SIAM J. NUMER. ANAL.
Vol. 59, No. 3, pp. 1325–1347

LOW-RANK UPDATES OF MATRIX FUNCTIONS II: RATIONAL KRYLOV METHODS*

BERNHARD BECKERMANN†, ALICE CORTINOVIS‡, DANIEL KRESSNER‡, AND MARCEL SCHWEITZER§

Abstract. This work develops novel rational Krylov methods for updating a large-scale matrix function $f(A)$ when A is subject to low-rank modifications. It extends our previous work in this context on polynomial Krylov methods, for which we present a simplified convergence analysis. For the rational case, our convergence analysis is based on an exactness result that is connected to work by Bernstein and Van Loan on rank-one updates of rational matrix functions. We demonstrate the usefulness of the derived error bounds for guiding the choice of poles in the rational Krylov method for the exponential function and Markov functions. Low-rank updates of the matrix sign function require additional attention; we develop and analyze a combination of our methods with a squaring trick for this purpose. A curious connection between such updates and existing rational Krylov subspace methods for Sylvester matrix equations is pointed out.

Key words. matrix function, low-rank update, rational Krylov subspace, tensorized Krylov subspace, sign function

AMS subject classifications. 15A16, 65D30, 65F30, 65F60

DOI. 10.1137/20M1362553

1. Introduction.

The need to compute matrix functions or associated quantities arises in a variety of applications, including network analysis [9, 21], signal processing [41], machine learning [45], and differential equations [31]. In many of these applications, slight changes of the problem setting, such as removing a vertex in a graph or changing a parameter in a differential equation, induce a low-rank change of the matrix. In this work, we discuss new methods for updating the matrix function under such changes. Specifically, assuming that a matrix function $f(A)$ has been computed and A is modified by a low-rank matrix D, we aim at computing the update

$$f(A+D)-f(A) \tag{1}$$

in a way that is cheaper than computing $f(A+D)$ from scratch. Such an update is also useful when only some quantities associated with $f(A)$, such as the trace or the diagonal entries, are of interest.

In [5], we have introduced and analyzed an algorithm for efficiently approximating (1) by projection onto polynomial Krylov subspaces. While this algorithm often

*Received by the editors August 26, 2020; accepted for publication (in revised form) February 22, 2021; published electronically May 13, 2021.
https://doi.org/10.1137/20M1362553
Funding: The work of the first author was partially supported by Labex CEMPI grant ANR-11-LABX-0007-01. The work of the second author was supported by SNSF research project "Fast Algorithms from Low-Rank Updates," grant 200020_178806. The work of the fourth author was partially supported by the SNSF research project "Low-Rank Updates of Matrix Functions and Fast Eigenvalue Solvers."

†Laboratoire Paul Painlevé UMR 8524, Département de Mathématiques, Université de Lille, F-59655 Villeneuve d'Ascq CEDEX, France (Bernhard.Beckermann@univ-lille1.fr).

‡Institute of Mathematics, EPF Lausanne, 1015 Lausanne, Switzerland (alice.cortinovis@epfl.ch, daniel.kressner@epfl.ch).

§Mathematisch-Naturwissenschaftliche Fakultät, Heinrich-Heine-Universität Düsseldorf, 40225 Düsseldorf, Germany (marcel.schweitzer@hhu.de).

shows satisfactory convergence, especially for entire functions like the matrix exponential and matrices with a "favorable" spectral distribution, convergence can also be very slow in other cases. In particular, this can happen when A has eigenvalues close to a singularity of f. A typical example is the matrix square root $A^{1/2}$ for a symmetric positive definite matrix A with eigenvalues close to zero. Rational Krylov spaces can lead to much faster convergence in such situations; at least this is indicated by existing work on approximating the action of the matrix function on a vector, $f(A)\boldsymbol{b}$, and solving matrix equations; see [17, 27, 28, 33, 42].

The main goal of this paper is thus to extend the techniques of [5] to incorporate *rational Krylov subspaces* and to analyze the convergence of the resulting algorithms. At the same time, we will also show that the original convergence analysis in [5] can be significantly simplified by using recent results from [16].

The work of Bernstein and Van Loan in [12] extends the Sherman–Morrison formula [40] for updating matrix inverses to rational matrix functions. In particular, Theorem 3 in [12] gives an analytic expression for the update (1) and shows that it has rank at most m if f is a rational function of degree m and D has rank one. In principle, it would be possible to exploit the good approximation properties of rational functions in the context of (1) by first replacing f with a suitable low-degree rational approximation r and then using the formula from [12, Theorem 3] to approximate the update

$$f(A+D) - f(A) \approx r(A+D) - r(A).$$

We will discuss the relation of this approach to our new method in section 3.3.

The remainder of this paper is organized as follows. We begin by briefly describing a general subspace projection approach for the computation of the update (1) in section 2. The particular choice of rational Krylov subspaces in this approach is then discussed in section 3. In addition, we show that the proposed rational Krylov method is exact when approximating updates of certain rational functions and discuss the connection of our approach to the generalized Sherman–Morrison formula for rational functions from [12]. In section 4, we analyze the convergence of our methods for several important matrix functions. Afterwards, in section 5, we specifically focus on the matrix sign function and its peculiarities in the context of approximating low-rank updates; we conclude by showing a connection to Sylvester equations.

2. A subspace projection approach for low-rank updates. In this section, we present a general subspace projection approach for approximating the update (1), which includes the algorithm from [5] as well as our newly proposed algorithm.

Let $A \in \mathbb{C}^{n\times n}$ and $D \in \mathbb{C}^{n\times n}$ be such that both $f(A)$ and $f(A+D)$ are well defined. In the following, we describe how an approximation to $f(A+D) - f(A)$ is extracted from two subspaces $\mathcal{U}_m, \mathcal{V}_m \subseteq \mathbb{C}^n$ of (low) dimension m_U and m_V, respectively. Considering orthonormal bases U_m, V_m of $\mathcal{U}_m, \mathcal{V}_m$, we let $G_m := U_m^* A U_m$ and $H_m := V_m^* A^* V_m$ denote the compressions of A and A^*, respectively. We then use an approximation of the form

$$f(A+D) - f(A) \approx U_m X_m(f) V_m^*,$$

where $X_m(f)$ is the (1,2)-block of the (small) matrix function

$$f\left(\begin{bmatrix} G_m & U_m^* D V_m \\ 0 & H_m^* + V_m^* D V_m \end{bmatrix}\right). \tag{2}$$

Note that $X_m(f)$ also depends on A and D, but we omit this dependence for notational convenience. In [5], this particular choice of $X_m(f)$ was motivated by a

polynomial exactness property for polynomial Krylov subspaces. We will see below, in Theorem 3.3, that an analogous property holds for rational Krylov subspaces. A more intuitive explanation, not tied to specific subspaces, is the observation [5, Lemma 2.2] that

$$f\left(\begin{bmatrix} A & D \\ 0 & A+D \end{bmatrix}\right) = \begin{bmatrix} f(A) & f(A+D)-f(A) \\ 0 & f(A+D) \end{bmatrix}. \tag{3}$$

Note that the compression onto $\mathcal{U}_m \oplus \mathcal{V}_m$ of the block matrix on the left-hand side of (3) corresponds to the matrix used in (2).

The described subspace projection approach is summarized in Algorithm 1, which encompasses Algorithm 2 from [5].

Algorithm 1 Subspace projection approach for approximating $f(A+D)-f(A)$.

1: Compute orthonormal bases $U_m \in \mathbb{C}^{n\times m_U}$, $V_m \in \mathbb{C}^{n\times m_V}$ of subspaces $\mathcal{U}_m, \mathcal{V}_m$.
2: Compute compressions $G_m = U_m^* A U_m$ and $H_m = V_m^* A^* V_m$.
3: Compute matrix function $F_m = f\left(\begin{bmatrix} G_m & U_m^* D V_m \\ 0 & H_m^* + V_m^* D V_m \end{bmatrix}\right)$.
4: Set $X_m(f) = F_m(1\!:\!m_U, m_U+1\!:\!m_U+m_V)$.
5: Return $U_m X_m(f) V_m^*$.

In the Hermitian case, $A = A^*$ and $D = D^*$, it is sensible to choose $\mathcal{U}_m = \mathcal{V}_m$, and thus $U_m = V_m$. In turn, $G_m = H_m^*$ and the computation of the update simplifies. Using the relation (3), one observes that

$$X_m(f) = f\big(U_m^*(A+D)U_m\big) - f\big(U_m^* A U_m\big) = f\big(G_m + U_m^* D U_m\big) - f(G_m). \tag{4}$$

The stopping criterion proposed in [5] uses the difference of two iterates as a simple error estimator, i.e.,

$$\|f(A+D) - U_m X_m(f) V_m^*\| \approx \|U_{m+d} X_{m+d}(f) V_{m+d}^* - U_m X_m(f) V_m^*\| \tag{5}$$

for some small integer $d \geq 1$, where $\|\cdot\|$ denotes the spectral norm of a matrix. When the subspaces are nested and, in turn, the orthonormal bases can be chosen to be nested (as is the case for Krylov subspaces and the Arnoldi method, for example), we have

$$\|U_{m+d} X_{m+d}(f) V_{m+d}^* - U_m X_m(f) V_m^*\| = \left\| X_{m+d}(f) - \begin{bmatrix} X_m(f) & 0 \\ 0 & 0 \end{bmatrix} \right\|.$$

Hence, there is no need to explicitly form $U_{m+d} X_{m+d}(f) V_{m+d}^*$ or $U_m X_m(f) V_m^*$. The heuristic (5) is often observed to give fairly accurate approximations to the exact error even for small values of d, say $d = 1$ or $d = 2$. A notable exception is when Algorithm 1 (almost) stagnates as m increases; in this case a small value of d might lead to severe underestimates; see [5, section 6.2] for an example.

3. Block rational Krylov subspace projection. In this section, we combine Algorithm 1 with rational Krylov subspaces. We assume that D is of rank ℓ and can thus be written as $D = \boldsymbol{B}\boldsymbol{C}^*$ for block vectors $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{C}^{n\times \ell}$ of full rank.

While a polynomial Krylov subspace with respect to A and $B = [\boldsymbol{b}_1, \ldots . \boldsymbol{b}_\ell]$ takes the form

$$\mathcal{K}_m(A, \boldsymbol{B}) = \operatorname{colspan}\big\{\boldsymbol{B}, A\boldsymbol{B}, \ldots, A^{m-1}\boldsymbol{B}\big\} = \mathcal{K}_m(A, \boldsymbol{b}_1) + \cdots + \mathcal{K}_m(A, \boldsymbol{b}_\ell),$$

the rational Krylov subspaces considered in this work take the form

$$q_m(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B}) = \operatorname{colspan}\{q_m(A)^{-1}\boldsymbol{B}, q_m(A)^{-1}A\boldsymbol{B}, \dots, q_m(A)^{-1}A^{m-1}\boldsymbol{B}\} \tag{6}$$

for a polynomial $q_m(z) = (z-\xi_1)(z-\xi_2)\cdots(z-\xi_m)$ of degree m and fixed poles $\xi_1, \dots, \xi_m \in \mathbb{C}$. Choosing some of the poles to be infinite corresponds to reducing the degree of q_m.

Remark 3.1. When choosing one pole to be infinite, our definition (6) coincides with the subspace $q_{m-1}(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B})$, which is more commonly found in the literature; see, e.g., [26]. Note that $\boldsymbol{B} \in q_{m-1}(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B})$ while this property fails to hold in general for $q_m(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B})$. One of the motivations for our choice (6) is that it nicely connects to the (generalized) Sherman–Morrison formula; see section 3.3 below.

Adapting the usual *rational Arnoldi method* [19] to (6), Algorithm 2 is used to compute an orthonormal basis $U_m = [\boldsymbol{U}_1, \dots, \boldsymbol{U}_m]$ of $q_m(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B})$. In the case of an infinite pole $\xi_j = \infty$, line 4 of Algorithm 2 is replaced by $\boldsymbol{W}_j \leftarrow A\boldsymbol{U}_{j-1}$ for $j > 1$ and line 1 is replaced by $\tilde{\boldsymbol{B}} \leftarrow A^{-1}\boldsymbol{B}$ for $j = 1$.

