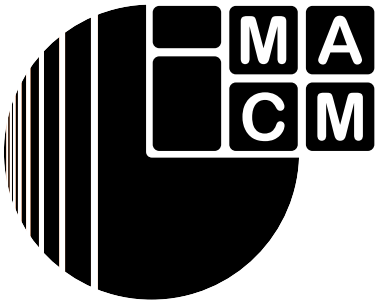

Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 16/07

Marcel Schweitzer

Error bounds and estimates for rational Krylov subspace approximations of Stieltjes matrix functions

March 3, 2016

http://www.math.uni-wuppertal.de

BIT manuscript No.
(will be inserted by the editor)

Error bounds and estimates for rational Krylov subspace approximations of Stieltjes matrix functions

Marcel Schweitzer

Received: date / Accepted: date

Abstract When approximating $f(A)\boldsymbol{b}$—the action of a matrix function on a vector—by a rational Krylov subspace method, one is in need of some measure for the norm of the error of the current iterate, to be able to decide when to terminate the iteration because the desired accuracy is reached. In this paper, we show how to generalize certain error estimators (based on Gauss quadrature) known for the polynomial Krylov case to rational Krylov subspaces. We investigate how it is possible to compute these error estimates with low computational cost, independent of the matrix size n, and compare them to other error estimates proposed in the literature in numerical experiments on several model problems. As a theoretical result, we prove that pairs of ℓ-point Gauss and $\ell+1$-point Gauss–Radau quadrature rules provide lower and upper bounds for the exact error norm when A is Hermitian positive definite, f is a Stieltjes function and the poles of the rational Krylov method are chosen on the negative real axis.

Keywords matrix function · Stieltjes function · rational Krylov subspaces · Lanczos method · error estimates · Gauss quadrature

Mathematics Subject Classification (2000) 65F60, 65F50, 65F30, 65D30, 65D32

1 Introduction

Approximating

$$f(A)\boldsymbol{b}, \tag{1.1}$$

the action of a matrix function $f(A) \in \mathbb{C}^{n\times n}$ on a vector $\boldsymbol{b} \in \mathbb{C}^n$ is necessary in many scientific computing applications, including exponential integrators for differential

This work was partially supported by Deutsche Forschungsgemeinschaft through Collaborative Research Centre SFB TRR55 “Hadron Physics from Lattice QCD”.

Department of Mathematics, Bergische Universität Wuppertal, 42097 Wuppertal, Germany, schweitzer@math.uni-wuppertal.de

equations [31–33, 49], network analysis [3, 17] and theoretical particle physics [7, 8, 10, 16]. Typical functions arising in this context are, e.g., the exponential $f(z) = \exp(z)$, the inverse square root $f(z) = z^{-1/2}$ or the sign function $f(z) = \operatorname{sign}(z)$. In the applications mentioned above (and also in countless other scenarios), the matrix A is typically *very large and sparse*. The matrix function $f(A)$, however, is in general a dense matrix, so that it is impossible to store it explicitly, notwithstanding the high computational cost. Therefore, one has to approximate the vector (1.1) directly by some iterative method. The most widely used and well-studied class of methods devoted to this task is the class of Krylov subspace methods [12, 15, 20, 21, 24, 44].

However, in recent years, *extended and rational* Krylov subspace methods have emerged as a powerful tool for the approximation of matrix functions [13, 27–29, 39], exploiting the fact that it is often possible to approximate a given function f by a low-degree rational function, while a polynomial approximation would require a very high degree. Instead of extracting an approximation from the (polynomial) Krylov subspace

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

these methods use a rational Krylov subspace

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

where

$$q_{m-1}(z) = \prod_{j=1}^{m-1}(1 - z/\xi_j) \tag{1.2}$$

is a polynomial of degree $m-1$, defined by the *poles* $\xi_j \in \big(\overline{\mathbb{C}} \setminus (\operatorname{spec}(A) \cup \{0\})\big)$, where $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ is the extended complex plane. Given a matrix $V_m = [\boldsymbol{v}_1, \dots, \boldsymbol{v}_m]$ whose columns form an orthonormal basis of $\mathcal{Q}_m(A, \boldsymbol{b})$, one can then compute the *rational Arnoldi approximation*

$$\boldsymbol{f}_m = V_m f(A_m) V_m^H \boldsymbol{b} \tag{1.3}$$

for $f(A)\boldsymbol{b}$, where $A_m = V_m^H A V_m$ is the projection of A onto the rational Krylov subspace. An important question arising in practical computations is whether the approximation $\boldsymbol{f}_m$ from (1.3) is an accurate enough approximation for $f(A)\boldsymbol{b}$, i.e., whether

$$\|f(A)\boldsymbol{b} - \boldsymbol{f}_m\|_2 \tag{1.4}$$

lies below some prescribed threshold `tol`. In [29], some estimators (of differing quality) for the error norm (1.4) were presented. In [2, 14], other a priori error estimates (or bounds) in rational Krylov subspace methods are derived in the context of choosing optimal poles. In this paper, we present some further error estimators, based on (rational) Gauss quadrature, similar to the results presented in [22, 23] for the polynomial Krylov subspace case. We will show to compute these error estimators efficiently and identify cases in which they provide upper or lower bounds for the exact error norm (1.4). Of particular interest in this context is the class of *completely monotonic functions*, and especially its subset of *Stieltjes functions*.

The remainder of this paper is organized as follows. In Section 2, we collect some basic results on rational Krylov subspaces. In Section 3 we introduce the class of Stieltjes functions and provide some results on rational Krylov approximations to

Stieltjes matrix functions. Section 4 is concerned with the computation of error estimators based on standard and rational Gauss quadrature. In Section 5 we present methods for efficiently computing the error estimates presented in the previous section, based on the *rational implicit Q theorem* from [6]. We briefly introduce other error estimators, presented in [29], in Section 6, before comparing them to our estimators in various numerical experiments in Section 7. Concluding remarks are given in Section 8.

2 Rational Krylov subspaces

To be able to use the rational Arnoldi approximation (1.3) for $f(A)\boldsymbol{b}$, one first needs to compute an orthonormal basis V_m of $\mathcal{Q}_m(A,\boldsymbol{b})$. This is typically done by the *rational Arnoldi algorithm*, first introduced in [43], given as Algorithm 1. We briefly remark that the algorithm in the presented form does not allow for poles at zero, but that this is no essential limitation. One can exclude any other finite pole σ instead, by running the algorithm with the shifted matrix $A-\sigma I$ and the shifted poles $\xi_j-\sigma$, see, e.g., [27].

Algorithm 1: Rational Arnoldi algorithm for computing an ONB of $\mathcal{Q}_m(A,\boldsymbol{b})$

Given: A, $\boldsymbol{b}$, $\{\xi_1,\ldots,\xi_m\}$
$\boldsymbol{v}_1 \leftarrow \boldsymbol{b}/\|\boldsymbol{b}\|$.
for $j=1,2,\ldots,m$ **do**
  $\boldsymbol{w}_j \leftarrow (I-A/\xi_j)^{-1}A\boldsymbol{v}_j$.
  for $i=1,\ldots,j$ **do**
    $h_{i,j} \leftarrow \boldsymbol{w}_j^H \boldsymbol{v}_i$.
    $\boldsymbol{w}_j \leftarrow \boldsymbol{w}_j - h_{i,j}\boldsymbol{v}_i$
  $h_{j+1,j} \leftarrow \|\boldsymbol{w}_j\|_2$.
  $\boldsymbol{v}_{j+1} \leftarrow 1/h_{j+1,j}\boldsymbol{w}_j$.

Associated with Algorithm 1 is a *rational Krylov decomposition*

$$AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m} \tag{2.1}$$

where $\underline{H_m} \in \mathbb{C}^{(m+1)\times m}$ is an unreduced upper Hessenberg matrix containing the orthogonalization coefficients $h_{i,j}$ and $\underline{K_m} \in \mathbb{C}^{(m+1)\times m}$ is given by

$$\underline{K_m} = \begin{bmatrix} I + H_m D_m \\ h_{m+1,m}\xi_m^{-1}\boldsymbol{e}_m^T \end{bmatrix}, \text{ where } D_m = \operatorname{diag}(\xi_1^{-1},\ldots,\xi_m^{-1})$$

and H_m denotes the upper $m\times m$ part of $\underline{H_m}$ (in the same way we denote $K_m = I + H_m D_m$ in the following), see, e.g., [27, 28]. If $\xi_m = \infty$, the last row of $\underline{K_m}$ is all zero, and (2.1) simplifies to the *reduced rational Krylov decomposition*

$$AV_m K_m = V_{m+1}\underline{H_m}.$$

In this case it is known [27, Lemma 5.6] that K_m is nonsingular and that the projected matrix A_m can be computed as

$$A_m = H_m K_m^{-1}, \tag{2.2}$$

so that it is not necessary to explicitly compute $V_m^H A V_m$ for evaluating (1.3). The next result, from [27], gives an interpolation characterization of the rational Arnoldi approximation (1.3), similar to the well-known connection between polynomial Krylov subspace methods and polynomial interpolation.

