Computation of generalized matrix functions with rational Krylov methods

Angelo A. Casulli [*] Igor Simunec [†]

Abstract

We present a class of algorithms based on rational Krylov methods to compute the action of a generalized matrix function on a vector. These algorithms incorporate existing methods based on the Golub-Kahan bidiagonalization as a special case. By exploiting the quasiseparable structure of the projected matrices, we show that the basis vectors can be updated using a short recurrence, which can be seen as a generalization to the rational case of the Golub-Kahan bidiagonalization. We also prove error bounds that relate the error of these methods to uniform rational approximation. The effectiveness of the algorithms and the accuracy of the bounds is illustrated with numerical experiments.

1 Introduction

Generalized matrix functions (GMFs) are an extension of the notion of matrix functions based on the singular value decomposition (SVD) instead of the spectral decomposition. They were introduced for the first time in [19], with the purpose of extending the definition of matrix functions to rectangular matrices. Although the introduction of GMFs dates to the Seventies [19], they have become a more active area of research only in recent years. For instance, theoretical aspects of generalized matrix functions have been investigated in [1, 6, 28], while efficient numerical methods have been developed in [3, 4]. GMFs have also been recently considered in the context of quantum algorithms [21]. For applications of generalized matrix functions, we direct the reader to [1, 3, 4] and the references therein.

In many applications that involve standard matrix functions, it is only required to compute matrix-vector products of the form $f(A)\boldsymbol{b}$, where the matrix A is usually large and sparse. In this case, the (expensive) computation of the whole matrix $f(A)$ can be bypassed by using methods that directly approximate the product $f(A)\boldsymbol{b}$, such as Krylov methods. These methods only require matrix-vector products and possibly the solution of shifted linear systems with the matrix A.

[*] Scuola Normale Superiore, Piazza dei Cavalieri, 7, 56126 Pisa, Italy, angelo.casulli@sns.it

[†] Scuola Normale Superiore, Piazza dei Cavalieri, 7, 56126 Pisa, Italy, igor.simunec@sns.it

A similar situation arises when GMFs are involved: indeed, it is often required to compute the action of a generalized matrix function on a vector [2, 3], and hence it is preferable to use a method that avoids the computation of the whole GMF by means of an SVD.

This problem was recently investigated in [3, 4], using methods based on the Golub-Kahan bidiagonalization in [3], and Chebyshev polynomial interpolation in [4].

In this paper, we propose a generalization of the method proposed in [3], using the interpretation of the Golub-Kahan bidiagonalization in terms of Krylov subspaces. Performing k steps of the Golub-Kahan bidiagonalization of a matrix A with starting vector $\boldsymbol{b}$ is equivalent to the simultaneous computation of orthonormal bases of the polynomial Krylov subspaces $\mathcal{P}_k(A^T A, \boldsymbol{b})$ and $\mathcal{P}_k(AA^T, A\boldsymbol{b})$. By replacing the polynomial Krylov subspaces with their rational counterparts, we obtain a rational Krylov method for the computation of the action of a GMF on a vector.

As can be expected by analogy with standard matrix functions, in the case of non-analytic functions and functions of low regularity these rational methods have a faster convergence than the method based on the Golub-Kahan bidiagonalization. However, their increased effectiveness comes at the cost of having to solve a linear system at each iteration, while the methods discussed in [3, 4] only require matrix-vector products.

The Golub-Kahan bidiagonalization of a matrix can be computed with a short recurrence, which relies on the fact that the projected matrix is bidiagonal. This structure is unfortunately not preserved in the rational case that we consider here. However, we are still able to construct a short recurrence to update the rational Krylov bases and the projected matrix, using the fact that the projected matrix is a quasiseparable matrix [15, 35]. We mention that structured projected matrices that lead to short recurrences also appear in the context of biorthogonal rational Krylov methods [36]. Rank structures of the projected rational Krylov matrices were also exploited in [22, 23].

We also prove error bounds that link the error of the method from [3] based on the Golub-Kahan bidiagonalization and the rational methods introduced in this paper with, respectively, the error of uniform polynomial and rational approximation of the function f. These bounds are a direct generalization of the bounds for standard matrix functions, and they can be proved with the same techniques. Although the connection of GMFs to standard matrix functions is well-known, to the best of our knowledge these error bounds for the approximation of GMFs have never appeared in previous literature.

The paper is organized as follows. In Section 2 we introduce the notation used throughout the paper. In Section 3 we recall the definition of standard matrix functions and GMFs and we present some of their properties. In Section 4 we briefly introduce the class of rational Krylov methods for standard matrix functions. The use of polynomial and rational Krylov methods in the context of generalized matrix functions is discussed in Section 5. Section 6 is dedicated to the proof of the error bounds and related discussion. Some numerical experiments to compare the different methods and illustrate the error bounds are presented in Section 7, and Section 8 contains concluding remarks.

2 Notation

We denote by $\mathbb{R}^{m\times n}$ the space of $m \times n$ real matrices. We use bold letters for vectors, e.g. $\boldsymbol{v} \in \mathbb{R}^n$. The entries of a vector $\boldsymbol{v}$ are given by $v_1, \dots, v_n$, and the entries of a matrix $A \in \mathbb{R}^{m\times n}$ are a_{ij}. We also use a MATLAB-like notation: $\mathrm{diag}(d_1, \dots, d_n)$ represents an $n \times n$ diagonal matrix with entries $d_1, \dots, d_n$ on the diagonal; for $i \le j$ and $h \le k$, we denote by $A(i:j, h:k)$ the submatrix of A corresponding to row indices from i to j and column indices from h to k.

We denote by $\mathrm{triu}(A)$ the upper triangular part of the matrix A, and more generally by $\mathrm{triu}(A, k)$ the matrix with all zeroes below the k-th diagonal whose other entries coincide with those of A. Diagonals above the main diagonal are represented with a positive index, so that $\mathrm{triu}(A, 1)$ indicates the strictly upper triangular part of A. Similarly, $\mathrm{tril}(A)$ and $\mathrm{tril}(A, -1)$ denote the lower triangular and strictly lower triangular part of A, respectively. We use the same notation also for rectangular matrices. We denote by A^+ the Moore-Penrose pseudoinverse of a matrix A.

3 Matrix functions

The goal of this section is to define generalized matrix functions (GMFs) and introduce their main properties. We begin by recalling some basic concepts about standard matrix functions, and then we introduce GMFs and some of their properties.

3.1 Standard matrix functions

The concept of matrix function is a natural way to extend the evaluation of a scalar function to square matrix arguments. For simplicity, we treat only the case of diagonalizable matrices. The general definitions and a thorough description of matrix functions can be found in the monograph [20].

Let A be an $n \times n$ matrix. Assume that A is diagonalizable, i.e. $A = VDV^{-1}$, where $D = \mathrm{diag}(d_1, \dots, d_n)$. Given a function f defined on the set $\{d_1, \dots, d_n\}$ the matrix function of f applied on A is defined as

$$f(A) = Vf(D)V^{-1},$$

where $f(D) = \mathrm{diag}(f(d_1), \dots, f(d_n))$.

3.2 Generalized matrix functions

Generalized matrix functions were first introduced in [19], with the purpose of extending the definition of matrix functions to rectangular matrices. They are defined in a similar way with respect to standard matrix functions, but the singular value decomposition is used instead of the diagonalization.

Let $A \in \mathbb{R}^{m\times n}$ and let $A = U\Sigma V^T$ be its SVD, where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal and $\Sigma \in \mathbb{R}^{m\times n}$ is defined as

$$\Sigma_{i,j} = \begin{cases} \sigma_i & \text{if } i = j \le r \\ 0 & \text{otherwise,} \end{cases}$$

where $r \le \min\{m, n\}$ is the rank of A and $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$ are the nonzero singular values of A.

Given a function f defined on the set $\{\sigma_1, \dots, \sigma_r\}$ the generalized matrix function of f applied on A is defined as

$$f^\diamond(A) = U f^\diamond(\Sigma) V^T,$$

where

$$f^\diamond(\Sigma)_{i,j} = \begin{cases} f(\sigma_i) & \text{if } i = j \le r \\ 0 & \text{otherwise.} \end{cases}$$

Observe that a GMF can be expressed in terms of the compact SVD of the matrix A, that is $A = U_r\Sigma_r V_r^T$, where $U_r \in \mathbb{R}^{m\times r}$ and $V \in \mathbb{R}^{n\times r}$ have orthonormal columns, and $\Sigma_r = \operatorname{diag}(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{r\times r}$. In such case, we have

$$f^\diamond(A) = U_r f(\Sigma_r) V_r^T.$$

Since the definition of a GMF only depends on the values of f on the nonzero singular values of A, we can always assume that f is an odd function, and in particular that $f(0) = 0$. Note also that although we only consider real matrices for simplicity, all of our results can be easily generalized to the complex case, since the singular values are always nonnegative real numbers.

Remark 3.1. [4, Theorem 2.1] If p is a polynomial that interpolates f in $\sigma_1, \dots, \sigma_r$ we have that $f^\diamond(A) = p^\diamond(A)$. Moreover, since $\sigma_i > 0$ for $i = 1, \dots, r$, we can always take p as an odd polynomial, i.e. $p(z) = q(z^2)z$ for some polynomial q.

Next, we list some properties of GMFs that will be required in the following sections. A discussion of additional properties of generalized matrix functions can be found in [3].

Lemma 3.2. *Let $S \in \mathbb{R}^{n\times n}$ be a symmetric matrix and let f be defined on the singular values of S. If S is positive definite, or S is positive semidefinite and $f(0) = 0$, we have*

$$f^\diamond(S) = f(S).$$

Proposition 3.3. *Let p be an odd polynomial, i.e. we can write $p(z) = q(z^2)z$ for some polynomial q. Then, for any matrix $A \in \mathbb{R}^{m\times n}$ the following holds:*

$$p^\diamond(A) = q(AA^T)A = Aq(A^TA).$$

For the proof of Proposition 3.3 we refer to [4, Section 2.2]. The following corollary can be proved in practically the same way.

Corollary 3.4. *Let r be a rational function with an odd numerator and an even denominator, i.e., $r(z) = q(z^2)^{-1}p(z^2)z$ for some polynomials p and q, such that the singular values of $A \in \mathbb{R}^{m\times n}$ and 0 are not roots of q. Then the following holds:*

$$r^\diamond(A) = q(AA^T)^{-1}p(AA^T)A = Aq(A^TA)^{-1}p(A^TA).$$

The following proposition gives a formulation of the above results for general functions. See also [3, Theorem 10] for an alternative formulation of the same identity.

Proposition 3.5. *Let $A \in \mathbb{R}^{m\times n}$ and let f be a function defined on the nonzero singular values of A. Defining $g(z) = \frac{f(\sqrt{z})}{\sqrt{z}}$, for $z \neq 0$, we have*

$$f^\diamond(A) = g^\diamond(AA^T)A = Ag^\diamond(A^TA).$$

Moreover, if $\lim\limits_{z\to 0} \frac{f(z)}{z} = 0$, we can define $g(0) = 0$. Then $g(AA^T)$ and $g(A^TA)$ are well defined, and we have

$$f^\diamond(A) = g(AA^T)A = Ag(A^TA).$$

Proof. Let $p(z) = q(z^2)z$ be an odd polynomial that interpolates f in the nonzero singular values of A. From Proposition 3.3 and Remark 3.1 we have

$$f^\diamond(A) = p^\diamond(A) = q(AA^T)A = Aq(A^TA).$$

Since p interpolates f in the nonzero singular values of A, q interpolates g in the squares of the nonzero singular values of A, which are the nonzero singular values of AA^T and A^TA. Hence $q^\diamond(AA^T) = g^\diamond(AA^T)$ and $q^\diamond(A^TA) = g^\diamond(A^TA)$.

If $g(0) = 0$, by Lemma 3.2 we also have $g^\diamond(A^TA) = g(A^TA)$ and $g^\diamond(AA^T) = g(AA^T)$. □

Remark 3.6. If AA^T is positive definite, by Lemma 3.2 we have that $f^\diamond(A) = g(AA^T)A$ without the assumption $\lim\limits_{z\to 0} \frac{f(z)}{z} = 0$, and similarly if A^TA is positive definite we have $f^\diamond(A) = Ag(A^TA)$. Note that we could also artificially define $g(0) = 0$ without any assumptions on f, since the definition of a GMF only depends on the nonzero singular values of the matrix. However, this would cause g to be discontinuous at 0 and it would not be very useful in practice.

The following proposition links $f^\diamond(A)$ with $f^\diamond(A^T)$, which will be useful when A is rectangular. The more general statement in Proposition 3.7 can be seen as a generalization of [3, Proposition 7(iv)] and [19, Theorem 4(d)].

Proposition 3.7. *Let $A \in \mathbb{R}^{m\times n}$ and let f be a function defined on the nonzero singular values of A. Then*

$$f^\diamond(A) = (A^+)^T f^\diamond(A^T)A.$$

More generally, assume that $f(z) = g(z)h(z)k(z)$, where g, h, k are functions defined on the nonzero singular values of A. Then

$$f^{\diamond}(A) = g^{\diamond}(A)h^{\diamond}(A^T)k^{\diamond}(A).$$

Proof. We directly prove the generalized version, since the first statement simply follows by taking $g(z) = z^{-1}$, $h(z) = f(z)$ and $k(z) = z$.

Let A have the singular value decomposition $A = U\Sigma V^T$. We have:

$$\begin{aligned} g^{\diamond}(A)h^{\diamond}(A^T)k^{\diamond}(A) &= Ug^{\diamond}(\Sigma)V^TVh^{\diamond}(\Sigma^T)U^TUk^{\diamond}(\Sigma)V^T \\ &= Ug^{\diamond}(\Sigma)h^{\diamond}(\Sigma)^Tk^{\diamond}(\Sigma)V^T \\ &= Uf^{\diamond}(\Sigma)V^T = f^{\diamond}(A), \end{aligned}$$

where we used that $g^{\diamond}(\Sigma)h^{\diamond}(\Sigma)^Tk^{\diamond}(\Sigma) = f^{\diamond}(\Sigma)$, which can be verified directly. □

Remark 3.8. The same proof of Proposition 3.7 can be used to show that, if $f(z)g(z) = h(z)k(z)$, then

$$f^{\diamond}(A^T)g^{\diamond}(A) = h^{\diamond}(A^T)k^{\diamond}(A).$$

In particular we have $A^Tf^{\diamond}(A) = f^{\diamond}(A^T)A$.