Algorithm 2 Block rational Arnoldi method.

1: $\tilde{\boldsymbol{B}} \leftarrow (A - \xi_1 I)^{-1}\boldsymbol{B}$
2: $\boldsymbol{U}_1 \leftarrow$ orthonormal basis of $\tilde{\boldsymbol{B}}$.
3: **for** $j = 2, 3, \dots, m$ **do**
4: $\quad \boldsymbol{W}_j \leftarrow (A - \xi_j I)^{-1} A \boldsymbol{U}_{j-1}$.
5: $\quad$ **for** $k = 1, \dots, j-1$ **do**
6: $\quad\quad \boldsymbol{\alpha}_{k,j-1} \leftarrow \boldsymbol{U}_k^* \boldsymbol{W}_j$.
7: $\quad\quad \boldsymbol{W}_j \leftarrow \boldsymbol{W}_j - \boldsymbol{U}_k \boldsymbol{\alpha}_{k,j-1}$
8: $\quad$ **end for**
9: $\quad \boldsymbol{U}_j \leftarrow$ orthonormal basis of $\boldsymbol{W}_j$.
10: **end for**

The description of Algorithm 2 assumes $\dim\left(q_m(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B})\right) = m\ell$, that is, all block vectors $\boldsymbol{W}_j$ have full rank. We will make this assumption from here on when discussing algorithms. Deflation techniques for removing linearly dependent columns are discussed, e.g., in [19, section 6].

We conclude our discussion of rational Krylov subspaces with a variation of an existing exactness result for rational matrix functions [26, Lemma 4.6].

LEMMA 3.2. *Let Π_{m-1}/q_m denote the space of all rational functions with numerator degree at most $m-1$ and denominator $q_m(z)$. Let U_m be an orthonormal basis of $q_m(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B})$. Then*

$$r(A)\boldsymbol{B} = U_m r(U_m^* A U_m) U_m^* \boldsymbol{B},$$

provided that $r(A)$ and $r(U_m^* A U_m)$ are well defined.

Proof. Consider $r = p/q_m$ for arbitrary $p \in \Pi_{m-1}$. We start by noting that $q_m(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B}) = \mathcal{K}_m(A,\boldsymbol{Q})$ with $\boldsymbol{Q} = q_m(A)^{-1}\boldsymbol{B}$. By existing results for (polynomial) Krylov subspaces (see [39, Lemma 3.1] and [26, Lemma 3.9]), which can be applied completely analogously in the block Krylov setting, we obtain

$$p(A)\boldsymbol{Q} = U_m p(U_m^* A U_m) U_m^* \boldsymbol{Q}, \tag{7}$$

as well as

$$U_m^* \boldsymbol{B} = U_m^* q_m(A)\boldsymbol{Q} = q_m(U_m^* A U_m) U_m^* \boldsymbol{Q}.$$

The latter relation is equivalent to $U_m^*\boldsymbol{Q} = q_m(U_m^*AU_m)^{-1}U_m^*\boldsymbol{B}$ and gives, when inserted into (7), the desired relation:

$$r(A)\boldsymbol{B} = p(A)\boldsymbol{Q} = U_m p(U_m^*AU_m) q_m(U_m^*AU_m)^{-1}U_m^*\boldsymbol{B} = U_m r(U_m^*AU_m)U_m^*\boldsymbol{B}. \quad \square$$

3.1. Algorithm. To compute an approximation of $f(A+\boldsymbol{B}\boldsymbol{C}^*)-f(A)$, we utilize Algorithm 2 to compute orthonormal bases U_m, V_m of rational Krylov subspaces

$$\mathcal{U}_m = q_m(A)^{-1}\mathcal{K}_m(A, \boldsymbol{B}), \qquad \mathcal{V}_m = \bar{q}_m(A^*)^{-1}\mathcal{K}_m(A^*, \boldsymbol{C}),$$

where $q_m(z) = (z-\xi_1)\cdots(z-\xi_m)$ and $\bar{q}_m(z) = (z-\overline{\xi}_1)\cdots(z-\overline{\xi}_m)$ are both determined by the same set of poles $\xi_1, \ldots, \xi_m$. Although it is in principle possible to choose a different set of poles for $\mathcal{V}_m$, we are not aware of any advantages of such a choice. Once U_m, V_m have been computed, we apply the general subspace projection approach, Algorithm 1, with these bases. For ease of reference, Algorithm 3 summarizes the resulting procedure.

Algorithm 3 Rational Krylov subspace approximation of $f(A + \boldsymbol{B}\boldsymbol{C}^*) - f(A)$.

1: Perform m steps of Algorithm 2 to compute an orthonormal basis U_m of $q_m(A)^{-1}\mathcal{K}_m(A, \boldsymbol{B})$ and set $G_m = U_m^*AU_m$.
2: Perform m steps of Algorithm 2 to compute an orthonormal basis V_m of $\bar{q}_m(A^*)^{-1}\mathcal{K}_m(A^*, \boldsymbol{C})$ and set $H_m = V_m^*A^*V_m$.
3: Compute matrix function $F_m = f\left(\begin{bmatrix} G_m & (U_m^*\boldsymbol{B})(V_m^*\boldsymbol{C})^* \\ 0 & H_m^* + (V_m^*\boldsymbol{B})(V_m^*\boldsymbol{C})^* \end{bmatrix}\right)$.
4: Set $X_m(f) = F_m(1:m, m+1:2m)$.
5: Return $U_m X_m(f) V_m^*$.

Several remarks concerning the implementation of Algorithm 3 are in order:

1. The efficient and stable implementation of rational Arnoldi methods requires some care, including the need for reorthogonalization; it is therefore advisable to build on available toolboxes, like, e.g., the `RKToolbox` by Berljafa, Elsworth, and Güttel [11].
2. In contrast to the (standard) Arnoldi method, the compressed matrices G_m and H_m do *not* contain the orthogonalization coefficients from Algorithm 2 explicitly. There are procedures which, possibly under additional conditions on the poles, circumvent the additional computation of the products $U_m^*AU_m$ and $V_m^*A^*V_m$ and compute G_m, H_m from $m\ell \times m\ell$ matrices containing the orthogonalization coefficients and the poles; see, e.g., [19, 26, 27] for details.
3. Assume that A is Hermitian and the rank-ℓ update can be written in the form $D = \boldsymbol{B}J\boldsymbol{B}^*$ for some $\boldsymbol{B} \in \mathbb{C}^{n\times\ell}$ and $J \in \mathbb{C}^{\ell\times\ell}$, that is, the columns of D and D^* span the same subspace of $\mathbb{C}^n$. In particular, this is the case when D is also Hermitian. Further, let us suppose that the poles are closed under complex conjugation, that is, ξ is a pole if and only if $\bar{\xi}$ is a pole, and both poles $\xi, \bar{\xi}$ have the same multiplicity. In particular, this holds when all poles are real. Then $q_m(A)^{-1}\mathcal{K}_m(A, \boldsymbol{B}) = \bar{q}_m(A^*)^{-1}\mathcal{K}_m(A^*, \boldsymbol{B}J^*)$. In turn, one can choose $V_m = U_m$ and step 2 in Algorithm 3 can be skipped and the corresponding remarks for Algorithm 1 apply. Specifically, we have the simplified expression $X_m(f) = f\big(G_m + U_m^*\boldsymbol{B}J\boldsymbol{B}^*U_m\big) - f(G_m)$.

4. When A is Hermitian, the (standard) block Arnoldi method reduces to the block Lanczos method [24]. Similarly, there exist short-term recurrences for extended Krylov subspaces, which only use the poles 0 and ∞ repeatedly; see, e.g., [17, 32, 42].
5. Each iteration of Algorithm 2 with a finite pole requires the solution of a shifted block linear system. The efficiency of Algorithm 3 largely depends on how efficiently these linear systems can be solved. When using a direct sparse factorization such as the sparse LU factorization, it is advantageous to use only a few different poles, allowing for the frequent reuse of factorizations when poles repeat. In the non-Hermitian case, *two* shifted linear systems—one with A and one with A^*—have to be solved at each iteration of the method. It is worth pointing out that it suffices to compute only *one* factorization, $A - \xi_j I = LU$, because this immediately gives the other factorization, $A^* - \overline{\xi}_j I = U^* L^*$.

3.2. Exactness properties. In [5], it was shown that the polynomial Krylov subspace approximation for the update $f(A + \boldsymbol{B}\boldsymbol{C}^*) - f(A)$ is exact when f is a polynomial of a certain degree. The following theorem extends this result to rational Krylov subspaces.

THEOREM 3.3. *Given $A \in \mathbb{C}^{n\times n}$, $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{C}^{n\times \ell}$, and $q_m(z) = (z-\xi_1)\cdots(z-\xi_m)$, with $\xi_1, \ldots, \xi_m \in \mathbb{C}$, the approximation returned by Algorithm 3 is exact for every $r \in \Pi_m/q_m$, that is,*

$$r(A + \boldsymbol{B}\boldsymbol{C}^*) - r(A) = U_m X_m(r) V_m^*,$$

provided that $r(A)$, $r(A + \boldsymbol{B}\boldsymbol{C}^*)$ as well as $r(G_m)$, $r(H_m^* + (V_m^*\boldsymbol{B})(V_m^*\boldsymbol{C})^*)$ are well defined.

Proof. By the partial fraction expansion, a rational function $r \in \Pi_m/q_m$ can be decomposed as the sum of a constant and scalar multiples of terms of the form $(z-\xi_s)^{-j}$, $j \le m_s$, where m_s denotes the multiplicity of ξ_s. By linearity, it suffices to show exactness for each of the terms individually. Exactness trivially holds for a constant function.

It remains to show exactness for $r_{\xi_s,j}(z) = (z-\xi_s)^{-j}$ for $j = 1, \ldots, m_s$. The matrix $X_m(r_{\xi_s,j})$ entering the rational Krylov approximation $U_m X_m(r_{\xi_s,j}) V_m^*$ is given by the $(1,2)$-block of the matrix

$$\begin{bmatrix} G_m - \xi_s I_m & U_m^*\boldsymbol{B}\boldsymbol{C}^*V_m \\ 0 & H_m^* - \xi_s I_m + V_m^*\boldsymbol{B}\boldsymbol{C}^*V_m \end{bmatrix}^{-j}. \tag{8}$$

For $j = 1$, we directly obtain

$$X_m(r_{\xi_s,1}) = -(G_m - \xi_s I_m)^{-1}(U_m^*\boldsymbol{B}\boldsymbol{C}^*V_m)(H_m^* - \xi_s I_m + V_m^*\boldsymbol{B}\boldsymbol{C}^*V_m)^{-1}. \tag{9}$$

For $j > 1$, (8) yields the recursive relation

$$X_m(r_{\xi_s,j}) = (G_m - \xi_s I_m)^{-(j-1)} X_m(r_{\xi_s,1}) + X_m(r_{\xi_s,j-1})(H_m^* - \xi_s I_m + V_m^*\boldsymbol{B}\boldsymbol{C}^*V_m)^{-1}.$$

Resolving this recursion and inserting (9) gives

$$\begin{aligned} X_m(r_{\xi_s,j}) &= \sum_{k=0}^{j-1} (G_m - \xi_s I_m)^{-(j-1-k)} X_m(r_{\xi_s,1})(H_m^* - \xi_s I_m + V_m^*\boldsymbol{B}\boldsymbol{C}^*V_m)^{-k} \\ &= -\sum_{k=0}^{j-1} (G_m - \xi_s I_m)^{-(j-k)}(U_m^*\boldsymbol{B}\boldsymbol{C}^*V_m)(H_m^* - \xi_s I_m + V_m^*\boldsymbol{B}\boldsymbol{C}^*V_m)^{-(k+1)}. \end{aligned} \tag{10}$$