Theorem 2.1 (Theorem 5.8 in [27]) *The approximation $\boldsymbol{f}_m$ from* (1.3) *satisfies*

$$\boldsymbol{f}_m = r_{m-1}(A)\boldsymbol{b}, \tag{2.3}$$

where $r_{m-1}(z) = p_{m-1}(z)/q_{m-1}(z)$ (with a polynomial p_{m-1} of degree at most $m-1$) interpolates f at the eigenvalues of A_m.

The characterization from Theorem 2.1 will be useful for deriving an error representation for the rational Arnoldi approximation by means of explicit formulas for interpolating rational functions, see Section 3.

We end this section by giving another important result, which generalizes a well-known property of polynomial Krylov subspaces and will be important throughout this paper.

Lemma 2.1 (Lemma 4.2 in [27]) *Let m^* be the Krylov invariance index corresponding to A and $\boldsymbol{b}$, i.e., the smallest integer such that $\mathcal{K}_{m^*}(A,\boldsymbol{b})$ is invariant under multiplication with A. Then*

$$\mathcal{Q}_1(A,\boldsymbol{b}) \subset \mathcal{Q}_2(A,\boldsymbol{b}) \subset \cdots \subset \mathcal{Q}_{m^*}(A,\boldsymbol{b}) = \mathcal{Q}_{m^*+1}(A,\boldsymbol{b}) = \mathcal{Q}_{m^*+2}(A,\boldsymbol{b}) = \ldots \tag{2.4}$$

We note that the nestedness property (2.4) clearly relies on the choice of denominator polynomials (1.2). If these polynomials are chosen in such a way that the sets of poles are not nested, then the corresponding rational Krylov subspaces will obviously also not be nested in general. In the remainder of this paper, we will for ease of presentation always assume that the number m of steps performed in the rational Arnoldi method is smaller than the invariance index m^* corresponding to A and $\boldsymbol{b}$.

3 Stieltjes functions

In this paper, we focus on the approximation of *Stieltjes functions* [4, 5, 30]. These functions can be defined by means of a Riemann–Stieltjes integral as

$$f(z) = \int_0^\infty \frac{1}{z+t}\,\mathrm{d}\mu(t), \quad z \in \mathbb{C} \setminus \mathbb{R}_0^-, \tag{3.1}$$

where $\mu(t)$ is a nonnegative, monotonically increasing function defined on $\mathbb{R}_0^+$ which satisfies

$$\int_0^\infty \frac{1}{z+1}\,\mathrm{d}\mu(t) < \infty. \tag{3.2}$$

Condition (3.2) guarantees that the integral on the right-hand side of (3.1) exists for all $z \in \mathbb{C} \setminus \mathbb{R}_0^-$. Examples of Stieltjes functions arising in applications are

$$z^{-\sigma} = \frac{\sin(\sigma\pi)}{\pi} \int_0^\infty \frac{t^{-\sigma}}{z+t}\,\mathrm{d}t \quad \text{for } \sigma \in (0,1)$$

and

$$\frac{\log(1+z)}{z} = \int_1^\infty \frac{t^{-1}}{z+t}\,\mathrm{d}t$$

as well as rational functions in partial fraction form,

$$f(z) = \sum_{i=0}^r \frac{\zeta_i}{t_i+z},$$

corresponding to a piecewise constant step function μ with positive jumps ζ_i at the points t_i. For further examples of Stieltjes functions, see, e.g., [4, 5, 19, 30].

Stieltjes functions are infinitely many times continuously differentiable on $\mathbb{C} \setminus \mathbb{R}_0^-$ with derivatives

$$f^{(k)}(z) = (-1)^k k! \int_0^\infty \frac{1}{(z+t)^{k+1}}\,\mathrm{d}\mu(t) \text{ for all } k \in \mathbb{N}_0. \tag{3.3}$$

From (3.3) it immediately follows that every Stieltjes function belongs to the class of completely monotonic functions [1,4] on $\mathbb{R}^+$, i.e., functions h which satisfy

$$(-1)^k h^{(k)}(t) \geq 0 \text{ for all } k, \text{ and } t \in \mathbb{R}^+. \tag{3.4}$$

Property (3.4) will prove useful in Section 4, as it allows the computation of bounds for the error norm (1.4) in certain situations. Another important tool which we will need in the following is an integral representation for the interpolating rational function $r_{m-1}(z)$ from (2.3) when f is a Stieltjes function, which can be found by an obvious modification of the Hermite–Walsh formula for rational interpolation [50, Theorem VIII.2], see also [29]. Defining the nodal polynomial

$$w_m(z) = (z+\theta_1)\cdot\cdots\cdot(z+\theta_m)$$

corresponding to the *rational Ritz values* $\operatorname{spec}(A_m) = \{\theta_1, \ldots, \theta_m\}$, we have

$$r_{m-1}(z) = \int_0^\infty \left(1 - \frac{w_m(-z)q_{m-1}(t)}{w_m(t)q_{m-1}(-z)}\right) \frac{1}{z+t}\,\mathrm{d}\mu(t). \tag{3.5}$$

The representation (3.5) can then be used to derive the error representation

$$f(A)\boldsymbol{b} - \boldsymbol{f}_m = f(A)\boldsymbol{b} - r_{m-1}(A)\boldsymbol{b} = \int_0^\infty \frac{q_{m-1}(t)}{w_m(t)}(A+tI)^{-1} w_m(-A)q_{m-1}(-A)^{-1}\boldsymbol{b}\,\mathrm{d}\mu(t). \tag{3.6}$$

It is known that $w_m(-A)q_{m-1}(-A)^{-1}\boldsymbol{b}$ is a scalar multiple of the next rational Arnoldi basis vector, i.e.,

$$w_m(-A)q_{m-1}(-A)^{-1}\boldsymbol{b} = \delta_m \boldsymbol{v}_{m+1},$$

for some $\delta_m \in \mathbb{C}$; see [29]. We can thus rewrite (3.6) as

$$f(A)\boldsymbol{b} - \boldsymbol{f}_m = \delta_m \widetilde{e}_m(A)\boldsymbol{v}_{m+1}, \text{ where } \widetilde{e}_m(z) = \int_0^\infty \frac{q_{m-1}(t)}{w_m(t)} \frac{1}{z+t}\,\mathrm{d}\mu(t). \tag{3.7}$$

When A is Hermitian positive definite and all poles ξ_j in the rational Arnoldi method are chosen to lie on the negative real axis (which is a natural choice, as a Stieltjes function has poles on the negative real axis itself), we can show that the *error function* $\widetilde{e}_m$ from (3.7) is again a Stieltjes function.

Theorem 3.1 *Let $A \in \mathbb{C}^{n\times n}$ be Hermitian positive definite, let f be a Stieltjes function, let $\boldsymbol{b} \in \mathbb{C}^n$, let $\xi_1, \ldots, \xi_m \in \mathbb{R}_0^-$ and let $\boldsymbol{f}_m$ be the rational Arnoldi approximation* (1.3)*. Then the error function $\widetilde{e}_m(z)$ from* (3.7) *is a Stieltjes function.*

Proof Define the function

$$\widetilde{\mu}(t) = \int_0^t \frac{q_{m-1}(\tau)}{w_m(\tau)}\,\mathrm{d}\mu(\tau). \tag{3.8}$$