4 Rational Krylov methods

The class of Krylov methods provides an efficient way to compute approximations to expressions of the form $f(A)\boldsymbol{b}$. The main idea behind these methods is to construct a low dimensional subspace $\mathcal{S}_k \subset \mathbb{R}^n$ for some integer $k \ll n$ using information from A and $\boldsymbol{b}$, and then to approximate $f(A)\boldsymbol{b}$ with an appropriate vector from $\mathcal{S}_k$.

A popular choice for the approximation subspace $\mathcal{S}_k$ is the *polynomial Krylov subspace*

$$\mathcal{P}_k(A, \boldsymbol{b}) = \operatorname{span}\{\boldsymbol{b}, A\boldsymbol{b}, \dots, A^{k-1}\boldsymbol{b}\} = \{p(A)\boldsymbol{b} : p \in \Pi_{k-1}\},$$

where Π_{k-1} denotes the set of polynomials of degree $\leq k-1$.

More generally, using a sequence of *poles* $\{\xi_k\}_{k\geq 1} \subseteq (\mathbb{C} \cup \{\infty\}) \setminus \sigma(A)$, one can define the *rational Krylov subspace*

$$\mathcal{Q}_k(A, \boldsymbol{b}) = q_{k-1}(A)^{-1}\mathcal{P}_k(A, \boldsymbol{b}) = \Big\{r(A)\boldsymbol{b} : r(z) = \frac{p_{k-1}(z)}{q_{k-1}(z)}, \text{with } p_{k-1} \in \Pi_{k-1}\Big\},$$

where $q_{k-1}(z) = \displaystyle\prod_{j=1}^{k-1}(1 - z/\xi_j)$. In the case when all poles are located at ∞, we have $q_{k-1}(z) \equiv 1$ and hence we recover the polynomial Krylov subspace $\mathcal{P}_k(A, \boldsymbol{b})$. It is easy to verify that the Krylov subspaces $\mathcal{Q}_k(A, \boldsymbol{b})$ form a nested sequence, and that $\dim \mathcal{Q}_k(A, \boldsymbol{b}) = k$ as long as k is smaller than the *invariance index* K of the sequence, i.e. the smallest integer such that $\mathcal{Q}_K(A, \boldsymbol{b}) = \mathcal{Q}_{K+1}(A, \boldsymbol{b})$ (or, equivalently, $\mathcal{P}_K(A, \boldsymbol{b}) = \mathcal{P}_{K+1}(A, \boldsymbol{b})$).

For $k \leq K$, an orthonormal basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ of $\mathcal{Q}_k(A, \boldsymbol{b})$ can be computed with the rational Arnoldi algorithm, introduced by Ruhe in [31]. In the basic algorithm, the first basis vector is chosen as $\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2$. Then, given a set of vectors $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_j\}$ which form an orthonormal basis of $\mathcal{Q}_j(A, \boldsymbol{b})$, the next basis vector $\boldsymbol{v}_{j+1}$ is computed by orthonormalizing the vector $(I - \frac{1}{\xi_j}A)^{-1}A\boldsymbol{v}_j$ against the previously computed basis vectors. To prevent the algorithm from failing, it is required that $(I - \frac{1}{\xi_j}A)^{-1}A\boldsymbol{v}_j \in \mathcal{Q}_{j+1}(A, \boldsymbol{b}) \setminus \mathcal{Q}_j(A, \boldsymbol{b})$; this property is almost always satisfied in practice, however there are no theoretical guarantees that it holds. A vector $\boldsymbol{w}_j$ that guarantees $(I - \frac{1}{\xi_j}A)^{-1}\boldsymbol{w}_j \in \mathcal{Q}_{j+1}(A, \boldsymbol{b}) \setminus \mathcal{Q}_j(A, \boldsymbol{b})$, hence ensuring that the Krylov subspace is enlarged, can be found with the approach recently discussed in [9], using the notion of *continuation pairs* $(\eta_j/\rho_j, \boldsymbol{t}_j)$: in general such a vector $\boldsymbol{w}_j$ is of the form $(\rho_m A - \eta_m I)V_j \boldsymbol{t}_j$, where $V_j = [\boldsymbol{v}_1 \dots \boldsymbol{v}_j] \in \mathbb{C}^{n \times j}$. From now on, we always assume that the dimension of the Krylov subspace actually increases at each iteration, i.e. that $\boldsymbol{v}_{j+1} \in \mathcal{Q}_{j+1}(A, \boldsymbol{b}) \setminus \mathcal{Q}_j(A, \boldsymbol{b})$ for all $j < K$.

Using the matrix with orthonormal columns $V_k = [\boldsymbol{v}_1 \dots \boldsymbol{v}_k] \in \mathbb{C}^{n \times k}$, we can compute the following approximation to $f(A)\boldsymbol{b}$ from the subspace $\mathcal{Q}_k(A, \boldsymbol{b})$:

$$\bar{\boldsymbol{y}}_k = V_k f(V_k^T A V_k) V_k^T \boldsymbol{b} = V_k f(A_k) \boldsymbol{e}_1,$$

where $A_k = V_k^T A V_k$ is the projection of A on the subspace $\mathcal{Q}_k(A, \boldsymbol{b})$, and $\boldsymbol{e}_1$ denotes the first vector of the canonical basis of $\mathbb{C}^n$. The accuracy of the approximation $\bar{\boldsymbol{y}}_k$ is largely dependent on the pole sequence $\{\xi_k\}_{k \geq 1}$. The problem of choosing a sequence of poles that is effective for a particular function f and a given set containing the spectrum of A has often been discussed in the literature: we refer, for instance, to [18] and the references therein.

Some specific sequences of poles lead to special cases of the rational Krylov subspace $\mathcal{Q}_k(A, \boldsymbol{b})$: if all the poles are equal to $\xi \in \mathbb{C} \setminus \sigma(A)$, then $\mathcal{Q}_k(A, \boldsymbol{b})$ is a *Shift-and-Invert Krylov subspace*,

$$Q_k(A, \boldsymbol{b}) = \mathcal{P}_k((I - A/\xi)^{-1}, \boldsymbol{b}),$$

that was first investigated for the computation of matrix functions in [25, 37]. If the poles alternate between 0 and ∞, we obtain the *extended Krylov subspace*, introduced in [13], which is of the form

$$\mathcal{Q}_{2k}(A, \boldsymbol{b}) = A^{-k}\mathcal{P}_{2k}(A, \boldsymbol{b}) = \mathcal{P}_{2k}(A, A^{-k}\boldsymbol{b}).$$

We refer to [17] for an extensive discussion on rational Krylov methods for the computation of matrix functions.

5 Krylov methods for GMFs

The computation of $f^\diamond(A)\boldsymbol{b}$ by using an SVD of A can be unfeasible if the size of A is large. A possible way to approximate the product of a generalized matrix function times a vector is to use two rectangular matrices with orthonormal columns to project the matrix A onto a smaller space and then to compute the generalized matrix function of the projected matrix: let $k \ll n$ and let $U_k, V_k \in \mathbb{R}^{n\times k}$ with orthonormal columns, $B_k \in \mathbb{R}^{k\times k}$ such that $A \approx U_k B_k V_k^T$, then

$$f^{\diamond}(A)\boldsymbol{b} \approx U_k f^{\diamond}(B_k) V_k^T \boldsymbol{b}, \tag{5.1}$$

where the matrix $f^{\diamond}(B_k)$ can be computed by means of the singular value decomposition.

In this section we describe how to compute such a projected matrix using the Golub-Kahan bidiagonalization, which is equivalent to using a polynomial Krylov method on the matrices A^TA and AA^T. This strategy corresponds to the "third approach" discussed in [3, Section 5.4]; the numerical results of [3] indicate that (5.1) is often more effective than the other approaches they propose, which are based on Gauss and Gauss-Radau quadrature formulas.

In Sections 5.2 and 5.3 we generalize this approach to the rational Krylov case and we show that a short recurrence like the one of the Golub-Kahan bidiagonalization can be obtained in the rational case too.

5.1 Golub-Kahan bidiagonalization

The first method we describe for the computation of a truncated SVD is the Golub-Kahan bidiagonalization introduced for the first time in 1965.

Theorem 5.1. *Let $A \in \mathbb{R}^{m\times n}$, with $m > n$. There exist orthogonal matrices $P \in \mathbb{R}^{m\times m}, Q \in \mathbb{R}^{n\times n}$ such that*

$$P^TAQ = B = \begin{bmatrix} \alpha_1 & \beta_1 & \dots & \dots & 0 \\ 0 & \alpha_2 & \beta_2 & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & 0 & \alpha_{n-1} & \beta_{n-1} \\ 0 & \dots & \dots & 0 & \alpha_n \\ 0 & \dots & \dots & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & \dots & \dots & \dots & 0 \end{bmatrix}.$$

Moreover the first column of Q can be chosen almost[1] *arbitrarily.*

The proof of Theorem 5.1 is constructive and it is usually called Householder bidiagonalization process. It can be found in [16, Section 5.4].

In the case of large matrices the full bidiagonalization is too expensive. The goal of the Golub-Kahan bidiagonalization is to extract good approximations

[1] In order to avoid breakdowns, the first column of Q must be chosen as a unit vector $\boldsymbol{q}_1$, such that $\mathcal{P}_n(A^TA, \boldsymbol{q}_1)$ has dimension n.

of singular values and singular vectors before the full bidiagonalization is completed.

Denote by $\boldsymbol{p}_j$ and $\boldsymbol{q}_j$ the columns of P and Q, respectively. By stopping the bidiagonalization process after k steps, we obtain the matrices $P_k = [\boldsymbol{p}_1|\dots|\boldsymbol{p}_k]$, $Q_k = [\boldsymbol{q}_1|\dots|\boldsymbol{q}_k]$ and

$$B_k = \begin{bmatrix} \alpha_1 & \beta_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \beta_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & \alpha_{k-1} & \beta_{k-1} \\ 0 & \dots & \dots & 0 & \alpha_k \end{bmatrix},$$

such that

$$\begin{aligned} AQ_k &= P_k B_k, \\ A^T P_k &= Q_k B_k^T + \boldsymbol{s}_k \boldsymbol{e}_k^T, \end{aligned}$$

where $\boldsymbol{s}_k = A^T \boldsymbol{p}_k - \alpha_k \boldsymbol{q}_k$. In particular $P_k^T A Q_k = B_k$. The columns of P_k, Q_k and the entries of B_k can be computed using the short recurrences

$$\begin{aligned} A\boldsymbol{q}_j &= \alpha_j \boldsymbol{p}_j + \beta_{j-1} \boldsymbol{p}_{j-1}, \qquad j \geq 2, \\ A^T \boldsymbol{p}_j &= \alpha_j \boldsymbol{q}_j + \beta_j \boldsymbol{q}_{j+1}, \qquad j \geq 1. \end{aligned} \tag{5.2}$$

It can be shown that

$$\begin{aligned} \operatorname{span}\{\boldsymbol{q}_1, \dots, \boldsymbol{q}_k\} &= \mathcal{P}_k(A^T A, \boldsymbol{q}_1), \\ \operatorname{span}\{\boldsymbol{p}_1, \dots, \boldsymbol{p}_k\} &= \mathcal{P}_k(AA^T, A\boldsymbol{q}_1), \end{aligned}$$

thus the convergence of the Golub-Kahan bidiagonalization follows from the convergence of the Lanczos method applied on $A^T A$ and AA^T.

For a given k, we can approximate A with $P_k B_k Q_k^T$, and hence we can compute an approximation of the SVD of A by computing the SVD of B_k. For further information on the Golub-Kahan bidiagonalization we refer to [16, Chapter 10].

The vector $\boldsymbol{y} = f^\diamond(A)\boldsymbol{b}$ can be approximated by the expression

$$\bar{\boldsymbol{y}}_k = P_k f^\diamond(B_k) Q_k^T \boldsymbol{b} = \|\boldsymbol{b}\|_2 P_k f^\diamond(B_k) \boldsymbol{e}_1 \in \mathcal{P}_k(AA^T, A\boldsymbol{b}). \tag{5.3}$$

We refer to this approximation as a polynomial Krylov method for GMFs.

5.2 Rational Krylov methods for GMFs

As we saw in the previous section, the Golub-Kahan bidiagonalization computes orthonormal bases for the polynomial Krylov subspaces $\mathcal{P}_k(A^T A, \boldsymbol{b})$ and $\mathcal{P}_k(AA^T, A\boldsymbol{b})$. By analogy with that approach, in this section we propose to compute an approximation to $f^\diamond(A)\boldsymbol{b}$ using the rational Krylov subspaces

$$\mathcal{Q}_k(A^T A, \boldsymbol{b}) \qquad \text{and} \qquad \mathcal{Q}_k(AA^T, A\boldsymbol{b}),$$

where $q_{k-1}(z) = \prod_{j=1}^{k-1}(1 - z/\xi_j)$, for a given pole sequence $\{\xi_j\}_{j\geq 1}$.

Assume that we have constructed two matrices with orthonormal columns P_k and Q_k, such that $\mathrm{span}(P_k) = \mathcal{Q}_k(AA^T, A\boldsymbol{b})$ and $\mathrm{span}(Q_k) = \mathcal{Q}_k(A^TA, \boldsymbol{b})$. Then, defining $B_k = P_k^T A Q_k$, by analogy with the polynomial Krylov approach we can introduce the vector

$$\bar{\boldsymbol{y}}_k = P_k f^\diamond(B_k) Q_k^T \boldsymbol{b} = P_k f^\diamond(B_k)\boldsymbol{e}_1, \tag{5.4}$$

which is an approximation to $f^\diamond(A)\boldsymbol{b}$ from the subspace $\mathcal{Q}_k(AA^T, A\boldsymbol{b})$.