Since $r_{\xi_s,s} \in \Pi_{m-1}/q_m$, we know from Lemma 3.2 that

$$U_m(G_m - \xi_s I_m)^{-d} U_m^* \boldsymbol{B} = (A - \xi_s I)^{-d} \boldsymbol{B} \quad \text{for all } d = 1, \ldots, m_s$$

and

$$\boldsymbol{C}^* V_m (H_m^* - \xi_s I_m + V_m^* \boldsymbol{B}\boldsymbol{C}^* V_m)^{-d} V_m^* = \boldsymbol{C}^* (A - \xi_s I + \boldsymbol{B}\boldsymbol{C}^*)^{-d} \quad \text{for all } d = 1, \ldots, m_s.$$

Combined with (10), these relations yield

$$U_m X_m(r_{\xi_s,j}) V_m^* = -\sum_{k=0}^{j-1} (A - \xi_s I)^{-(j-k)} \boldsymbol{B}\boldsymbol{C}^* (A - \xi_s I + \boldsymbol{B}\boldsymbol{C}^*)^{-(k+1)}. \tag{11}$$

We now use the matrix identity

$$M^j - N^j = \sum_{k=0}^{j-1} N^{j-1-k}(M - N)M^k$$

(see [5, Proposition 3.1]), with $M = (A - \xi_s I + \boldsymbol{B}\boldsymbol{C}^*)^{-1}$ and $N = (A - \xi_s I)^{-1}$. This yields

(12)

$$\begin{aligned}
&(A - \xi_s I + \boldsymbol{B}\boldsymbol{C}^*)^{-j} - (A - \xi_s I)^{-j} \\
&= \sum_{k=0}^{j-1} (A - \xi_s I)^{-(j-1-k)} ((A - \xi_s I + \boldsymbol{B}\boldsymbol{C}^*)^{-1} - (A - \xi_s I)^{-1})(A - \xi_s I + \boldsymbol{B}\boldsymbol{C}^*)^{-k} \\
&= -\sum_{k=0}^{j-1} (A - \xi_s I)^{-(j-k)} \boldsymbol{B}\boldsymbol{C}^* (A - \xi_s I + \boldsymbol{B}\boldsymbol{C}^*)^{-(k+1)},
\end{aligned}$$

where the latter equality utilizes the second resolvent identity. Comparing (12) with (11) establishes the desired exactness property for $r_{\xi_s,j}$ for $j \le m_s$. ∎

Remark 3.4. Although the statement of Theorem 3.3 assumes the poles to be finite, the result also holds in the presence of infinite poles. To see this, let $\widetilde{m} \le m$ be the multiplicity of ∞ as a pole of the rational Krylov subspace, that is, $\deg q_m = m - \widetilde{m}$. We can then decompose a rational function $r \in \Pi_m/q_m$ as $r = p + \tilde{r}$ with $p \in \Pi_{\widetilde{m}}$ and $\tilde{r} \in \Pi_{m-\widetilde{m}-1}/q_m$. By linearity, it suffices to show exactness for p and $\tilde{r}$ individually. Because of $\mathcal{K}_{\widetilde{m}}(A, \boldsymbol{B}) \subset q_m(A)^{-1}\mathcal{K}_m(A, \boldsymbol{B})$, exactness for p can be shown along the lines of the proof of Theorem 3.2 in [5]. Exactness for $\tilde{r}$ follows directly from the proof of Theorem 3.3.

3.3. Connection to the Sherman–Morrison formula and its generalization to rational functions. It is instructive to rederive the Sherman–Morrison formula for rank-one updates from Algorithm 3. Let A, $\boldsymbol{b} \neq 0$, $\boldsymbol{c} \neq 0$ be such that A and $A + \boldsymbol{b}\boldsymbol{c}^*$ are invertible. By Theorem 3.3, one step of Algorithm 3 with pole 0 should produce the exact update $(A + \boldsymbol{b}\boldsymbol{c}^*)^{-1} - A^{-1}$. In this situation, $U_1 = A^{-1}\boldsymbol{b}/\beta$, $V_1 = A^{-*}\boldsymbol{c}/\gamma$ with $\beta = \|A^{-1}\boldsymbol{b}\|$ and $\gamma = \|A^{-*}\boldsymbol{c}\|$. Therefore,

$$\begin{bmatrix} G_1 & (U_1^*\boldsymbol{b})(V_1^*\boldsymbol{c})^* \\ 0 & H_1^* + (V_1^*\boldsymbol{b})(V_1^*\boldsymbol{c})^* \end{bmatrix} = \begin{bmatrix} \frac{1}{\beta^2}\boldsymbol{b}^*A^{-*}\boldsymbol{b} & \frac{1}{\beta\gamma}(\boldsymbol{b}^*A^{-*}\boldsymbol{b})(\boldsymbol{c}^*A^{-*}\boldsymbol{c}) \\ 0 & \frac{1}{\gamma^2}(\boldsymbol{c}^*A^{-*}\boldsymbol{c})(1 + \boldsymbol{c}^*A^{-1}\boldsymbol{b}) \end{bmatrix}.$$

Provided that $\boldsymbol{b}^*A^{-*}\boldsymbol{b} \neq 0$, $\boldsymbol{c}^*A^{-*}\boldsymbol{c} \neq 0$, and $1+\boldsymbol{c}^*A^{-1}\boldsymbol{b} \neq 0$, this matrix is invertible and the $(1,2)$ entry of its inverse is given by $-\beta\gamma/(1+\boldsymbol{c}^*A^{-1}\boldsymbol{b})$. Hence, Algorithm 3 returns the exact update

$$-\frac{\beta\gamma}{1+\boldsymbol{c}^*A^{-1}\boldsymbol{b}}U_1V_1^* = -\frac{A^{-1}\boldsymbol{b}\boldsymbol{c}^*A^{-1}}{1+\boldsymbol{c}^*A^{-1}\boldsymbol{b}}.$$

Two observations can be made. On the one hand, the Sherman–Morrison formula is nicely reproduced by Algorithm 3. On the other hand, two assumptions ($\boldsymbol{b}^*A^{-*}\boldsymbol{b} \neq 0$, $\boldsymbol{c}^*A^{-*}\boldsymbol{c} \neq 0$) need to be made that are not necessary, neither for the existence of $(A+\boldsymbol{b}\boldsymbol{c}^*)^{-1} - A^{-1}$ nor for the validity of the Sherman–Morrison formula. Note that the violation of the conditions, $(\boldsymbol{b}^*A^{-*}\boldsymbol{b})(\boldsymbol{c}^*A^{-*}\boldsymbol{c}) = 0$, implies that the numerical range $W(A) := \{\boldsymbol{x}^*A\boldsymbol{x} : \|\boldsymbol{x}\| = 1\}$ of A contains 0, a singularity of the matrix function. In general, it is not advisable to use Algorithm 3 in such situations, and we will discuss in section 5, for a different scenario, how this can sometimes be circumvented.

In [12], Bernstein and Van Loan provide a generalization of the Sherman–Morrison formula for rational functions. The following theorem recalls their main result.

THEOREM 3.5 (Theorem 3 in [12]). *Let $r(z) = p(z)/q(z)$ with polynomials $p(z) = \sum_{i=0}^{m_p}\alpha_i z^i$ and $q(z) = \sum_{i=0}^{m_q}\beta_i z^i$ and set $m = \max\{m_p, m_q\}$. Let $H(\alpha)$ be the $m \times m$ Hankel matrix containing the coefficients α_i, i.e.,*

$$H(\alpha) = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{m_p} & 0 & \cdots & 0 \\ \alpha_2 & & \iddots & \iddots & & \iddots & \\ \vdots & \iddots & \iddots & & \iddots & & \\ \alpha_{m_p} & \iddots & & \iddots & & & \\ 0 & & \iddots & & & & \\ \vdots & \iddots & & & & & \\ 0 & & & & & & 0 \end{bmatrix} \in \mathbb{C}^{m\times m}$$

and define $H(\beta) \in \mathbb{C}^{m\times m}$ analogously. Suppose that $A \in \mathbb{C}^{n\times n}$, $\boldsymbol{b} \in \mathbb{C}^n$, $\boldsymbol{c} \in \mathbb{C}^n$ are such that $r(A)$ and $r(A+\boldsymbol{b}\boldsymbol{c}^*)$ are well defined. Set

$$\begin{aligned} K_m &= [\boldsymbol{b}, A\boldsymbol{b}, \ldots, A^{m-1}\boldsymbol{b}], \\ L_m &= [\boldsymbol{c}, (A^*+\boldsymbol{c}\boldsymbol{b}^*)\boldsymbol{c}, \ldots, (A^*+\boldsymbol{c}\boldsymbol{b}^*)^{m-1}\boldsymbol{c}], \\ Y_\alpha &= L_m H(\alpha)^*,\ Y_\beta = L_m H(\beta)^*. \end{aligned}$$

Then

$$r(A+\boldsymbol{b}\boldsymbol{c}^*) - r(A) = XY^*, \tag{13}$$

where the $n \times m$ matrices X, Y are defined by $X = q(A)^{-1}K_m$ and $Y^* = Y_\alpha^* - M^{-1}Y_\beta^*(r(A) + XY_\alpha^*)$ with $M = I + Y_\beta^*X$.

Note that it is also stated in [12] that the result of Theorem 3.5 can be extended to general rank-ℓ updates, but the technical details are omitted.

Consider a rational function r of the form stated in Theorem 3.5 with $\beta_{m_q} \neq 0$. Then Theorem 3.3 and Remark 3.4 state that Algorithm 3 is exact when choosing m_q poles equal to the zeros of q and, additionally, $\max\{m_p - m_q, 0\}$ infinite poles. In turn, the low-rank updates produced by Algorithm 3 and Theorem 3.5 have the same rank and yield *mathematically* the same result, up to normalization of the low-rank

factors. Also, the cost of an algorithm based on Theorem 3.5 is comparable to the cost of Algorithm 3. However, there are a number of important differences between these two approaches:

- Obviously, Algorithm 3 is more general as it applies to general functions, while Theorem 3.5 is restricted to rational functions. As discussed in the introduction, Theorem 3.5 could still be used to address a general function f by constructing a priori a rational approximation $r \approx f$. While Algorithm 3 also requires choosing the poles a priori, the numerator polynomial p_m is determined automatically by the method. In turn, significantly less knowledge about the spectra of A and $A + \boldsymbol{bc}^*$ is needed in order to obtain effective approximations. Another advantage of Algorithm 3 is that it easily combines with existing adaptive pole selection strategies for rational Krylov methods [28].
- In contrast to Algorithm 3, Theorem 3.5 makes explicit use of the *non-orthogonal Krylov bases* K_m, L_m. These bases are prone to ill-conditioning as m increases; see [3] for the case of a Hermitian matrix A leading to numerical instabilities. Thus, when the degree of the rational function/approximation is rather high, we expect Algorithm 3 to be more accurate in the presence of round-off error.
- Reiterating what we already observed for the classic Sherman–Morrison formula, Algorithm 3 requires two additional conditions not needed in Theorem 3.5: $f(G_m)$, $f(H_m^* + (V_m^*\boldsymbol{b})(V_m^*\boldsymbol{c})^*)$ need to be well defined. Note, however, that these conditions are always met when the numerical ranges of A and $A + \boldsymbol{bc}^*$ do not contain a singularity of f.

Thus, we conclude that although our approach is related to the work in [12], it differs significantly in key aspects and seems to be the preferred approach in many situations of practical interest.

4. Convergence analysis. This section is concerned with the convergence analysis, and its purpose is twofold. We first show how the polynomial case can be treated in an elegant and, compared to our previous work [5], much simpler fashion by using a result from [16]. Unfortunately, it is not clear how this technique extends to the rational case, which will therefore be treated separately in the second part.

In the following, we let

$$E_m(f) := f(A + \boldsymbol{BC}^*) - f(A) - U_m X_m(f) V_m^* \tag{14}$$

denote the error of the approximation returned by Algorithm 1.

