Generalized rational Krylov decompositions with an application to rational approximation

Berljafa, Mario and Güttel, Stefan

2014

MIMS EPrint: **2014.59**

ISSN 1749-9097

GENERALIZED RATIONAL KRYLOV DECOMPOSITIONS WITH AN APPLICATION TO RATIONAL APPROXIMATION

MARIO BERLJAFA* AND STEFAN GÜTTEL*

Abstract. Generalized rational Krylov decompositions are matrix relations which, under certain conditions, are associated with rational Krylov spaces. We study the algebraic properties of such decompositions and present an implicit Q theorem for rational Krylov spaces. Transformations on rational Krylov decompositions allow for changing the poles of a rational Krylov space without recomputation, and two algorithms are presented for this task. Using such transformations we develop a rational Krylov method for rational least squares fitting. Numerical experiments indicate that the proposed method converges fast and robustly. A MATLAB toolbox with implementations of the presented algorithms and experiments is provided.

Key words. rational Krylov decomposition, inverse eigenvalue problem, rational approximation

AMS subject classifications. 15A22, 65F15, 65F18, 30E10

1. Introduction.

Numerical methods based on rational Krylov spaces have become an indispensable tool of scientific computing. Rational Krylov spaces were initially proposed by Ruhe in the 1980s for the purpose of solving large sparse eigenvalue problems [29, 31, 32]. Since then many more applications have been found in model order reduction [18, 14], large-scale matrix functions and matrix equations [12, 1, 21, 22], and nonlinear eigenvalue problems [33, 24, 39, 23], to name just a few.

In this paper we study various algebraic properties of rational Krylov spaces, using as starting point a *generalized rational Krylov decomposition*

$$AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}, \tag{1.1}$$

where $A \in \mathbb{C}^{N\times N}$ is a given matrix, and the matrices $V_{m+1} \in \mathbb{C}^{N\times(m+1)}$ and $\{\underline{K_m}, \underline{H_m}\} \subset \mathbb{C}^{(m+1)\times m}$ are of maximal rank. Throughout this paper the underlined matrices have one row more than they have columns. Our matrix decomposition approach is inspired by the work of Stewart [35, 37] who studied transformations on a *(polynomial) Krylov decomposition*

$$AV_m = V_{m+1}\underline{H_m}, \tag{1.2}$$

which is a special case of (1.1) with $\underline{K_m} = \underline{I_m}$, the $m \times m$ identity matrix with an appended row of zeros. Indeed, all results in this paper apply to polynomial Krylov spaces as well.

The outline of this work is as follows: in section 2 we study algebraic properties of rational Arnoldi decompositions (a special case of (1.1) where $(\underline{H_m}, \underline{K_m})$ is an unreduced upper-Hessenberg pencil) and relate these decompositions to the poles and the starting vector of a rational Krylov space. Section 3 provides a rational implicit Q theorem about the uniqueness of rational Arnoldi decompositions. We also show how the rational functions associated with the rational Krylov space can be evaluated at any point $z \in \mathbb{C}$ by computing a full QR factorization of $z\underline{K_m} - \underline{H_m}$. In section 4 we show that when the lower $m \times m$ part of the pencil $(\underline{H_m}, \underline{K_m})$ is regular

* School of Mathematics, The University of Manchester, Alan Turing Building, Oxford Road, M13 9PL Manchester, United Kingdom, mario.berljafa@manchester.ac.uk, stefan.guettel@manchester.ac.uk

then the range of V_{m+1} in (1.1) is a rational Krylov space. Via transformations on such decompositions we are able to move the poles of a rational Krylov space to arbitrary positions (even to the eigenvalues of A), and we give two algorithms for this task. Finally, in section 5 we incorporate one of these algorithms into an iterative method, called RKFIT, for finding a rational function $R_m(z)$ of type (m, m) such that $\|F\boldsymbol{v} - R_m(A)\boldsymbol{v}\|_2$ is minimal, where $\{A, F\} \subset \mathbb{C}^{N\times N}$ are given matrices and $\boldsymbol{v} \in \mathbb{C}^N$.

All algorithms and numerical experiments presented in this paper are contained in a MATLAB toolbox [2] available for download.[1]

Notation. Matrices are labeled with uppercase Latin letters (e.g., A or H) and their elements with the corresponding lowercase Latin letters and indices (e.g., a_{ij} or h_{ij}). (Underlined) lowercase Latin letters in bold with an index k (e.g., $\boldsymbol{q}_k, \underline{\boldsymbol{h}_k}, \boldsymbol{h}_k$) denote the kth column of the corresponding matrix, or just its leading ($k+1$ or) k components. Vectors are also labeled by (underlined) lowercase Latin letters in bold (e.g., $\boldsymbol{b}$ or $\boldsymbol{v}_k$). (Underlined) uppercase Latin letters ($Q_k, \underline{H_k}, H_k$) with an index k denote the corresponding submatrix made of the leading k columns, or their leading ($k+1$ or) k rows, as appropriate. By $\mathcal{R}V$ we denote the range of a matrix V.

2. Rational Arnoldi decompositions. Given a matrix $A \in \mathbb{C}^{N\times N}$, a starting vector $\boldsymbol{v} \in \mathbb{C}^N$, and an integer $m < N$, the associated *polynomial Krylov space of order* $m+1$ is defined as $\mathcal{K}_{m+1} = \mathcal{K}_{m+1}(A, \boldsymbol{v}) = \operatorname{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^m\boldsymbol{v}\}$. There exists an integer $M \le N$, called the *invariance index* for $(A, \boldsymbol{v})$, such that

$$\mathcal{K}_1 \subset \mathcal{K}_2 \subset \cdots \subset \mathcal{K}_{M-1} \subset \mathcal{K}_M = \mathcal{K}_{M+1}.$$

Throughout this work we assume that $0 < m < M$, in which case the space $\mathcal{K}_{m+1}$ is isomorphic to $\mathcal{P}_m$, the linear space of polynomials of degree at most m, i.e., any $\boldsymbol{w} \in \mathcal{K}_{m+1}$ corresponds to a polynomial $p \in \mathcal{P}_m$ satisfying $\boldsymbol{w} = p(A)\boldsymbol{v}$, and vice versa.

Given a nonzero polynomial $q_m \in \mathcal{P}_m$ with roots disjoint from the spectrum $\Lambda(A)$, we define the associated *rational Krylov space* as

$$\mathcal{Q}_{m+1} = \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) := q_m(A)^{-1}\mathcal{K}_{m+1}(A, \boldsymbol{v}). \tag{2.1}$$

Note that $q_m(A)$ is nonsingular since no root of q_m is an eigenvalue of A and therefore $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ is well defined. Clearly, we have $\boldsymbol{v} \in \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$. Moreover, $\dim(\mathcal{Q}_{m+1}) = \dim(\mathcal{K}_{m+1})$ so that $\mathcal{Q}_{m+1}$ is A-variant (i.e., $A\mathcal{Q}_{m+1} \not\subseteq \mathcal{Q}_{m+1}$) if and only if $m+1 < M$.

The roots of q_m are called *poles* of the rational Krylov space and denoted by $\xi_1, \xi_2, \ldots, \xi_m$. Infinity is an allowed root and it symbolises that q_m is not of exact degree m, i.e., $\deg(q_m) < m$. We now show that the poles of a rational Krylov space are uniquely determined by the starting vector and vice versa.

LEMMA 2.1. *Let $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ be a given A-variant rational Krylov space, i.e., $m+1 < M$. Then the poles of $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ are uniquely determined by the starting vector $\boldsymbol{v}$, or equivalently, the starting vector of $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ is uniquely, up to scaling, determined by the roots of q_m.*

Proof. We first show that given an A-variant polynomial Krylov space $\mathcal{K}_{m+1}(A, \boldsymbol{q})$, all vectors $\boldsymbol{w} \in \mathcal{K}_{m+1}(A, \boldsymbol{q})$ that satisfy $\mathcal{K}_{m+1}(A, \boldsymbol{q}) = \mathcal{K}_{m+1}(A, \boldsymbol{w})$ are of the form $\boldsymbol{w} = \alpha\boldsymbol{q}$, $\alpha \neq 0$. Assume to the contrary that there exists a polynomial p_j with $1 \le \deg(p_j) = j \le m$ such that $\boldsymbol{w} = p_j(A)\boldsymbol{q}$. Then $A^{m+1-j}\boldsymbol{w} \in \mathcal{K}_{m+1}(A, \boldsymbol{w})$, but

[1] `http://guettel.com/rktoolbox` as of November 2014.

for the same vector we have $A^{m+1-j}\boldsymbol{w} = A^{m+1-j}p_j(A)\boldsymbol{q} \notin \mathcal{K}_{m+1}(A, \boldsymbol{q})$. This is a contradiction to $\mathcal{K}_{m+1}(A, \boldsymbol{q}) = \mathcal{K}_{m+1}(A, \boldsymbol{w})$.