First of all, notice that it is sufficient to compute only one of the two rational Krylov subspaces: indeed, since we have $A\mathcal{Q}_k(A^TA, \boldsymbol{b}) = \mathcal{Q}_k(AA^T, A\boldsymbol{b})$, we can compute an orthonormal basis of the subspace $\mathcal{Q}_k(AA^T, A\boldsymbol{b})$ simply by orthonormalizing the columns of AQ_k. This is equivalent to computing a QR decomposition $AQ_k = W_kR_k$, where W_k has orthonormal columns and R_k is upper triangular, so we can set $P_k = W_k$. Moreover, notice that we also have

$$B_k = P_k^T A Q_k = W_k^T W_k R_k = R_k,$$

i.e. with the QR decomposition we also recover the matrix B_k, without the need to project A explicitly. The basis Q_k of the subspace $\mathcal{Q}_k(A^TA, \boldsymbol{b})$ can be computed by applying the rational Arnoldi algorithm to the matrix A^TA with initial vector $\boldsymbol{b}$. This procedure is summarized in Algorithm 1.

Algorithm 1: Rational Krylov approximation of $f^\diamond(A)\boldsymbol{b}$

```
Input: $A \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^n, f, \{\xi_1, \dots, \xi_{k-1}\}$
Output: $\bar{\boldsymbol{y}}_k \in \mathcal{Q}_k(AA^T, A\boldsymbol{b})$ s.t. $\bar{\boldsymbol{y}}_k \approx f^\diamond(A)\boldsymbol{b}$
$\boldsymbol{q}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2$
for $j = 1, \dots, k-1$ do
    $\boldsymbol{w}_j = (I - A^TA/\xi_j)^{-1}A^TA\boldsymbol{q}_j$        // other choices can be used
    Compute $\boldsymbol{q}_{j+1}$ by orthogonalizing $\boldsymbol{w}_j$ against $[\boldsymbol{q}_1, \dots, \boldsymbol{q}_j]$
$Q_k = [\boldsymbol{q}_1, \dots, \boldsymbol{q}_k]$                    // orthonormal basis of $\mathcal{Q}_k(A^TA, \boldsymbol{b})$
Compute the QR decomposition $P_kB_k = Q_k$
Compute $f^\diamond(B_k)$, e.g. via an SVD of $B_k$
$\bar{\boldsymbol{y}}_k = P_k f^\diamond(B_k)\boldsymbol{e}_1$
```

5.3 Short recurrence for rational Golub-Kahan algorithm

In the Golub-Kahan bidiagonalizazion (without reorthogonalization), we can compute the last column of P_k, Q_k and B_k just by knowing a few previous columns of P_k and Q_k, by means of the equations (5.2). This short recurrence is possible because of the bidiagonal structure of the matrix B_k, that unfortunately is not preserved when we perform a rational Krylov method.

In this section we show that, if a rational Krylov method is used, the matrix B_k is a quasiseparable matrix (see Definition 5.3). This structure extends the bidiagonal form that is obtained during the Golub-Kahan bidiagonalization. Using such structure we build a short recurrence that allows us to update the matrix P_k and the matrix B_k avoiding the full orthogonalization related to the computation of the QR factorization of the matrix AQ_k.

We note that rank structures of rational Krylov matrices have already been investigated in the literature, see for instance [10, 15, 35].

Definition 5.2. A matrix $S \in \mathbb{R}^{n\times n}$ is called a semiseparable matrix if all the submatrices extracted from the lower and upper triangular part of the matrix have rank at most 1, that is

$$\operatorname{rank} S(i:n, 1:i) \leq 1 \quad \text{and} \quad \operatorname{rank} S(1:i, i:n) \leq 1,$$

for every $i = 1, \ldots, n$. Moreover, S is called a generator representable semiseparable matrix if the lower and upper triangular parts of the matrix are derived from a rank 1 matrix, that is

$$\operatorname{tril}(S) = \operatorname{tril}(\boldsymbol{u}\boldsymbol{v}^T) \quad \text{and} \quad \operatorname{triu}(S) = \operatorname{triu}(\boldsymbol{p}\boldsymbol{q}^T),$$

for $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^n$.

Definition 5.3. A matrix S is called a quasiseparable matrix if all the submatrices extracted from the strictly lower and strictly upper triangular part of the matrix are of rank at most 1, that is

$$\operatorname{rank} S(i+1:n, 1:i) \leq 1 \quad \text{and} \quad \operatorname{rank} S(1:i, i+1:n) \leq 1,$$

for every $i = 1, \ldots, n$.

The following theorem from [38, Section 1.5.2] gives us a complete characterization of invertible semiseparable matrices.

Theorem 5.4. *The inverse of an invertible tridiagonal matrix is a semiseparable matrix, and vice versa. Moreover, the inverse of an invertible irreducible tridiagonal matrix is a generator representable semiseparable matrix and vice versa.*

As it has been proved in [17, Section 5.2], if we perform a rational Krylov algorithm on a symmetric matrix we obtain a particular equivalence called rational Arnoldi decomposition. See also [8, Section 2] and [30, eq. (2.2)].

Theorem 5.5. *Given a symmetric matrix $A \in \mathbb{R}^{n\times n}$ and a vector $\boldsymbol{b} \in \mathbb{R}^n$, let $\mathcal{Q}_{k+1}(A, \boldsymbol{b})$ be the rational Krylov space with nonzero poles $\{\xi_1, \ldots, \xi_k\}$ and assume that k is less than the invariance index of the Krylov subspace. Let $Q_{k+1} \in \mathbb{R}^{n\times(k+1)}$ be the matrix with orthonormal columns generated by the Arnoldi algorithm, such that* $\operatorname{span}(Q_{k+1}) = \mathcal{Q}_{k+1}(A, \boldsymbol{b})$. *Then the following relation holds:*

$$AQ_{k+1}\underline{K_k} = Q_{k+1}\underline{H_k}, \quad \textit{with} \quad \underline{K}_k = \underline{I}_k + D_k\underline{H}_k, \tag{5.5}$$

where $\underline{H_k} \in \mathbb{R}^{(k+1)\times k}$ is a full rank tridiagonal irreducible symmetric matrix, $\underline{I}_k$ is the $(k+1)\times k$ identity matrix and $D_k = diag(0, \frac{1}{\xi_1}, \dots, \frac{1}{\xi_{k-1}})$, where $\frac{1}{\infty} = 0$.

Starting form (5.5), we are going to prove that the projection of the symmetric matrix A on the Krylov subspace (i.e., $Q_{k+1}^T A Q_{k+1}$) is the sum of a diagonal matrix and a semiseparable matrix. A similar result has been proved by Fasino, see [15, Theorem 1].

Theorem 5.6. *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix and let $Q_{k+1} \in \mathbb{R}^{n\times(k+1)}$ be the matrix with orthonormal columns generated by the rational Arnoldi algorithm using poles $\{\xi_1, \dots, \xi_k\}$ different from zero and infinity. Assuming that $k+1$ is less than the invariance index of the Krylov subspace, we have*

$$J_{k+1} := Q_{k+1}^T A Q_{k+1} = S + \tilde{D}_k,$$

where $\tilde{D}_k = \operatorname{diag}(0, \xi_1, \dots, \xi_k)$ and S is a symmetric generator representable semiseparable matrix.

Proof. Using $\xi_{k+1} = \infty$, from Theorem 5.5 we obtain the relation

$$A Q_{k+2} \underline{K}_{k+1} = Q_{k+2} \underline{H}_{k+1}, \tag{5.6}$$

where $\underline{K}_{k+1}$ and $\underline{H}_{k+1}$ are tridiagonal, $\boldsymbol{e}_{k+2}^T \underline{K}_{k+1} = \mathbf{0}^T$ and $\boldsymbol{e}_{k+2}^T \underline{H}_{k+1}$ is a multiple of $\boldsymbol{e}_{k+1}^T$. Hence, multiplying (5.6) on the left by Q_{k+2}^T and taking the first $k+1$ columns and rows, we have

$$J_{k+1} = H_{k+1} K_{k+1}^{-1}.$$

Let us define $\hat{D}_k = \operatorname{diag}(\gamma, \xi_1, \dots, \xi_k)$, for $\gamma \in \mathbb{R}$. Notice that, since the first column of Q_{k+1} is not an eigenvector of A, the entry in position (1,2) of J_{k+1} has to be different from 0. From this it can be noticed that $J_{k+1} - \tilde{D}_k$ is a symmetric generator representable semiseparable matrix if and only if $J_{k+1} - \hat{D}_k$ is so. Moreover, taking $\gamma \neq h_{1,1}$ and $\gamma \neq 0$ we have that $J_{k+1} - \hat{D}_k$ is invertible and its inverse can be computed as follows:

$$\begin{aligned}\left(H_{k+1} K_{k+1}^{-1} - \hat{D}_k\right)^{-1} &= \left(-\hat{D}_k (K_{k+1} - \hat{D}_k^{-1} H_{k+1}) K_{k+1}^{-1}\right)^{-1} = \\ &= -K_{k+1} (K_{k+1} - \hat{D}_k^{-1} H_{k+1})^{-1} \hat{D}_k^{-1}.\end{aligned}$$

Since $K_{k+1} = I_{k+1} + D_k H_{k+1}$ where $D_k = \operatorname{diag}(0, \frac{1}{\xi_1}, \dots, \frac{1}{\xi_k})$, we have

$$\begin{aligned}\left(H_{k+1} K_{k+1}^{-1} - \hat{D}_k\right)^{-1} &= -K_{k+1} (I_{k+1} + (D_k - \hat{D}_k^{-1}) H_{k+1})^{-1} \hat{D}_k^{-1} \\ &= -K_{k+1} (I_{k+1} - \frac{1}{\gamma} \boldsymbol{e}_1 \boldsymbol{e}_1^T H_{k+1})^{-1} \hat{D}_k^{-1} \\ &= -K_{k+1} (I_{k+1} + \frac{1}{\gamma - h_{1,1}} \boldsymbol{e}_1 \boldsymbol{e}_1^T H_{k+1}) \hat{D}_k^{-1} = \\ &= -(K_{k+1} + \frac{1}{\gamma - h_{1,1}} K_{k+1} (\boldsymbol{e}_1 \boldsymbol{e}_1^T) H_{k+1}) \hat{D}_k^{-1}.\end{aligned}$$

The third equality follows from the Sherman-Morrison formula, using the fact that $\gamma \neq h_{1,1}$. This also shows that the matrix $J_{k+1} - \hat{D}_k$ is indeed invertible.

The obtained matrix is an irreducible tridiagonal matrix since K_{k+1} and H_{k+1} have such structure and $\gamma \neq 0$. Hence, using Theorem 5.4, we have that $J_{k+1} - \hat{D}_k$ is a generator representable semiseparable matrix, and therefore also $J_{k+1} - \tilde{D}_k$ is so. □

The following corollary generalizes the statement of Theorem 5.6 to the case with poles at ∞. To simplify its proof we introduce a lemma.

Lemma 5.7. *Let $\{\xi_i^{(j)}\}_{j\in\mathbb{N}}$ be a sequence of real numbers outside of the convex hull of $\sigma(A)$ that tends to infinity. Assuming that $k+1$ is less than the invariance index of the Krylov subspace, let $Q_{k+1}^{(j)}$ be the orthonormal basis computed by the Arnoldi algorithm using poles $\{\xi_1, \ldots, \xi_{i-1}, \xi_i^{(j)}, \xi_{i+1}, \ldots, \xi_k\}$ and let $J_{k+1}^{(j)} = (Q_{k+1}^{(j)})^T A Q_{k+1}^{(j)}$ be the associated projected matrix. We have that*

$$\lim_{j\to\infty} Q_{k+1}^{(j)} = Q_{k+1} \qquad \textit{and} \qquad \lim_{j\to\infty} J_{k+1}^{(j)} = J_{k+1},$$

where Q_{k+1} is the orthonormal basis computed by the Arnoldi algorithm using poles $\{\xi_1, \ldots, \xi_{i-1}, \infty, \xi_{i+1}, \ldots, \xi_k\}$ and $J_{k+1} = Q_{k+1}^T A Q_{k+1}$.

Proof. Denote by $\boldsymbol{q}_\ell$ and $\boldsymbol{q}_\ell^{(j)}$, $1 \leq \ell \leq k+1$, the columns of Q_{k+1} and $Q_{k+1}^{(j)}$, respectively. Since the first i columns of the matrices $Q_{k+1}^{(j)}$ and Q_{k+1} are the same, we denote them by $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_i$. Let us prove that $\boldsymbol{q}_{i+1}^{(j)}$ converges to $\boldsymbol{q}_{i+1}$ as $j \to \infty$. To compute $\boldsymbol{q}_{i+1}^{(j)}$ by the rational Arnoldi algorithm, we first compute $\boldsymbol{w}_i^{(j)} = (I - A/\xi_i^{(j)})^{-1} A\boldsymbol{q}_i$. Note that this is a continuous operation when $\xi_i^{(j)}$ tends to infinity, indeed

$$\lim_{j\to\infty} (I - A/\xi_i^{(j)})^{-1} A\boldsymbol{q}_i = A\boldsymbol{q}_i.$$

Since the orthonormalization of $\boldsymbol{q}_{i+1}^{(j)}$ against $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_i$ is a continuous operation, we have that

$$\lim_{j\to\infty} \boldsymbol{q}_{i+1}^{(j)} = \boldsymbol{q}_{i+1}.$$

We can now prove by induction that $\lim_{j\to\infty} \boldsymbol{q}_{\ell+1}^{(j)} = \boldsymbol{q}_{\ell+1}$ for all $i+1 \leq \ell \leq k$. By writing the orthogonalization step explicitly, we have

$$\boldsymbol{q}_{\ell+1}^{(j)} = (I - Q_\ell^{(j)}(Q_\ell^{(j)})^T)\boldsymbol{w}_\ell^{(j)}, \qquad \text{where} \qquad \boldsymbol{w}_\ell^{(j)} = (I - A/\xi_\ell)^{-1} A\boldsymbol{q}_\ell^{(j)},$$

and likewise

$$\boldsymbol{q}_{\ell+1} = (I - Q_\ell Q_\ell^T)\boldsymbol{w}_\ell, \qquad \text{where} \qquad \boldsymbol{w}_\ell = (I - A/\xi_\ell)^{-1} A\boldsymbol{q}_\ell.$$

By the inductive hypothesis $\lim_{j\to\infty} Q_\ell^{(j)} = Q_\ell$, so we also have $\lim_{j\to\infty} \boldsymbol{w}_\ell^{(j)} = \boldsymbol{w}_\ell$, and therefore

$$\lim_{j\to\infty} \boldsymbol{q}_{\ell+1}^{(j)} = \boldsymbol{q}_{\ell+1}.$$

The statement for J_{k+1} follows immediately from the definitions of $J_{k+1}^{(j)}$ and J_{k+1}. □

Corollary 5.8. *Under the same assumptions of Theorem 5.6, but allowing some poles to be equal to infinity, the projected matrix $J_{k+1} = Q_{k+1}^T A Q_{k+1}$ is still a quasiseparable matrix.*

Proof. Let us prove the statement by induction on the number of poles at infinity. If there are no poles equal to infinity the thesis follows from Theorem 5.6.