4.1. Simpler convergence analysis for polynomial Krylov subspaces. In this section, we consider the case in which $\mathcal{U}_m$ and $\mathcal{V}_m$ are block polynomial Krylov subspaces, and we obtain a convergence result for Algorithm 1 based on polynomial approximation of the derivative of f; see Remark 4.4 below for a comparison with the convergence analysis in [5].

The following lemma is key to our analysis; its proof uses a recent bound on the Fréchet derivative from [16]. We recall that $W(A)$ denotes the numerical range of A and $\|\cdot\|_F$ denotes the Frobenius matrix norm.

LEMMA 4.1. *Let* $\mathcal{B} = \left[\begin{smallmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{smallmatrix}\right]$, *let* $\mathbb{E}$ *be a compact convex set containing* $W(B_{11})$ *and* $W(B_{22})$, *and let* f *be analytic in* $\mathbb{E}$. *Then*

$$\|[f(\mathcal{B})]_{1,2}\|_F \le (1+\sqrt{2})^2 \|f'\|_{\mathbb{E}} \|B_{12}\|_F,$$

where $[f(\mathcal{B})]_{1,2}$ *denotes the* $(1,2)$*-block of* $f(\mathcal{B})$ *and* $\|\cdot\|_{\mathbb{E}}$ *denotes the supremum norm on* $\mathbb{E}$.

Proof. For $n \times n$ matrices A and B, let $L_f(A,B)$ denote the Fréchet derivative of f at A applied to matrix B and let $L_f(A,\cdot)$ denote the corresponding linear operator represented as an $n^2 \times n^2$ matrix. By [30, Theorem 4.12],

$$f(\mathcal{B}) = f(\mathcal{D}) + L_f(\mathcal{D},\mathcal{N}), \text{ where } \mathcal{D} := \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22} \end{bmatrix} \text{ and } \mathcal{N} := \begin{bmatrix} 0 & B_{12} \\ 0 & 0 \end{bmatrix}.$$

Because $f(\mathcal{D})$ is block diagonal, we have that

$$\|[f(\mathcal{B})]_{1,2}\|_F = \|L_f(\mathcal{D},\mathcal{N})\|_F \le \|L_f(\mathcal{D},\cdot)\| \cdot \|B_{12}\|_F.$$

Corollary 5.1 in [16] states that $\|L_f(\mathcal{D},\cdot)\| \le (1+\sqrt{2})^2 \|f'\|_{W(\mathcal{D})}$, which concludes the proof because $W(\mathcal{D})$, as the convex hull of $W(B_{11})$ and $W(B_{22})$, is contained in $\mathbb{E}$. ∎

Lemma 4.1 applied to the matrix $\left[\begin{smallmatrix} A & D \\ 0 & A+D \end{smallmatrix}\right]$ from (3) gives the following result, which might be of independent interest.

COROLLARY 4.2. *Let* $A, D \in \mathbb{C}^{n\times n}$*, let* $\mathbb{E}$ *be a compact convex set containing the union of* $W(A)$ *and* $W(A+D)$*, and let* f *be analytic in* $\mathbb{E}$*. Then*

$$\|f(A+D) - f(A)\|_F \le (1+\sqrt{2})^2 \|f'\|_{\mathbb{E}} \|D\|_F. \tag{15}$$

When A and D are Hermitian, it is well known that the inequality (15) holds without the constant $(1+\sqrt{2})^2$; see, e.g., [44, Proposition 3.1.5]. For general diagonalizable matrices A and $A+D$, Corollary 2.4 in [23] states that

$$\|f(A+D) - f(A)\|_F \le \kappa_A \kappa_{A+D} \max |f'| \cdot \|D\|_F,$$

where κ_A, κ_{A+D} are the condition numbers of any eigenvector matrices of A and $A+D$, respectively. The maximum of $|f'|$ is taken over the convex hull of the spectra of $A+D$ and A. Corollary 4.2 instead holds for any matrix and does not feature the potentially large constant $\kappa_A \kappa_{A+D}$, at the cost of bounding f' on a larger domain $\mathbb{E}$.

We are now prepared to state a convergence result for Algorithm 1 when using block polynomial Krylov subspaces.

THEOREM 4.3. *Let* $A \in \mathbb{C}^{n\times n}$ *and let* f *be analytic in a compact convex set* $\mathbb{E}$ *containing* $W(A)$ *and* $W(A + \boldsymbol{B}\boldsymbol{C}^*)$*. Let* U_m*,* V_m *be orthonormal bases of* $\mathcal{U}_m = \mathcal{K}_m(A, \boldsymbol{B})$*,* $\mathcal{V}_m = \mathcal{K}_m(A^*, \boldsymbol{C})$*. Then the error of Algorithm* 1 *satisfies*

$$\|E_m(f)\|_F \le 2(1+\sqrt{2})^2 \|\boldsymbol{B}\boldsymbol{C}^*\|_F \inf_{p \in \Pi_{m-1}} \|f' - p\|_{\mathbb{E}}.$$

Proof. The first part of the proof is the same as in Theorem 4.2 in [5]: The exactness property [5, Theorem 3.2]—which also holds in the block case—implies that for all $q \in \Pi_m$ we have $E_m(f) = E_m(f-q)$; therefore

$$\begin{aligned} \|E_m(f)\|_F &= \|(f-q)(A+\boldsymbol{B}\boldsymbol{C}^*) - (f-q)(A) - U_m X_m (f-q) V_m^*\|_F \\ &\le \|(f-q)(A+\boldsymbol{B}\boldsymbol{C}^*) - (f-q)(A)\|_F + \|U_m X_m (f-q) V_m^*\|_F \\ &\le \|(f-q)(A+\boldsymbol{B}\boldsymbol{C}^*) - (f-q)(A)\|_F + \|X_m(f-q)\|_F. \end{aligned} \tag{16}$$

Moreover, by definition (line 4 in Algorithm 1), we have $X_m(f-q) = [(f-q)(\tilde{\mathcal{A}})]_{1,2}$, where $\tilde{\mathcal{A}} := \left[\begin{smallmatrix} U_m^* A U_m & U_m^* \boldsymbol{B}\boldsymbol{C}^* V_m \\ 0 & V_m^*(A+\boldsymbol{B}\boldsymbol{C}^*)V_m \end{smallmatrix}\right]$. We can now use Corollary 4.2 to get

$$\|(f-q)(A+\boldsymbol{B}\boldsymbol{C}^*) - (f-q)(A)\|_F \le (1+\sqrt{2})^2 \|(f-q)'\|_{\mathbb{E}} \|\boldsymbol{B}\boldsymbol{C}^*\|_F \tag{17}$$

and Lemma 4.1 to get

$$\|X_m(f-q)\|_F \le (1+\sqrt{2})^2\|(f-q)'\|_{\mathbb{E}}\|U_m^*\boldsymbol{B}\boldsymbol{C}^*V_m\|_F \le (1+\sqrt{2})^2\|(f-q)'\|_{\mathbb{E}}\|\boldsymbol{B}\boldsymbol{C}^*\|_F,$$

because of the inclusions $W(U_m^*AU_m) \subseteq W(A)$ and $W(V_m^*(A+\boldsymbol{B}\boldsymbol{C}^*)V_m) \subseteq W(A+\boldsymbol{B}\boldsymbol{C}^*)$. Combining these with (16) gives the result of the theorem, because $q' \in \Pi_{m-1}$ can be chosen arbitrarily. ∎

Remark 4.4. Let us compare the result of Theorem 4.3 with Theorem 4.2 in [5], which establishes the upper bound $2(1+\sqrt{2})\inf_{p\in\Pi_m}\|f-p\|_{\tilde{\mathbb{E}}}$ for the error in the non-Hermitian case. While this bound features a somewhat smaller constant and the approximation of f instead of f', it comes with the major disadvantage that $\tilde{\mathbb{E}}$ needs to contain the numerical range of $\mathcal{A} = \left[\begin{smallmatrix} A & \boldsymbol{B}\boldsymbol{C}^* \\ 0 & A+\boldsymbol{B}\boldsymbol{C}^* \end{smallmatrix}\right]$, which can be critically larger than the convex hull of $W(A)$ and $W(A+\boldsymbol{B}\boldsymbol{C}^*)$. Indeed, there are situations [5, Figure 6.2] in which $W(\mathcal{A})$ contains a singularity of f (and hence the bound becomes void) but the assumptions of Theorem 4.3 are still satisfied. In order to deal with these situations, specialized techniques had to be developed to address the issue (see [5, section 5]), which can now be bypassed by Theorem 4.3.

4.2. Convergence analysis for rational Krylov subspaces. In this section, we analyze the convergence of the proposed rational Krylov subspace method for updating matrix functions, both in the Hermitian and non-Hermitian cases for certain classes of functions.

4.2.1. Convergence analysis in the Hermitian case. We first discuss the Hermitian case, that is, $A = A^*$ and $D = D^*$. The following theorem links this error to a rational approximation problem. We omit its proof because it follows from Theorem 3.3 in a manner entirely analogous to the proof of Theorem 4.1 in [5].

THEOREM 4.5. *Let A and $D = \boldsymbol{B}J\boldsymbol{B}^*$ be Hermitian, let the set of poles be closed under complex conjugation, and let $U_m = V_m$ be an orthonormal basis of $q_m(A)^{-1}\mathcal{K}_m(A,\boldsymbol{B})$. Furthermore, let f be analytic in a domain $\mathbb{E}$ containing the union of $W(A)$ and $W(A+D)$. Then the error* (14) *returned by Algorithm* 3 *satisfies*

$$\|E_m(f)\| \le 4 \min_{r\in\Pi_m/q_m} \|f-r\|_{\mathbb{E}}, \tag{18}$$

where $\|\cdot\|_{\mathbb{E}}$ denotes the supremum norm on $\mathbb{E}$.

Theorem 4.5 allows us to derive convergence bounds for Algorithm 3 by considering rational uniform approximation problems on intervals $\mathbb{E}$ containing $[\widetilde{\lambda}_{\min}, \widetilde{\lambda}_{\max}]$, where

$$\widetilde{\lambda}_{\min} := \min\{\lambda_{\min}(A), \lambda_{\min}(A+D)\}, \quad \widetilde{\lambda}_{\max} := \max\{\lambda_{\max}(A), \lambda_{\max}(A+D)\}.$$

This problem has been addressed numerous times in the literature, e.g., in the context of analyzing rational Krylov subspace methods for approximating $f(A)b$; see, e.g., [6, 26, 27] and the references therein. In the following, we give several examples for the bounds obtained this way.

The exponential function. Under the assumptions of Theorem 4.5, consider the exponential function $f(z) = \exp(x)$. We will suppose in the following that the spectra of A and $A+D$ (and the corresponding poles) have already been shifted[1] such

[1] Such a shift would lead to an additional factor $\exp(\widetilde{\lambda}_{\max})$ in (18).

that A and $A + D$ are negative semidefinite, and thus one can choose $\mathbb{E} = (-\infty, 0]$ in Theorem 4.5.

From the seminal work of Gonchar and Rakhmanov [25] and its improvements established by Aptekarev [2] it is known that for every integer m there exists an optimal denominator $q_m \in \Pi_m$ such that

$$\min_{r \in \Pi_m / q_m} \| \exp - r \|_{(-\infty,0]} \leq C \, \kappa^{-m}, \quad \kappa \approx 9.28903\ldots,$$

for some constant C independent of m. The numerical values of the optimal poles (that is, the roots of q_m) are known.

We now consider the case of a single, repeated pole, which bears the advantage that only one sparse factorization needs to be computed when using a direct solver in Algorithm 3. Andersson [1] showed that, for $q_m(z) = (z - m/\sqrt{2})^m$,

$$\limsup_{m \to \infty} \Big(\min_{r \in \Pi_m / q_m} \| \exp - r \|_{L^\infty((-\infty,0])} \Big)^{1/m} = \frac{1}{1 + \sqrt{2}}.$$

This agrees with observations from [34, 46] that a well-chosen single pole ξ repeated m times already yields good convergence.

