2}$$

Setting $E_m := -R_m \underline{K_m}^\dagger Q_{m+1}^*$, we observe that we have computed an exact Arnoldi decomposition

$$(A + E_m) Q_{m+1} \underline{K_m} = Q_{m+1} \underline{H_m}$$

for the matrix $A + E_m$ [5]. The Rayleigh quotient $\widetilde{A}_m$ computed from the data $\underline{K_m}$ and $\underline{H_m}$ satisfies

$$\begin{aligned} \widetilde{A}_m &= Q_m^*(A + E_m) Q_m \\ &= Q_m^* A Q_m + Q_m^*(-R_m \underline{K_m}^\dagger Q_{m+1}^*) Q_m \\ &= \widehat{A}_m - Q_m^* R_m \underline{K_m}^\dagger \underline{I_m}. \end{aligned}$$

Here, $\widehat{A}_m := Q_m^* A Q_m$ is referred to as the *corrected Rayleigh quotient*, because it is a compression of A instead of $A + E_m$. It can be computed from $\widetilde{A}_m$ without explicit projection, only by additional inner-products $Q_m^* R_m$.

We now decompose the error

$$\| f(A)\boldsymbol{b} - \boldsymbol{f}_m \| \leq \underbrace{\| f(A)\boldsymbol{b} - f(A + E_m)\boldsymbol{b} \|}_{\text{sensitivity error}} + \underbrace{\| f(A + E_m)\boldsymbol{b} - \boldsymbol{f}_m \|}_{\text{approximation error}},$$

and estimate

$$\text{sensitivity error} \approx \| f(\widetilde{A}_m) Q_m^* \boldsymbol{b} - f(\widehat{A}_m) Q_m^* \boldsymbol{b} \|.$$

It is advisable to terminate the rational Arnoldi method if the approximation error falls below the sensitivity error, because after this happens we only improve approximations to a sequence of "wrong" problems $(f(A + E_m)\boldsymbol{b})$.

Together with an estimator for the approximation error we hence obtain a practical stopping criterion. One such estimator results by adding ℓ interpolation nodes to the Rayleigh-Ritz approximant, i.e., constructing an approximation

$$\widetilde{\boldsymbol{f}}_{m+\ell} = q_{m-1}(A)^{-1} p_{m-1+\ell}(A)\boldsymbol{b},$$

where $p_{m-1+\ell}$ interpolates $(q_{m-1} \cdot f)$ at the rational Ritz values $\theta_1, \ldots, \theta_m$ *and* some auxiliary nodes $\vartheta_1, \ldots, \vartheta_\ell$ (cf. [1]). Then

$$\text{approximation error} \approx \| \boldsymbol{f}_m - \widetilde{\boldsymbol{f}}_{m+\ell} \|.$$