Assume now that the statement holds in the case of s poles equal to infinity and assume that we are using $s+1$ infinite poles. Let $\xi_i = \infty$ be the last infinite pole and let $\{\xi_i^{(j)}\}_{j\in\mathbb{N}}$ be a sequence of real numbers outside of the convex hull of $\sigma(A)$ that converges to ξ_i, that is

$$\lim_{j\to\infty} \xi_i^{(j)} = \infty.$$

Let $J_{k+1}^{(j)}$ be the projected matrix obtained by using poles equal to

$$\{\xi_1, \dots, \xi_{i-1}, \xi_i^{(j)}, \xi_{i+1}, \dots, \xi_k\}.$$

By Lemma 5.7, we have that $J_{k+1}^{(j)}$ converges to the projected matrix J_{k+1} obtained with the poles $\{\xi_1, \dots, \xi_i, \dots, \xi_k\}$, that is

$$\lim_{j\to\infty} J_{k+1}^{(j)} = J_{k+1}.$$

From the inductive hypothesis we have that the matrices $J_{k+1}^{(j)}$ are quasiseparable for all j. Since the quasiseparable matrices are a closed set [38, Section 1.4.1], we have the thesis. □

Notice that, if the matrix A is symmetric positive semidefinite and $k+1$ is less than the invariance index, the projected matrix $J_k = Q_k^T A Q_k$ has to be positive definite. Indeed, if there exists a vector $\boldsymbol{x} \neq \boldsymbol{0}$ such that $J_k \boldsymbol{x} = \boldsymbol{0}$, we have that $AQ_k\boldsymbol{x} = \boldsymbol{0}$. In particular, since $Q_k\boldsymbol{x} \in q_{k-1}(A)^{-1}\mathcal{P}_k(A, \boldsymbol{b})$, where q_{k-1} is as defined in Section 4, there exist $\alpha_0, \dots, \alpha_j$, $j \leq k-1$, with $\alpha_j \neq 0$ such that

$$Q_k\boldsymbol{x} = q_{k-1}(A)^{-1} \sum_{i=0}^{j} \alpha_i A^i \boldsymbol{b},$$

and so

$$\boldsymbol{0} = AQ_k\boldsymbol{x} = q_{k-1}(A)^{-1} \sum_{i=0}^{j} \alpha_i A^{i+1} \boldsymbol{b}.$$

This implies that $A^{j+1}\boldsymbol{b} \in \mathcal{P}_j(A, \boldsymbol{b})$, but this is impossible because $k+1$ is less than the invariance index. In particular, this implies the existence of the Cholesky factorization of the matrix J_k if the matrix A is symmetric positive semidefinite.

The matrix $B_k = P_k^T A Q_k$ we are interested in when using rational Krylov methods for GMFs is exactly the transpose of the Cholesky factor of the matrix $J_k = Q_k^T(A^T A)Q_k$. Indeed B_k is upper triangular and

$$J_k = (AQ_k)^T AQ_k = (P_k B_k)^T (P_k B_k) = B_k^T B_k.$$

The following proposition gives us that the matrix B_k is the upper triangular part of a rank one plus diagonal matrix, hence B_k is quasiseparable. We could not find a reference for this fact, so we provide a brief proof sketch.

Proposition 5.9. *Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite and let $L \in \mathbb{R}^{n\times n}$ be its Cholesky factor (i.e., L is lower triangular and $A = LL^T$). If there exist $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ such that* $\operatorname{tril}(A, -1) = \operatorname{tril}(\boldsymbol{v}\boldsymbol{u}^T, -1)$*, then* $\operatorname{tril}(L, -1) = \operatorname{tril}(\boldsymbol{v}\boldsymbol{x}^T, -1)$ *for $\boldsymbol{x} \in \mathbb{R}^n$.*

Proof. It can be easily proved that the last row of $\operatorname{tril}(L, -1)$ is equal to

$$v_n \cdot \begin{bmatrix} u_1 & \dots & u_{n-1} & 0 \end{bmatrix} \begin{bmatrix} L_{n-1}^{-T} & \\ & 1 \end{bmatrix},$$

where L_{n-1} is the leading principal submatrix of L of size $n-1$. Using this fact the thesis can be proved recursively. □

Remark 5.10. Notice that if $\operatorname{triu}(J_k, 1) = \operatorname{triu}(\boldsymbol{u}\boldsymbol{v}^T, 1)$, the vector $\boldsymbol{v}$ cannot have zero entries: indeed if there exists $s \le k$ such that $v_s = 0$, then, as a consequence of the proof of Theorem 5.6, the matrix $J_s - \operatorname{diag}(\gamma, \xi_1, \dots, \xi_{s-1})$ has the last column equal to zero for all $\gamma \in \mathbb{R}$, but this is impossible since for an appropriate choice of γ this matrix has to be invertible.

Exploiting the quasiseparable structure of the matrix B_k, we can compute the matrices B_k and P_k by only performing a few scalar products. Indeed, if we let

$$B_k = \begin{bmatrix} d_1 & \beta_1 & \gamma_1 & & & \\ & d_2 & \beta_2 & \gamma_2 & * & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & d_{k-2} & \beta_{k-2} & \gamma_{k-2} \\ & & & & d_{k-1} & \beta_{k-1} \\ & & & & & d_k \end{bmatrix},$$

and we define $\boldsymbol{x}_k = [P_{k-1}\, \boldsymbol{0}]B_k \boldsymbol{e}_k$, we have that

$$A\boldsymbol{q}_k = AQ_k\boldsymbol{e}_k = P_k B_k \boldsymbol{e}_k = d_k \boldsymbol{p}_k + \boldsymbol{x}_k.$$

Using the fact that the submatrix of B_k that involves the last two columns and all except for the last two rows has rank at most 1, we can compute $\boldsymbol{x}_k$ with the recursive relation

$$\boldsymbol{x}_k = \frac{\gamma_{k-2}}{\beta_{k-2}}\boldsymbol{x}_{k-1} + \beta_{k-1}\boldsymbol{p}_{k-1}.$$

Algorithm 2: k-th step of rational Golub-Kahan algorithm

```
Input: $A, \boldsymbol{q}_k, \boldsymbol{p}_{k-1}, \boldsymbol{p}_{k-2}, \boldsymbol{x}_{k-1}$
Output: $\boldsymbol{p}_k, d_k, \beta_{k-1}, \gamma_{k-2}, \boldsymbol{x}_k$
$\boldsymbol{w} = A\boldsymbol{q}_k$
$\beta_{k-1} = \boldsymbol{w}^T \boldsymbol{p}_{k-1}$
$\gamma_{k-2} = \boldsymbol{w}^T \boldsymbol{p}_{k-2}$
$\boldsymbol{x}_k = \frac{\gamma_{k-2}}{\beta_{k-2}} \boldsymbol{x}_{k-1} + \beta_{k-1} \boldsymbol{p}_{k-1}$
$\boldsymbol{w} = \boldsymbol{w} - \boldsymbol{x}_k$
$d_k = \|\boldsymbol{w}\|_2$
$\boldsymbol{p}_k = \boldsymbol{w}/d_k$
```

This allows us to compute $d_k, \beta_{k-1}, \gamma_{k-2}$ and $\boldsymbol{p}_k$ with only two scalar products. The k-th step of the procedure is summarized in Algorithm 2.

Notice that during the k-th step of the procedure, we do not require the first $k-1$ columns of Q_k. Moreover, for the computation of the projected solution $\bar{\boldsymbol{y}}_k$ defined in (5.4), we do not need the matrix Q_k. For this reason, the computation of the matrix Q_k can be performed by using a short recurrence rational Lanczos algorithm, as the one presented in [17, Section 5.2], and we can keep in memory only the last two columns of Q_k. After the k-th step of the algorithm, the k-th column of the matrix B_k can be computed using the newly computed quantities and the previous column, by exploiting the rank structure of the matrix B_k.

Note that the approximation of $f^\diamond(A)\boldsymbol{b}$ with (5.4) still requires storing the whole basis P_k in order to perform the product $P_k f^\diamond(B_k)\boldsymbol{e}_1$. Nevertheless, the low memory requirements of the short recurrence can be exploited in full when the goal is the computation of the bilinear form $\boldsymbol{w}^T f^\diamond(A)\boldsymbol{b}$ for some vector $\boldsymbol{w}$: in this setting, the approximation $\boldsymbol{w}^T P_k f^\diamond(B_k)\boldsymbol{e}_1$ can be computed by updating the vector $\boldsymbol{w}^T P_k$ while the Krylov basis is being constructed, bypassing the need to store the whole matrix P_k.

The whole procedure for the computation of $f^\diamond(A)\boldsymbol{b}$ is summarized in Algorithm 3. The algorithm reduces to the standard Golub-Kahan bidiagonalization if all the poles are chosen equal to infinity.

Remark 5.11. In the algorithm it is implicitly assumed that $\beta_i \neq 0$ for each i. In practice this hypothesis is always satisfied, however, as observed in Remark 5.10, if k is less than the invariance index there is at least one nonzero off-diagonal entry in the k-th column of B_k. Hence we could modify the algorithm to avoid the issue of $\beta_i = 0$.

Remark 5.12. The algorithm presented in this section also works if some of the poles are equal to zero. However, the proof of this fact requires slightly different tools, and hence we omitted it for brevity.

Algorithm 3: Short recurrence rational Krylov approximation of $f^\diamond(A)\boldsymbol{b}$

```
**Input:** $A \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^n, f, \{\xi_1, \dots, \xi_{k-1}\}$
**Output:** $\bar{\boldsymbol{y}}_k \in \mathcal{Q}_k(AA^T, A\boldsymbol{b})$ s.t. $\bar{\boldsymbol{y}}_k \approx f^\diamond(A)\boldsymbol{b}$
$\boldsymbol{q}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2$
$\boldsymbol{w}_1 = (I - A^TA/\xi_1)^{-1}A^TA\boldsymbol{q}_1$          // other choices can be used
Compute $\boldsymbol{q}_2$ by orthogonalizing $\boldsymbol{w}_1$ against $\boldsymbol{q}_1$
Compute the QR decomposition $[\boldsymbol{p}_1, \boldsymbol{p}_2] \begin{bmatrix} d_1 & \beta_1 \\ 0 & d_2 \end{bmatrix} = [\boldsymbol{q}_1, \boldsymbol{q}_2]$
Define $B_2 = \begin{bmatrix} d_1 & \beta_1 \\ 0 & d_2 \end{bmatrix}$ and $\boldsymbol{x}_2 = \beta_1\boldsymbol{p}_1$
**for** $j = 2, \dots, k-1$ **do**
    $\boldsymbol{w}_j = (I - A^TA/\xi_j)^{-1}A^TA\boldsymbol{q}_j$      // other choices can be used
    Compute $\boldsymbol{q}_{j+1}$ by orthogonalizing $\boldsymbol{w}_j$ against $[\boldsymbol{q}_1, \dots, \boldsymbol{q}_j]$
    Compute $\boldsymbol{p}_{j+1}, d_{j+1}, \beta_j, \gamma_{j-1}, \boldsymbol{x}_{j+1}$ by Algorithm 2 with $k = j+1$
    $B_{j+1} = \begin{bmatrix} B_j & \boldsymbol{s}_{j+1} \\ 0 & d_{j+1} \end{bmatrix}$, where $s_{j+1} = \begin{bmatrix} \frac{\gamma_{j-1}}{\beta_{j-1}}(B_j)_{1:j-1,j} \\ \beta_j \end{bmatrix}$
$P_k = [\boldsymbol{p}_1, \dots, \boldsymbol{p}_k]$
Compute $f^\diamond(B_k)$, e.g. via an SVD of $B_k$
$\bar{\boldsymbol{y}}_k = P_k f^\diamond(B_k)\boldsymbol{e}_1$
```

6 Error bounds

In this section we prove some error bounds for the approximation of $f^\diamond(A)\boldsymbol{b}$ using the polynomial and rational Krylov methods described above. These bounds link the approximation error with the error of polynomial and rational approximation of f on an interval containing the singular values of A. Our results are the analogue of the ones that hold for standard matrix functions, and they can be proved in a similar way.

We first find an upper and lower bound for the singular values of B_k. For convenience, given a matrix $A \in \mathbb{R}^{m\times n}$, throughout this section we are going to use an extended notation for singular values, defining $\sigma_j := 0$ for all j such that $\min\{m, n\} < j \le \max\{m, n\}$.

Lemma 6.1. *Let σ_1 and σ_n be the first and n-th singular value of $A \in \mathbb{R}^{m\times n}$, respectively. Then the singular values of B_k belong to the interval $[\sigma_n, \sigma_1]$.*

Proof. We have

$$B_k^TB_k = Q_k^TA^TP_kP_k^TAQ_k = Q_k^T(A^TA)Q_k, \tag{6.1}$$

where we used the fact that $P_kP_k^TAQ_k = AQ_k$, since the columns of AQ_k span the Krylov subspace $\mathcal{Q}_k(AA^T, A\boldsymbol{b})$.

Therefore, by the Cauchy Interlacing Theorem the eigenvalues of $B_k^TB_k$ are contained in the interval $[\sigma_n^2, \sigma_1^2]$, which containes the eigenvalues of $A^TA \in$

$\mathbb{R}^{n\times n}$, and thus the singular values of B_k are contained in the interval $[\sigma_n, \sigma_1]$. □

Observe that if $A \in \mathbb{R}^{m\times n}$ is rectangular with $n > m$, we always have $\sigma_n = 0$, and hence B_k may have singular values arbitrarily close to 0 even if $\sigma_{\min\{m,n\}}(A) > 0$. This fact is going to affect the error bounds in Theorem 6.3 and Theorem 6.7.