Strategies for choosing poles (adaptively) for finite intervals are surveyed in [27, section 4.2].

Markov functions. Under the assumptions of Theorem 4.5, let us now consider a *Markov function*

$$f(x) = \int_\alpha^\beta \frac{\mathrm{d}\mu(z)}{x - z}, \tag{19}$$

where μ is a positive measure with support in the interval $[\alpha, \beta]$ with $-\infty \leq \alpha < \beta < \infty$. Important examples of Markov functions are inverse fractional powers

$$f(z) = z^{-\gamma} = \frac{\sin(\gamma\pi)}{\pi} \int_{-\infty}^0 \frac{(-x)^{-\gamma} \, \mathrm{d}x}{z - x} \tag{20}$$

for $\gamma \in (0, 1)$, or

$$f(z) = \frac{1}{z} \log(1 + z) = \int_{-\infty}^{-1} \frac{(-1/x) \, \mathrm{d}x}{z - x}. \tag{21}$$

For more details on Markov functions and further examples, we refer the reader to [10, 29]. A detailed discussion of rational approximation of Markov functions can be found in [6, section 6]. From [6, Theorem 6.1(b)] we quote the following estimate.

THEOREM 4.6. *Let $\mathbb{E}$ be a compact convex set, symmetric with respect to the real axis, and let f be a Markov function* (19) *such that*

$$\beta < \omega := \min \mathbb{E} \cap \mathbb{R}. \tag{22}$$

Let ψ denote the conformal map from $\overline{\mathbb{C}} \setminus \mathbb{D}$ onto $\overline{\mathbb{C}} \setminus \mathbb{E}$ normalized such that $\psi(\infty) = \infty, \psi'(\infty) > 0$, where $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ denotes the extended complex plane and $\mathbb{D}$ denotes the closed unit disk, and let ϕ denote its inverse map from $\overline{\mathbb{C}} \setminus \mathbb{E}$ onto $\overline{\mathbb{C}} \setminus \mathbb{D}$. Then

$$\min_{r \in \Pi_m / q_m} \| f - r \|_{L^\infty(\mathbb{E})} \leq \frac{2 \| f \|_{L^\infty(\mathbb{E})}}{|\phi(\beta)|} \cdot \eta_m, \quad \eta_m := \max_{x \in [\phi(\alpha), \phi(\beta)]} \frac{1}{|B_m(x)|},$$

with the Blaschke product

$$B_m(x) := \prod_{j=1}^{m} \frac{1 - x\overline{\phi(\xi_j)}}{x - \phi(\xi_j)}.$$

For estimating $\|E_m(f)\|$ for Markov functions f, we may therefore combine Theorem 4.5 with Theorem 4.6 for $\mathbb{E} = [\widetilde{\lambda}_{\min}, \widetilde{\lambda}_{\max}]$, as long as $\beta < \omega = \widetilde{\lambda}_{\min}$. In this case, explicit formulas for the conformal maps ϕ, ψ are available. Noting that only the convergence factor η_m depends on the poles $\xi_1, \ldots, \xi_m$, it remains to derive upper bounds on η_m for particular choices of poles.

According to [6, Corollary 6.4], we may minimize η_m among all single, repeated poles $\xi = \xi_1 = \cdots = \xi_m$ by setting

$$\sigma = \frac{\phi(\beta) - \phi(\alpha)}{\phi(\beta)\phi(\alpha) - 1}, \quad y_{\text{opt}} = -\frac{1}{\sigma} - \sqrt{\frac{1}{\sigma^2} - 1}, \quad w = \frac{1 + \phi(\alpha) y_{\text{opt}}}{\phi(\alpha) + y_{\text{opt}}},$$

resulting in the optimal pole $\xi = \psi(w)$ and $\eta_m = |y_{\text{opt}}|^{-m}$.

In the important special case $\alpha = -\infty, \beta = 0$, which occurs, e.g., for inverse fractional powers (20), the above formulas simplify and we obtain the pole $\xi = -\sqrt{\widetilde{\lambda}_{\max} \cdot \widetilde{\lambda}_{\min}}$ and the corresponding convergence rate

$$\eta_m = \left(\frac{\sqrt[4]{\widetilde{\lambda}_{\max}/\widetilde{\lambda}_{\min}} - 1}{\sqrt[4]{\widetilde{\lambda}_{\max}/\widetilde{\lambda}_{\min}} + 1} \right)^m. \tag{23}$$

Let us note that, asymptotically, the convergence rate (23) is also attained when alternatingly choosing the poles 0 and ∞, i.e., when using *extended Krylov subspaces* [6, 33].

Example 4.7. We illustrate the above results by a simple numerical experiment, using a diagonal matrix $A \in \mathbb{C}^{200\times 200}$ with logarithmically spaced eigenvalues in the interval $[10^{-3}, 10^3]$ and $D = \boldsymbol{b}\boldsymbol{b}^*$, where $\boldsymbol{b}$ is a random vector with $\|\boldsymbol{b}\| = 100$. This leads to $\widetilde{\lambda}_{max} \approx 1.0078 \cdot 10^4$, and thus $\widetilde{\lambda}_{max}/\widetilde{\lambda}_{min}) \approx 1.0078 \cdot 10^7$. Figure 1 displays the convergence of Algorithm 3 with all poles equal to $-\sqrt{\widetilde{\lambda}_{max} \cdot \widetilde{\lambda}_{min}}$ for approximating $(A + \boldsymbol{b}\boldsymbol{b}^*)^{-1/2} - A^{-1/2}$. In the initial phase, the error reduces linearly and the convergence rate of the method is predicted quite accurately by (23). The superlinear convergence phase starting around iteration 120 can of course not be captured by (23). ◇

We now turn to rational approximations using several different poles. In [6, section 6.2], quasi-optimal poles are constructed that admit closed formulas in terms of Jacobi elliptic functions. Using these poles,

$$\eta_m \le 2 \exp\left(-m \frac{\pi^2}{\log(16\widetilde{\lambda}_{\max}/\widetilde{\lambda}_{\min})} \right). \tag{24}$$

Thus, the rate of convergence now depends on the *logarithm* of the ratio $\widetilde{\lambda}_{\max}/\widetilde{\lambda}_{\min}$ instead of the fourth root. The corresponding poles are mutually distinct and, in turn, the rational Arnoldi method requires the computation of a new Cholesky decomposition in each of the m iterations. As already mentioned in section 3, it is preferable

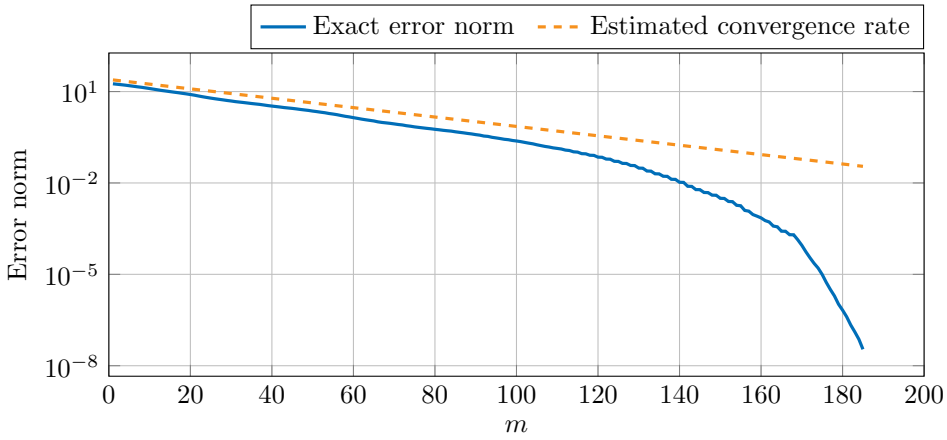

FIG. 1. *Convergence of* $\|E_m(f)\|$ *for Algorithm* 3 *with a single, repeated, asymptotically optimal pole, and estimated convergence rate* (23) *for approximating* $(A+\boldsymbol{b}\boldsymbol{b}^*)^{-1/2}-A^{-1/2}$ *with* $A, \boldsymbol{b}$ *as in Example* 4.7.

in practice to use a smaller number of poles and repeat them (typically cyclically) in order to limit the number of matrix factorizations that need to be computed. When using $\widetilde{m}$ quasi-optimal poles and repeating each of them k times, the error bound (24) changes to

$$\eta_m \le 2^k \exp\left(-k\widetilde{m}\frac{\pi^2}{\log(16\widetilde{\lambda}_{\max}/\widetilde{\lambda}_{\min})}\right), \quad m = k\widetilde{m}. \tag{25}$$

Thus, compared to using m (mutually distinct) quasi-optimal poles, the error bound worsens by a factor of 2^{k-1}.

We repeat the experiment from Example 4.7, now using ten cyclically repeated, quasi-optimal poles in Leja ordering [37]. Figure 2 displays the resulting convergence. The overall convergence rate is again predicted quite accurately, although the actual convergence curve shows a staircase-like behavior (which is typical for rational Krylov methods with poles in Leja ordering).

Other, practically relevant functions like the *matrix square root* are obtained as slight modifications of Markov functions.

Example 4.8. Let us consider functions of the form

$$f(z) = z\widehat{f}(z), \tag{26}$$

where $\widehat{f}$ is a Markov function (19). This includes the square root $z^{1/2} = zz^{-1/2}$ as well as the logarithm $\log(1+z) = z\frac{\log(1+z)}{z}$. The following simple trick allows us to apply Theorem 4.6 to this setting. Fixing the pole $\xi_m = \infty$, which gives $q_m = q_{m-1} \in \Pi_{m-1}$, and setting $p_1(z) = z$, we obtain

$$\begin{aligned}\min_{r\in\Pi_m/q_m} \|f-r\|_{L^\infty(\mathbb{E})} &\le \min_{r\in\Pi_{m-1}/q_m} \|\widehat{f}-r\|_{L^\infty(\mathbb{E})}\,\|p_1\|_{L^\infty(\mathbb{E})} \\ &= \min_{r\in\Pi_{m-1}/q_{m-1}} \|\widehat{f}-r\|_{L^\infty(\mathbb{E})}\,\|p_1\|_{L^\infty(\mathbb{E})}.\end{aligned}$$

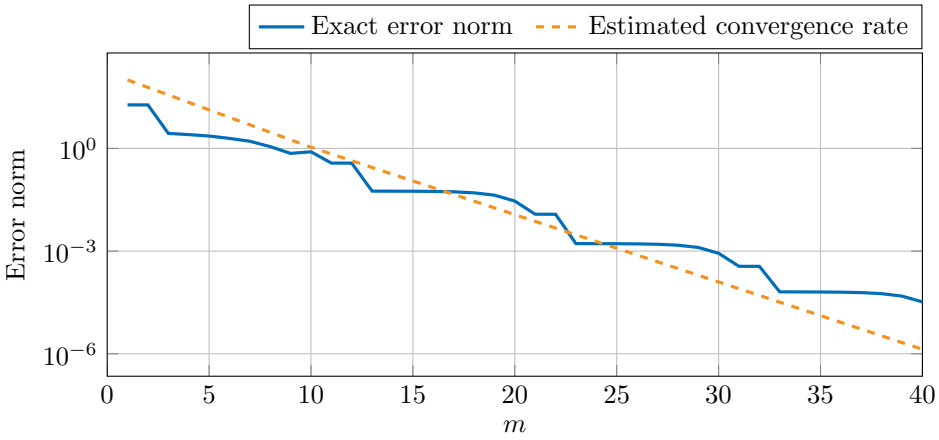

FIG. 2. *Convergence of* $\|E_m(f)\|$ *for Algorithm* 3 *with* 10 *quasi-optimal, cyclically repeated poles, and estimated convergence rate* (25) *for approximating* $(A+\boldsymbol{b}\boldsymbol{b}^*)^{-1/2}-A^{-1/2}$ *with* $A, \boldsymbol{b}$ *as in Example* 4.7.