To show that the poles are uniquely determined by the starting vector $\boldsymbol{v}$, assume that $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) = \mathcal{Q}_{m+1}(A, \boldsymbol{v}, \hat{q}_m)$. Using the definition of a rational Krylov space (2.1), this is equivalent to $\mathcal{K}_{m+1}(A, \hat{q}_m(A)\boldsymbol{v}) = \mathcal{K}_{m+1}(A, q_m(A)\boldsymbol{v})$. This space is A-variant, hence by the above argument we know that $q_m(A)\boldsymbol{v} = \alpha\hat{q}_m(A)\boldsymbol{v}$, $\alpha \neq 0$. This vector is an element of $\mathcal{K}_{m+1}(A, \boldsymbol{v})$ which is isomorphic to $\mathcal{P}_m$. Therefore $q_m = \alpha\hat{q}_m$ and hence q_m and $\hat{q}_m$ have identical roots. Similarly one shows that if $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) = \mathcal{Q}_{m+1}(A, \hat{\boldsymbol{v}}, q_m)$, then $\boldsymbol{v} = \alpha\hat{\boldsymbol{v}}$ with $\alpha \neq 0$. □

In the following we aim to establish a one-to-one correspondence between rational Krylov spaces and a particular type of matrix decompositions. As a consequence we are able to study the algebraic properties of rational Krylov spaces using these decompositions. Recall that a matrix $\underline{H_m} \in \mathbb{C}^{(m+1)\times m}$ is called *upper-Hessenberg* if all the elements below the first subdiagonal are zero, i.e., if $i > j+1$ implies $h_{ij} = 0$. Further, we say that $\underline{H_m}$ is *unreduced* if none of the elements on the first subdiagonal are zero, i.e., $h_{i+1,i} \neq 0$. For convenience, we now generalize this terminology from matrices to pencils $(\underline{H_m}, \underline{K_m})$.

DEFINITION 2.2. *Let* $\{\underline{K_m}, \underline{H_m}\} \subset \mathbb{C}^{(m+1)\times m}$ *be upper-Hessenberg matrices. We say that* $(\underline{H_m}, \underline{K_m})$ *is an* unreduced upper-Hessenberg pencil *if* $|h_{j+1,j}| + |k_{j+1,j}| \neq 0$ *for all* $j = 1, \ldots, m$.

We are now ready to introduce the notion of a rational Arnoldi decomposition.

DEFINITION 2.3. *Let* $A \in \mathbb{C}^{N\times N}$ *be a given matrix. A relation of the form*

$$AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m} \tag{2.2}$$

is called a rational Arnoldi decomposition (RAD) *if* $V_{m+1} \in \mathbb{C}^{N\times(m+1)}$ *is of full column rank,* $(\underline{H_m}, \underline{K_m})$ *is an unreduced upper-Hessenberg pencil of size* $(m+1)\times m$*, and the quotients* $h_{j+1,j}/k_{j+1,j}$*, called* poles *of the decomposition, are outside the spectrum* $\Lambda(A)$ *for* $j = 1, \ldots, m$.

The columns of V_{m+1} *are called the* basis *of the decomposition and they span the* space *of the decomposition. If* V_{m+1} *is orthonormal, we say that* (2.2) *is an* orthonormal RAD.

It is noteworthy that both $\underline{H_m}$ and $\underline{K_m}$ in the RAD (2.2) are of full rank. To see this take any $\xi \in \mathbb{C}$ and subtract $\xi V_{m+1}\underline{K_m}$ from both sides of (2.2). This leads to

$$(A - \xi I)V_{m+1}\underline{K_m} = V_{m+1}(\underline{H_m} - \xi\underline{K_m}). \tag{2.3}$$

Since $(\underline{H_m}, \underline{K_m})$ is unreduced there are at most m numbers ξ such that $\underline{H_m} - \xi\underline{K_m}$ is not unreduced. For any other ξ the right-hand side in (2.3) is of full rank and so must be the left-hand side. Therefore $\underline{K_m}$ is of full rank. If A is nonsingular, comparing the ranks of the left- and right-hand side in (2.2) we now see that $\underline{H_m}$ is of full rank as well. If A is singular, then zero is not an allowed pole and therefore $\underline{H_m}$ is unreduced and hence of full rank.

Furthermore, any RAD (2.2) can be transformed into an orthonormal RAD using the thin QR decomposition $V_{m+1} = QR$. Setting $\breve{V}_{m+1} = Q$, $\underline{\breve{K}_m} = R\underline{K_m}$, and $\underline{\breve{H}_m} = R\underline{H_m}$, we obtain the decomposition $A\breve{V}_{m+1}\underline{\breve{K}_m} = \breve{V}_{m+1}\underline{\breve{H}_m}$, spanning $\mathcal{R}V_{m+1}$ (i.e., $\mathcal{R}\breve{V}_{m+1} = \mathcal{R}V_{m+1}$) and satisfying $h_{j+1,j}/k_{j+1,j} = \breve{h}_{j+1,j}/\breve{k}_{j+1,j}$ for all $j = 1, \ldots, m$. We call these two RADs equivalent.

DEFINITION 2.4. *Two RADs with the same matrix* $A \in \mathbb{C}^{N\times N}$ *are* equivalent *if they span the same space and have the same poles.*

From now on we assume all RADs to be orthonormal. In Theorem 2.5 we show that for every rational Krylov space $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ there exists an RAD (2.2) spanning $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ and conversely, if such a decomposition exists it spans a rational Krylov space. To proceed it is convenient to write the polynomial q_m in factored form, and to label separately all the leading factors

$$q_j(z) = \prod_{i=1}^{j} \left(h_{i+1,i} - k_{i+1,i} z\right), \quad j = 0, 1, \ldots, m, \tag{2.4}$$

with some scalars $\{h_{i+1,i}, k_{i+1,i}\}_{i=1}^{m} \subset \mathbb{C}$ such that $\xi_i = h_{i+1,i}/k_{i+1,i}$. Since (2.1) is independent of the scaling of q_m any choice of the scalars $h_{i+1,i}$ and $k_{i+1,i}$ is valid as long as their ratio is ξ_i. When we make use of (2.4) without specifying the order of appearance of the poles, we mean any order. The fact that $q_j \,|\, q_{j+1}$ gives rise to a sequence of nested rational Krylov spaces

$$\mathcal{Q}_1 \subset \mathcal{Q}_2 \subset \cdots \subset \mathcal{Q}_{m+1},$$

where $\mathcal{Q}_{j+1} = \mathcal{Q}_{j+1}(A, \boldsymbol{v}, q_j)$ for $j = 0, 1, \ldots, m$.

THEOREM 2.5. *Let $\mathcal{V}_{m+1}$ be a vector space of dimension $m+1$. Then $\mathcal{V}_{m+1}$ is a rational Krylov space with starting vector $\boldsymbol{v} \in \mathcal{V}_{m+1}$ and poles $\xi_1, \ldots, \xi_m$ if and only if there exists an RAD* (2.2) *with $\mathcal{R}V_{m+1} = \mathcal{V}_{m+1}$, $\boldsymbol{v}_1 = \boldsymbol{v}$, and poles $\xi_1, \ldots, \xi_m$.*

Proof. Let (2.2) hold and define the polynomials $\{q_j\}_{j=0}^{m}$ as in (2.4). Note that these are nonzero polynomials since the pencil $\left(\underline{H_m}, \underline{K_m}\right)$ is unreduced. We show by induction that

$$\mathcal{V}_{j+1} := \operatorname{span}\left\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{j+1}\right\} = q_j(A)^{-1} \mathcal{K}_{j+1}(A, \boldsymbol{v}), \tag{2.5}$$

for $j = 1, \ldots, m$, and with $\boldsymbol{v} = \boldsymbol{v}_1$. In particular, for $j = m$ we obtain $\mathcal{V}_{m+1} = q_m(A)^{-1} \mathcal{K}_{m+1}(A, \boldsymbol{v})$. Consider $j = 1$. Reading (2.2) column-wise, first column only, and rearranging the terms yields

$$q_1(A) \boldsymbol{v}_2 = \left(h_{21} I - k_{21} A\right) \boldsymbol{v}_2 = \left(k_{11} A - h_{11} I\right) \boldsymbol{v}_1. \tag{2.6}$$

Therefore, $\boldsymbol{v}_2 = q_1(A)^{-1} \left(k_{11} A - h_{11} I\right) \boldsymbol{v}_1 \in q_1(A)^{-1} \mathcal{K}_2(A, \boldsymbol{v})$ which together with the fact $\boldsymbol{v}_1 \in q_1(A)^{-1} \mathcal{K}_2(A, \boldsymbol{v})$ proves (2.5) for $j = 1$. Let us assume that (2.5) holds for $j = 1, \ldots, n-1 < m$. We now consider the case $j = n$. Comparing the nth column on the left- and the right-hand side in (2.2) and rearranging the terms gives

$$\left(h_{n+1,n} I - k_{n+1,n} A\right) \boldsymbol{v}_{n+1} = \sum_{i=1}^{n} \left(k_{in} A - h_{in} I\right) \boldsymbol{v}_i, \tag{2.7}$$

$$\text{and hence} \qquad q_n(A) \boldsymbol{v}_{n+1} = \sum_{i=1}^{n} \left(k_{in} A - h_{in} I\right) q_{n-1}(A) \boldsymbol{v}_i. \tag{2.8}$$

By the induction hypothesis $\boldsymbol{v}_i \in q_{n-1}(A)^{-1} \mathcal{K}_n(A, \boldsymbol{v})$, therefore

$$\left(k_{in} A - h_{in} I\right) q_{n-1}(A) \boldsymbol{v}_i \in \mathcal{K}_{n+1}(A, \boldsymbol{v}), \quad i = 1, \ldots, n. \tag{2.9}$$

It follows from (2.8) and (2.9) that $\boldsymbol{v}_{n+1} \in q_n(A)^{-1} \mathcal{K}_{n+1}(A, \boldsymbol{v})$. The induction hypothesis asserts $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \subseteq q_n(A)^{-1} \mathcal{K}_{n+1}(A, \boldsymbol{v})$ which concludes this direction.