As all rational Ritz values are real and positive when A is Hermitian positive definite and all poles ξ_i are real and negative by assumption, the function $\frac{q_{m-1}(\tau)}{w_m(\tau)}$ is real and nonnegative for all $\tau \geq 0$. As μ is nonnegative and monotonically increasing, the integral on the right-hand side of (3.8) must be nonnegative. Therefore, $\widetilde{\mu}(t) \geq 0$ for $t \geq 0$. For $t_1 > t_0 \geq 0$, we have

$$\begin{aligned}
\widetilde{\mu}(t_1) &= \int_0^{t_1} \frac{q_{m-1}(\tau)}{w_m(\tau)}\,\mathrm{d}\mu(\tau) \\
&= \int_0^{t_0} \frac{q_{m-1}(\tau)}{w_m(\tau)}\,\mathrm{d}\mu(\tau) + \int_{t_0}^{t_1} \frac{q_{m-1}(\tau)}{w_m(\tau)}\,\mathrm{d}\mu(\tau) \\
&= \widetilde{\mu}(t_0) + \int_{t_0}^{t_1} \frac{q_{m-1}(\tau)}{w_m(\tau)}\,\mathrm{d}\mu(\tau) \\
&\geq \widetilde{\mu}(t_0),
\end{aligned}$$

so that $\widetilde{\mu}$ is monotonically increasing. It remains to check whether the condition (3.2) holds for $\widetilde{\mu}$. Note that (3.8) implies

$$\mathrm{d}\widetilde{\mu}(t) = \frac{q_{m-1}(t)}{w_m(t)}\,\mathrm{d}\mu(t). \tag{3.9}$$

As $\deg q_{m-1} \leq m-1 < m = \deg w_m$, the function q_{m-1}/w_m is bounded on $\mathbb{R}_0^+$, i.e., there exists a constant $d \geq 0$ such that

$$\int_0^\infty \frac{1}{1+t}\,\mathrm{d}\widetilde{\mu}(t) = \int_0^\infty \frac{q_{m-1}(t)}{w_m(t)} \frac{1}{1+t}\,\mathrm{d}\mu(t) \leq d \int_0^\infty \frac{1}{1+t}\,\mathrm{d}\mu(t) < \infty,$$

where the last inequality holds because f is a Stieltjes function and μ thus satisfies (3.2). Summarizing, we have shown that the function

$$\int_0^\infty \frac{1}{z+t}\,\mathrm{d}\widetilde{\mu}(t) \tag{3.10}$$

is a Stieltjes function, and by (3.9) it follows that (3.10) coincides with $\widetilde{e}_m(z)$, thus proving the assertion. □

The result of Theorem 3.1 is of importance for two major reasons. First, it guarantees the existence of the integral representation (3.7), and second, it allows to compute bounds for the error norm (1.4) in certain situations when A is Hermitian positive definite, cf. Section 4.

4 Error estimates based on (rational) Gauss quadrature

In this section, we investigate how to approximate the error norm (1.4) by (rational) Gauss quadrature rules. We first consider the case that A is Hermitian positive definite, and briefly comment on the non-Hermitian case afterwards.

From (3.7), we find

$$\|f(A)\boldsymbol{b} - \boldsymbol{f}_m\|_2^2 = \delta_m^2 \boldsymbol{v}_{m+1}^H \widetilde{e}_m(A)^2 \boldsymbol{v}_{m+1}. \tag{4.1}$$

The function $\widetilde{e}_m(z)^2$ is completely monotonic, as it is the product of two Stieltjes (and thus completely monotonic) functions. This directly follows from the Leibniz rule for product differentiation. The quadratic form on the right-hand side of (4.1) can be interpreted as a Riemann–Stieltjes integral

$$\delta_m^2 \boldsymbol{v}_{m+1}^H \widetilde{e}_m(A)^2 \boldsymbol{v}_{m+1} = \delta_m^2 \int_{\lambda_{\min}}^{\lambda_{\max}} \widetilde{e}_m(z)^2 \,\mathrm{d}\alpha(z), \tag{4.2}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote the smallest and largest eigenvalue of A, respectively, and α is a nonnegative, monotonically increasing step function depending on $\mathrm{spec}(A)$ and the coefficients of $\boldsymbol{v}_{m+1}$ in the eigenvector basis of A, see, e.g., [23, 25, 26]. By a classical result, Gauss and Gauss–Radau quadrature rules can be used to compute upper and lower bounds for integrals of completely monotonic functions, and thus for quadratic forms like (4.1), see, e.g., [11, 26].

Theorem 4.1 *Let h be completely monotonic on $[a,b]$ and let α be nonnegative and monotonically increasing on $[a,b]$. Let $z_i, \omega_i, i = 1,\dots,\ell$ be the nodes and weights of the ℓ-point Gauss quadrature rule for α on $[a,b]$ and let $\widetilde{z}_i, \widetilde{\omega}_i, i = 1,\dots,\ell+1$ be the nodes and weights of the $(\ell+1)$-point Gauss–Radau quadrature rule for α on $[a,b]$ with one node fixed at $\widetilde{z}_1 \le a$. Then*

$$\sum\nolimits_{i=1}^{\ell} \omega_i h(z_i) \le \int_a^b h(z)\,\mathrm{d}\alpha(z) \le \sum\nolimits_{i=1}^{\ell+1} \widetilde{\omega}_i h(\widetilde{z}_i).$$

It is known, see, e.g. [26], that the ℓ-point Gauss quadrature rule for (4.2) is given by

$$\delta_m^2 \boldsymbol{e}_1^H \widetilde{e}_m(T_\ell)^2 \boldsymbol{e}_1, \tag{4.3}$$

where T_ℓ is the tridiagonal matrix obtained by performing ℓ steps of the Lanczos process [40] for A and $\boldsymbol{v}_{m+1}$. Similarly, the $(\ell+1)$-point Gauss–Radau quadrature rule can be evaluated by considering a suitable rank-one modification of $T_{\ell+1}$, see, e.g., [23, 26], provided that (a lower bound for) the smallest eigenvalue $\lambda_{\min}$ of A is known. Precisely, given $z_0 \le \lambda_{\min}$, one constructs the matrix

$$\widetilde{T}_{\ell+1} = \begin{bmatrix} T_\ell & t_{\ell+1,\ell}\boldsymbol{e}_\ell \\ t_{\ell+1,\ell}\boldsymbol{e}_\ell^H & z_0 + d_\ell \end{bmatrix}, \text{ where } \boldsymbol{d} = t_{\ell+1,\ell}^2 (T_\ell - z_0 I)^{-1} \boldsymbol{e}_\ell \tag{4.4}$$

and substitutes it for T_ℓ in (4.3). Thus, a first (naive) approach for computing error bounds in the rational Arnoldi method is given by Algorithm 2. We just briefly mention that the scaling factor δ_m needed for evaluating (4.3) can cheaply be obtained during the execution of the rational Arnoldi method, see, e.g., [29].

Remark 4.1 For evaluating the error function $\widetilde{e}_m(\widetilde{T}_\ell)^2$, one typically needs to use numerical quadrature, as it is only known by its integral representation. This topic is covered in depth in [23]; see also [21]. In [23, Proposition 2], it is also shown how one can choose this "inner" quadrature rule in such a way that the computed estimates are guaranteed bounds. As this transfers completely analogously to the rational Krylov case, we omit a detailed discussion of this topic here.

Algorithm 2: Rational Arnoldi algorithm with naive computation of quadrature-based error bounds

Given: A, $\boldsymbol{b}$, m, ℓ, `tol`, lower bound $\theta < \lambda_{\min}$, pole sequence $\xi_1, \xi_2, \ldots$

$i \leftarrow 1$
Perform m steps of the rational Arnoldi method, Algorithm 1, and compute $\boldsymbol{f}_m$ via (1.3).
Perform ℓ steps of the Lanczos process for A and $\boldsymbol{v}_{m+1}$, yielding T_ℓ.
Modify T_ℓ according to (4.4), yielding $\widetilde{T}_{\ell+1}$.
Evaluate $\varepsilon \leftarrow \delta_m^2 \boldsymbol{e}_1^H \widetilde{e}_m(\widetilde{T}_\ell)^2 \boldsymbol{e}_1$.
while $\varepsilon >$ `tol` **do**
  $i \leftarrow i+1$
  Perform m further steps of the rational Arnoldi method and compute $\boldsymbol{f}_{im}$ via (1.3).
  Perform ℓ steps of the Lanczos process for A and $\boldsymbol{v}_{im+1}$, yielding T_ℓ.
  Modify T_ℓ according to (4.4), yielding $\widetilde{T}_{\ell+1}$.
  Evaluate $\varepsilon \leftarrow \delta_{im}^2 \boldsymbol{e}_1^H \widetilde{e}_{im}(\widetilde{T}_\ell)^2 \boldsymbol{e}_1$

In Figure 4.1, we report the results of an artificial numerical experiment for validating our theory. We compute the inverse square root of a diagonal matrix $A \in \mathbb{R}^{10{,}000 \times 10{,}000}$ with Chebyshev points in $[10^{-2}, 10^2]$ as eigenvalues (i.e., $\kappa(A) = 10^4$) times a random vector of unit length. The poles of the rational Arnoldi method are chosen adaptively via the approach of [29]. For purpose of illustration, we compute error bounds after each step in Algorithm 2, i.e., we choose $m = 1$. We observe that the estimates computed by Gauss and Gauss–Radau quadrature are lower and upper bounds, respectively, as predicted by our theory. For higher numbers of quadrature nodes, the estimates lie closer to the exact error norm (as one would expect).