That is, besides the additional factor $\|p_1\|_{L^\infty(\mathbb{E})}$, we obtain an upper bound for $E_m(f)$ by combining Theorem 4.5 for m, f with Theorem 4.6 for $m-1, \widehat{f}$. A similar technique has been used in [22] in the context of convergence theory for restarted (polynomial) Krylov methods for $f(A)\boldsymbol{b}$ when A is Hermitian positive definite. In that situation, $\|p_1\|_{L^\infty(\mathbb{E})} = \lambda_{\max}$. ◇

4.2.2. Convergence analysis for Markov functions in the non-Hermitian case. We now turn to the more difficult task of analyzing the convergence for general A, $D = \boldsymbol{B}\boldsymbol{C}^\star$ in terms of a convex and compact set $\mathbb{E}$ containing both numerical ranges $W(A)$ and $W(A+\boldsymbol{B}\boldsymbol{C}^*)$, and f being analytic in $\mathbb{E}$. In principle, Theorem 4.3 also holds for rational Krylov subspaces by replacing p with the derivative of a function in Π_m/q_m. However, due to the special form of such a derivative, the resulting optimization problem appears to be too exotic to be of assistance in getting practical convergence bounds. Therefore, inspired by [5, section 5.1], we consider the shifted (block) linear systems

$$(zI-A)\boldsymbol{X}(z)=\boldsymbol{B} \quad \text{and} \quad (zI-A-\boldsymbol{B}\boldsymbol{C}^*)^*\,\boldsymbol{Y}(z)=\boldsymbol{C}, \tag{27}$$

together with the rational block FOM (full orthogonalization method) approximations for (27), given by

$$\begin{aligned}\boldsymbol{X}_m(z) &:= U_m(zI-G_m)^{-1}U_m^*\boldsymbol{B},\\ \boldsymbol{Y}_m(z) &:= V_m(\bar{z}I-H_m-V_m^*\boldsymbol{C}\boldsymbol{B}^*V_m)^{-1}V_m^*\boldsymbol{C}.\end{aligned}$$

The following result links these quantities to the approximation error of low-rank updates.

LEMMA 4.9. *With* Γ *a contour surrounding* $\mathbb{E}$ *once and sufficiently close to* $\mathbb{E}$, *the error defined in* (14) *satisfies*

$$E_m(f)=\frac{1}{2\pi i}\int_\Gamma f(z)\big(\boldsymbol{X}(z)\,\boldsymbol{Y}(z)^*-\boldsymbol{X}_m(z)\,\boldsymbol{Y}_m(z)^*\big)\,\mathrm{d}z.$$

Proof. This result has been derived in [5, section 5.1] in the context of polynomial Krylov subspaces, but it is straightforward to verify that the derivations are valid for general choices of subspaces. ∎

Lemma 4.9 shows that $\|E_m(f)\|$ is small if the rational FOM approximation errors $\boldsymbol{X}(z) - \boldsymbol{X}_m(z)$ and $\boldsymbol{Y}(z) - \boldsymbol{Y}_m(z)$ are small, uniformly for $z \in \Gamma$. The analysis is complicated by this dependence on Γ. Therefore, in what follows we will only consider the particular case (19) of a Markov function f, which allows us to switch from Γ to the interval $[\alpha, \beta]$.

THEOREM 4.10. *Let $\mathbb{E}$ be a convex and compact set, symmetric with respect to the real axis, and containing both numerical ranges $W(A)$ and $W(A + \boldsymbol{B}\boldsymbol{C}^*)$. Let ω and η_m be defined as in Theorem* 4.6. *Then for a Markov function f satisfying* (22)*, the error* (14) *returned by Algorithm* 3 *satisfies*

$$\|E_m(f)\| \le 8\,|f'(\omega)|\,\frac{\eta_m}{1-\eta_m}\,\|\boldsymbol{B}\|\,\|\boldsymbol{C}\|.$$

Proof. In the same way as in the proof of [5, Theorem 5.7], we obtain from Lemma 4.9 and the Fubini theorem the bound

$$\|E_m(f)\| \le \int_\alpha^\beta (\|\boldsymbol{X}(t)\|\|\boldsymbol{Y}(t) - \boldsymbol{Y}_m(t)\| + \|\boldsymbol{Y}_m(t)\|\|\boldsymbol{X}(t) - \boldsymbol{X}_m(t))\|)\,\mathrm{d}\mu(t). \tag{28}$$

We have

$$\|(tI - A)^{-1}\| \le \frac{1}{\operatorname{dist}(t, W(A))} \le \frac{1}{\operatorname{dist}(t, \mathbb{E})} \le \frac{1}{\omega - t}, \tag{29}$$

where the last inequality follows for all $t \in [\alpha, \beta]$ from condition (22). Analogously,

$$\|(tI - H_m - V_m^*\boldsymbol{C}\boldsymbol{B}^*V_m)^{-1}\| \le \frac{1}{\operatorname{dist}(t, W(A + \boldsymbol{B}\boldsymbol{C}^*))} \le \frac{1}{\omega - t}.$$

In particular, these bounds imply

$$\frac{\|\boldsymbol{X}(t)\|}{\|\boldsymbol{B}\|} \le \frac{1}{\operatorname{dist}(t, \mathbb{E})} = \frac{1}{\omega - t}, \quad \frac{\|\boldsymbol{Y}_m(t)\|}{\|\boldsymbol{C}\|} \le \frac{1}{\omega - t}$$

for all $t \in [\alpha, \beta]$. We claim that, for $t \in [\alpha, \beta]$,

$$\frac{\|\boldsymbol{X}(t) - \boldsymbol{X}_m(t)\|}{\|\boldsymbol{B}\|} \le \frac{4}{\omega - t} \cdot \frac{\eta_m}{1 - \eta_m}, \quad \frac{\|\boldsymbol{Y}(t) - \boldsymbol{Y}_m(t)\|}{\|\boldsymbol{C}\|} \le \frac{4}{\omega - t} \cdot \frac{\eta_m}{1 - \eta_m}. \tag{30}$$

Inserting these bounds into (28) leads to

$$\|E_m(f)\| \le 8\|\boldsymbol{B}\|\|\boldsymbol{C}\|\frac{\eta_m}{1-\eta_m} \int \frac{d\mu(t)}{(\omega - t)^2},$$

with the integral being equal to $|f'(\omega)| = \|f'\|_{L^\infty(\mathbb{E})}$. Hence, we arrive at the assertion of the theorem.

It remains to show the first inequality of (30); the proof of the second is entirely analogous. Theorem 3.4 in [4] establishes the existence of a rational function $R \in \Pi_m/q_m$ depending only on q_m and $\mathbb{E}$ such that

$$\begin{aligned} &\|R(\widetilde{A})\| \le 2 \quad \text{for all square matrices } \widetilde{A} \text{ with } W(\widetilde{A}) \subset \mathbb{E}, \\ &|R(z)| \le 2 \quad \text{for all } z \in \mathbb{E}, \\ &|R(t)| \ge |B_m(\phi(t))| - 1 \text{ for all } t \notin \mathbb{E}. \end{aligned} \tag{31}$$

Let $t \in [\alpha, \beta]$ be fixed, and consider the rational function

$$z \mapsto r_t(z) = \frac{1}{z-t} - \frac{1}{z-t} \cdot \frac{R(z)}{R(t)}.$$

Since $r_t \in \Pi_{m-1}/q_m$, the exactness property of Lemma 3.2 allows us to conclude that $r_t(A)\boldsymbol{B} = U_m r_t(G_m) U_m^* \boldsymbol{B}$, and thus

$$\boldsymbol{X}(t) - \boldsymbol{X}_m(t) = (tI - A)^{-1} \frac{R(A)}{R(t)} \boldsymbol{B} - U_m (tI - G_m)^{-1} \frac{R(G_m)}{R(t)} U_m^* \boldsymbol{B}. \tag{32}$$

Using the properties of R from (31) and the bound (29), we have

$$\left\| (tI - A)^{-1} \frac{R(A)}{R(t)} \boldsymbol{B} \right\| \leq \frac{\|\boldsymbol{B}\|}{\omega - t} \frac{\|R(A)\|}{|R(t)|} \leq \frac{2\|\boldsymbol{B}\|}{\omega - t} \frac{\eta_m}{1 - \eta_m}$$

and the same upper bound if one replaces $\boldsymbol{B}$ and A by $U_m^* \boldsymbol{B}$ and G_m, respectively. Inserting these bounds into (32) shows the claim (30) and completes the proof. ∎

Remark 4.11. For polynomial Krylov subspaces, $\xi_1 = \cdots = \xi_m = \infty$. In turn, $B_m(\phi(x)) = \phi(x)^m$ and $\eta_m = 1/|\phi(\beta)|^m$. Thus, up to the factor $1/(1 - \eta_m)$, our Theorem 4.10 reduces to [5, Theorem 5.7]. We mention in passing that this factor can be removed, using the techniques of [5, Lemma 5.1], if at least two of the poles $\xi_1, \ldots, \xi_m$ are infinite. We should also mention that, once a suitable set $\mathbb{E}$ with more explicit conformal map ϕ (as, for instance, an ellipse or a teardrop set) is found, we may use some of the estimates for η_m in terms of ϕ, α, β, as stated in section 4.2.1.

5. The matrix sign function. When the numerical range of A or $A + \boldsymbol{B}\boldsymbol{C}^*$ contains a singularity of f, none of the convergence results from section 4 applies. For the matrix sign function, a notorious example for this situation, we discuss a potential remedy.

Letting

$$\operatorname{sign} : \mathbb{C} \setminus \mathrm{i}\mathbb{R} \to \mathbb{C}, \qquad \operatorname{sign}(z) = \begin{cases} -1, & \operatorname{Re}(z) < 0, \\ 1, & \operatorname{Re}(z) > 0, \end{cases}$$

where $\operatorname{Re}(z)$ denotes the real part of z, the matrix sign function $\operatorname{sign}(A)$ is defined whenever A has no purely imaginary eigenvalue. This function plays an important role in, e.g., linear-quadratic optimal control [38], quantum chromodynamics [14, 20], and eigenvalue solvers [13, 35].

5.1. Low-rank updates. Except for trivial situations ($\operatorname{sign}(A) = \pm I$), the sign function is usually *not* defined on the numerical range of $W(A)$, which poses a severe problem for Krylov subspace techniques, not only in theory but also in practice. In the context of approximating $\operatorname{sign}(A)\boldsymbol{b}$, Krylov subspace methods have been observed to exhibit slow, irregular or erratic convergence [20]. As a remedy, it has been proposed to exploit the relation

$$\operatorname{sign}(A) = (A^2)^{-1/2} A \tag{33}$$

and approximate $\operatorname{sign}(A)\boldsymbol{b}$ in the Krylov space $\mathcal{K}_m(A^2, A\boldsymbol{b})$; see, e.g., [15, 20]. For an invertible Hermitian matrix A the advantage of (33) is obviously that the numerical range of A^2 does not contain a singularity of the inverse square root.

In the following, we will discuss an approach based on (33) for approximating low-rank updates (1) of the matrix sign function. Because (33) offers a clear advantage

only for the Hermitian case, we now assume that $A = A^*$ and $D = \boldsymbol{B}J\boldsymbol{B}^*$ with $J = J^*$. Let us mention, however, that the construction readily extends to the non-Hermitian case.