Alternatively, let $\mathcal{V}_{m+1} = q_m(A)^{-1}\mathcal{K}_{m+1}(A, \boldsymbol{v})$ be a rational Krylov space with a basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m+1}\}$ satisfying (2.5). Thus for $n = 1, \ldots, m$ there holds

$$\boldsymbol{v}_{n+1} \in q_n(A)^{-1}\mathcal{K}_{n+1}(A, \boldsymbol{v}) \Leftrightarrow \left(h_{n+1,n}I - k_{n+1,n}A\right) \boldsymbol{v}_{n+1} \in q_{n-1}(A)^{-1}\mathcal{K}_{n+1}(A, \boldsymbol{v}).$$

Since $\mathcal{K}_{n+1}(A, \boldsymbol{v}) = \mathcal{K}_n(A, \boldsymbol{v}) \oplus A\mathcal{K}_n(A, \boldsymbol{v})$ we have $q_{n-1}(A)^{-1}\mathcal{K}_{n+1}(A, \boldsymbol{v}) = \mathcal{Q}_n \oplus A\mathcal{Q}_n$. Consequently, there exist numbers $\{h_{in}, k_{in}\}_{i=1}^n \subset \mathbb{C}$ such that (2.7) holds. These relations can be merged into matrix form to get (2.2) with the pencil $\left(\underline{H_m}, \underline{K_m}\right)$ being unreduced as a consequence of q_m being a nonzero polynomial. ∎

By an inductive argument, using the decomposition (2.2), we arrive at the explicit formula [29]

$$\boldsymbol{v}_{j+1} = p_j(A)q_j(A)^{-1}\boldsymbol{v}_1, \quad j = 0, 1, \ldots, m, \tag{2.10}$$

where $p_0(z) := 1$ and for $j > 0$ we have $p_j(z) = \det\left(zK_j - H_j\right)$. The polynomials q_j are given by (2.4). Note that $p_j(z)$ is the determinant of the upper $j \times j$ submatrix of $z\underline{K_j} - \underline{H_j}$, whilst $(-1)^j q_j(z)$ is the determinant of its lower $j \times j$ submatrix. We only give a brief outline for a proof of (2.10). From (2.6) the relation (2.10) follows for $j = 1$. Assume (2.10) has been established for $j = 1, \ldots, n < m$ and insert it into (2.8). Then (2.10) follows for $j = n + 1$ by noticing that the right-hand side of (2.8) represents the Laplace expansion of $\det\left(zK_n - H_n\right)$ along the last column.

3. A rational implicit Q theorem. A special case of an orthonormal rational Arnoldi decomposition (2.2) is the polynomial Arnoldi decomposition (1.2). The corresponding polynomial q_m is constant and $\mathcal{R}V_{m+1} = \mathcal{K}_{m+1}(A, \boldsymbol{v}_1)$. The implicit Q theorem, see [36, Theorem 3.3], states that once the first column of V_{m+1} is fixed, so is, up to column scaling, the whole matrix V_{m+1}. Since $\underline{H_m} = V_{m+1}^* A V_m$ any scaling of V_{m+1} also affects $\underline{H_m}$. If V_{m+1} is rescaled to $\breve{V}_{m+1} = V_{m+1}D_{m+1}$, with $|D_{m+1}| = I_{m+1}$, then $\underline{\breve{H}_m} = D_{m+1}^* \underline{H_m} D_m$. There is no essential difference between V_{m+1} and $\underline{H_m}$ on one side and $\breve{V}_{m+1}$ and $\underline{\breve{H}_m}$ on the other. In this sense we say that both V_{m+1} and $\underline{H_m}$ are essentially uniquely determined by the first column of V_{m+1}. With Theorem 3.1 we now generalize this result to RADs. Our proof uses the proof of [36, Theorem 3.3] as a template.

Apart from the column scaling of V_{m+1}, in the rational case the decomposition (2.2) is also invariant (in the sense that it spans the same space, the poles remain unchanged, and the upper-Hessenberg structure is preserved) under right-multiplication by upper-triangular nonsingular matrices R_m. Therefore V_{m+1} and $(\underline{H_m}, \underline{K_m})$ are essentially the same as $\breve{V}_{m+1}$ and $(\underline{\breve{H}_m}R_m, \underline{\breve{K}_m}R_m)$, where $\underline{\breve{K}_m} = D_{m+1}^* \underline{K_m} D_m$.

THEOREM 3.1. *Let $A \in \mathbb{C}^{N\times N}$ satisfy the orthonormal rational Arnoldi decomposition $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with poles $\xi_j = h_{j+1,j}/k_{j+1,j}$. For every $j = 1, \ldots, m$ the orthonormal matrix V_{j+1} and the pencil $\left(\underline{H_j}, \underline{K_j}\right)$ are essentially uniquely determined by the first column of V_{m+1} and the poles $\xi_1, \ldots, \xi_j$.*

Proof. The proof goes by induction on j. We assume without loss of generality that $h_{j+1,j} \neq 0$ for all $j = 1, \ldots, m$. Otherwise, if $h_{j+1,j} = 0$ for some j, then $0 = \xi_j \notin \Lambda(A)$ and we can consider $V_{m+1}\underline{K_m} = A^{-1}V_{m+1}\underline{H_m}$ at that step j, thus interchanging the roles of $\underline{H_m}$ and $\underline{K_m}$ and using A^{-1} instead of A. Since $\left(\underline{H_m}, \underline{K_m}\right)$ is unreduced, $k_{j+1,j} \neq 0$ if $h_{j+1,j} = 0$. The relation $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ can be shifted for all $\xi \in \overline{\mathbb{C}}^* \setminus \Lambda(A)$ to provide

$$A^{(\xi)}V_{m+1}\underline{L_m^{(\xi)}} = V_{m+1}\underline{H_m}, \tag{3.1}$$

where $A^{(\xi)} := (I - A/\xi)^{-1} A$ and $\underline{L_m^{(\xi)}} := \big(\underline{K_m} - \underline{H_m}/\xi\big)$. We make frequent use of this relation, reading it column-wise. It is worth noticing that the jth column of $\underline{L_m^{(\xi)}}$ has all but eventually the leading j components equal to zero, and that $\underline{L_j^{(\xi)}}$ is of full rank for all j and ξ. We are now ready to prove the statement.

The first column in (3.1) for $\xi = \xi_1$ yields

$$\ell_{11}^{(\xi_1)} A^{(\xi_1)} \boldsymbol{v}_1 = h_{11}\boldsymbol{v}_1 + h_{21}\boldsymbol{v}_2.$$

Since $\boldsymbol{v}_1^*\boldsymbol{v}_1 = 1$ and $\boldsymbol{v}_1^*\boldsymbol{v}_2 = 0$, we have

$$h_{11} = \ell_{11}^{(\xi_1)} \boldsymbol{v}_1^* A^{(\xi_1)} \boldsymbol{v}_1 =: \ell_{11}^{(\xi_1)} \breve{h}_{11}.$$

We then have

$$\begin{aligned} h_{21}\boldsymbol{v}_2 &= \ell_{11}^{(\xi_1)} A^{(\xi_1)} \boldsymbol{v}_1 - h_{11}\boldsymbol{v}_1 \\ &= \ell_{11}^{(\xi_1)} \big(A^{(\xi_1)}\boldsymbol{v}_1 - \breve{h}_{11}\boldsymbol{v}_1\big). \end{aligned}$$

Since $\|\boldsymbol{v}_2\|_2 = 1$ and $h_{21} \neq 0$ by assumption, we have $\ell_{11}^{(\xi_1)} \neq 0$ and may take $h_{21} = \ell_{11}^{(\xi_1)} \|A^{(\xi_1)}\boldsymbol{v}_1 - \breve{h}_{11}\boldsymbol{v}_1\|_2 =: \ell_{11}^{(\xi_1)}\breve{h}_{21}$, which determines $\boldsymbol{v}_2$ uniquely up to a factor of modulus one. Note that $k_{11} = \ell_{11}^{(\xi_1)}\big(1 + \breve{h}_{11}/\xi_1\big) =: \ell_{11}^{(\xi_1)}\breve{k}_{11}$ and $k_{21} = \ell_{11}^{(\xi_1)}\breve{h}_{21}/\xi_1 =: \ell_{11}^{(\xi_1)}\breve{k}_{21}$. The vector $\underline{\breve{\boldsymbol{h}}_1}$ is essentially unique, as is $\underline{\breve{\boldsymbol{k}}_1} = \underline{\boldsymbol{e}_1} + \underline{\breve{\boldsymbol{h}}_1}/\xi_1$. Clearly, the pencil $\big(\underline{\breve{\boldsymbol{h}}_1}, \underline{\breve{\boldsymbol{k}}_1}\big)$ is essentially unique and $Z_1 = \big[\ell_{11}^{(\xi_1)}\big]$ is nonsingular and such that $\big(\underline{\boldsymbol{h}_1}, \underline{\boldsymbol{k}_1}\big) = \big(\underline{\breve{\boldsymbol{h}}_1}Z_1, \underline{\breve{\boldsymbol{k}}_1}Z_1\big)$. Therefore V_2 and the pencil $\big(\underline{H_1}, \underline{K_1}\big) = \big(\underline{\boldsymbol{h}_1}, \underline{\boldsymbol{k}_1}\big)$ are essentially uniquely defined by $\boldsymbol{v}_1$ and ξ_1.

Suppose that, for $2 \leq j \leq m$, the matrix V_j and the pencil $\big(\underline{H_{j-1}}, \underline{K_{j-1}}\big) = \big(\underline{\breve{H}_{j-1}}Z_{j-1}, \underline{\breve{K}_{j-1}}Z_{j-1}\big)$, for an upper-triangular and nonsingular Z_{j-1}, are essentially uniquely defined by $\boldsymbol{v}_1$ and $\xi_1, \ldots, \xi_{j-1}$.