The main drawback of Algorithm 2 is that the evaluation of the error bounds requires ℓ additional matrix-vector products with A. Depending on the cost for solving the shifted linear systems in the rational Arnoldi method, this additional cost may not be too severe (especially in contrast to the situation one is facing when using a similar approach for computing error bounds in the polynomial Lanczos method, see [22,23]), but it is nonetheless additional computational work which does not help in advancing the iteration, in particular if one wants to check the error norm every m steps for a small number m. In Section 5, we therefore show how it is possible to modify Algorithm 2 in such a way that all matrix-vector products and linear system

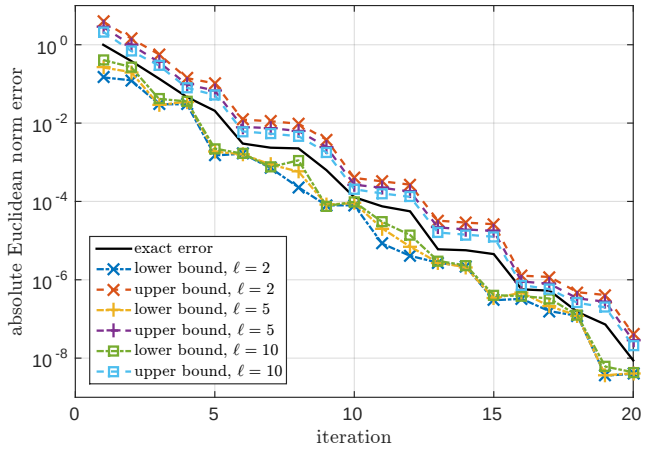

Fig. 4.1 Upper and lower bounds computed by Algorithm 2 when approximating the inverse square root of a diagonal matrix with Chebyshev eigenvalues in $[10^{-2}, 10^{2}]$ for varying number of quadrature nodes.

solves can be used to advance the iteration (i.e., increase the dimension of the computed rational Krylov subspace) but still allow for the computation of error bounds (without any additional operations with cost in the order of n).

Before doing so, we explore another related approach for computing error estimates in the rational Arnoldi method: An obvious idea when using a rational Krylov subspace method for approximating $f(A)\boldsymbol{b}$ is to use *rational Gauss rules* instead of standard Gauss rules to compute error estimates. Various rational Gauss quadrature rules have been introduced and analyzed in, e.g., [37, 38, 41].

Following [41], we call an ℓ-point quadrature rule a rational Gauss rule corresponding to the denominator

$$s_k(z) = (z-\theta_1)\cdot\cdots\cdot(z-\theta_k), \text{ where } k \leq 2\ell-1, \tag{4.5}$$

if it integrates exactly all rational functions with denominator (4.5) and arbitrary numerator of degree at most $2\ell-1$. For the special case that k is even and

$$s_k(z) = (z-\theta_1)^2\cdot\cdots\cdot(z-\theta_{\frac{k}{2}})^2$$

one way to compute such quadrature rules is by performing ℓ steps of the rational Arnoldi method with poles $\xi_j = \theta_j, j = 1,\ldots,\frac{k}{2}$ (and $\xi_{\frac{k}{2}+1} = \cdots = \xi_\ell = \infty$, in case $\frac{k}{2} < \ell$) and then, analogously to (4.3), evaluate

$$\boldsymbol{e}_1^H h(A_m)\boldsymbol{e}_1 \approx \boldsymbol{v}^H h(A)\boldsymbol{v},$$

see [28, Remark 3.2] or, for the special case of extended Krylov methods, [37]. By this approach, the computation of rational Gauss rules is much more costly than that of standard Gauss rules, as it does not only require ℓ matrix-vector products, but an additional $\frac{k}{2}$ linear system solves. In return, one can expect these quadrature rules to

yield more accurate results in many cases (depending of course, e.g., on the function h to be approximated and the choice of poles).

One can also construct rational Gauss–Radau rules and prove that pairs of rational Gauss and Gauss–Radau rules yield lower and upper bounds for the exact value of the integral in certain situations. For this to be the case, the function $s_k(z) \cdot h(z)$ must be completely monotonic, a property which is in general not fulfilled for $h(z) = \widetilde{e}_m(z)^2$ in our setting. We therefore do not pursue this approach any further here and refer the reader to [37, 41] for a detailed treatment of this topic.

One could, of course, incorporate the computation of rational Gauss quadrature based bounds into the rational Arnoldi method similar to what was done in Algorithm 2. This would, however, require performing additional linear system solves which do not contribute to extending the rational Krylov subspace (instead of just additional matrix-vector products). Therefore, in this setting it is even more crucial to be able to somehow re-use the quantities needed for computing the error estimates also for the primary rational Krylov iteration (or the other way around). This will be the topic of the next section.

Remark 4.2 So far, we assumed that A is Hermitian positive definite. Most of the concepts discussed so far can also be extended to non-Hermitian matrices. When dealing with polynomial Krylov methods, one arrives at the so-called *Arnoldi quadrature rules* [9, 18] in this case, see [23, Section 5]. When replacing $\widetilde{e}_m(A)^2$ by $\widetilde{e}_m(A)^H \widetilde{e}_m(A)$ in (4.1), one can use the same techniques as described before, resulting in error estimates one can interpret as obtained by *rational Arnoldi quadrature rules*. The main drawback of these estimates in comparison to those obtained in the Hermitian case is that in general they are neither lower nor upper bounds for the error.

5 Efficient computation of error estimates

In [22, 23], the concept of *Lanczos restart recovery* was introduced in the context of computation of error estimates in polynomial Krylov subspace methods. It relies on the nestedness properties of Krylov subspaces, in particular the relation

$$\mathcal{K}_\ell(A, \boldsymbol{v}_{m+1}) \subseteq \mathcal{K}_{m+\ell+1}(A, \boldsymbol{b}),$$

and the implicit Q theorem [48, Chapter 5, Theorem 1.3] in order to compute *retrospective* error estimates in the Lanczos method. Precisely, it allows to use the matrix-vector products which are performed *both* for advancing the iteration and for the computation of error bounds. This way, error estimates based on ℓ-point Gauss quadrature (and/or $\ell+1$ point Gauss–Radau quadrature) for the error $\|f(A)\boldsymbol{b} - \boldsymbol{f}_m\|_2$ of the mth Lanczos iterate can be computed in step $m+\ell+1$, with computational cost independent of m and n.

In this section, we consider the generalization of this concept to rational Krylov subspaces. As rational Krylov subspaces also fulfill the nestedness property (2.4) and recently, a *rational implicit Q theorem* was proven in [6], the two most important tools from the polynomial setting are generalizable to the rational Krylov case. Therefore,

it is reasonable to assume that similar restart recovery techniques can be used in the computation of error estimates in rational Krylov methods.

We begin by stating the rational implicit Q theorem. In the statement of the theorem, two rational Krylov decompositions $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ and $A\widetilde{V}_{m+1}\underline{\widetilde{K}_m} = \widetilde{V}_{m+1}\underline{\widetilde{H}_m}$ are called *essentially equal* if there exist a unitary diagonal matrix $D_{m+1} \in \mathbb{C}^{(m+1)\times(m+1)}$ and a nonsingular upper triangular matrix $T_m \in \mathbb{C}^{m\times m}$ such that

$$\widetilde{V}_{m+1} = V_{m+1}D_{m+1}, \quad \underline{\widetilde{H}_m} = D_{m+1}^H \underline{H_m} T_m, \text{ and } \underline{\widetilde{K}_m} = D_{m+1}^H \underline{K_m} T_m. \tag{5.1}$$

Theorem 5.1 (Theorem 3.2 in [6]) *Let A satisfy an orthonormal rational Krylov decomposition $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with poles $\xi_j = h_{j+1,j}/k_{j+1,j}$. For every $j = 1,\ldots,m$, the matrix V_{j+1} and the pencil $(\underline{H_j}, \underline{K_j})$ are essentially uniquely determined by the first column of V_{m+1} and the poles $\xi_1,\ldots,\xi_m$.*

Remark 5.1 We remark that the rational Arnoldi approximations (1.3) corresponding to two essentially equal rational Krylov decomposition are equal. Due to (5.1), the rational Arnoldi approximation corresponding to $A\widetilde{V}_{m+1}\underline{\widetilde{K}_m} = \widetilde{V}_{m+1}\underline{\widetilde{H}_m}$ is

$$\widetilde{V}_m f(\widetilde{V}_m^H A \widetilde{V}_m)\widetilde{V}_m^H \boldsymbol{b} = V_m D_m f(D_m^H V_m^H A V_m D_m) D_m^H V_m^H \boldsymbol{b} = V_m f(A_m) V_m^H \boldsymbol{b}.$$

Using the result of Theorem 5.1, we are in a position to prove the main result of this section.