Using (33), it follows that

$$\begin{aligned}\operatorname{sign}(A+D) - \operatorname{sign}(A) &= (A+D)((A+D)^2)^{-1/2} - A(A^2)^{-1/2} \\ &= (A+D)\big((A^2+\widetilde{D})^{-1/2} - (A^2)^{-1/2}\big) + \boldsymbol{B}J\boldsymbol{B}^*(A^2)^{-1/2} \end{aligned} \tag{34}$$

with $\widetilde{D} := A\boldsymbol{B}J\boldsymbol{B}^* + \boldsymbol{B}J\boldsymbol{B}^*(A + \boldsymbol{B}J\boldsymbol{B}^*)$. A rank-$\ell$ update of the sign function is thus performed by computing a rank-2ℓ update of $(A^2)^{-1/2}$ and the action of $(A^2)^{-1/2}$ on $\boldsymbol{B}$. Because the range and co-range of $\widetilde{D}$ are contained in the span of $[\boldsymbol{B}, A\boldsymbol{B}]$, it is natural to choose the rational Krylov subspace

$$\mathcal{U}_m := q_m(A^2)^{-1}\mathcal{K}_m(A^2, [\boldsymbol{B}, A\boldsymbol{B}]) \tag{35}$$

with suitably chosen poles $\xi_1, \dots, \xi_m$ for approximating the rank-2ℓ update. To approximate the second term in (34), we utilize the usual block Krylov approximation

$$(A^2)^{-1/2}\boldsymbol{B} \approx U_m G_m^{-1/2} U_m^* \boldsymbol{B}$$

for an orthonormal basis U_m of $\mathcal{U}_m$. Algorithm 4 summarizes the described approach for approximating (34).

Algorithm 4 Rational block Krylov subspace approximation of sign matrix function update (34) for Hermitian A, D.

1: Choose poles $\xi_1, \dots, \xi_m \in \mathbb{C} \cup \{\infty\}$ closed under complex conjugation.
2: Perform m steps of Algorithm 2 to compute an orthonormal basis U_m of $\mathcal{U}_m = q_m(A^2)^{-1}\mathcal{K}_m(A^2, [\boldsymbol{B}, A\boldsymbol{B}])$ and set $G_m = U_m^* A^2 U_m$.
3: Set $X_m(z^{-1/2}) = \big(G_m + U_m^*\big(A\boldsymbol{B}J\boldsymbol{B}^* + \boldsymbol{B}J\boldsymbol{B}(A + \boldsymbol{B}J\boldsymbol{B}^*)\big)U_m\big)^{-1/2} - G_m^{-1/2}$.
4: Compute $\boldsymbol{f}_m = U_m G_m^{-1/2} U_m^* \boldsymbol{B}$.
5: Return $(A + \boldsymbol{B}J\boldsymbol{B}^*)(U_m X_m(z^{-1/2}) U_m^*) + \boldsymbol{B}J\boldsymbol{f}_m^*$.

Remark 5.1. The rational Krylov space (35) used in Algorithm 4 has a very specific structure, and its polynomial part is actually identical to an ordinary block Krylov space of order $2m$ for A. Precisely

$$\mathcal{U}_m = q_m(A^2)^{-1}\mathcal{K}_m(A^2, [\boldsymbol{B}, A\boldsymbol{B}]) = q_m(A^2)^{-1}\mathcal{K}_{2m}(A, \boldsymbol{B}). \tag{36}$$

This is different from the situation arising when approximating $\operatorname{sign}(A)\boldsymbol{b}$, where the polynomial part of the subspace corresponds only to odd powers of A. When $\boldsymbol{B}$ is a vector, this observation could in principle be used to implement Algorithm 4 such that it avoids block arithmetic.

The convergence of Algorithm 4 can be analyzed by combining the results from section 4 with known convergence results for Krylov subspace methods.

THEOREM 5.2. *Let A and $D = \boldsymbol{B}J\boldsymbol{B}^*$ be Hermitian such that A and $A + D$ are invertible. Then the error of the approximation returned by Algorithm 4 satisfies*

$$\begin{aligned}&\|\operatorname{sign}(A+D) - \operatorname{sign}(A) - (A+D)(U_m X_m(z^{-1/2})U_m^*) + \boldsymbol{B}J\boldsymbol{f}_m^*\| \\ &\quad\le (4\|A+D\| + 2\|\boldsymbol{B}J\|\,\|\boldsymbol{B}\|) \min_{r \in \Pi_m/q_m} \|f - r\|_{\mathbb{E}},\end{aligned} \tag{37}$$

where $\mathbb{E} = \big[\min\{\lambda_{\min}(A^2), \lambda_{\min}((A+D)^2)\}, \max\{\lambda_{\max}(A^2), \lambda_{\max}((A+D)^2)\}\big]$ *and* $f(z) = z^{-1/2}$.

Proof. Using (34) and setting $M = (A^2 + \widetilde{D})^{-1/2} - (A^2)^{-1/2}$, it follows that (37) is bounded by

$$\begin{aligned}&\|(A+D)\big(M - U_m X_m(f) U_m^*\big) + \boldsymbol{B}J\big(\boldsymbol{B}^*(A^2)^{-1/2} - \boldsymbol{f}_m^*\big)\| \\ &\le \|A+D\|\,\|M - U_m X_m(f) U_m^*\| + \|\boldsymbol{B}J\|\,\|\boldsymbol{B}^*(A^2)^{-1/2} - \boldsymbol{f}_m^*\|.\end{aligned}$$

Using Theorem 4.5, the first term is bounded via

$$\|M - U_m X_m(f) U_m^*\| \le 4 \min_{r \in \Pi_m/q_m} \|f - r\|_{\mathbb{E}}. \tag{38}$$

For the second term, we can estimate

$$\|(A^2)^{-1/2}\boldsymbol{B} - \boldsymbol{f}_m^*\| \le 2\|\boldsymbol{B}\| \min_{r \in \Pi_m/q_m} \|f - r\|_{\widetilde{\mathbb{E}}} \le 2\|\boldsymbol{B}\| \min_{r \in \Pi_m/q_m} \|f - r\|_{\mathbb{E}} \tag{39}$$

with $\widetilde{\mathbb{E}} = [\lambda_{\min}(A^2), \lambda_{\max}(A^2)] \subseteq \mathbb{E}$. For the case that $\boldsymbol{B}$ is a vector, (39) is shown in [26, Theorem 4.10] (see also the proof of [6, Theorem 5.2]), and the proof of this result carries over to the block case (and the nonstandard rational Krylov space that we are using) completely analogously, using the exactness property from Lemma 3.2 as a basis. Further note that the estimate (39) is actually valid for the smaller subspace $q_m(A^2)^{-1}\mathcal{K}_m(A^2, \boldsymbol{B}) \subseteq q_m(A^2)^{-1}\mathcal{K}_m(A^2, [\boldsymbol{B}, A\boldsymbol{B}])$. Combining (38) and (39) gives the desired result. ∎

As $f(z) = z^{-1/2}$ is a Markov function, we can, e.g., apply Theorem 4.6 to obtain bounds for $\|f - r\|_{\mathbb{E}}$ in Theorem 5.2.

Example 5.3. Consider the diagonal, indefinite matrix $A \in \mathbb{C}^{200\times 200}$ with 100 linearly spaced eigenvalues in each of the intervals $[-1, -10^{-2}]$ and $[10^{-2}, 1]$. Let $\boldsymbol{b} \in \mathbb{C}^{200}$ be a random vector of unit norm. We compare Algorithm 4 to the straightforward application of Algorithm 3 to perform the update $\operatorname{sign}(A + \boldsymbol{b}\boldsymbol{b}^*) - \operatorname{sign}(A)$. We use the poles of the Zolotarev approximation of degrees 2 and 10 for the inverse square root in Algorithm 4 and the poles of the corresponding Zolotarev approximation of the sign function in Algorithm 3; see [36, 47]. Again, the poles are in Leja ordering and cyclically repeated. The resulting convergence curves are depicted in Figure 3. As expected, the convergence curve of Algorithm 4 is much smoother than that of Algorithm 3. In addition, the subspace dimension required to reach the target accuracy 10^{-6} by Algorithm 4 is smaller: When using 10 different poles, it needs 24 vs. 34 iterations, i.e., a reduction of about 30%. For only 2 different poles, the difference becomes a lot more pronounced, and Algorithm 4 requires 44 iterations, while Algorithm 3 fails to converge in a reasonable number of iterations.

Concerning the computation cost of the algorithms, several things have to be taken into account: On the one hand, the number of nonzeros in A^2 is typically larger than in A, which leads to higher expenses when factoring $A^2 + \xi_i I$. On the other hand, the poles of the Zolotarev approximation for the sign function are complex, so that Algorithm 3 requires complex arithmetic even though A and $\boldsymbol{b}$ are real (note, however, that only half the number of Cholesky factorizations needs to be computed, as the Zolotarev shifts come in complex conjugate pairs). ◇

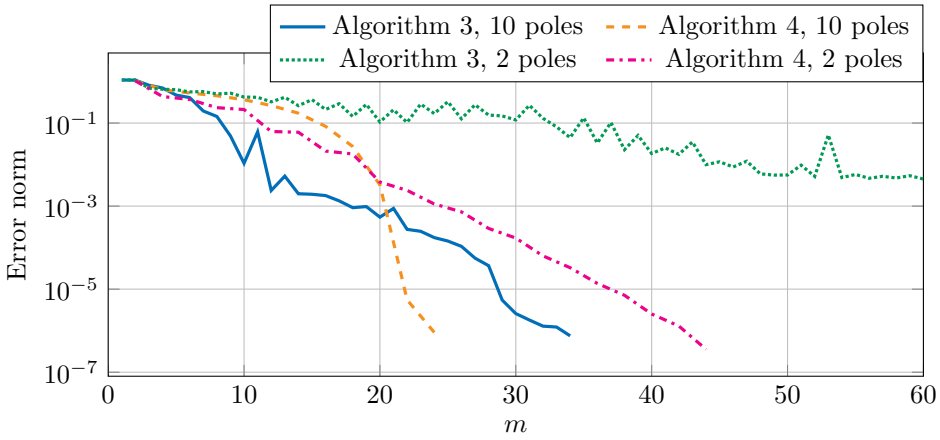

FIG. 3. *Convergence curves of Algorithms* 3 *and* 4 *using the poles of a Zolotarev approximation of degree* 2 *or* 10 *for approximating* $\operatorname{sign}(A + \boldsymbol{b}\boldsymbol{b}^*) - \operatorname{sign}(A)$, *where* $\Lambda(A) \subseteq [-1, -10^{-2}] \cup [10^{-2}, 1]$, $\|\boldsymbol{b}\| = 1$.

5.2. Connection to Krylov subspace methods for linear matrix equations. We conclude this work by pointing out a curious connection to Krylov subspace methods for the matrix Sylvester equation

$$A_1 Z - Z A_2 + \boldsymbol{B}_1 \boldsymbol{C}_2^* = 0, \tag{40}$$

with coefficients $A_1 \in \mathbb{C}^{n_1 \times n_1}$, $A_2 \in \mathbb{C}^{n_2 \times n_2}$ and $\boldsymbol{B}_1 \in \mathbb{C}^{n_1 \times \ell}$, $\boldsymbol{C}_2 \in \mathbb{C}^{n_2 \times \ell}$ such that $\ell \ll \min\{n_1, n_2\}$. We refer the reader to [43] for an overview of applications and numerical algorithms for this and similar equations.

We assume that $W(A_1), W(-A_2)$ are contained in the open right-half plane, which implies that (40) has a unique solution Z. Moreover, it is well known that

$$\operatorname{sign}\left(\begin{bmatrix} A_1 & \boldsymbol{B}_1 \boldsymbol{C}_2^* \\ 0 & A_2 \end{bmatrix}\right) = \begin{bmatrix} I_{n_1} & 2Z \\ 0 & -I_{n_2} \end{bmatrix}.$$

In turn,

$$\operatorname{sign}\left(\begin{bmatrix} A_1 & \boldsymbol{B}_1 \boldsymbol{C}_2^* \\ 0 & A_2 \end{bmatrix}\right) - \operatorname{sign}\left(\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}\right) = \begin{bmatrix} 0 & 2Z \\ 0 & 0 \end{bmatrix}, \tag{41}$$

showing that the solution Z of (40) can be obtained from a rank-ℓ update of the matrix sign function. After setting

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_1 \\ 0 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 0 \\ \boldsymbol{C}_2 \end{bmatrix},$$

the left-hand side of (41) takes the familiar form $\operatorname{sign}(A + \boldsymbol{B}\boldsymbol{C}^*) - \operatorname{sign}(A)$.