The jth column in (3.1) for $\xi = \xi_j$ gives

$$A^{(\xi_j)} V_j \boldsymbol{l}_j^{(\xi_j)} = V_{j+1}\underline{\boldsymbol{h}_j}.$$

Since $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j+1}$ are orthonormal we have

$$h_{ij} = \boldsymbol{v}_i^* A^{(\xi_j)} V_j \boldsymbol{l}_j^{(\xi_j)}, \qquad i = 1, 2, \ldots, j.$$

Rearranging the two equations above we deduce

$$\begin{aligned} h_{j+1,j}\boldsymbol{v}_{j+1} &= A^{(\xi_j)} V_j \boldsymbol{l}_j^{(\xi_j)} - V_j \boldsymbol{h}_j \\ &= A^{(\xi_j)} V_j \boldsymbol{l}_j^{(\xi_j)} - V_j V_j^* A^{(\xi_j)} V_j \boldsymbol{l}_j^{(\xi_j)} \\ &= \big(I - V_j V_j^*\big) A^{(\xi_j)} V_j \boldsymbol{l}_j^{(\xi_j)}. \end{aligned}$$

Expanding $\boldsymbol{l}_j^{(\xi_j)}$ as $\boldsymbol{l}_j^{(\xi_j)} =: \underline{L_{j-1}^{(\xi_j)}}\boldsymbol{z}_{j-1} + \boldsymbol{q}_j$, where $\boldsymbol{q}_j^* \underline{L_{j-1}^{(\xi_j)}} = 0$, gives

$$\begin{aligned} h_{j+1,j}\boldsymbol{v}_{j+1} &= \big(I - V_j V_j^*\big) A^{(\xi_j)} V_j \left(\underline{L_{j-1}^{(\xi_j)}}\boldsymbol{z}_{j-1} + \boldsymbol{q}_j\right) \\ &= \big(I - V_j V_j^*\big) A^{(\xi_j)} V_j \underline{L_{j-1}^{(\xi_j)}}\boldsymbol{z}_{j-1} + \big(I - V_j V_j^*\big) A^{(\xi_j)} V_j \boldsymbol{q}_j \\ &= \big(I - V_j V_j^*\big) A^{(\xi_j)} V_j \boldsymbol{q}_j. \end{aligned} \tag{3.2}$$

To obtain the last equality we have used $A^{(\xi_j)}V_j\underline{L_{j-1}^{(\xi_j)}} = V_j\underline{H_{j-1}}$, which are the first $j-1$ columns in (3.1) with $\xi = \xi_j$. Now, to see that (3.2) defines $\boldsymbol{v}_{j+1}$ uniquely up to scaling we note that since $h_{j+1,j} \neq 0$ the vector $\boldsymbol{q}_j$ is also nonzero. Further, $\boldsymbol{q}_j$ is essentially unique since $\underline{L_{j-1}^{(\xi_j)}} \in \mathbb{C}^{j\times(j-1)}$ is of full column rank and essentially unique by hypothesis. We can set $h_{j+1,j} = \|\left(I - V_jV_j^*\right)A^{(\xi_j)}V_j\boldsymbol{q}_j\|_2$.

It remains to prove the essential uniqueness of $\left(\underline{H_j}, \underline{K_j}\right)$. Let us normalize $\boldsymbol{q}_j = \rho_j\breve{\boldsymbol{q}}_j$ so that $\|\breve{\boldsymbol{q}}_j\|_2 = 1$. Recalling $\boldsymbol{l}_j^{(\xi_j)} = \underline{L_{j-1}^{(\xi_j)}}\boldsymbol{z}_{j-1} + \boldsymbol{q}_j = \underline{L_{j-1}^{(\xi_j)}}\boldsymbol{z}_{j-1} + \rho_j\breve{\boldsymbol{q}}_j$ we have

$$\begin{aligned} h_{ij} &= \boldsymbol{v}_i^* A^{(\xi_j)}V_j\boldsymbol{l}_j^{(\xi_j)} \\ &= \boldsymbol{e}_i^*\underline{H_{j-1}}\boldsymbol{z}_{j-1} + \rho_j\overbrace{\boldsymbol{v}_i^*A^{(\xi_j)}V_j\breve{\boldsymbol{q}}_j}^{\breve{h}_{ij}}, \end{aligned}$$

for $i = 1, 2, \ldots, j$. Defining $\breve{\boldsymbol{h}}_j := V_j^*A^{(\xi_j)}V_j\breve{\boldsymbol{q}}_j$ and $\breve{h}_{j+1,j} := \rho_j^{-1}h_{j+1,j}$, the vectors $\underline{\breve{\boldsymbol{h}}_j}$ and $\underline{\breve{\boldsymbol{k}}_j} := \underline{\breve{\boldsymbol{q}}_j} + \underline{\breve{\boldsymbol{h}}_j}/\xi_j$, where $\underline{\breve{\boldsymbol{q}}_j} = \begin{bmatrix}\breve{\boldsymbol{q}}_j^* & 0\end{bmatrix}^*$, are essentially unique. Further,

$$\underline{\boldsymbol{h}_j} = \begin{bmatrix} \underline{H_{j-1}} & \breve{\boldsymbol{h}}_j \\ \boldsymbol{0}^* & \breve{h}_{j+1,j} \end{bmatrix}\begin{bmatrix}\boldsymbol{z}_{j-1} \\ \rho_j\end{bmatrix} = \begin{bmatrix} \underline{\breve{H}_{j-1}} & \breve{\boldsymbol{h}}_j \\ \boldsymbol{0}^* & \breve{h}_{j+1,j} \end{bmatrix}\begin{bmatrix}Z_{j-1}\boldsymbol{z}_{j-1} \\ \rho_j\end{bmatrix} =: \underline{\breve{H}_j}\boldsymbol{z}_j,$$

and similarly for $\underline{\boldsymbol{k}_j} = \underline{\boldsymbol{l}_j^{(\xi_j)}} + \underline{\boldsymbol{h}_j}/\xi_j$ we find

$$\underline{\boldsymbol{k}_j} = \begin{bmatrix} \underline{K_{j-1}} & \breve{\boldsymbol{k}}_j \\ \boldsymbol{0}^* & \breve{k}_{j+1,j} \end{bmatrix}\begin{bmatrix}\boldsymbol{z}_{j-1} \\ \rho_j\end{bmatrix} = \begin{bmatrix} \underline{\breve{K}_{j-1}} & \breve{\boldsymbol{k}}_j \\ \boldsymbol{0}^* & \breve{k}_{j+1,j} \end{bmatrix}\begin{bmatrix}Z_{j-1}\boldsymbol{z}_{j-1} \\ \rho_j\end{bmatrix} =: \underline{\breve{K}_j}\boldsymbol{z}_j.$$

Define $\underline{Z_{j-1}} = \begin{bmatrix}Z_{j-1}^* & \boldsymbol{0}\end{bmatrix}^*$, and $Z_j = \begin{bmatrix}\underline{Z_{j-1}} & \boldsymbol{z}_j\end{bmatrix}$. Since Z_j is upper-triangular and nonsingular, $\left(\underline{H_j}, \underline{K_j}\right) = \left(\underline{\breve{H}_j}Z_j, \underline{\breve{K}_j}Z_j\right)$ is essentially unique as $\left(\underline{\breve{H}_j}, \underline{\breve{K}_j}\right)$ is. ∎

A further comment for the case $m = N-1$ is required. For the polynomial case, i.e., $\underline{K_{N-1}} = \underline{I_{N-1}}$, we have $AV_{N-1} = V_N\underline{H_{N-1}}$. The vector $\boldsymbol{h}_N = V_N^*AV_N\boldsymbol{e}_N$ is uniquely defined by the starting vector and A and $AV_N = V_NH_N$ holds. This last decomposition is usually stated as the (polynomial) implicit Q theorem and essential uniqueness of H_N is rightfully claimed. Let us consider a more general RAD $AV_N\underline{K_{N-1}} = V_N\underline{H_{N-1}}$. Left-multiplying it with V_N^* and right-multiplying with $\boldsymbol{e}_N$ gives $V_N^*AV_N\boldsymbol{k}_N = \boldsymbol{h}_N$, and for any $\boldsymbol{k}_N$ we can find the corresponding $\boldsymbol{h}_N$ such that $AV_NK_N = V_NH_N$ holds. Therefore we cannot say that $\left(H_N, K_N\right)$ is essentially unique. In fact, for the polynomial case $\boldsymbol{k}_N = \boldsymbol{e}_N$ is tacitly fixed.

As already mentioned, a polynomial Krylov space $\mathcal{K}_{m+1}(A, \boldsymbol{v})$ with orthonormal basis V_{m+1} is related to a decomposition of the form

$$AV_m = V_{m+1}\underline{H_m} = V_mH_m + h_{m+1,m}\boldsymbol{v}_{m+1}\boldsymbol{e}_m^T, \tag{3.3}$$

where H_m is upper-Hessenberg. For a rational Krylov space we have an RAD (2.2) with an upper-Hessenberg pencil $\left(\underline{H_m}, \underline{K_m}\right)$ rather than a single upper-Hessenberg matrix $\underline{H_m}$. It has been shown in [13, 38, 40, 27] that decompositions of the form (3.3) with H_m being *semiseparable plus diagonal*[2] are related to rational Krylov spaces in the same way as RADs are. Corresponding implicit Q theorems have been developed.

[2] A matrix S is called *(upper) semiseparable* if all submatrices consisting of elements in the lower-triangular part of S only are of rank at most 1. Examples of semiseparable matrices are (inverse) upper-Hessenberg matrices. The *diagonal* matrix D in $H_m = S + D$ carries the finite poles ξ_j whilst the infinite ones can be replaced by any finite number. The structure of S captures the difference between finite and infinite poles.