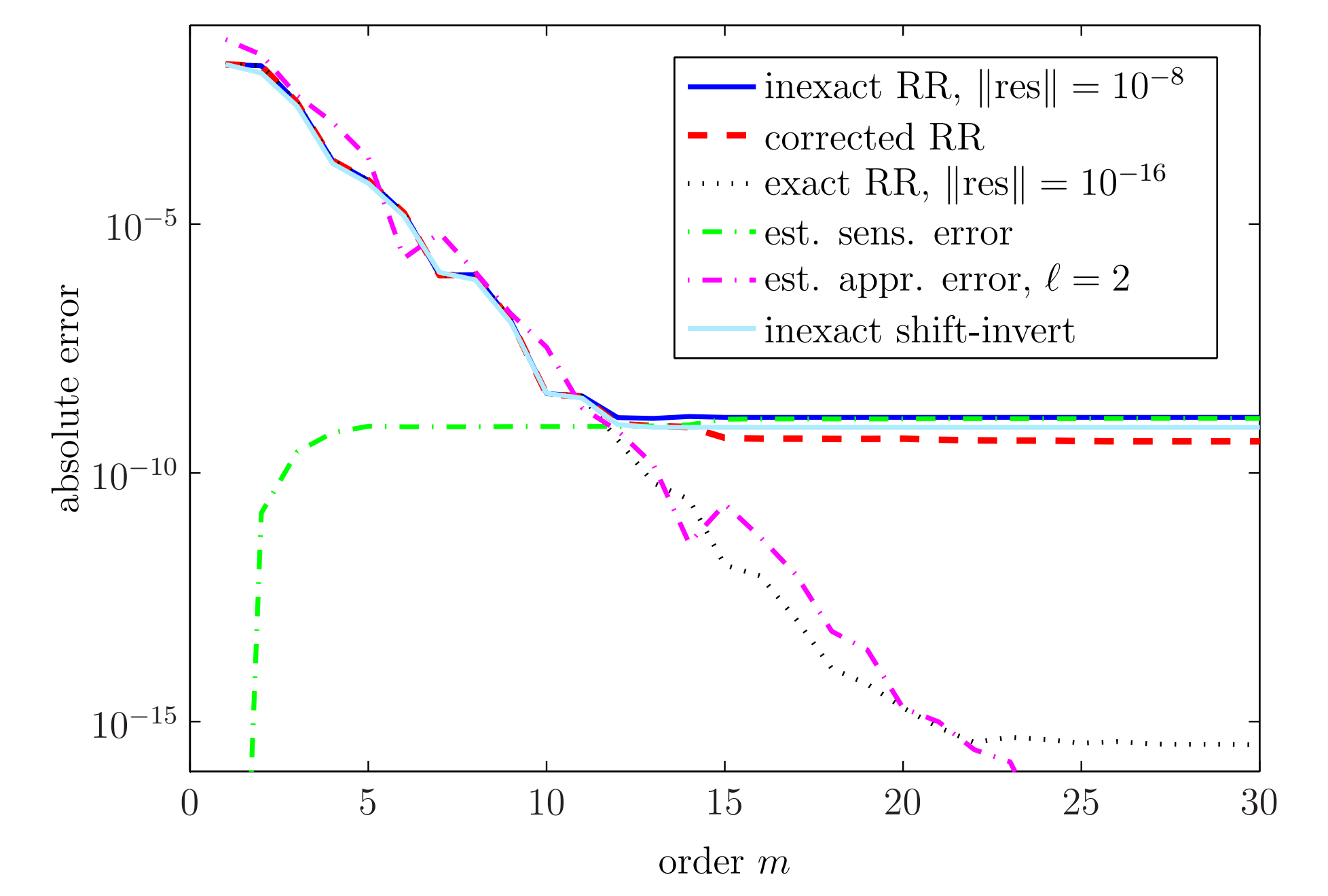

Figure 3: Convergence of the inexact, corrected and exact Rayleigh-Ritz approximations and error estimators for solving the 2D heat equation on the unit square with 100 interior grid points, $f(z) = \exp(-tz)$, $t = 0.1$, all $\xi_j = 1$. We also show the convergence curve of the shift-invert method (see, e.g., [7]) with inexact solves.

References

[1] M. Afanasjew, M. Eiermann, O. G. Ernst, and S. Güttel, *Implementation of a restarted Krylov subspace method for the evaluation of matrix functions*, Linear Algebra Appl., 429 (2008), pp. 2293–2314.

[2] B. Beckermann, S. Güttel, and R. Vandebril, *On the convergence of rational Ritz values*, submitted to SIAM J. Matrix Anal. Appl., (2009).

[3] S. Güttel, *On rational Krylov methods for the approximation of operator functions*, PhD thesis, TU Bergakademie Freiberg, to appear.

[4] A.-K. Kassam and L. N. Trefethen, *Fourth-order time-stepping for stiff PDEs*, SIAM J. Sci. Comput., 26 (2005), pp. 1214–1233.

[5] R. Lehoucq and K. Meerbergen, *Using generalized Cayley transformations within an inexact rational Krylov sequence method*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 131–148.

[6] A. Ruhe, *Rational Krylov sequence methods for eigenvalue computation*, Linear Algebra Appl., 58 (1984), pp. 391–405.

[7] J. van den Eshof and M. Hochbruck, *Preconditioning Lanczos approximations to the matrix exponential*, SIAM J. Sci. Comput., 27 (2006), pp. 1438–1457.