As an example, consider the 1×2 matrix $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and the vector $\boldsymbol{b} = \begin{bmatrix} \epsilon \\ 1 \end{bmatrix}$, for small $\epsilon > 0$. For $k = 1$, we have $Q_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2 = \frac{1}{\sqrt{1+\epsilon^2}}\boldsymbol{b}$, and $P_1 = A\boldsymbol{b}/\|A\boldsymbol{b}\|_2 = 1$. So we have $B_1 = P_1^T A Q_1 = \frac{\epsilon}{\sqrt{1+\epsilon^2}} \in \mathbb{R}^{1\times 1}$, and hence B_1 can have an arbitrarily small singular value even if $\sigma_1(A) = 1$.

6.1 Polynomial error bounds

We first prove the error bounds in the polynomial case. Recall that a polynomial Krylov method computes an approximation to $\boldsymbol{y} = f^{\diamond}(A)\boldsymbol{b}$ from the subspace $\mathcal{P}_k(AA^T, A\boldsymbol{b})$ as

$$\bar{\boldsymbol{y}}_k = P_k f^{\diamond}(B_k) Q_k^T \boldsymbol{b} = \|\boldsymbol{b}\|_2 P_k f^{\diamond}(B_k)\boldsymbol{e}_1, \tag{6.2}$$

where $B_k = P_k^T A Q_k$, and P_k and Q_k are the matrices computed in the Golub-Kahan bidiagonalization of A, satisfying $\operatorname{span}(P_k) = \mathcal{P}_k(AA^T, A\boldsymbol{b})$ and $\operatorname{span}(Q_k) = \mathcal{P}_k(A^T A, \boldsymbol{b})$.

A key observation for proving the bounds is the exactness of the approximation (6.2) when f is an odd polynomial, stated in the following lemma.

Lemma 6.2. *Assume that $f = p_{2k-1}$ is an odd polynomial of degree $\leq 2k-1$. Then the approximation $\bar{\boldsymbol{y}}_k$ given by (6.2) is exact, i.e., we have $\boldsymbol{y} = \bar{\boldsymbol{y}}_k$.*

Proof. Let $p_{2k-1}(z) = zq(z^2)$, where $\deg q \leq k-1$. Using Proposition 3.3 and recalling (6.1), we have

$$\begin{aligned} \bar{\boldsymbol{y}}_k &= P_k B_k q(B^T B_k) Q_k^T \boldsymbol{b} \\ &= P_k P_k^T A Q_k q(Q_k^T A^T A Q_k) Q_k^T \boldsymbol{b}. \end{aligned}$$

Due to the exactness property of the polynomial Krylov approximation for standard matrix functions [17, Lemma 3.9], we have

$$Q_k q(Q_k^T A^T A Q_k) Q_k^T \boldsymbol{b} = q(A^T A)\boldsymbol{b},$$

and hence we get

$$\begin{aligned} \bar{\boldsymbol{y}}_k &= P_k P_k^T A q(A^T A)\boldsymbol{b} \\ &= A q(A^T A)\boldsymbol{b} = p_{2k-1}^{\diamond}(A)\boldsymbol{b} = \boldsymbol{y}, \end{aligned}$$

where we used the fact that $\boldsymbol{w} = Aq(A^T A)\boldsymbol{b} \in \mathcal{P}_k(AA^T, A\boldsymbol{b})$, and therefore we have $P_k P_k^T \boldsymbol{w} = \boldsymbol{w}$. □

Using Lemma 6.2, we can prove the following theorem.

Theorem 6.3. *Let $A \in \mathbb{R}^{m\times n}$, and let σ_1, σ_n and σ_m be the first, n-th and m-th singular value of A, respectively. Let $\bar{\boldsymbol{y}}_k$ be the approximation to $\boldsymbol{y} = f^\diamond(A)\boldsymbol{b}$ given by* (6.2)*. Then the following inequality holds:*

$$\|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 \leq 2\|\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|f(z) - p(z^2)z\|_{\infty,[\sigma_n,\sigma_1]}. \tag{6.3}$$

Moreover, if A is square with $\sigma_m = \sigma_n > 0$, or if $\lim\limits_{z\to 0} \frac{f(z)}{z} = 0$, we also have

$$\|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 \leq 2\|A\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|f(z)/z - p(z^2)\|_{\infty,[\sigma_{\max\{m,n\}},\sigma_1]}. \tag{6.4}$$

Proof. Let p be a polynomial of degree $\leq k-1$. Then $p_{2k-1}(z) = p(z^2)z$ is an odd polynomial of degree $\leq 2k-1$, and by Lemma 6.2 we have

$$p^\diamond_{2k-1}(A)\boldsymbol{b} = P_k p^\diamond_{2k-1}(B_k)Q_k^T\boldsymbol{b}. \tag{6.5}$$

By adding and subtracting the quantity in (6.5) to $f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k$, we get

$$f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k = [f^\diamond(A) - p^\diamond_{2k-1}(A)]\boldsymbol{b} - P_k[f^\diamond(B_k) - p^\diamond_{2k-1}(B_k)]Q_k^T\boldsymbol{b}. \tag{6.6}$$

By invariance of the 2-norm under unitary transformations, we have

$$\|f^\diamond(A) - p^\diamond_{2k-1}(A)\|_2 = \|f - p_{2k-1}\|_{\infty,\sigma_{\text{sing}}(A)} \leq \|f - p_{2k-1}\|_{\infty,[\sigma_{\min\{m,n\}},\sigma_1]},$$

and similarly, by Lemma 6.1,

$$\|f^\diamond(B_k) - p^\diamond_{2k-1}(B_k)\|_2 \leq \|f - p_{2k-1}\|_{\infty,[\sigma_n,\sigma_1]}.$$

Combining the above inequalities with (6.6), we get

$$\|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 \leq 2\|\boldsymbol{b}\|_2\|f - p_{2k-1}\|_{\infty,[\sigma_n,\sigma_1]},$$

and by taking the minimum over $p \in \Pi_{k-1}$ we obtain (6.3).

To prove (6.4), recall that if $\sigma_m > 0$ or $\lim\limits_{z\to 0} \frac{f(z)}{z} = 0$, by Proposition 3.5 we have $f^\diamond(A) = g(AA^T)A$, where $g(z) = f(\sqrt{z})/\sqrt{z}$, and similarly, if $\sigma_n > 0$ or $\lim\limits_{z\to 0} \frac{f(z)}{z} = 0$, then $f^\diamond(B_k) = g(B_kB_k^T)B_k$. Therefore, by also using Proposition 3.3, we can rewrite (6.6) in the form

$$f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k = [g(AA^T) - p(AA^T)]A\boldsymbol{b} - P_k[g(B_kB_k^T) - p(B_kB_k^T)]B_kQ_k^T\boldsymbol{b}.$$

Given that the eigenvalues of $B_kB_k^T$ are the squares of the singular values of B_k, with a similar argument as before we obtain

$$\begin{aligned}\|f^\diamond(A) - \bar{\boldsymbol{y}}_k\|_2 &\leq \|A\boldsymbol{b}\|_2 \left(\|g-p\|_{\infty,[\sigma_m,\sigma_1]} + \|g-p\|_{\infty,[\sigma_n^2,\sigma_1^2]}\right)\\ &\leq 2\|A\boldsymbol{b}\|_2\left\|f(z)/z - p(z^2)\right\|_{\infty,[\sigma_{\max\{m,n\}},\sigma_1]}.\end{aligned}$$

As before, (6.4) follows by taking the minimum over $p \in \Pi_{k-1}$. □

Remark 6.4. Observe that if the matrix A is not square, we have $\sigma_{\max\{m,n\}} = 0$, and hence the bound (6.4) always involves a polynomial approximation over the whole interval $[0, \sigma_1]$, even when $\sigma_{\min\{m,n\}} > 0$. If $A \in \mathbb{R}^{m \times n}$ is rectangular with $m < n$, then the bound (6.3) also involves the whole interval $[0, \sigma_1]$.

A possible strategy to overcome this issue is to use Proposition 3.7 and write

$$\boldsymbol{y} = f^{\diamond}(A)\boldsymbol{b} = (A^{+})^T f^{\diamond}(A^T)A\boldsymbol{b}.$$

The vector $\boldsymbol{w} = f^{\diamond}(A^T)A\boldsymbol{b}$ can be approximated using a Krylov method on A^T, and then $\boldsymbol{y}$ can be recovered by solving the least squares problem

$$\boldsymbol{y} = (A^{+})^T \boldsymbol{w} = \arg\min_{\boldsymbol{y}} \|A^T \boldsymbol{y} - \boldsymbol{w}\|_2. \tag{6.7}$$

By rewriting the problem in this form, if $m < n$ and $\sigma_m > 0$ we get a bound involving approximation on the smaller interval $[\sigma_m, \sigma_1]$ for the approximation of $\boldsymbol{w}$, which translates to a bound for the approximation of $\boldsymbol{y}$.

The bound (6.3) can be manipulated to obtain a more explicit bound. Assume that $\sigma_n > 0$, and let $I = [\sigma_n, \sigma_1]$. The polynomial $p(z^2)z$ is odd, and we can assume that f is also odd, so we have

$$\begin{aligned} \min_{p \in \Pi_{k-1}} \|f(z) - p(z^2)z\|_{\infty, I} &= \min_{p \in \Pi_{k-1}} \|f(z) - p(z^2)z\|_{\infty, (-I) \cup I} \\ &= \min_{q \in \Pi_{2k-1}} \|f(z) - q(z)\|_{\infty, (-I) \cup I}, \end{aligned} \tag{6.8}$$

where we used the fact that the polynomial of best approximation on $(-I) \cup I$ for an odd function is itself odd. Bounds on the asymptotic rate of convergence for the polynomial approximation of a function on the union of disjoint intervals $(-I) \cup I$ have been developed in [11]. We remark that using [11, Theorem 1] to bound (6.8) would lead to the same asymptotic rate of convergence as (6.9), with a slightly larger constant.

The bound (6.3) can also be related to a polynomial approximation problem on the interval $[\sigma_n^2, \sigma_1^2]$. Indeed, we have

$$\begin{aligned} \min_{p \in \Pi_{k-1}} \|f(z) - p(z^2)z\|_{\infty, [\sigma_n, \sigma_1]} &= \sigma_1 \min_{p \in \Pi_{k-1}} \|f(z)/z - p(z^2)\|_{\infty, [\sigma_n, \sigma_1]} \\ &\le \sigma_1 \min_{p \in \Pi_{k-1}} \|f(\sqrt{z})/\sqrt{z} - p(z)\|_{\infty, [\sigma_n^2, \sigma_1^2]}. \end{aligned}$$

Let $1 < \rho \le \dfrac{\sigma_1 + \sigma_n}{\sigma_1 - \sigma_n}$, and denote by E_ρ the ellipse with vertices at $\pm\frac{1}{2}(\rho + \frac{1}{\rho})$ and foci at ± 1, and by $\tilde{E}_\rho$ its image under the linear function that maps $[-1, 1]$ to the interval $[\sigma_n^2, \sigma_1^2]$. The ellipse $\tilde{E}_\rho$ has vertices at $\frac{1}{2}(\sigma_n^2 + \sigma_1^2) \pm \frac{1}{4}(\rho + \frac{1}{\rho})(\sigma_n^2 - \sigma_1^2)$ and foci at σ_n^2 and σ_1^2. Note that for $\rho = \dfrac{\sigma_1 + \sigma_n}{\sigma_1 - \sigma_n}$, the ellipse $\tilde{E}_\rho$ has a vertex at 0. We recall the following classical result from approximation theory.

Theorem 6.5. *[34, Theorem 8.2] Let the function g be analytic in the interior of the ellipse $\tilde{E}_\rho$, and assume that $\max_{z \in \tilde{E}_\rho} |g(z)| \le M$. Then*

$$\min_{p \in \Pi_k} \|g(z) - p(z)\|_{\infty, [\sigma_n^2, \sigma_1^2]} \le \frac{2M}{\rho - 1} \rho^{-k}.$$

If the function $f(\sqrt{z}/\sqrt{z})$ is analytic in the interior of $\tilde{E}_\rho$, applying Theorem 6.5 to the bound (6.3) gives us

$$\begin{aligned} \|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 &\le 2\sigma_1\|\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|f(\sqrt{z})/\sqrt{z} - p(z)\|_{\infty,[\sigma_n^2,\sigma_1^2]} \\ &\le 4M\sigma_1\|\boldsymbol{b}\|_2 \frac{\rho}{\rho-1}\rho^{-k}, \end{aligned} \tag{6.9}$$

where $M = \max\limits_{z\in\tilde{E}_\rho}|f(\sqrt{z}/\sqrt{z})|$ and $1 < \rho \le \dfrac{\sigma_1+\sigma_n}{\sigma_1-\sigma_n}$.

Remark 6.6. Note that if the function $f(\sqrt{z})/\sqrt{z}$ is unbounded on $\tilde{E}_{\bar{\rho}}$ for a certain $\bar{\rho}$, Theorem 6.5 can only be used for $\rho < \bar{\rho}$. In such a situation, the bound (6.9) only makes sense for $\rho < \bar{\rho}$.

6.2 Rational error bounds

Next, we prove similar error bounds for the rational approximation (5.4). We start by stating the result analogous to Theorem 6.3.

Recall that the denominator in the rational Krylov space $\mathcal{Q}_k(AA^T, A\boldsymbol{b})$ is given by the polynomial $q_{k-1}(z) = \prod_{j=1}^{k-1}(1 - z/\xi_j)$, where $\{\xi_j\}_{j\ge1}$ is a sequence of poles in $(\mathbb{C}\cup\{\infty\})\setminus\sigma(AA^T)$.