Algorithm 3.1 Rational Krylov method [29, 31, 32]. RK Toolbox: `rat_krylov`

Input: $A \in \mathbb{C}^{N\times N}$, $\boldsymbol{v} \in \mathbb{C}^N$ and poles $\{\xi_j\}_{j=1}^m \subset \overline{\mathbb{C}} \setminus \Lambda(A)$ with $m < M$.
Output: RAD $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ spanning $\mathcal{Q}_{m+1}(A, \boldsymbol{v})$ with poles $\{\xi_j\}_{j=1}^m$.

1. Set $\boldsymbol{v}_1 := \boldsymbol{v}/\|\boldsymbol{v}\|_2$, and $\boldsymbol{q}_1 := [1]$.
2. **for** $j = 1, \dots, m$ **do**
3. Compute $\boldsymbol{w} := A^{(\xi_j)} V_j \boldsymbol{q}_j$.
4. Project $\boldsymbol{h}_j := V_j^* \boldsymbol{w}$, and compute $h_{j+1,j} := \|\boldsymbol{w} - V_j \boldsymbol{h}_j\|_2$.
5. Orthogonalize $\boldsymbol{v}_{j+1} := \left(\boldsymbol{w} - V_j \boldsymbol{h}_j\right)/h_{j+1,j}$.
6. **if** $\xi_j \neq 0$ **then** Set $\underline{\boldsymbol{k}_j} := \underline{\boldsymbol{q}_j} + \underline{\boldsymbol{h}_j}/\xi_j$.
   else Set $\underline{\boldsymbol{k}_j} := \underline{\boldsymbol{h}_j}$, and replace $\underline{\boldsymbol{h}_j} := \underline{\boldsymbol{q}_j}$. **end if**
7. For $j \neq m$ set $\boldsymbol{q}_{j+1} := \boldsymbol{q}_{j+1}^{(\xi_{j+1})}$, where $\underline{L_j^{(\xi_{j+1})}} =: Q_{j+1}^{(\xi_{j+1})} \underline{R_j}$ is a QR factorization.
8. **end for**

We prefer to work with the pencil $\left(\underline{H_m}, \underline{K_m}\right)$ instead of the semiseparable plus diagonal representation since the former is widely used in practice. In fact, this pencil is a by-product of the rational Krylov method used to construct rational Krylov bases, and which is stated in Algorithm 3.1. We use the notation

$$A^{(\xi)} = \begin{cases} (I - A/\xi)^{-1} A, & \text{if } \xi \in \overline{\mathbb{C}}^* \setminus \Lambda(A), \\ A^{-1}, & \text{if } \xi = 0 \text{ and } A \text{ is nonsingular,} \end{cases}$$

and correspondingly

$$\underline{L_m^{(\xi)}} = \begin{cases} \underline{K_m} - \underline{H_m}/\xi, & \text{if } \xi \in \overline{\mathbb{C}}^*, \\ \underline{H_m}, & \text{otherwise.} \end{cases}$$

The vector $\boldsymbol{q}_j$ in Alg. 3.1 is often called *continuation combination* as it specifies onto which linear combination of the previously computed basis vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_j$ the operator $A^{(\xi)}$ is applied to, cf. line 3, Alg. 3.1. The choice made in line 7 is due to Ruhe [32] and it guarantees that the new vector $\boldsymbol{w}$ will be linearly independent of the previous vectors and hence expand the space, provided that we have not reached the invariance index. This can be seen from the proof of Theorem 3.1 or equivalently from (3.1) with m replaced by j.

In (2.10) we have given explicit formulas for the orthonormal vectors $\boldsymbol{v}_{j+1}$, implicitly defined by the starting vector $\boldsymbol{v}_1$ and the poles $\xi_1, \dots, \xi_j$. Note that the determinants appearing in the formulas do not change if the pencil $\left(\underline{H_j}, \underline{K_j}\right)$ is right-multiplied by an upper-triangular nonsingular matrix. We now give a formula for (a multiple of) the continuation vector $\boldsymbol{q}_{j+1} = \boldsymbol{q}_{j+1}^{(\xi_{j+1})}$ used in Alg. 3.1.

THEOREM 3.2. *Let* (2.2) *be an RAD associated with an A-variant space $\mathcal{R}V_{m+1}$, and let the polynomials $p_j, q_j \in \mathcal{P}_j$ be as in* (2.10). *Define the polynomials*

$$p_j^{[m]}(z) := q_m(z) q_j(z)^{-1} p_j(z) \in \mathcal{P}_m, \quad j = 0, 1, \dots, m.$$

Equivalently, $p_j^{[m]}(z)$ is the determinant of the $m \times m$ minor of $z\underline{K_m} - \underline{H_m}$ resulting from the removal of the jth row. Then for any $\xi \in \mathbb{C}$ there holds

$$\boldsymbol{q}_{m+1}^{(\xi)} \neq \boldsymbol{0} \quad \textit{and} \quad \left(\boldsymbol{q}_{m+1}^{(\xi)}\right)^* \underline{L_m^{(\xi)}} = \boldsymbol{0}^*,$$

where $\boldsymbol{q}_{m+1}^{(\xi)} := \begin{bmatrix} p_0^{[m]}(\xi) & p_1^{[m]}(\xi) & \dots & p_m^{[m]}(\xi) \end{bmatrix}^*$.

Proof. Let $\xi \in \mathbb{C}$ be an arbitrary scalar. Label the roots of $p_j(z)$ as $\vartheta_1^{[j]}, \dots, \vartheta_j^{[j]}$, for $j = 1, \dots, m$, and the roots of q_m as $\xi_1, \dots, \xi_m$, eventual roots at infinity included. It follows from (2.10) that for all $j = 1, \dots, m$ and $i = 1, \dots, j$ we have $\xi_j \neq \vartheta_i^{[j]}$. Otherwise we would have $\boldsymbol{v}_{j+1} \in \mathcal{Q}_j(A, \boldsymbol{v}_1, q_{j-1})$ and V_{j+1} would not be of full column rank. We are ready to prove that $\boldsymbol{q}_{m+1}^{(\xi)} \neq \boldsymbol{0}$. Assume that $\boldsymbol{q}_{m+1}^{(\xi)} = \boldsymbol{0}$. In particular, the first component of $\boldsymbol{q}_{m+1}^{(\xi)}$ is zero, and thus so is $p_0^{[m]}(\xi) = 0$. Hence $\xi \in \{\xi_1, \dots, \xi_m\}$. Since $p_1^{[m]}(\xi) = 0$ and $\xi_1 \neq \vartheta_1^{[1]}$ we may further restrict $\xi \in \{\xi_2, \dots, \xi_m\}$. Looking at $p_j^{[m]}(\xi) = 0$ for the remaining $j = 2, \dots, m$ we exclude one by one all the ξ_j and have $\xi \in \emptyset$. Hence there is no ξ such that $\boldsymbol{q}_{m+1}^{(\xi)} = \boldsymbol{0}$.

Let us now prove that $\left(\boldsymbol{q}_{m+1}^{(\xi)}\right)^* \underline{L_m^{(\xi)}} = \boldsymbol{0}^*$. Note that $\mathcal{R}V_{m+1} = \mathcal{Q}_{m+1}(A, \boldsymbol{v}_1, q_m)$ implies $\mathcal{R}q_m(A)V_{m+1} = \mathcal{K}_{m+1}(A, \boldsymbol{v}_1)$, and further, $\mathcal{K}_{m+1}(A, \boldsymbol{v}_1)$ is A-variant since $\mathcal{R}V_{m+1}$ is. Left-multiplying the RAD (2.2) with $q_m(A)$ yields $Aq_m(A)V_{m+1}\underline{K_m} = q_m(A)V_{m+1}\underline{H_m}$. The columns of $\breve{V}_{m+1} := q_m(A)V_{m+1}$ satisfy $\breve{\boldsymbol{v}}_j = p_{j-1}^{[m]}(A)\boldsymbol{v}_1$ and span the A-variant space $\mathcal{K}_{m+1}(A, \boldsymbol{v}_1)$. The natural isomorphism between $\mathcal{K}_{m+1}(A, \boldsymbol{v}_1)$ and $\mathcal{P}_m$ allows us to write the decomposition in scalar form as

$$z \begin{bmatrix} p_0^{[m]}(z) & p_1^{[m]}(z) & \dots & p_m^{[m]}(z) \end{bmatrix} \underline{K_m} = \begin{bmatrix} p_0^{[m]}(z) & p_1^{[m]}(z) & \dots & p_m^{[m]}(z) \end{bmatrix} \underline{H_m},$$

for all $z \in \mathbb{C}$. If $\xi = 0$ the result follows by setting $z = 0$. Otherwise, using $z = \xi$, and subtracting $\begin{bmatrix} p_0^{[m]}(\xi) & p_1^{[m]}(\xi) & \dots & p_m^{[m]}(\xi) \end{bmatrix} \underline{H_m}$, gives the result for $\xi \neq 0$. $\square$

REMARK 3.3 (Evaluating rational functions). *The introduction of polynomials $p_j^{[m]}$ is necessary only for the case when $q_m(\xi) = 0$, since then $q_m(\xi)^{-1}$ is infinite. When $q_m(\xi) \neq 0$, we can replace equivalently the evaluation $p_j^{[m]}(\xi)$ with the evaluation of the rational functions $r_j(\xi) := q_j(\xi)^{-1}p_j(\xi)$, for all $j = 0, \dots, m$. As $r_0(\xi) = 1$, we see that $\overline{r_j(\xi)} = q_{j+1,m+1}^{(\xi)}/q_{1,m+1}^{(\xi)}$, the ratio of the $(j+1)$st and first element of the vector $\boldsymbol{q}_{m+1}^{(\xi)}$. The rational functions r_j are such that $V_{m+1} = \begin{bmatrix} r_0(A)\boldsymbol{v}_1 & r_1(A)\boldsymbol{v}_1 & \dots & r_m(A)\boldsymbol{v}_1 \end{bmatrix}$, and Theorem 3.2 can be used to evaluate the functions $r_j(z)$ at arbitrary points $z \in \mathbb{C}$.*

4. Moving the poles. Let us give a brief resume. For a fixed rational Krylov space $\mathcal{Q}_{m+1} = \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ the poles are uniquely defined by the starting vector $\boldsymbol{v}$ and, up to scaling of $\boldsymbol{v}$, the reverse is true. Further, by Theorem 2.5, there exists an orthonormal RAD (2.2) spanning $\mathcal{Q}_{m+1}$ with starting vector $\boldsymbol{v}$ and poles q_m. Upon fixing the order of appearance of the poles, Theorem 3.1 guarantees the RAD to be essentially unique.