Theorem 5.2 *Let $V_{m+\ell+1}$ be the orthonormal basis of $\mathcal{Q}_{m+\ell+1}(A,\boldsymbol{b})$ computed by Algorithm 1, corresponding to the pole sequence $\xi_1,\ldots,\xi_{m+l+1}$ with $\xi_{m+\ell+1} = \infty$, let $A_{m+\ell+1} = V_{m+\ell+1}^H A V_{m+\ell+1}$ and assume that $A_{m+\ell+1} - \xi_j I$ is nonsingular for $j = m+1,\ldots,m+\ell+1$. Further, let the following two decompositions be given:*

(i) $A\widetilde{V}_\ell\widetilde{K}_\ell = \widetilde{V}_{\ell+1}\underline{\widetilde{H}_\ell}$, corresponding to ℓ steps of Algorithm 1 applied to $A, \boldsymbol{v}_{m+1}$ and the pole sequence $\xi_{m+1},\ldots,\xi_{m+\ell+1}$ and

(ii) $A_{m+\ell+1}\widehat{V}_\ell\widehat{K}_\ell = \widehat{V}_{\ell+1}\underline{\widehat{H}_\ell}$, corresponding to ℓ steps of Algorithm 1 applied to $A_{m+\ell+1}$, $\boldsymbol{e}_{m+1}$ and the pole sequence $\xi_{m+1},\ldots,\xi_{m+\ell+1}$.

Then the pencil $(\underline{\widetilde{H}_\ell}, \underline{\widetilde{K}_\ell})$ is essentially equal to the pencil $(\underline{\widehat{H}_\ell}, \underline{\widehat{K}_\ell})$ in the sense of (5.1).

Proof Note that we will not make a distinction in notation between rational Krylov spaces corresponding to different pole sequences. The choice of poles will always be clear from the choice of starting vector. Due to the nestedness of the pole sequences, we have

$$\mathcal{Q}_{\ell+1}(A,\boldsymbol{v}_{m+1}) \subseteq \mathcal{Q}_{m+\ell+2}(A,\boldsymbol{b}),$$

so that there exists a matrix $U_{\ell+1} \in \mathbb{C}^{(m+\ell+2)\times(\ell+1)}$ such that

$$V_{m+\ell+2}U_{\ell+1} = \widetilde{V}_{\ell+1}. \tag{5.2}$$

As both $V_{m+\ell+2}$ and $\widetilde{V}_{\ell+1}$ have orthonormal columns, so has $U_{\ell+1}$. Inserting the relation (5.2) into the decomposition from (i) gives

$$AV_{m+\ell+1}U_\ell\widetilde{K}_\ell = V_{m+\ell+2}U_{\ell+1}\underline{\widetilde{H}_\ell}. \tag{5.3}$$

By using the reduced rational Krylov decomposition for $\mathcal{Q}_{m+\ell+1}(A, \boldsymbol{b})$,

$$AV_{m+\ell+1}K_{m+\ell+1} = V_{m+\ell+2}\underline{H_{m+\ell+1}},$$

and the fact that $K_{m+\ell+1}$ is nonsingular, equation (5.3) can be recast into

$$V_{m+\ell+2}\underline{H_{m+\ell+1}}K_{m+\ell+1}^{-1}U_\ell\widetilde{K}_\ell = V_{m+\ell+2}U_{\ell+1}\underline{\widetilde{H}_\ell}.$$

As $V_{m+\ell+2}$ is of full rank, this implies

$$\underline{H_{m+\ell+1}}K_{m+\ell+1}^{-1}U_\ell\widetilde{K}_\ell = U_{\ell+1}\underline{\widetilde{H}_\ell}. \tag{5.4}$$

Noting further that $H_{m+\ell+1}K_{m+\ell+1}^{-1} = A_{m+\ell+1}$ and dropping the last row in (5.4), we find

$$A_{m+\ell+1}U_\ell'\widetilde{K}_\ell = U_{\ell+1}'\underline{\widetilde{H}_\ell}, \tag{5.5}$$

where $U_{\ell+1}'$ is just $U_{\ell+1}$ without its last row. Due to (5.2), the first column of $U_{\ell+1}$ is just $\boldsymbol{e}_{m+1}$, so that (5.5) is a rational Krylov decomposition corresponding to the pole sequence $\xi_{m+1}, \ldots, \xi_{m+\ell+1}$ and starting vector $\boldsymbol{e}_{m+1}$. Thus, by Theorem 5.1, it is essentially equal to the decomposition from (ii), from which the assertion of the theorem follows. □

The result of Theorem 5.2 allows to compute retrospective error estimates for the error at iteration m based on ℓ-point rational Gauss quadrature with pole sequence $\xi_{m+1}, \ldots, \xi_{m+\ell+1}$ without investing additional matrix-vector products or linear system solves with A. Instead, the necessary operations have to be performed with the compressed matrix $A_{m+\ell+1}$. One situation in which it is guaranteed that this is possible is when A is Hermitian positive definite and all poles lie outside of $[\lambda_{\min}, \lambda_{\max}]$, as all eigenvalues of $A_{m+\ell+1}$ also lie in $[\lambda_{\min}, \lambda_{\max}]$.

Remark 5.2 When A is Hermitian, the additional work necessary for computing error bounds via Lanczos restart recovery in the polynomial Krylov case is independent both of n and m, as one can exploit that large parts of the matrix $U_{\ell+1}$ from the proof of Theorem 5.2 are zero (due to the short recurrence for the Lanczos basis vectors) and one therefore only needs to perform the secondary Lanczos process with the bottom right $(2\ell+1)\times(2\ell+1)$ submatrix of $A_{m+\ell+1}$, see [22,23]. As there is no short recurrence for the basis vectors for general rational Krylov subspace methods, one typically has to invest additional work depending on m for multiplications and linear system solves with $A_{m+\ell+1}$. One notable exception from this are *extended* Krylov subspaces, i.e., rational Krylov subspaces for which only the poles 0 and ∞ are used (often alternatingly), as there exist short (five-term) recurrences for the basis vectors in this case, see, e.g., [35, 36, 42, 46]. One can then perform the necessary operations with the bottom right $(4\ell+1)\times(4\ell+1)$ submatrix of $A_{m+\ell+1}$; see [45] for details.

One important assumption in Theorem 5.2 is that $\xi_{m+\ell+1} = \infty$, i.e., that the matrix $A_{m+\ell+1}$ from which the error bounds are computed corresponds to a reduced decomposition. When the rational Arnoldi algorithm is implemented as in Algorithm 1, one can thus not compute error estimates in each iteration (as otherwise, all poles would need to be chosen at ∞, and the method would thus reduce to a polynomial

Krylov method). As an alternative, one can implement the rational Arnoldi method as in [29, Algorithm 1], where a "temporary" rational Arnoldi basis with the last pole at infinity is constructed (to be able to cheaply compute A_m via (2.2)), which is then afterwards updated to a rational Arnoldi basis with the last vector corresponding to a different pole. This increases the amount of work needed for orthogonalization, but this work is often negligible in comparison to the cost of linear system solves in the rational Krylov method.

Algorithm 3: Rational Arnoldi algorithm with computation of error estimates (prototype)

Given: A, $\boldsymbol{b}$, ℓ, f, `tol`, pole sequence $\xi_1, \xi_2, \ldots$,
$\boldsymbol{v}_1 \leftarrow \boldsymbol{b}/\|\boldsymbol{b}\|$.
for $j = 1, 2, \ldots, m$ **do**
  Set $\boldsymbol{w}_j := (I - A/\xi_j)^{-1} A \boldsymbol{v}_j$.
  for $i = 1, \ldots, j$ **do**
    $h_{i,j} \leftarrow \boldsymbol{w}_j^H \boldsymbol{v}_i$.
    $\boldsymbol{w}_j \leftarrow \boldsymbol{w}_j - h_{i,j} \boldsymbol{v}_i$
  $h_{j+1,j} \leftarrow \|\boldsymbol{w}_j\|_2$.
  $\boldsymbol{v}_{j+1} \leftarrow 1/h_{j+1,j} \boldsymbol{w}_j$.
  if $\xi_j = \infty$ *and* $j > \ell + 1$ **then**
    Compute $A_j \leftarrow H_j K_j^{-1}$
    Perform ℓ steps of rational Arnoldi for A_j and $\boldsymbol{e}_{j-\ell}$, yielding $\widetilde{A}_\ell$.
    Evaluate $\varepsilon \leftarrow \delta_j^2 \boldsymbol{e}_1^H \widetilde{e}_{j-\ell-1}(\widetilde{A}_\ell) \boldsymbol{e}_1$.
    if $\varepsilon <$ `tol` **then**
      Return $\boldsymbol{f}_j \leftarrow V_j f(A_j) V_j^H \boldsymbol{b}$.