As we will see below, the particular structure of the update implies that the squaring trick from the previous subsection is not needed for (41). Applying Algorithm 3 directly to (41) involves the rational Krylov subspaces

$$q_m(A)^{-1}\mathcal{K}_m(A, \boldsymbol{B}) = \left\{ \begin{bmatrix} u \\ 0 \end{bmatrix} : u \in q_m(A_1)^{-1}\mathcal{K}_m(A_1, \boldsymbol{B}_1) \right\},$$
$$\bar{q}_m(A)^{-*}\mathcal{K}_m(A^*, \boldsymbol{C}) = \left\{ \begin{bmatrix} 0 \\ v \end{bmatrix} : v \in q_m(A_2)^{-*}\mathcal{K}_m(A_2^*, \boldsymbol{C}_2) \right\}.$$

Thus, we obtain orthonormal bases $U_m = \left[\begin{smallmatrix} U_{1,m} \\ 0 \end{smallmatrix}\right]$, $V_m = \left[\begin{smallmatrix} 0 \\ V_{2,m} \end{smallmatrix}\right]$ by letting $U_{1,m}$ and $V_{2,m}$ contain orthonormal bases of $q_m(A_1)^{-1}\mathcal{K}_m(A_1, \boldsymbol{B}_1)$ and $q_m(A_2)^{-*}\mathcal{K}_m(A_2^*, \boldsymbol{C}_2)$, respectively. The compressions of A and A^* to these bases take the form

$$G_m := U_m^* A U_m = U_{1,m}^* A_1 U_{1,m}, \quad H_m := V_m^* A^* V_m = V_{2,m}^* A_2^* V_{2,m}.$$

We recall that the matrix $X_m(\mathrm{sign})$ in Algorithm 3 is extracted from the (1,2)-block of the matrix (2). In the described setting, this matrix takes the form

$$\operatorname{sign}\left(\begin{bmatrix} G_m & U_{1,m}^* \boldsymbol{B}_1 \boldsymbol{C}_2^* V_{2,m} \\ 0 & H_m^* \end{bmatrix}\right) = \begin{bmatrix} I & 2\tilde{Z}_m \\ 0 & -I \end{bmatrix},$$

where $\tilde{Z}_m$ satisfies the Sylvester equation $G_m \tilde{Z}_m - \tilde{Z}_m H_m^* + U_{1,m}^* \boldsymbol{B}_1 \boldsymbol{C}_2^* U_{2,m} = 0$, which has a unique solution because of $W(G_m) \subset W(A_1)$, $W(H_m^*) \subset W(A_2^*)$.

In summary, Algorithm 3 applied to (41) reduces to the following procedure:

1. Apply Algorithm 2 to compute orthogonal basis $U_{1,m}$ of $q_m(A_1)^{-1}\mathcal{K}_m(A_1, \boldsymbol{B}_1)$ and $G_m = U_{1,m}^* A_1 U_{1,m}$.
2. Apply Algorithm 2 to compute orthogonal basis $V_{2,m}$ of $\bar{q}_m(A_2)^{-*}\mathcal{K}_m(A_2, \boldsymbol{C}_2)$ and $H_m = V_{2,m}^* A_2^* V_{2,m}$.
3. Solve Sylvester equation $G_m \tilde{Z}_m - \tilde{Z}_m H_m^* + U_{1,m}^* \boldsymbol{B}_1 \boldsymbol{C}_2^* V_{2,m} = 0$.
4. Return approximate solution $Z_m = U_{1,m} \tilde{Z}_m V_{2,m}^*$.

This procedure turns out to be identical to existing rational Krylov subspace methods for Sylvester equations; see [8, 18] as well as [43] for additional references. In turn, the theory developed in this work can be used to bound the convergence of these methods via the best rational approximation of the sign function on $W(A_1) \cup W(-A_2)$. However, the bounds resulting from such an approach do not seem to offer advantages compared to existing bounds [4, 7, 18], and we will therefore skip the details.

6. Conclusions. The rational Krylov methods developed in this work constitute a fast way to approximate low-rank updates of the form $f(A + \boldsymbol{B}\boldsymbol{C}^*) - f(A)$, provided that shifted inverses with A can be applied efficiently. Their computational cost is comparable to the application of existing rational Krylov methods for approximating $f(A)\boldsymbol{B}$ and $f(A^*)\boldsymbol{C}$. This work has focused on theoretical and algorithmic foundations. Future work will explore the application and the adaptation of our methods to specific problems in scientific computing and data science.

Acknowledgments. The authors gratefully acknowledge inspiring discussions with Stefano Massei, Vanni Noferini, and Ana Šušnjara.

REFERENCES

[1] J.-E. Andersson, *Approximation of e^{-x} by rational functions with concentrated negative poles*, J. Approx. Theory, 32 (1981), pp. 85–95, https://doi.org/10.1016/0021-9045(81)90106-4.

[2] A. I. Aptekarev, *Sharp constants for rational approximations of analytic functions*, Mat. Sb., 193 (2002), pp. 3–72, https://doi.org/10.1070/SM2002v193n01ABEH000619.

[3] B. Beckermann, *The condition number of real Vandermonde, Krylov and positive definite Hankel matrices*, Numer. Math., 85 (2000), pp. 553–577.

[4] B. Beckermann, *An error analysis for rational Galerkin projection applied to the Sylvester equation*, SIAM J. Numer. Anal., 49 (2011), pp. 2430–2450, https://doi.org/10.1137/110824590.

[5] B. Beckermann, D. Kressner, and M. Schweitzer, *Low-rank updates of matrix functions*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 539–565, https://doi.org/10.1137/17M1140108.

[6] B. Beckermann and L. Reichel, *Error estimation and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883, https://doi.org/10.1137/080741744.

[7] B. Beckermann and A. Townsend, *Bounds on the singular values of matrices with displacement structure*, SIAM Rev., 61 (2019), pp. 319–344, https://doi.org/10.1137/19M1244433.
[8] P. Benner, R.-C. Li, and N. Truhar, *On the ADI method for Sylvester equations*, J. Comput. Appl. Math., 233 (2009), pp. 1035–1045, https://doi.org/10.1016/j.cam.2009.08.108.
[9] M. Benzi and P. Boito, *Matrix functions in network analysis*, GAMM-Mitt., 43 (2020), e202000012, https://doi.org/10.1002/gamm.202000012.
[10] C. Berg and G. Forst, *Potential Theory on Locally Compact Abelian Groups*, Springer, Berlin, Heidelberg, 1975.
[11] M. Berljafa, S. Elsworth, and S. Güttel, *A Rational Krylov Toolbox for MATLAB*, Tech. report 2014.56, Manchester Institute for Mathematical Sciences, The University of Manchester, 2014.
[12] D. S. Bernstein and C. F. Van Loan, *Rational matrix functions and rank-1 updates*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 145–154, https://doi.org/10.1137/S0895479898333636.
[13] G. Beylkin, N. Coult, and M. J. Mohlenkamp, *Fast spectral projection algorithms for density-matrix computations*, J. Comput. Phys., 152 (1999), pp. 32–54.
[14] 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.
[15] A. Boriçi, *On the Neuberger overlap operator*, Phys. Lett. B, 453 (1999), pp. 46–53.
[16] M. Crouzeix and D. Kressner, *A Bivariate Extension of the Crouzeix-Palencia Result with an Application to Fréchet Derivatives of Matrix Functions*, preprint, https://arxiv.org/abs/2007.09784, 2020.
[17] 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. 755–771, https://doi.org/10.1137/S0895479895292400.
[18] V. Druskin, L. Knizhnerman, and V. Simoncini, *Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1875–1898, https://doi.org/10.1137/100813257.
[19] S. Elsworth and S. Güttel, *The block rational Arnoldi method*, SIAM J. Matrix Anal. Appl., 41 (2020), pp. 365–388, https://doi.org/10.1137/19M1245505.
[20] J. van den Eshof, A. Frommer, Th. 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.
[21] E. Estrada and D. J. Higham, *Network properties revealed through matrix functions*, SIAM Rev., 52 (2010), pp. 696–714, https://doi.org/10.1137/090761070.
[22] 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, https://doi.org/10.1137/140973463.
[23] M. I. Gil', *Perturbations of functions of diagonalizable matrices*, Electron. J. Linear Algebra, 20 (2010), pp. 303–313, https://doi.org/10.13001/1081-3810.1375.
[24] G. H. Golub and R. Underwood, *The block Lanczos method for computing eigenvalues*, in Mathematical Software, III (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, WI, 1977), Publ. Math. Res. Center 39, Academic Press, New York, 1977, pp. 361–377.
[25] A. A. Gonchar and E. A. Rakhmanov, *Equilibrium distributions and the rate of rational approximation of analytic functions*, Mat. Sb. (N.S.), 134 (1987), pp. 306–352, 447.
[26] S. Güttel, *Rational Krylov Methods for Operator Functions*, Ph.D. thesis, Fakultät für Mathematik und Informatik der Technischen Universität Bergakademie Freiberg, 2010.
[27] S. Güttel, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, GAMM-Mitt., 36 (2013), pp. 8–31.
[28] S. Güttel and L. Knizhnerman, *A black-box rational Arnoldi variant for Cauchy–Stieltjes matrix functions*, BIT, 53 (2013), pp. 595–616.
[29] P. Henrici, *Applied and Computational Complex Analysis*, Vol. 2, John Wiley & Sons, New York, 1977.
[30] N. J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008, https://doi.org/10.1137/1.9780898717778.
[31] M. Hochbruck and A. Ostermann, *Exponential integrators*, Acta Numer., 19 (2010), pp. 209–286.
[32] C. Jagels and L. Reichel, *Recursion relations for the extended Krylov subspace method*, Linear Algebra Appl., 434 (2011), pp. 1716–1732.
[33] 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.

[34] I. Moret and P. Novati, *RD-rational approximations of the matrix exponential*, BIT, 44 (2004), pp. 595–615.
[35] Y. Nakatsukasa and N. J. Higham, *Stable and efficient spectral divide and conquer algorithms for the symmetric eigenvalue decomposition and the SVD*, SIAM J. Sci. Comput., 35 (2013), pp. A1325–A1349, https://doi.org/10.1137/120876605.
[36] P. P. Petrushev and V. A. Popov, *Rational Approximation of Real Functions*, Encyclopedia Math. Appl., Cambridge University Press, 1988.
[37] L. Reichel, *Newton interpolation at Leja points*, BIT, 30 (1990), pp. 332–346.
[38] J. D. Roberts, *Linear model reduction and solution of the algebraic Riccati equation by use of the sign function*, Internat. J. Control, 32 (1980), pp. 677–687.
[39] Y. Saad, *Analysis of some Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 29 (1992), pp. 209–228, https://doi.org/10.1137/0729014.
[40] J. Sherman and W. J. Morrison, *Adjustment of an inverse matrix corresponding to a change in one element of a given matrix*, Ann. Math. Statist., 21 (1950), pp. 124–127.
[41] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, *The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains*, IEEE Signal Process. Mag., 30 (2013), pp. 83–98, https://doi.org/10.1109/MSP.2012.2235192.
[42] V. Simoncini, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comput., 29 (2007), pp. 1268–1288, https://doi.org/10.1137/06066120X.
[43] V. Simoncini, *Computational methods for linear matrix equations*, SIAM Rev., 58 (2016), pp. 377–441, https://doi.org/10.1137/130912839.
[44] A. Skripka and A. Tomskova, *Multilinear Operator Integrals*, Lecture Notes in Math. 2250, Springer, Cham, 2019, https://doi.org/10.1007/978-3-030-32406-3.
[45] M. Stoll, *A literature survey of matrix methods for data science*, GAMM-Mitt., 43 (2020), e202000013, https://doi.org/10.1002/gamm.202000013.
[46] J. van den Eshof and M. Hochbruck, *Preconditioning Lanczos approximations to the matrix exponential*, SIAM J. Sci. Comput., 27 (2006), pp. 1438–1457, https://doi.org/10.1137/040605461.
[47] G. Zolotarev, *Application of elliptic functions to the problem of functions which vary the least or the most from zero*, Abh. St. Petersb., 30 (1877), pp. 1–59.