Observe that $\mathcal{Q}_{m+1}$ can be interpreted as a rational Krylov space with starting vector being almost any vector from $\mathcal{Q}_{m+1}$. Indeed, let $\breve{q}_m \in \mathcal{P}_m^*$ have roots disjoint from $\Lambda(A)$, then

$$\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) = \mathcal{Q}_{m+1}(A, \breve{q}_m(A)q_m(A)^{-1}\boldsymbol{v}, \breve{q}_m). \tag{4.1}$$

We are now interested in transforming an RAD (2.2) for $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ into one for $\mathcal{Q}_{m+1}(A, \breve{q}_m(A)q_m(A)^{-1}\boldsymbol{v}, \breve{q}_m)$. For the moment this is only of theoretical importance, however, in subsection 4.3 we show a connection with implicit filtering and provide references to the literature. Further, an application of the ideas developed here to rational approximation is given in section 5.

To get an RAD for $\mathcal{Q}_{m+1}(A, \breve{q}_m(A)q_m(A)^{-1}\boldsymbol{v}, \breve{q}_m)$ one can focus on either the "new starting vector" or the "new poles". The result is essentially the same and both the starting vector and poles change. We first look at the case when the new starting vector is given as $\breve{\boldsymbol{v}} = V_{m+1}\boldsymbol{c}$, for a nonzero $\boldsymbol{c} \in \mathbb{C}^{m+1}$, and later we focus on the case when the new poles $\breve{q}_m$ are prescribed.

4.1. Moving the poles implicitly. Let $\breve{\boldsymbol{v}} = V_{m+1}\boldsymbol{c} \in \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ be a nonzero vector and take any nonsingular matrix P of size $m+1$ such that $P\boldsymbol{e}_1 = \boldsymbol{c}$. Then

$$A\breve{V}_{m+1}\underline{\breve{K}_m} = \breve{V}_{m+1}\underline{\breve{H}_m}, \tag{4.2}$$

where $\breve{V}_{m+1} = V_{m+1}P$, $\underline{\breve{H}_m} = P^{-1}\underline{H_m}$, and $\underline{\breve{K}_m} = P^{-1}\underline{K_m}$. This construction guarantees the first column $\breve{\boldsymbol{v}}_1$ of $\breve{V}_{m+1}$ to be $\breve{\boldsymbol{v}}$, however, the pencil $\left(\underline{\breve{H}_m}, \underline{\breve{K}_m}\right)$ may loose the upper-Hessenberg structure. In the following we aim at recovering this structure in (4.2) without affecting $\breve{\boldsymbol{v}}_1$. For that purpose we generalize the notion of RADs by first giving a technical definition. For a matrix $\underline{X_m} \in \mathbb{C}^{(m+1)\times m}$ the notation $\underline{X_{-m}}$ is used to denote its lower $m \times m$ submatrix.

Definition 4.1. *Let* $\{\underline{\breve{K}_m}, \underline{\breve{H}_m}\} \subset \mathbb{C}^{(m+1)\times m}$ *be matrices. We say that the pencil* $\left(\underline{\breve{H}_m}, \underline{\breve{K}_m}\right)$ *is* regular *if the lower* $m\times m$ *subpencil* $\left(\underline{\breve{H}_{-m}}, \underline{\breve{K}_{-m}}\right)$ *is regular, i.e.,* $\breve{q}_m(z) = \det\left(z\underline{\breve{K}_{-m}} - \underline{\breve{H}_{-m}}\right)$ *is not identically equal to zero.*

Note that an upper-Hessenberg pencil of size $(m+1)\times m$ is unreduced if and only if it is regular. We are now ready to introduce decompositions of the form (4.2).

Definition 4.2. *A relation of the form* (4.2) *where* $\breve{V}_{m+1}$ *is of full column rank and* $\left(\underline{\breve{H}_m}, \underline{\breve{K}_m}\right)$ *is regular is called a* generalized rational Krylov decomposition. *The generalized eigenvalues of* $\left(\underline{\breve{H}_{-m}}, \underline{\breve{K}_{-m}}\right)$ *are called* poles *of the decomposition. If the poles of* (4.2) *are outside the spectrum* $\Lambda(A)$*, then* (4.2) *is called a* rational Krylov decomposition (RKD).

The notion of *(orthonormal) basis, space* and *equivalent decompositions* are the same as for RADs. We call a generalized RKD with an upper-Hessenberg pencil a *generalized RAD*. The two definitions above let us speculate that the unique poles associated with $\breve{\boldsymbol{v}}$ are the eigenvalues of $\left(\underline{\breve{H}_{-m}}, \underline{\breve{K}_{-m}}\right)$. The justification follows from Theorem 2.5 (or Theorem 3.1) and the following result.

Theorem 4.3. *Any generalized RKD is equivalent to a generalized RAD with the same starting vector.*

Proof. Let (4.2) be a generalized RKD. We need to bring both $\underline{\breve{H}_m}$ and $\underline{\breve{K}_m}$ into upper-Hessenberg form. To achieve this it suffices to bring $\left(\underline{\breve{H}_{-m}}, \underline{\breve{K}_{-m}}\right)$ into generalized Schur form. The existence of unitary $Q_m, Z_m \in \mathbb{C}^{m\times m}$ such that $Q_m^*\underline{\breve{H}_{-m}}Z_m$ and $Q_m^*\underline{\breve{K}_{-m}}Z_m$ are both upper-triangular follows from [15, Theorem 7.7.1]. Multiplying $A\breve{V}_{m+1}\underline{\breve{H}_m} = \breve{V}_{m+1}\underline{\breve{K}_m}$ from the right with Z_m and "inserting" $I_{m+1} = Q_{m+1}Q_{m+1}^*$, where $Q_{m+1} = \mathrm{blkdiag}(1, Q_m)$, we obtain the generalized RAD

$$A\underbrace{\left(\breve{V}_{m+1}Q_{m+1}\right)}_{V_{m+1}}\underbrace{Q_{m+1}^*\underline{\breve{K}_m}Z_m}_{\underline{K_m}} = \underbrace{\left(\breve{V}_{m+1}Q_{m+1}\right)}_{V_{m+1}}\underbrace{Q_{m+1}^*\underline{\breve{H}_m}Z_m}_{\underline{H_m}}.$$

Note that $\mathcal{R}\breve{V}_{m+1} = \mathcal{R}V_{m+1}$ with the poles of $\left(\underline{\breve{H}_m}, \underline{\breve{K}_m}\right)$ and $\left(\underline{H_m}, \underline{K_m}\right)$ being identical. The first vector $\breve{\boldsymbol{v}}_1 = \boldsymbol{v}_1$ is unaffected. $\square$

Algorithm 4.1 Implicit pole placement. RK Toolbox: `move_poles_impl`

Input: Generalized RAD $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ and unit 2-norm $\boldsymbol{c} \in \mathbb{C}^{m+1}$.
Output: Generalized RAD (4.2) spanning $\mathcal{R}V_{m+1}$ with $\breve{\boldsymbol{v}}_1 = V_{m+1}\boldsymbol{c}$.

1. Define $P := I_{m+1} - 2\boldsymbol{u}\boldsymbol{u}^*$, where $\boldsymbol{u} := \boldsymbol{c} \pm \boldsymbol{e}_1$.
2. Find unitary $Q = \text{blkdiag}(1, Q_m)$ and Z, of order $m+1$ and m respectively, such that $Q^*P\underline{H_m}Z$ and $Q^*P\underline{K_m}Z$ are both upper-Hessenberg.
3. Define $\breve{V}_{m+1} := V_{m+1}PQ$, $\underline{\breve{H}_m} := Q^*P\underline{H_m}Z$ and $\underline{\breve{K}_m} := Q^*P\underline{K_m}Z$.

This discussion is summarized in Algorithm 4.1, used to replace the starting vector $\boldsymbol{v}$ with $\breve{\boldsymbol{v}} = V_{m+1}\boldsymbol{c}$. Note that there is no guarantee that by transforming an RAD the resulting decomposition is also an RAD, i.e., some poles may be moved to eigenvalues of A. We prove later (cf. Theorem 4.4) that if $\breve{\boldsymbol{v}} = V_{m+1}\boldsymbol{c} = p_m(A)q_m(A)^{-1}\boldsymbol{v}$ then the poles of the decomposition are always the roots of p_m, even if they coincide with eigenvalues of A.

4.2. Moving the poles explicitly. If the vector $\breve{\boldsymbol{v}}$ is not given as a linear combination $\breve{\boldsymbol{v}} = V_{m+1}\boldsymbol{c}$ of the basis vectors V_{m+1} but rather by specifying the new poles $\breve{q}_m$ one can compute $\boldsymbol{c} = V_{m+1}^*\breve{\boldsymbol{v}}$, where $\breve{\boldsymbol{v}} = \breve{q}_m(A)q_m(A)^{-1}\boldsymbol{v}$, and still use Alg. 4.1 to recover the new decomposition. The vector $\breve{\boldsymbol{v}} = \breve{q}_m(A)q_m(A)^{-1}\boldsymbol{v}$ can be computed cheaply as a rational Arnoldi approximation $\breve{\boldsymbol{v}} = V_{m+1}\breve{q}_m(A_{m+1})q_m(A_{m+1})^{-1}V_{m+1}^*\boldsymbol{v}$, where $A_{m+1} := V_{m+1}^*AV_{m+1}$, see for instance [22]. In the following we present an approach that works directly with the pencil $(\underline{H_m}, \underline{K_m})$, changing the poles iteratively one by one, and thence requires no information about the reduced matrix A_{m+1}.