The poles of the rational Gauss rule used for computing the error estimates coincide with the last ℓ poles used in advancing the rational Arnoldi method. Therefore, if one wants to compute guaranteed bounds (in the Hermitian case) based on standard ℓ-point Gauss quadrature, one needs to choose the pole ∞ a total of ℓ times in the primary iteration. But instead of being "lost" like in Algorithm 2, the corresponding matrix-vector products advance the dimension of the rational Krylov subspace.

A prototype version of a rational Arnoldi method with computation of error estimates is given in Algorithm 3. According to the comments above, there are several obvious modifications possible, for which we do not give algorithmic details.

6 Comparison to other error estimators

In this section, we compare the error estimates developed in this manuscript to two of the error estimators presented in [29], which we will briefly describe in the following.

The first error estimator, named *approximate error bound* in [29] is derived by replacing A by A_m and taking norms in (3.7), giving

$$\|f(A)\boldsymbol{b} - \boldsymbol{f}_m\|_2 \approx \delta_m \|\widetilde{e}_m(A_m)\|_2, \tag{6.1}$$

where the matrix function $\widetilde{e}_m(A_m)$ is then approximated by quadrature. If A_m is diagonalizable, then this can be done by evaluating m scalar quadrature rules.

The second error estimator from [29] which we compare our estimates to is the so-called *residual-based estimator*. Using the intimate relation of Stieltjes functions to shifted linear systems, the authors define the "residual" of a Stieltjes matrix function f as

$$\text{residual}(f,m) := h_{m+1,m} \int_0^\infty e_m^H (tK_m - H_m)^{-1} e_1 \, \mathrm{d}\mu(t) v_{m+1}. \tag{6.2}$$

The error estimator is then obtained by taking the Euclidean norm of (6.2), where the right-hand side is again evaluated by using a suitable quadrature rule.

Before comparing the quality of the different error estimates, we give a few remarks concerning their computational cost. For simplicity, we assume that all involved matrices are diagonalizable, so that all integrals can be computed via scalar quadrature. Furthermore, we assume that all these quadrature rules require $\mathcal{O}(k)$ quadrature points for reaching the desired accuracy (which is a reasonable assumption and in line with what is observed in numerical experiments).

1. The approximate error bound (6.1) requires computing an eigendecomposition of A_m, which typically requires $\mathcal{O}(m^3)$ operations, the evaluation of m scalar quadrature rules with $\mathcal{O}(k)$ nodes and the computation of the norm of the error function. This results in an overall asymptotic cost of $\mathcal{O}(m^3 + mk)$.
2. The estimator based on the residual (6.2) requires the solution of a linear system with $t_i K_m - H_m$ for each quadrature node t_i. As both K_m and H_m are upper Hessenberg, the solution of these systems can be computed with cost $\mathcal{O}(m^2)$, resulting in an overall cost of $\mathcal{O}(m^2 k)$.
3. Our Gauss-quadrature based error estimator requires performing ℓ steps of a secondary rational Arnoldi algorithm with the compressed matrix A_m. Using the factored form $A_{m+\ell+1} = H_{m+\ell+1} K_{m+\ell+1}^{-1}$, both matrix-vector products and linear system solves with A_m can be computed with cost $\mathcal{O}((m+\ell)^2)$, so that the secondary rational Arnoldi process requires $\mathcal{O}((m+\ell)^2\ell + m\ell)$ operations in total. Approximating the error function $\widetilde{e}_m(\widetilde{A}_\ell)$ by scalar quadrature then costs $\mathcal{O}(\ell^3)$ operations for computing an eigendecomposition and $\mathcal{O}(k\ell)$ operations for the scalar quadrature rules. Summarizing, the computation of the quadrature-based error estimate requires $\mathcal{O}((m+\ell)^2\ell + \ell^3 + k\ell)$ operations. Assuming $\ell = \mathcal{O}(1)$, as it will typically be the case in practice, this reduces to $\mathcal{O}(m^2 + k)$.

Under the assumption that $\ell = \mathcal{O}(1)$, our proposed error bounds are therefore less costly than those from [29], at least in $\mathcal{O}$-sense. The drawback, however, is that our error estimates become available only in retrospect, i.e., the error estimate for step m is available at iteration $m+\ell+1$, while the error estimates from [29] are available right away.

Remark 6.1 In the extended Krylov case, the cost of computing the quadrature-based bounds reduces to $\mathcal{O}(\ell^3 + k\ell)$, as it suffices to perform the secondary Lanczos process with a matrix of size $(4\ell+1) \times (4\ell+1)$; cf. Remark 5.2.

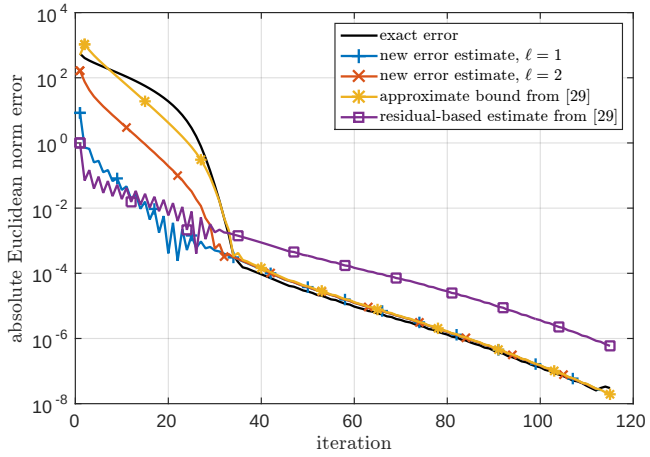

Fig. 6.1 Comparison of different error estimators when approximating the action of the inverse square root of the discrete 1D Laplacian with $n = 2000$ on the normalized vector of all ones. The rational Arnoldi algorithm uses the single repeated pole $\xi = -\sqrt{\lambda_{\min}\lambda_{\max}}$.

7 Numerical experiments

In this section, we compare the quality of the different error estimators in several numerical experiments. In the first experiment, we approximate the inverse square root of the one-dimensional Laplace operator discretized by standard finite differences (with $n = 2000$ grid points), applied to the normalized vector of all ones. We compare our error estimators based on rational Gauss quadrature with $\ell = 1$ and $\ell = 2$ quadrature nodes to the two estimators from [29]. As pole sequence, we choose the single, repeated, asymptotically optimal pole $\xi = -\sqrt{\lambda_{\min}\lambda_{\max}}$. The resulting error norms and estimators are depicted in Figure 6.1. Both quadrature based estimators and the approximate error bound lie very close to the exact error norm in the *linear* convergence phase of the method. In the initial superlinear phase, the approximate bound most closely resembles the actual behavior of the method, and the quadrature based estimate corresponding to $\ell = 2$ is much more accurate than the one corresponding to $\ell = 1$. The residual-based estimator overestimates the exact error norm by about two orders of magnitude, and it does not resolve the convergence slope completely accurately. While this first experiment does not show a superiority of our approach when compared to the approximate bound from [29], it at least shows that we are able to compute bounds of the same quality, with (asymptotically) lower computational work (as the approximate upper bound was found to be the most expensive estimator in $\mathcal{O}$-sense in Section 6).

In the next experiment, we approximate the inverse square root of a diagonal matrix with Chebyshev points in $[10^{-4}, 10^4]$ as eigenvalues, i.e. $\kappa(A) = 10^8$, applied to a normalized random vector. We again compare the same error estimates as in the previous experiment and report the results in Figure 7.1. The poles in the rational

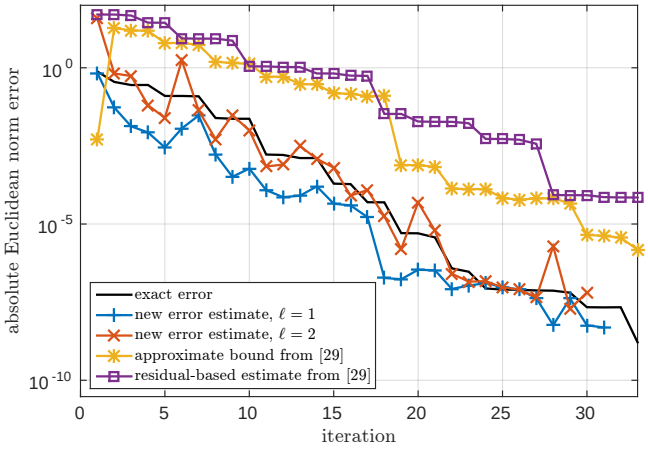

Fig. 7.1 Comparison of different error estimators when approximating the action of the inverse square root of a diagonal matrix $A \in \mathbb{R}^{10,000\times 10,000}$ on a normalized random vector. The poles in the rational Arnoldi algorithm are chosen adaptively according to [29].