Theorem 6.7. *Let $A \in \mathbb{R}^{m\times n}$, and let σ_1, σ_n and σ_m be the first, n-th and m-th singular value of A, respectively. Let $\bar{\boldsymbol{y}}_k$ be the approximation to $\boldsymbol{y} = f^\diamond(A)\boldsymbol{b}$ from $\mathcal{Q}_k(AA^T, A\boldsymbol{b})$ given by* (5.4). *Then the following inequality holds:*

$$\|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 \le 2\|\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|f(z) - q_{k-1}(z^2)^{-1}p(z^2)z\|_{\infty,[\sigma_n,\sigma_1]}. \tag{6.10}$$

Moreover, if A is square with $\sigma_m = \sigma_n > 0$, or if $\lim_{z\to0}\frac{f(z)}{z} = 0$, we also have

$$\|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 \le 2\|A\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|f(z)/z - q_{k-1}(z^2)^{-1}p(z^2)\|_{\infty,[\sigma_{\max\{m,n\}},\sigma_1]}. \tag{6.11}$$

Note that the same issues discussed after Theorem 6.3 in the case of rectangular matrices also arise in the rational case, and the same approach proposed in Remark 6.4 can be used to address them.

Similarly to the polynomial case, the bound (6.10) can be rewritten by exploiting the fact that f can be assumed to be odd and hence that the best approximant on a symmetric interval is odd, yielding

$$\|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 \le 2\|\boldsymbol{b}\|_2 \min_{p\in\Pi_{2k-1}} \|f(z) - q_{k-1}(z^2)^{-1}p(z)\|_{\infty,(-I)\cup I}, \tag{6.12}$$

where again $I = [\sigma_n, \sigma_1]$. However, rational approximation on disjoint intervals is a complicated problem, and hence this formulation might be less useful in practice.

A more practical way to rewrite the bound (6.10) is the following:

$$\begin{aligned}\|f^{\diamond}(A)\boldsymbol{b}-\bar{\boldsymbol{y}}_k\|_2 &\le 2\|\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|f(z)-q_{k-1}(z^2)^{-1}p(z^2)z\|_{\infty,[\sigma_n,\sigma_1]}\\ &= 2\|\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|\sqrt{z}\big(f(\sqrt{z})/\sqrt{z}-q_{k-1}(z)^{-1}p(z)\big)\|_{\infty,[\sigma_n^2,\sigma_1^2]}\\ &\le 2\sigma_1\|\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|\big(f(\sqrt{z})/\sqrt{z}-q_{k-1}(z)^{-1}p(z)\big)\|_{\infty,[\sigma_n^2,\sigma_1^2]}.\end{aligned} \tag{6.13}$$

Although we get an additional factor σ_1, this bound relates the error with a uniform rational approximation problem on a real interval. This approximation problem is well-studied in the literature, and it is the same that appears when computing standard matrix functions with rational Krylov methods, so it can be a viable tool for selecting good poles.

Similarly to the polynomial case, to prove Theorem 6.7 we require the following auxiliary lemma, which shows the exactness of the rational approximation on rational functions of the form $f(z) = q_{k-1}(z^2)^{-1}p(z^2)z$, where p is any polynomial in Π_{k-1}.

Lemma 6.8. *Assume that f is a function of the form $f(z) = q_{k-1}(z^2)^{-1}p(z^2)z$, where $p \in \Pi_{k-1}$. Then the approximation $\bar{\boldsymbol{y}}_k$ from the rational Krylov subspace $\mathcal{Q}_k(AA^T, A\boldsymbol{b})$ given by (5.4) is exact, i.e., we have $\boldsymbol{y} = \bar{\boldsymbol{y}}_k$.*

Proof. Using Corollary 3.4 and (6.1), we have

$$\begin{aligned}\bar{\boldsymbol{y}}_k &= P_k B_k q_{k-1}(B^T B_k)^{-1} p(B_k^T B_k) Q_k^T \boldsymbol{b}\\ &= P_k P_k^T A Q_k q_{k-1}(Q_k^T A^T A Q_k)^{-1} p(Q_k^T A^T A Q_k) Q_k^T \boldsymbol{b}.\end{aligned}$$

Due to the exactness property of the rational Krylov approximation for standard matrix functions [18, Lemma 3.1], we have

$$Q_k q_{k-1}(Q_k^T A^T A Q_k) p(Q_k^T A^T A Q_k) Q_k^T \boldsymbol{b} = q_{k-1}(A^T A)^{-1} p(A^T A)\boldsymbol{b},$$

and hence we get

$$\begin{aligned}\bar{\boldsymbol{y}}_k &= P_k P_k^T A q_{k-1}(A^T A)^{-1} p(A^T A)\boldsymbol{b}\\ &= A q_{k-1}(A^T A)^{-1} p(A^T A)\boldsymbol{b} = f^{\diamond}(A)\boldsymbol{b} = \boldsymbol{y},\end{aligned}$$

where we used the fact that $\boldsymbol{w} = A q_{k-1}(A^T A)^{-1} p(A^T A)\boldsymbol{b} \in \mathcal{Q}(AA^T, A\boldsymbol{b})$, and therefore $P_k P_k^T \boldsymbol{w} = \boldsymbol{w}$. □

We are now ready to prove Theorem 6.7. The proof follows the same strategy as the proof of Theorem 6.3.

Proof of Theorem 6.7. Let p be a polynomial of degree $\le k-1$. Then the rational function $r(z) = q_{k-1}(z^2)^{-1}p(z^2)z$ satisfies the assumptions of Lemma 6.8, and hence we have

$$r^{\diamond}(A)\boldsymbol{b} = P_k r^{\diamond}(B_k) Q_k^T \boldsymbol{b}. \tag{6.14}$$

By adding and subtracting the quantity in (6.14) to $f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k$, we get

$$f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k = [f^\diamond(A) - r^\diamond(A)]\boldsymbol{b} - P_k[f^\diamond(B_k) - r^\diamond(B_k)]Q_k^T\boldsymbol{b}. \tag{6.15}$$

Since by Lemma 6.1 the nonzero singular values of B_k are contained in the interval $[\sigma_n, \sigma_1]$, by invariance of the 2-norm under unitary transformations, we have

$$\|f^\diamond(A) - r^\diamond(A)\|_2 \leq \|f - r\|_{\infty,[\sigma_{\min\{m,n\}},\sigma_1]},$$
$$\|f^\diamond(B_k) - r^\diamond(B_k)\|_2 \leq \|f - r\|_{\infty,[\sigma_n,\sigma_1]}.$$

Combining the above inequalities with (6.15), we get

$$\|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 \leq 2\|\boldsymbol{b}\|_2\|f - r\|_{\infty,[\sigma_n,\sigma_1]},$$

and by taking the minimum over $p \in \Pi_{k-1}$ we obtain (6.10).

To prove (6.11), if $\sigma_n = \sigma_m > 0$ or $\lim\limits_{z\to 0} \frac{f(z)}{z} = 0$, by Proposition 3.5 we can write $f^\diamond(A) = g(AA^T)A$ and $f^\diamond(B_k) = g(B_kB_k^T)B_k$, where $g(z) = f(\sqrt{z})/\sqrt{z}$. Thus, using also Corollary 3.4, we can rewrite (6.15) in the form

$$f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k = h(AA^T)A\boldsymbol{b} - P_kh(B_kB_k^T)B_kQ_k^T\boldsymbol{b},$$

where $h(z) = g(z) - q_{k-1}(z)^{-1}p(z)$.

Given that the eigenvalues of $B_kB_k^T$ are the squares of the singular values of B_k, using Lemma 6.1 and proceeding as above we obtain

$$\begin{aligned}\|f^\diamond(A) - \bar{\boldsymbol{y}}_k\|_2 &\leq \|A\boldsymbol{b}\|_2\left(\|h\|_{\infty,[\sigma_m^2,\sigma_1^2]} + \|h\|_{\infty,[\sigma_n^2,\sigma_1^2]}\right)\\ &= 2\|A\boldsymbol{b}\|_2\big\|f(z)/z - q_{k-1}(z^2)^{-1}p(z^2)\big\|_{\infty,[\sigma_{\max\{m,n\}},\sigma_1]}.\end{aligned}$$

As before, (6.11) follows by taking the minimum over $p \in \Pi_{k-1}$. □

We mention that the results of this section can also be obtained by exploiting the link betweeen GMFs of A and standard functions of the matrix $\mathcal{A} = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$. Indeed, it was observed in [3] that for an odd function f we have

$$f(\mathcal{A}) = \begin{bmatrix} 0 & f^\diamond(A) \\ f^\diamond(A^T) & 0 \end{bmatrix}.$$

If we define the orthogonal matrix $\mathcal{U}_{2k} = \begin{bmatrix} P_k & 0 \\ 0 & Q_k \end{bmatrix}$ and the vector $\boldsymbol{c} = \begin{bmatrix} 0 \\ \boldsymbol{b} \end{bmatrix} \in \mathbb{R}^{m+n}$, we have that $\mathcal{U}_{2k}^T\mathcal{A}\mathcal{U}_{2k} = \begin{bmatrix} 0 & B_k \\ B_k^T & 0 \end{bmatrix}$ and

$$f(\mathcal{A})\boldsymbol{c} = \begin{bmatrix} f^\diamond(A)\boldsymbol{b} \\ 0 \end{bmatrix},$$

$$\mathcal{U}_{2k}f(\mathcal{U}_{2k}^T\mathcal{A}\mathcal{U}_{2k})\mathcal{U}_{2k}^T\boldsymbol{c} = \begin{bmatrix} P_kf^\diamond(B_k)Q_k^T\boldsymbol{b} \\ 0 \end{bmatrix}.$$

Moreover, it can be proved that the columns of $\mathcal{U}_{2k}$ are an orthonormal basis for a rational Krylov subspace $\mathcal{Q}_{2k}(\mathcal{A}, \boldsymbol{c})$, whose poles consist of a single pole at ∞, and $\pm\theta_j$, $j = 1, \ldots, k-1$, where $\theta_j^2 = \xi_j$ for each j. This fact is straightforward to prove in the polynomial case, where all θ_j are equal to ∞.

An alternative derivation of the error bounds in Theorem 6.3 and Theorem 6.7 could then be obtained by combining the above fact with Lemma 6.1 and error bounds concerning rational Krylov approximation of standard matrix functions (see, e.g., [18, Corollary 3.4]).

6.3 An optimal pole for the Shift-and-Invert method

In this section we use the error bounds in Theorem 6.7 combined with a known result from approximation theory to find a pole that optimizes the bounds in the case of a single repeated pole (Shift-and-Invert method) located on the negative real line.

We consider the case of a nonsingular square matrix $A \in \mathbb{R}^{n\times n}$, with singular values contained in the interval $[\sigma_{\min}, \sigma_{\max}]$, with $\sigma_{\min} > 0$.

Note that with a change of variables the bound (6.11) can be rewritten in the form

$$\|f^{\diamond}(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2 \leq 2\|A\boldsymbol{b}\|_2 \min_{p\in\Pi_{k-1}} \|g(z) - q_{k-1}(z)^{-1}p(z)\|_{\infty,[\sigma_{\min}^2,\sigma_{\max}^2]},$$

where the function g is defined as $g(z) = f(\sqrt{z})/\sqrt{z}$.

In the case of a single repeated pole $\xi < 0$, we have $q_{k-1}(z) = (z-\xi)^{k-1}$ and

$$\Big\{(z-\xi)^{-k+1}p(z) \,:\, p \in \Pi_{k-1}\Big\} = \Big\{p((z-\xi)^{-1}) \,:\, p \in \Pi_{k-1}\Big\}.$$

By defining $h(z) = g(z^{-1} + \xi)$, so that $g(z) = h((z-\xi)^{-1})$, we get

$$\min_{p\in\Pi_{k-1}} \|g(z) - \frac{p(z)}{(z-\xi)^{k-1}}\|_{\infty,[\sigma_{\min}^2,\sigma_{\max}^2]} = \min_{p\in\Pi_{k-1}} \|h(z) - p(z)\|_{\infty,[\mu_{\min},\mu_{\max}]}, \tag{6.16}$$

where $\mu_{\min} := (\sigma_{\max}^2 - \xi)^{-1}$ and $\mu_{\max} := (\sigma_{\min}^2 - \xi)^{-1}$. Notice that for $\xi < 0$ we indeed have $0 < \mu_{\min} \leq \mu_{\max} \leq (-\xi)^{-1}$.

The minimum in (6.16) can be bounded using the following result in approximation theory, adapted from [26, Proposition 3.1]. Its proof relies on classical bounds for Faber series, see [14, Corollary 2.2].

Proposition 6.9. *Let $\xi < 0$, and assume that $h(z) = g(z^{-1} + \xi)$ is analytic in the strip $0 < \operatorname{Re} z < (-\xi)^{-1}$ and continuous in $[0, (-\xi)^{-1}]$. Then, for any integer $k \geq 1$ the following inequality holds,*

$$\min_{p\in\Pi_{k-1}} \|h(z) - p(z)\|_{\infty,[\mu_{\min},\mu_{\max}]} \leq 2M\frac{\rho^k}{1-\rho}, \tag{6.17}$$

where $M = \|h(z)\|_{\infty,[0,(-\xi)^{-1}]}$ and

$$\rho = \max\left\{ \frac{\sqrt{\sigma_{\max}^2 - \xi} - \sqrt{\sigma_{\min}^2 - \xi}}{\sqrt{\sigma_{\max}^2 - \xi} + \sqrt{\sigma_{\min}^2 - \xi}}, \frac{\sigma_{\max}\sqrt{\sigma_{\min}^2 - \xi} - \sigma_{\min}\sqrt{\sigma_{\max}^2 - \xi}}{\sigma_{\max}\sqrt{\sigma_{\min}^2 - \xi} + \sigma_{\min}\sqrt{\sigma_{\max}^2 - \xi}} \right\}.$$

It follows from the analysis after [26, Proposition 3.1] that the bound (6.17) is optimized by choosing $\xi = -\sigma_{\min}\sigma_{\max}$. This choice leads to the following bound for the Shift-and-Invert iterates:

$$\|\boldsymbol{y} - \bar{\boldsymbol{y}}_k\|_2 \le 2\|\boldsymbol{b}\|_2 M \sqrt{\frac{\sigma_{\max}}{\sigma_{\min}}} \exp\Big(- 2k\sqrt{\frac{\sigma_{\min}}{\sigma_{\max}}}\Big). \tag{6.18}$$

We remark that the original result (see [26, equation (3.4)]) exihibited an error like $\exp\big(- 2k\sqrt[4]{\frac{a}{b}}\big)$ for a symmetric matrix A with spectrum in $[a, b] \subset (0, +\infty)$, when using the Shift-and-Invert method with the optimal pole $\xi = -\sqrt{ab}$. The fact that in (6.18) we have $\sqrt{\frac{\sigma_{\min}}{\sigma_{\max}}}$ instead of $\sqrt[4]{\frac{\sigma_{\min}}{\sigma_{\max}}}$ is not surprising, since we are essentially applying the result from [26] to the matrix A^TA, whose spectrum is contained in $[\sigma_{\min}^2, \sigma_{\max}^2]$.