Moving the first pole. The poles are the ratios of the subdiagonal elements of $(\underline{H_m}, \underline{K_m})$. Applying a Givens rotation G acting on planes $(1, 2)$ from the left of the pencil does not destroy the upper-Hessenberg structure and, as we show, can move the first pole anywhere. We now derive the formulas for $s = e^{i\phi}\sin\vartheta$ and $c = \cos\vartheta$ satisfying $c^2 + |s|^2 = 1$ and such that the Givens rotation

$$G = \text{blkdiag}\left(\begin{bmatrix} c & -s \\ \overline{s} & c \end{bmatrix}, I_{m-1}\right)$$

replaces the pole ξ_1 with $\breve{\xi}_1$ when applied to the pencil from the left. Define $\underline{\breve{H}_m} = G\underline{H_m}$ and $\underline{\breve{K}_m} = G\underline{K_m}$. This gives

$$\begin{aligned} \breve{h}_{11} &= ch_{11} - sh_{21}, \quad & \breve{k}_{11} &= ck_{11} - sk_{21}, \\ \breve{h}_{21} &= \overline{s}h_{11} + ch_{21}, \quad & \breve{k}_{21} &= \overline{s}k_{11} + ck_{21}. \end{aligned} \tag{4.3}$$

Additionally, G is chosen so that $\breve{\xi}_1 = \breve{h}_{21}/\breve{k}_{21}$. Using the notation $t = \overline{s}/c$, we derive

$$t = \begin{cases} k_{21}/k_{11}, & \breve{\xi}_1 = \infty, \\ (\breve{\xi}_1 k_{21} - h_{21})/(h_{11} - \breve{\xi}_1 k_{11}), & \breve{\xi}_1 \neq \infty. \end{cases}$$

Using standard trigonometric relations we arrive at

$$s = \frac{t}{\sqrt{1 + |t|^2}}, \qquad c = \frac{1}{\sqrt{1 + |t|^2}}$$

if $t \neq \infty$, and otherwise, $s = 1$ and $c = 0$.

Formula (4.1) asserts (with the roots of $\breve{q}_m$ being $\breve{\xi}_1, \xi_2, \ldots, \xi_m$) that this process replaces the starting vector $\boldsymbol{v}_1$ with a multiple of $\big(A - \breve{\xi}_1 I\big)\big(A - \xi_1 I\big)^{-1}\boldsymbol{v}_1$, where for notational convenience only we assume both ξ_1 and $\breve{\xi}_1$ to be finite. Let us verify that. Define $\breve{V}_{n+1} = V_{n+1}G^*$. In particular,

$$\breve{\boldsymbol{v}}_1 = c\boldsymbol{v}_1 - \overline{s}\boldsymbol{v}_2. \tag{4.4}$$

Recall that (2.6) reads $\big(h_{21}I - k_{21}A\big)\boldsymbol{v}_2 = \big(k_{11}A - h_{11}I\big)\boldsymbol{v}_1$. Hence, using the relation (2.6) within (4.4) together with (4.3) provides

$$\big(h_{21}I - k_{21}A\big)\breve{\boldsymbol{v}}_1 = \big[c\big(h_{21}I - k_{21}A\big) - \overline{s}\big(k_{11}A - h_{11}I\big)\big]\,\boldsymbol{v}_1 = \big(\breve{h}_{21}I - \breve{k}_{21}A\big)\boldsymbol{v}_1. \tag{4.5}$$

Note that (4.5) holds even if $\breve{\xi}_1 \in \Lambda(A)$ and/or $\xi_1 \in \Lambda(A)$, as long as the starting decomposition exists. If however $\xi_1 \notin \Lambda(A)$ we can further write

$$\breve{\boldsymbol{v}}_1 = \big(\breve{h}_{21}I - \breve{k}_{21}A\big)\big(h_{21}I - k_{21}A\big)^{-1}\boldsymbol{v}_1.$$

Moving all poles. Changing the other ratios with Givens rotations results in the loss of the upper-Hessenberg structure. However, the poles are the eigenvalues of the pencil $\big(\underline{H_{-m}}, \underline{K_{-m}}\big)$ which is (already) in generalized Schur form. After changing the first pole, using the Givens rotation approach just described, the poles can be reordered (see for instance [25, 26]) with the aim of bringing an unchanged pole to the front of the decomposition so that it can be changed using a Givens rotation. This process is formalized in Algorithm 4.2 and an illustration is presented in Figure 4.1.

Let us now consider Alg. 4.2 when $k = m$. For notational convenience only, we assume the poles to be finite. As we have shown in (4.5), after applying the first Givens rotation the starting vector $\boldsymbol{v}_1$ gets replaced with $\boldsymbol{v}_1^{[1]}$ satisfying

$$\big(A - \xi_1 I\big)\boldsymbol{v}_1^{[1]} = \gamma_1\big(A - \breve{\xi}_m I\big)\boldsymbol{v}_1, \tag{4.6}$$

where $\gamma_1 \in \mathbb{C}^*$ is a scaling factor. By reordering the poles we do not affect the "new starting vector" $\boldsymbol{v}_1^{[1]}$ and bring ξ_2 to the leading positions, i.e., second row, first column, where the next Givens rotation acts. Thus, for $j = 2$ the Givens rotation replaces $\boldsymbol{v}_1^{[1]}$ with $\boldsymbol{v}_1^{[2]}$ satisfying $\big(A - \xi_2 I\big)\boldsymbol{v}_1^{[2]} = \gamma_2\big(A - \breve{\xi}_{m-1}I\big)\boldsymbol{v}_1^{[1]}$, for some $\gamma_2 \in \mathbb{C}^*$. Using (4.6) we obtain

$$\big(A - \xi_1 I\big)\big(A - \xi_2 I\big)\boldsymbol{v}_1^{[2]} = \gamma_1\gamma_2\big(A - \breve{\xi}_{m-1}I\big)\big(A - \breve{\xi}_m I\big)\boldsymbol{v}_1.$$

Reasoning inductively we deduce

$$q_m(A)\breve{\boldsymbol{v}}_1 = \gamma\breve{q}_m(A)\boldsymbol{v}_1, \tag{4.7}$$

where $\gamma \in \mathbb{C}^*$ is a scalar, $\breve{\boldsymbol{v}}_1 = \boldsymbol{v}_1^{[m]}$, q_m is given by (2.4), and $\breve{q}_m$ is defined in an analogous manner. The above discussion is the gist of the following result.

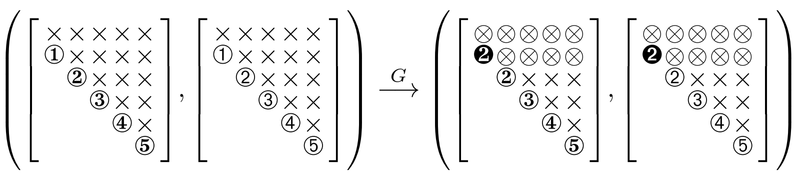
THEOREM 4.4. *Let $\mathcal{Q}_{m+1} = \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ be A-variant. If the generalized RKD $A\breve{V}_{m+1}\underline{\breve{K}_m} = \breve{V}_{m+1}\underline{\breve{H}_m}$ with poles $\breve{q}_m$ spans $\mathcal{Q}_{m+1}$ then $\breve{\boldsymbol{v}}_1 = \gamma\breve{q}_m(A)q_m(A)^{-1}\boldsymbol{v}$ with a scalar $\gamma \in \mathbb{C}^*$. Alternatively, if $\breve{\boldsymbol{v}}_1 = \breve{q}_m(A)q_m(A)^{-1}\boldsymbol{v}$ then there exists a generalized RKD $A\breve{V}_{m+1}\underline{\breve{K}_m} = \breve{V}_{m+1}\underline{\breve{H}_m}$ with poles $\breve{q}_m$ spanning $\mathcal{Q}_{m+1}$.*

Proof. If $A\breve{V}_{m+1}\underline{\breve{K}_m} = \breve{V}_{m+1}\underline{\breve{H}_m}$ spans $\mathcal{Q}_{m+1}$ we can transform it into an equivalent generalized RAD (cf. Theorem 4.3) and then, using Alg. 4.2, into $AV_{m+1}\underline{K_m} =$

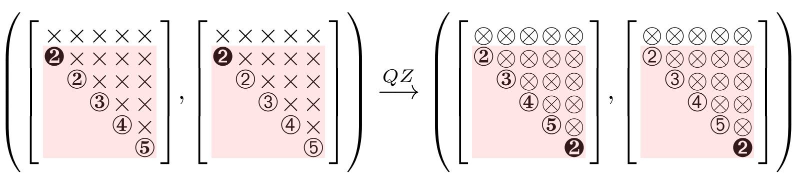
Algorithm 4.2 Explicit pole placement. RK Toolbox: `move_poles_expl`

Input: Generalized RAD $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ and $\breve{\boldsymbol{\xi}} = \{\breve{\xi}_j\}_{j=1}^{k} \subset \overline{\mathbb{C}}$, $1 \le k \le m$.
Output: Generalized RAD (4.2) spanning $\mathcal{R}V_{m+1}$ and having poles $\breve{\boldsymbol{\xi}} \cup \{\xi_j\}_{j=k+1}^{m}$.