Arnoldi method are chosen adaptively according to the strategy proposed in [29]. For this matrix, which has a very large condition number, the different error estimators are of highly varying quality. Both estimators from [29] severely overestimate the magnitude of the error (by two or more orders of magnitude), while the quadrature-based estimate for $\ell = 2$ lies very close to the exact error norm in most iterations. We observe, however, that there are certain "spikes" in the error estimate at places where convergence of the rational Arnoldi method (almost) stagnates. Still, this example demonstrates that our new approach may lead to results of much better quality than previously proposed error estimators.

Before we turn our attention to an application problem, we perform a last artificial experiment. We choose A as a single Jordan block of size 4000×4000 and approximate the action of the inverse square root on a normalized random vector again. As in the previous experiment, we choose the poles in the rational Arnoldi method adaptively. This time, the matrix A is non-Hermitian and non-diagonalizable and its condition number is $\kappa(A) = 8 \cdot 10^3$. We report results of this experiment in Figure 7.2. This time, all estimates yield results of comparable quality. The quadrature-based estimates show similar peaks as in the previous experiment in iterations where the rational Arnoldi method stagnates. While our new estimates do not clearly outperform the estimates from [29], they are at least of comparable accuracy.

To demonstrate the quality of our estimates for a problem coming from a real-world application, we consider a problem from stochastics, namely *sampling from a Gaussian Markov random field*, see, e.g., [34, 47]. Given a set of n points $s_i \in \mathbb{R}^2, i = 1,\dots,n$, we define the so-called *precision matrix* $A \in \mathbb{R}^{n\times n}$ with respect to

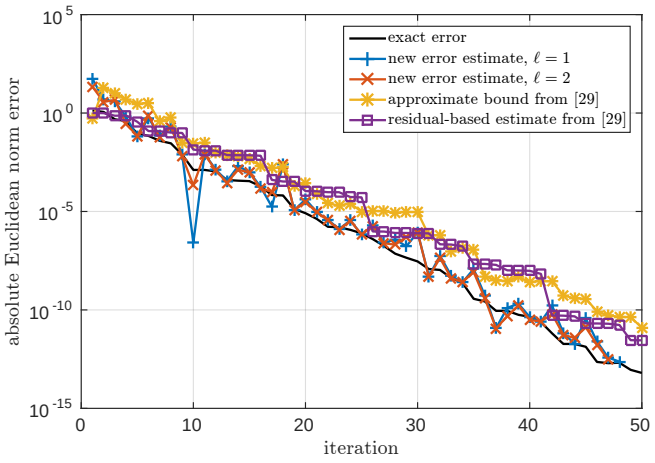

Fig. 7.2 Comparison of different error estimators when approximating the action of the inverse square root of a Jordan block $A \in \mathbb{R}^{4{,}000\times 4{,}000}$ on a normalized random vector. The poles in the rational Arnoldi algorithm are chosen adaptively according to [29].

two parameters δ, ϕ

$$a_{ij} = \begin{cases} 1 + \phi \sum_{k=1, k\neq i}^{n} \chi_{ik}^{\delta} & \text{if } i = j, \\ -\phi \chi_{ij}^{\delta} & \text{otherwise,} \end{cases}$$

where χ^{δ} is given by

$$\chi_{ij}^{\delta} = \begin{cases} 1 & \text{if } \|s_i - s_j\|_2 < \delta, \\ 0 & \text{otherwise.} \end{cases}$$

This matrix is Hermitian and strictly diagonally dominant and its smallest eigenvalue is 1. A sample from a Gaussian Markov random field—a collection of random variables x_i corresponding to the points s_i—is obtained by computing $A^{-1/2}\boldsymbol{z}$, where $\boldsymbol{z}$ is a vector of independently and identically distributed standard normal random variables. We generate the precision matrix for $n = 50{,}000$ pseudo-random points which are uniformly distributed in the unit square with parameters $\phi = 3, \delta = 0.01$, resulting in $\operatorname{spec}(A) \subset [1, 109.6]$ (i.e., a rather well-conditioned matrix) and 830,626 nonzeros in A. As the nonzero entries of A are related to points with a small spatial distance, A can be reordered to have rather small bandwidth, so that linear systems with A can be solved efficiently and it is attractive to use rational Krylov methods.

We again test the same error estimators as in the previous experiments and choose the shifts in the rational Arnoldi method adaptively, and we observe a similar behavior also for this application-oriented problem. The residual-based estimator tends to overestimate the order of magnitude of the error the most (while quite accurately predicting the shape of the convergence curve) while the other estimates lie rather close to the exact error norm, he quadrature-based estimate for $\ell = 2$ being the most accurate, slightly outperforming the more costly approximate bound from [29].

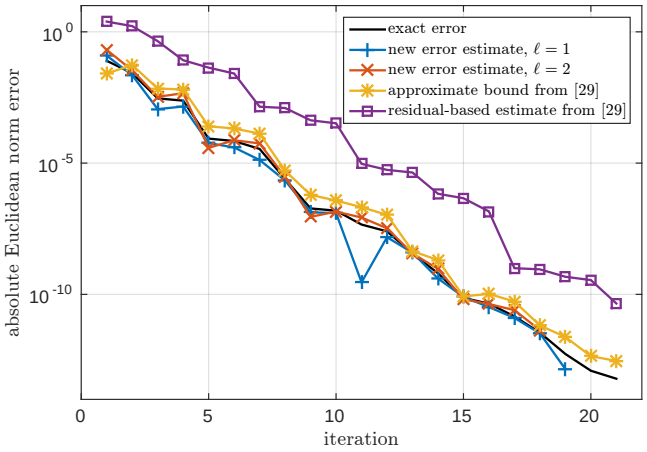

Fig. 7.3 Comparison of different error estimators when approximating the action of the inverse square root for sampling from a Gaussian Markov random field with precision matrix $A \in \mathbb{R}^{50,000\times 50,000}$. The poles in the rational Arnoldi algorithm are chosen adaptively according to [29].

Remark 7.1 In [23], where quadrature-based bounds and estimates for polynomial Krylov subspace approximations to Stieltjes matrix functions are presented, the authors also investigate the influence of loss of orthogonality of the Lanczos basis due to propagated round-off error on the quality of the estimates. We refrain from doing so here for different reasons. On the one hand, the experiments and reasoning in [23] demonstrate that one can expect that the error estimates are not negatively influenced by round-off error, and on the other hand loss of orthogonality is typically not a problem in rational Krylov subspace methods due to the use of long recurrences and the rather small number of iterations that is typically performed.

8 Conclusions

We investigated the possibility of using (rational) Gauss quadrature rules for computing error estimates in rational Krylov subspace methods for approximating $f(A)\boldsymbol{b}$ when f is a Stieltjes function. To do so, we generalized the concept of Lanczos restart recovery to the rational Krylov case to allow the cost-efficient computation of these quadrature rules. We proved that Gauss and Gauss–Radau rules give lower and upper bounds for the error in the rational Arnoldi method when A is Hermitian positive definite and all poles are chosen on the negative real axis. In numerical experiments, we compared our new error estimates to other error indicators proposed in the literature.

Interesting topics for future research could include the extension of the results to other classes of functions and, in particular, identification of situations in which rational Gauss rules provide error bounds while standard Gauss rules do not.