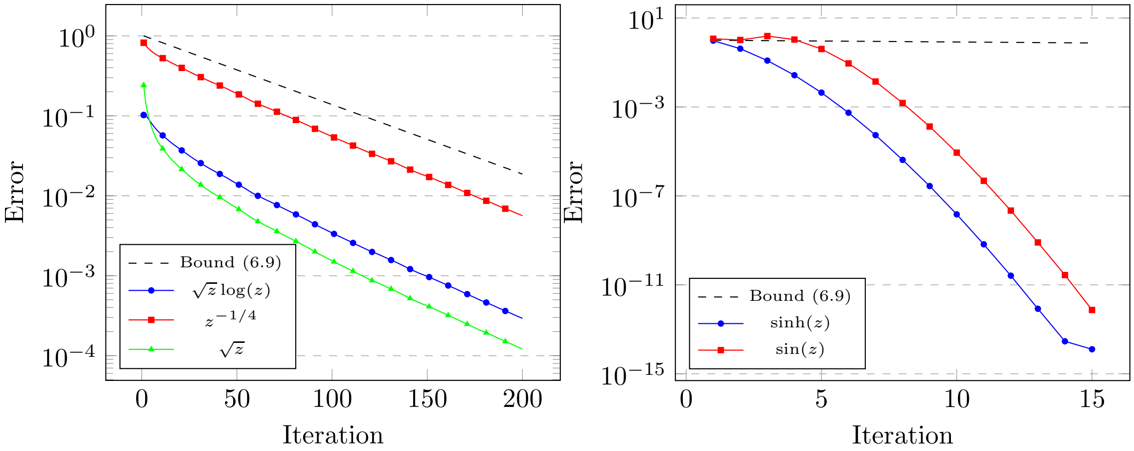
7 Numerical results

In this section we present some numerical experiments with the purpose of illustrating the error bounds and comparing the different methods proposed in the previous sections. We first test the methods on randomly generated matrices with a prescribed distribution of singular values, obtained by taking two random orthogonal matrices $U, V \in \mathbb{R}^{n\times n}$ and constructing $A = U\Sigma V^T$, where $\Sigma = \operatorname{diag}(\sigma_1, \dots, \sigma_n) \in \mathbb{R}^{n\times n}$. The random orthogonal matrices are obtained by taking a matrix $B \in \mathbb{R}^{n\times n}$ with entries from the normal distribution $\mathcal{N}(0, 1)$ and computing the QR factorization $B = QR$. If the diagonal entries of R are nonnegative, then Q is a random orthogonal matrix from the Haar distribution, a natural uniform probability distribution on the manifold of $n \times n$ orthogonal matrices [33]. In our last experiment we use the adjacency matrix of a directed graph from the Sparse Matrix Collection [12]. For simplicity, we assume that the interval of singular values $[\sigma_n, \sigma_1]$ is known. In a practical situation, one would first need to compute a (rough) approximation of the extremal singular values.

The experiments were done with MATLAB, using the `rat_krylov` function from the Rational Krylov Toolbox [7] for the implementation of the rational Arnoldi algorithm. The plots display the relative 2-norm error $\|f^\diamond(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2$, where $\bar{\boldsymbol{y}}_k$ is the approximation defined in (5.3) or (5.4), depending on the Krylov method that was used.

7.1 Error bounds

We start by illustrating in Figure 1 the sharpness of the bound (6.9) for the polynomial Krylov method. Under the assumptions of Theorem 6.5, the rate of convergence in the bound only depends on the interval $[\sigma_n, \sigma_1]$, and hence we can expect it to be pessimistic for most functions. Indeed, for entire functions such as $\sinh(z)$ and $\sin(z)$ (see Figure 1(b)) the convergence is much faster



(a) (b)

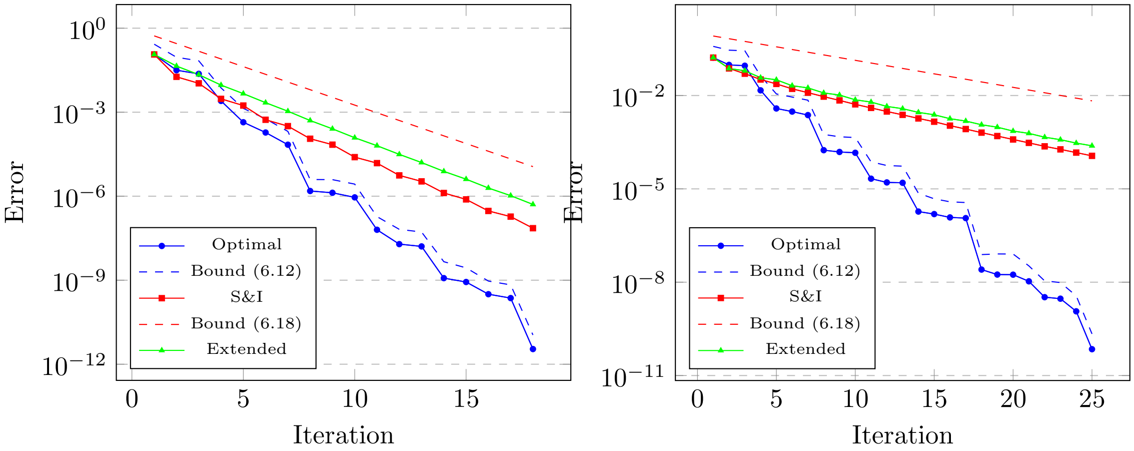
Figure 1: Convergence of the polynomial Krylov method for the approximation of $f^{\diamond}(A)\boldsymbol{b}$, where A is a 2000×2000 matrix whose singular values are Chebyshev points of the second kind for the interval $[10^{-1}, 10]$, and $\boldsymbol{b}$ is a random vector. Left: functions with an asymptotic convergence rate predicted by the bound (6.9). Right: entire functions with fast convergence.

than the bound (6.9); however, the bound can capture the asymptotic rate of convergence for certain functions with lower regularity such as $\sqrt{z}$, for suitable singular value distributions (see Figure 1(a)). This implies that, under the same assumptions of Theorem 6.5, it is only possible to improve the multiplicative constant in the bound (6.9). Note that in Figure 1 the multiplicative constant in the bound (6.9) was ignored for better visualization.

In Figure 2 we compare the convergence of the rational Krylov methods and we test the sharpness of the bounds (6.12) and (6.18). We use the Shift-and-Invert method with the pole $\xi = -\sigma_{\min}\sigma_{\max}$, the extended Krylov method [13], that alternates poles at ∞ with poles at 0, and a general Krylov method with an asymptotically optimal pole sequence for Laplace-Stieltjes and Cauchy-Stieltjes functions, developed in [24]. The poles were selected using the interval $[\sigma_n^2, \sigma_1^2]$, with reference to the bound (6.13). The function $f(z) = \sqrt{z}\log(1+\sqrt{z})$ is such that the function

$$\frac{f(\sqrt{z})}{\sqrt{z}} = \frac{\log(1+\sqrt[4]{z})}{\sqrt[4]{z}}$$

is Laplace-Stieltjes, or equivalently, completely monotonic [32, Definition 1.3]. This follows from [32, Theorem 3.7] and the fact that $\log(1+z)/z$ is completely monotonic. An approximation from above to the bound (6.12) was evaluated using a quasi-optimal polynomial p computed by replacing the uniform norm with the 2-norm on a discrete set of points in $I \cup (-I)$. We can see in Figure 2 that the convergence of the rational Krylov method with asymptotically optimal poles closely follows the bound (6.12), and that the convergence rate of the Shift-



(a)

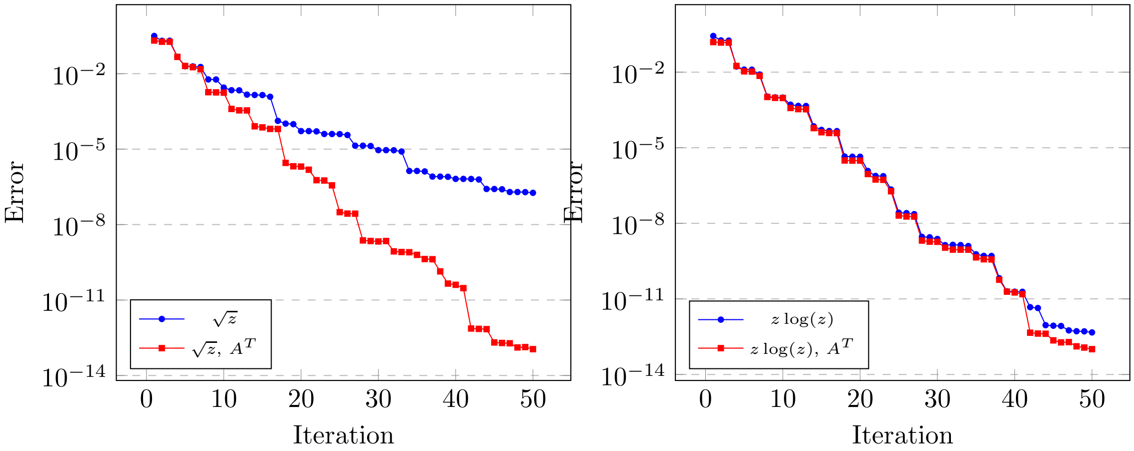
(b)

Figure 2: Convergence of different rational Krylov methods for the approximation of $f^{\diamond}(A)\boldsymbol{b}$, where A is a 2000×2000 matrix with logspaced singular values in the interval $[1, 10]$ (left) or $[10^{-1}, 10]$ (right), $f(z) = \sqrt{z}\log(1 + \sqrt{z})$, and $\boldsymbol{b}$ is a random vector.

and-Invert method is correctly predicted by the bound (6.18). The convergence speed of the extended Krylov method is comparable to the one of the Shift-and-Invert method. Note that, as in the polynomial case, the bound (6.18) displayed in Figure 2 does not include the multiplicative constant.

7.2 Rectangular case

Next, we investigate the performance of the methods in the case of a rectangular matrix $A \in \mathbb{R}^{m \times n}$, and the effectiveness of the strategy proposed in Remark 6.4 when $m < n$ to reduce the computation of $f^{\diamond}(A)\boldsymbol{b}$ to the computation of $f^{\diamond}(A^T)A\boldsymbol{b}$. We report in Figure 3 the convergence plots of the rational Krylov method with asymtotically optimal poles, for the functions $f(z) = \sqrt{z}$ and $f(z) = z\log(z)$. We can observe that the convergence is similar for the function $z\log(z)$ (Figure 3(b)), while there is a large benefit in using the alternative expression (6.7) in the case of the function $f(z) = \sqrt{z}$. This is likely due to the fact that $\sqrt{z}$ has a large derivative close to zero, and hence roundoff errors in the smallest computed singular values of the matrix B_k are extremely amplified when applying the function f. Indeed, we can see in Figure 3(a) that it is not possible to get below a relative accuracy of 10^{-8} if we directly approximate $f^{\diamond}(A)\boldsymbol{b}$, while we can reach a relative error of about 10^{-13} if we use the connection with $f^{\diamond}(A^T)A\boldsymbol{b}$, since in this case the projected matrix B_k has no singular values close to zero.



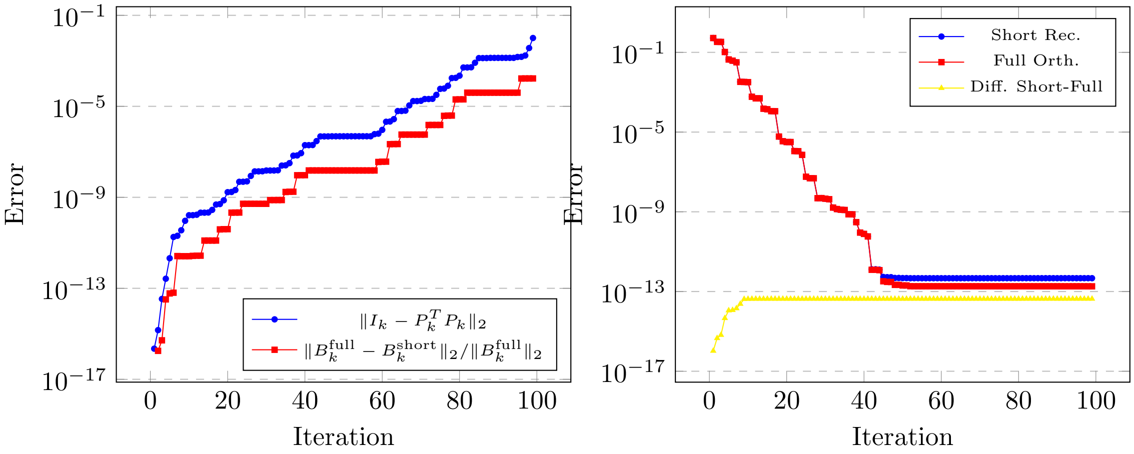
(a) (b)

Figure 3: Convergence of the rational Krylov methods with asymptotically optimal poles for the approximation of $f^\diamond(A)\boldsymbol{b}$, where A is a rectangular 1000×1500 matrix whose singular values are Chebyshev points of the second kind for the interval $[10^{-2}, 10]$. The red line shows the convergence of the method described in Remark 6.4, which computes $f^\diamond(A)\boldsymbol{b}$ by first computing $f^\diamond(A^T)A\boldsymbol{b}$ and then solving a least squares problem.

7.3 Finite precision issues

In finite precision, one of the main practical problems of Krylov methods based on a short recurrence (such as, for instance, the Lanczos method) is the loss of orthogonality in the computed basis vector. This phenomenon has been studied for the polynomial Lanczos case in [29]. A brief study of the problem for the rational Lanczos case can be found in [30].