References

1. H. ALZER AND C. BERG, *Some classes of completely monotonic functions*, Ann. Acad. Sci. Fenn. Math., 27 (2002), pp. 445–460.
2. B. BECKERMANN AND L. REICHEL, *Error estimates and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883.
3. M. BENZI AND P. BOITO, *Quadrature rule-based bounds for functions of adjacency matrices*, Linear Algebra Appl., 433 (2010), pp. 637–652.
4. C. BERG, *Stieltjes-Pick-Bernstein-Schoenberg and their connection to complete monotonicity*, in Positive Definite Functions. From Schoenberg to Space-Time Challenges, S. Mateu and E. Porcu, eds., Dept. of Mathematics, University Jaume I, Castellón de la Plana, Spain, 2008.
5. C. BERG AND G. FORST, *Potential Theory on Locally Compact Abelian Groups*, Springer-Verlag, Heidelberg, 1975.
6. M. BERLJAFA AND S. GÜTTEL, *Generalized rational Krylov decompositions with an application to rational approximation*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 894–916.
7. J. BLOCH, A. FROMMER, B. LANG, AND T. WETTIG, *An iterative method to compute the sign function of a non-Hermitian matrix and its application to the overlap Dirac operator at nonzero chemical potential*, Comput. Phys. Commun., 177 (2007), pp. 933–943.
8. J. BRANNICK, A. FROMMER, K. KAHL, B. LEDER, M. ROTTMANN, AND A. STREBEL, *Multigrid preconditioning for the overlap operator in lattice QCD*, Numer. Math., (2015), p. to appear.
9. D. CALVETTI, S. KIM, AND L. REICHEL, *Quadrature rules based on the Arnoldi process*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 765–781.
10. N. CUNDY, J. VAN DEN ESHOF, A. FROMMER, S. KRIEG, T. LIPPERT, AND K. SCHÄFER, *Numerical methods for the QCD overlap operator. III: Nested iterations*, Comput. Phys. Commun., 165 (2005), pp. 221–242.
11. P. J. DAVIS AND P. RABINOWITZ, *Methods of Numerical Integration*, Academic Press, New York, 1975.
12. V. DRUSKIN AND L. KNIZHNERMAN, *Two polynomial methods of calculating functions of symmetric matrices*, Comput. Math. Math. Phys., 29 (1989), pp. 112–121.
13. V. DRUSKIN AND L. KNIZHNERMAN, *Extended Krylov subspaces: Approximation of the matrix square root and related functions*, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 775–771.
14. V. DRUSKIN, L. KNIZHNERMAN, AND M. ZASLAVSKY, *Solution of large scale evolutionary problems using rational Krylov subspaces with optimized shifts*, SIAM J. Sci. Comput., 31 (2009), pp. 3760–3780.
15. M. EIERMANN AND O. G. ERNST, *A restarted Krylov subspace method for the evaluation of matrix functions*, SIAM J. Numer. Anal., 44 (2006), pp. 2481–2504.
16. J. VAN DEN ESHOF, A. FROMMER, T. LIPPERT, K. SCHILLING, AND H. A. VAN DER VORST, *Numerical methods for the QCD overlap operator, I. Sign-function and error bounds*, Comput. Phys. Commun., 146 (2002), pp. 203–224.
17. E. ESTRADA AND D. HIGHAM, *Network properties revealed through matrix functions*, SIAM Rev., 52 (2010), pp. 696–714.
18. R. W. FREUND AND M. HOCHBRUCK, *Gauss quadrature associated with the Arnoldi process and the Lanczos algorithm*, in Linear algebra for large scale and real-time applications, M. S. Moonen, G. H. Golub, and B. De Moor, eds., Kluwer, Dordrecht, 1993, pp. 377–380.
19. A. FROMMER, *Monotone convergence of the Lanczos approximations to matrix functions of Hermitian matrices*, Electron. Trans. Numer. Anal., 35 (2009), pp. 118–128.
20. A. FROMMER, S. GÜTTEL, AND M. SCHWEITZER, *Convergence of restarted Krylov subspace methods for Stieltjes functions of matrices*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 1602–1624.
21. A. FROMMER, S. GÜTTEL, AND M. SCHWEITZER, *Efficient and stable Arnoldi restarts for matrix functions based on quadrature*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 661–683.
22. A. FROMMER, K. KAHL, T. LIPPERT, AND H. RITTICH, *2-norm error bounds and estimates for Lanczos approximations to linear systems and rational matrix functions*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1046–1065.
23. A. FROMMER AND M. SCHWEITZER, *Error bounds and estimates for Krylov subspace approximations of Stieltjes matrix functions*, BIT, (2015). to appear.
24. E. GALLOPOULOS AND Y. SAAD, *Efficient solution of parabolic equations by Krylov approximation methods*, SIAM J. Sci. Stat. Comp., 13 (1992), pp. 1236–1264.

25. G. H. GOLUB AND G. MEURANT, *Matrices, moments and quadrature*, in Numerical Analysis 1993, Pitman Research Notes in Mathematics Series, Vol. 303, D. F. Griffiths and G. A. Watson, eds., Essex, England, 1994, pp. 105–156.
26. ———, *Matrices, Moments and Quadrature with Applications*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton and Oxford, 2010.
27. S. GÜTTEL, *Rational Krylov Methods for Operator Functions*, PhD thesis, Fakultät für Mathematik und Informatik der Technischen Universität Bergakademie Freiberg, 2010.
28. S. GÜTTEL, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, GAMM-Mitteilungen, 36 (2013), pp. 8–31.
29. S. GÜTTEL AND L. KNIZHNERMAN, *A black-box rational Arnoldi variant for Cauchy–Stieltjes matrix functions*, BIT, 53 (2013), pp. 595–616.
30. P. HENRICI, *Applied and Computational Complex Analysis, Vol. 2*, John Wiley & Sons, New York, 1977.
31. M. HOCHBRUCK AND C. LUBICH, *On Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 34 (1997), pp. 1911–1925.
32. M. HOCHBRUCK, C. LUBICH, AND H. SELHOFER, *Exponential integrators for large systems of differential equations*, SIAM J. Sci. Comput., 19 (1998), pp. 1552–1574.
33. M. HOCHBRUCK AND A. OSTERMANN, *Exponential integrators*, Acta Numer., 19 (2010), pp. 209–286.
34. M. ILIĆ, I. W. TURNER, AND D. P. SIMPSON, *A restarted Lanczos approximation to functions of a symmetric matrix*, IMA J. Numer. Anal., 30 (2010), pp. 1044–1061.
35. C. JAGELS AND L. REICHEL, *The extended Krylov subspace method and orthogonal Laurent polynomials*, Linear Algebra Appl., 431 (2009), pp. 441–458.
36. ———, *Recursion relations for the extended Krylov subspace method*, Linear Algebra Appl., 434 (2011), pp. 1716–1732.
37. ———, *The structure of matrices in rational Gauss quadrature*, Math. Comp., 82 (2013), pp. 2035–2060.
38. ———, *On the computation of gauss quadrature rules for measures with a monomial denominator*, J. Comput. Appl. Math., 286 (2015), pp. 102–113.
39. L. KNIZHNERMAN AND V. SIMONCINI, *A new investigation of the extended Krylov subspace method for matrix function evaluations*, Numer. Linear Algebra Appl., 17 (2010), pp. 615–638.
40. C. LANCZOS, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, J. Res. Nat. Bur. Stand., 45 (1950), pp. 255–282.
41. G. LÓPEZ LAGOMASINO, L. REICHEL, AND L. WUNDERLICH, *Matrices, moments, and rational quadrature*, Linear Algebra Appl., 429 (2008), pp. 2540–2554.
42. C. MERTENS AND R. VANDEBRIL, *Short recurrences for computing extended Krylov bases for Hermitian and unitary matrices*, Numer. Math., 131 (2015), pp. 303–328.
43. A. RUHE, *Rational Krylov sequence methods for eigenvalue computation.*, Linear Algebra Appl., 58 (1984), pp. 391–405.
44. Y. SAAD, *Analysis of some Krylov subspace approximations to the exponential operator*, SIAM J. Numer. Anal., 29 (1992), pp. 209–228.
45. M. SCHWEITZER, *Restarting and error estimation in polynomial and extended Krylov subspace methods for the approximation of matrix functions*, PhD thesis, Bergische Universität Wuppertal, 2016.
46. V. SIMONCINI, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comput., 29 (2007), pp. 1268–1288.
47. D. P. SIMPSON, *Krylov subspace methods for approximating functions of symmetric positive definite matrices with applications to applied statistics and anomalous diffusion*, PhD thesis, Queensland University of Technology, 2008.
48. G. W. STEWART, *Matrix Algorithms Volume II: Eigensystems*, SIAM, 2001.
49. J. VAN DEN ESHOF AND M. HOCHBRUCK, *Preconditioning Lanczos approximations to the matrix exponential*, SIAM J. Sci. Comput., 27 (2006), pp. 1438–1457.
50. J. L. WALSH, *Interpolation and Approximation by Rational Functions in the Complex Domain*, American Mathematical Society, Providence, RI, 1969.