As can be expected, the algorithm presented in Section 5.3 also suffers from this numerical instability. However, our experiments show that this loss of orthogonality deteriorates only slightly the accuracy of the algorithm: if the poles are chosen to guarantee a moderate number of iterations for convergence, it appears that the error produced by comparing the short recurrence algorithm with the one that uses full ortogonalization remains rather small, and it stops growing after a few iterations (see Figure 4). The effect of finite precision arithmetic and the subsequent loss of orthogonality in the Krylov basis have been already studied in [27] for the approximation of the product between a standard matrix function and a vector by means of the polynomial Lanczos algorithm; on the other hand, a theoretical analysis of the loss of orthogonality in the short-recurrence rational Lanczos algorithm, even in the context of standard matrix functions, is a challenging problem: see for example [30, Section 4].



(a) (b)

Figure 4: Effects of the loss of orthogonality in the rational Golub-Kahan algorithm for the approximation of $f^{\diamond}(A)\boldsymbol{b}$, where $f(z) = \sqrt{z}$ and A is a 2000×2000 matrix with logspaced singular values in the interval $[10^{-1}, 10^{2}]$, for the rational Krylov method with asymptotically optimal poles. Left: loss of orthogonality and error in the projected matrix when using the short recurrence. Right: comparison of the error in the approximation of $f^{\diamond}(A)\boldsymbol{b}$ when using the short recurrence or full orthogonalization of the basis vectors. In yellow we reported the norm of the difference between the two approximations.

7.4 A practical example

Our last experiment compares the accuracy and execution times of polynomial and rational Krylov methods in a more practical scenario. We consider the computation of $f^{\diamond}(A)\boldsymbol{b}$, where A is the 8490×8490 adjacency matrix of the largest strongly connected component of the directed graph `p2p-Gnutella30` from the Sparse Matrix Collection [12] and $\boldsymbol{b}$ is the vector of all ones. This kind of expression arises when computing functions of the adjacency matrix of the associated bipartite graph. We consider the functions $f_1(z) = \sinh(z)$ and $f_2(z) = z^{1/3}$. The function f_1 is an entire function that appears in the computation of hub and authority communicabilities of nodes in a directed graph [3, 5], and the function f_2 is such that $f_2(\sqrt{z})/\sqrt{z}$ is Cauchy-Stieltjes, in order to use the asymptotically optimal poles developed in [24].

In order to simulate a real situation where the exact solution is not available, we used as a simple stopping criterion the relative difference between two consecutive approximations, i.e. we stopped the algorithm as soon as

$$\frac{\|\bar{\boldsymbol{y}}_k - \bar{\boldsymbol{y}}_{k+1}\|_2}{\|\bar{\boldsymbol{y}}_k\|_2} \leq \texttt{tol}, \tag{7.1}$$

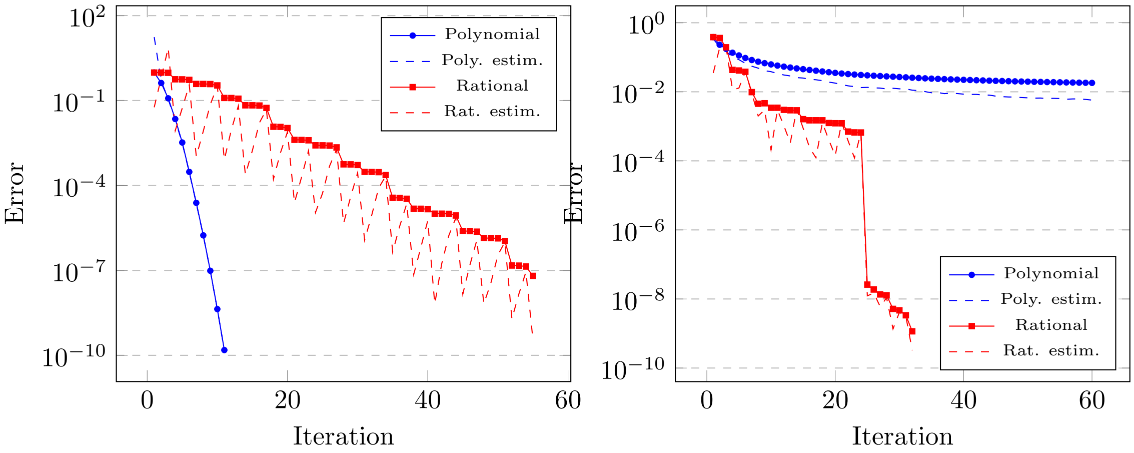
where `tol` is the requested error tolerance. We remark that although this kind of stopping criterion is widely used in practice, it does not provide any guarantees and it often underestimates the actual error $\|f^{\diamond}(A)\boldsymbol{b} - \bar{\boldsymbol{y}}_k\|_2$, especially if the method is stagnating.

Our results are summarized in Figure 5 and Table 1. Since f_1 is an entire function, the polynomial method converges very quickly and thus outperforms any rational Krylov method[2]. On the other hand, the function f_2 is less regular, and the performance of the rational Krylov method with asymptotically optimal poles is much more competitive, both in terms of number of iterations and execution time. We can see that the error estimate (7.1) is quite accurate when the method converges quickly, but it stops the algorithm too early if the convergence is slower.

8 Conclusions

In this paper we have proposed the use of rational Krylov methods in the computation of the action of a generalized matrix function on a vector. We have developed an extension of the Golub-Kahan bidiagonalization to the rational case, that uses a short recurrence to compute the basis vectors of the rational Krylov subspace. We have proved error bounds for the computation of GMFs with polynomial and rational Krylov methods, that relate the error of approximating $f^{\diamond}(A)\boldsymbol{b}$ with the best uniform polynomial or rational approximation of the function f on a real interval containing the singular values of A, and we

[2] We remark that the number of iterations of the rational Krylov method for f_1 could be reduced with a more careful choice of poles. However, it would still be outperformed by the polynomial Krylov method.



(a)

(b)

Figure 5: Comparison between polynomial and rational Krylov for the approximation of $f^{\diamond}(A)\boldsymbol{b}$, where A is the 8490×8490 adjacency matrix of the largest strongly connected component of the directed graph `p2p-Gnutella30` and $\boldsymbol{b}$ is the vector of all ones. We have set `tol` $= 10^{-9}$ in the stopping criterion (7.1). The dashed lines are the error estimates (7.1). Left: $f_1(z) = \sinh(z)$. Right: $f_2(z) = z^{1/3}$.

function	polynomial			rational		
	k	t_k	E_k	k	t_k	E_k
$\sinh(z)$	11	0.0103	1.54e-10	55	10.0469	6.41e-08
$z^{1/3}$	2000	12.6563	2.48e-03	32	5.6455	1.16e-09

Table 1: Number of iterations k, execution time t_k in seconds required to achieve tolerance `tol` $= 10^{-9}$, and actual error E_k at iteration k. The execution times are for the short recurrence implementations, obtained as an average over 10 runs. In the case of $f_2(z) = z^{1/3}$, after 2000 iterations of the polynomial method the error estimate was still `1.99e-03`.

have conducted experiments to investigate the sharpness of such bounds. The experiments we performed also show that rational Krylov methods are particularly effective compared to polynomial Krylov methods when the function f or its derivatives have singularities close to the singular values of A.

Acknowledgements

The authors would like to thank Michele Benzi for his support and advice, as well as the anonymous reviewers for their useful comments and suggestions.

References

[1] Fredrik Andersson, Marcus Carlsson, and Karl-Mikael Perfekt. Operator-Lipschitz estimates for the singular value functional calculus. *Proc. Amer. Math. Soc.*, 144(5):1867–1875, 2016.

[2] Francesca Arrigo and Michele Benzi. Edge modification criteria for enhancing the communicability of digraphs. *SIAM J. Matrix Anal. Appl.*, 37(1):443–468, 2016.

[3] Francesca Arrigo, Michele Benzi, and Caterina Fenu. Computation of generalized matrix functions. *SIAM J. Matrix Anal. Appl.*, 37(3):836–860, 2016.

[4] Jared L. Aurentz, Anthony P. Austin, Michele Benzi, and Vassilis Kalantzis. Stable computation of generalized matrix functions via polynomial interpolation. *SIAM J. Matrix Anal. Appl.*, 40(1):210–234, 2019.

[5] Michele Benzi, Ernesto Estrada, and Christine Klymko. Ranking hubs and authorities using matrix functions. *Linear Algebra Appl.*, 438(5):2447–2474, 2013.

[6] Michele Benzi and Ru Huang. Some matrix properties preserved by generalized matrix functions. *Spec. Matrices*, 7:27–37, 2019.

[7] Mario Berljafa, Steven Elsworth, and Stefan Güttel. A rational Krylov toolbox for MATLAB. *MIMS EPrint 2014.56, Manchester Institute for Mathematical Sciences, University of Manchester, Manchester, UK*, 2014.

[8] Mario Berljafa and Stefan Güttel. Generalized rational Krylov decompositions with an application to rational approximation. *SIAM J. Matrix Anal. Appl.*, 36(2):894–916, 2015.

[9] Mario Berljafa and Stefan Güttel. Parallelization of the rational Arnoldi algorithm. *SIAM J. Sci. Comput.*, 39(5):S197–S221, 2017.

[10] Daan Camps, Karl Meerbergen, and Raf Vandebril. An implicit filter for rational Krylov using core transformations. *Linear Algebra Appl.*, 561:113–140, 2019.

[11] Charles K. Chui and Maurice Hasson. Degree of uniform approximation on disjoint intervals. *Pacific J. Math.*, 105(2):291–297, 1983.

[12] Timothy A. Davis and Yifan Hu. The University of Florida Sparse Matrix Collection. *ACM Trans. Math. Software*, 38(1), 2011. Art. 1.

[13] Vladimir Druskin and Leonid Knizhnerman. Extended Krylov subspaces: approximation of the matrix square root and related functions. *SIAM J. Matrix Anal. Appl.*, 19(3):755–771, 1998.

[14] S. W. Ellacott. Computation of Faber series with application to numerical polynomial approximation in the complex plane. *Math. Comp.*, 40(162):575–587, 1983.

[15] Dario Fasino. Rational Krylov matrices and QR steps on Hermitian diagonal-plus-semiseparable matrices. *Numer. Linear Algebra Appl.*, 12(8):743–754, 2005.

[16] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

[17] Stefan Güttel. *Rational Krylov Methods for Operator Functions*. PhD thesis, Technische Universität Bergakademie Freiberg, Germany, 2010. Dissertation available as MIMS Eprint 2017.39.

[18] Stefan Güttel. Rational Krylov approximation of matrix functions: numerical methods and optimal pole selection. *GAMM-Mitt.*, 36(1):8–31, 2013.

[19] J. B. Hawkins and Adi Ben-Israel. On generalized matrix functions. *Linear and Multilinear Algebra*, 1(2):163–171, 1973.

[20] Nicholas J. Higham. *Functions of Matrices. Theory and Computation*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[21] Lin Lin. Lecture Notes on Quantum Algorithms for Scientific Computation, 2022. arXiv:2201.08309.

[22] Thomas Mach, Miroslav S. Pranić, and Raf Vandebril. Computing approximate extended Krylov subspaces without explicit inversion. *Electron. Trans. Numer. Anal.*, 40:414–435, 2013.

[23] Thomas Mach, Miroslav S. Pranić, and Raf Vandebril. Computing approximate (block) rational Krylov subspaces without explicit inversion with extensions to symmetric matrices. *Electron. Trans. Numer. Anal.*, 43:100–124, 2014/15.

[24] Stefano Massei and Leonardo Robol. Rational Krylov for Stieltjes matrix functions: convergence and pole selection. *BIT*, 61(1):237–273, 2021.

[25] Igor Moret and Paolo Novati. RD-rational approximations of the matrix exponential. *BIT*, 44(3):595–615, 2004.

[26] Igor Moret and Paolo Novati. Krylov subspace methods for functions of fractional differential operators. *Math. Comp.*, 88(315):293–312, 2019.

[27] Cameron Musco, Christopher Musco, and Aaron Sidford. Stability of the Lanczos method for matrix function approximation. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1605–1624. SIAM, Philadelphia, PA, 2018.

[28] Vanni Noferini. A formula for the Fréchet derivative of a generalized matrix function. *SIAM J. Matrix Anal. Appl.*, 38(2):434–457, 2017.

[29] Christopher C. Paige. Error analysis of the Lanczos algorithm for tridiagonalizing a symmetric matrix. *J. Inst. Math. Appl.*, 18(3):341–349, 1976.

[30] Davide Palitta, Stefano Pozza, and Valeria Simoncini. The short-term rational Lanczos method and applications, 2021. arXiv:2103.04054.

[31] Axel Ruhe. Rational Krylov algorithms for nonsymmetric eigenvalue problems. In *Recent Advances in Iterative Methods*, volume 60 of *IMA Vol. Math. Appl.*, pages 149–164. Springer, New York, 1994.

[32] René L. Schilling, Renming Song, and Zoran Vondraček. *Bernstein Functions. Theory and Applications*, volume 37 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, second edition, 2012.

[33] G. W. Stewart. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM J. Numer. Anal.*, 17(3):403–409, 1980.

[34] Lloyd N. Trefethen. *Approximation Theory and Approximation Practice.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2013.

[35] Marc Van Barel, Dario Fasino, Luca Gemignani, and Nicola Mastronardi. Orthogonal rational functions and structured matrices. *SIAM J. Matrix Anal. Appl.*, 26(3):810–829, 2005.

[36] Niel Van Buggenhout, Marc Van Barel, and Raf Vandebril. Biorthogonal rational Krylov subspace methods. *Electron. Trans. Numer. Anal.*, 51:451–468, 2019.

[37] Jasper van den Eshof and Marlis Hochbruck. Preconditioning Lanczos approximations to the matrix exponential. *SIAM J. Sci. Comput.*, 27(4):1438–1457, 2006.

[38] Raf Vandebril, Marc Van Barel, and Nicola Mastronardi. *Matrix Computations and Semiseparable Matrices: Linear Systems. Vol. 1.* Johns Hopkins University Press, Baltimore, MD, 2008.