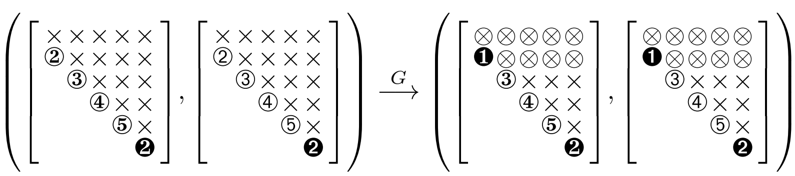
1. Set $\breve{V}_{m+1} := V_{m+1}$, $\underline{\breve{H}_m} := \underline{H_m}$, and $\underline{\breve{K}_m} := \underline{K_m}$.
2. Label $\xi_j := h_{j+1,j}/k_{j+1,j}$ for $j = 1, \ldots, k$.
3. **for** $j = 1, \ldots, k$ **do**
4. Find Givens rotation G to replace the pole ξ_j with $\breve{\xi}_\ell$ where $\ell = k - j + 1$.
5. Update $\breve{V}_{m+1} := \breve{V}_{m+1}G^*$, $\underline{\breve{H}_m} := G\underline{\breve{H}_m}$, and $\underline{\breve{K}_m} := G\underline{\breve{K}_m}$.
6. Find unitary $Q_{m+1} = \mathrm{blkdiag}(1, Q_\ell, I_{m-\ell})$ and $Z_m = \mathrm{blkdiag}(Z_\ell, I_{m-\ell})$, to circularly shift the ℓ poles from position $(2,1)$ to position $(\ell+1,\ell)$ for one place forward so that $\breve{\xi}_\ell$ gets pushed back to position $(\ell+1,\ell)$.
7. Update $\breve{V}_{m+1} := \breve{V}_{m+1}Q_{m+1}$, $\underline{\breve{H}_m} := Q_{m+1}^*\underline{\breve{H}_m}Z_m$, and $\underline{\breve{K}_m} := Q_{m+1}^*\underline{\breve{K}_m}Z_m$.
8. **end for**



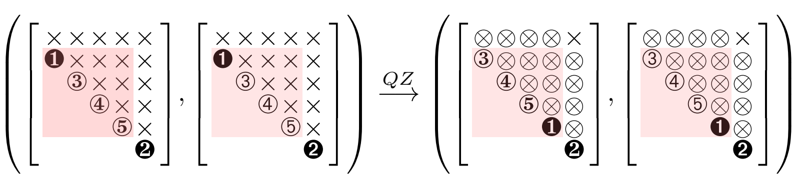
(a) Appling the first Givens rotation to replace ξ_1 = ①/① with $\breve{\xi}_2$ = ❷/❷.



(b) Reordering the generalized Schur form of the lower 5×5 part.



(c) Appling the second Givens rotation to replace ξ_2 = ②/② with $\breve{\xi}_1$ = ❶/❶.



(d) Reordering the generalized Schur form of size 4.

Fig. 4.1: Looking at the 6×5 upper-Hessenberg pencil while Algorithm 4.2 is applied on the corresponding RAD with $k = 2$. The original poles are the ratios ①/①,…⑤/⑤. The first two poles are replaced with ❷/❷ and ❶/❶. The transition from $\times$ to $\otimes$ symbolizes that the element potentially changes.

$V_{m+1}\underline{H_m}$, having poles q_m and still spanning $\mathcal{Q}_{m+1}$. According to Lemma 2.1, $\boldsymbol{v}_1$ is collinear with $\boldsymbol{v}$. Therefore, it follows from (4.7) that $\breve{\boldsymbol{v}}_1 = \gamma \breve{q}_m(A) q_m(A)^{-1}\boldsymbol{v}$ for some scalar $\gamma \in \mathbb{C}^*$. The other direction follows from Theorem 2.5 and (4.7) after using Alg. 4.2. ∎

Theorem 4.4 shows that Alg. 4.1 and Alg. 4.2 are equivalent, provided that equivalent input data are given. It also shows, together with Theorem 2.5 and Theorem 4.3, that an $m+1$-dimensional space $\mathcal{V}_{m+1}$ is a rational Krylov space if and only if there exist a generalized RKD spanning $\mathcal{V}_{m+1}$.

REMARK 4.5 (Recovering the polynomial Krylov space). *With $\breve{q}_m(z) = 1$, there holds $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m) = \mathcal{Q}_{m+1}(A, q_m(A)^{-1}\boldsymbol{v}, \breve{q}_m) = \mathcal{K}_{m+1}(A, q_m(A)^{-1}\boldsymbol{v})$, and we can recover a polynomial Arnoldi decomposition for $\mathcal{K}_{m+1}(A, q_m(A)^{-1}\boldsymbol{v})$ from an RAD for $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ using Alg. 4.2 with all poles $\breve{\xi}_j$ set to infinity. In this particular case, a simpler method is to bring the pencil $(\underline{H_m}, \underline{K_m})$ from upper-Hessenberg–upper-Hessenberg to upper-Hessenberg–upper-triangular form using just Givens rotations; see, e.g., [34, p. 495]. Bringing the pencil to upper-triangular–upper-Hessenberg form would move all the poles to zero.*

4.3. Implicit filters in the rational Krylov method. Implicit filtering aims at compressing the space $\mathcal{Q}_{m+1}(A, \boldsymbol{v}_1, q_m)$ into $\mathcal{Q}_{m+1-k}(A, p_k(A)q_k(A)^{-1}\boldsymbol{v}_1, \breve{q}_{m-k})$, where $1 \le k \le m$, $q_m = q_k \cdot \breve{q}_{m-k}$, and $p_k \in \mathcal{P}_k$ is a polynomial with roots (infinity allowed) in the region we want to filter out. In applications this technique is usually used to deal with large memory requirements or orthogonalization costs for V_{m+1}, or to purge unwanted or spurious eigenvalues (see, e.g., [4, 7, 8] and the references given therein). Implicit filtering for RADs was first introduced in [8] and further studied in [7]. Alg. 4.2 can easily be used for implicit filtering. In fact, applying Alg. 4.2 with the k poles $\breve{\xi}_j$ being the roots of p_k implicitly applies the filter $p_k(A)q_k(A)^{-1}$ to the RAD. The k "new" poles correspond to the rightmost k columns in $\breve{V}_{m+1}$, $\underline{\breve{K}_m}$ and $\underline{\breve{H}_m}$, cf. Figure 4.1. Hence, truncating the decomposition to the leading $m+1-k$ columns completes the process. The derivation and algorithms in [7, 8] are different, and it would perhaps be interesting to compare them. This is, however, not done here. Pertinent ideas for polynomial Krylov methods have recently appeared in [4] where the authors relate implicit filtering in the Krylov–Schur algorithm [35, 37] with partial eigenvalue assignment.

As an alternative to Alg. 3.1 for a Hermitian matrix A, it was proposed in [28] to use the spectral transformation Lanczos method with change of (the repeated) pole. The approach for changing poles taken here is different and more general.

5. An application to rational least squares approximation. Given matrices $\{A, F\} \subset \mathbb{C}^{N\times N}$ and a unit 2-norm vector $\boldsymbol{v} \in \mathbb{C}^N$, we consider in this section the following rational least squares problem: find a rational function $R_m^{\min}(z)$ of type (m,m), with $m < M$, such that

$$\|F\boldsymbol{v} - R_m(A)\boldsymbol{v}\|_2 \to \min. \tag{5.1}$$

This is a nonlinear approximation problem as the denominator of $R_m^{\min}$ is unknown. Hence an iterative algorithm is required.

Let $q_m \in \mathcal{P}_m$ be a given polynomial and consider the linear space of rational functions of type (m,m) with denominator q_m, denoted by $\mathcal{P}_m/q_m$. Each element $R_m \in \mathcal{P}_m/q_m$ is in a one-to-one correspondence with an element $R_m(A)\boldsymbol{v}$ of the rational Krylov space $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$. Instead of (5.1) we now consider a linear

Algorithm 5.1 Rational Krylov fitting (RKFIT). RK Toolbox: `rkfit`

Input: $\{A, F\} \subset \mathbb{C}^{N\times N}$, $\boldsymbol{v} \in \mathbb{C}^N$ and poles $\{\xi_j\}_{j=1}^m \subset \overline{\mathbb{C}} \setminus \Lambda(A)$ with $m < M$.
Output: Poles $\xi_1, \xi_2, \dots, \xi_m$ and coefficient vector $\boldsymbol{c} \in \mathbb{C}^{m+1}$.

1. **repeat**
2. Compute RAD $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with $\boldsymbol{v}_1 = \boldsymbol{v}/\|\boldsymbol{v}\|_2$ and poles $\{\xi_j\}_{j=1}^m$.
3. Compute a right singular vector $\boldsymbol{c} \in \mathbb{C}^{m+1}$ of $(FV_{m+1} - V_{m+1}V_{m+1}^*FV_{m+1})$ corresponding to a smallest singular value $\sigma_{\min}$.
4. Use Alg. 4.1 to form $A\breve{V}_{m+1}\underline{\breve{H}_m} = \breve{V}_{m+1}\underline{\breve{K}_m}$ spanning $\mathcal{R}V_{m+1}$ with $\breve{\boldsymbol{v}}_1 = V_{m+1}\boldsymbol{c}$.
5. Obtain new poles $\xi_1, \xi_2, \dots, \xi_m$ as the poles of $(\underline{\breve{H}_m}, \underline{\breve{K}_m})$.
6. **until** $\sigma_{\min}$ is small enough or a maximal iteration count is exceeded.
7. Compute RAD $AV_{m+1}\underline{K_m} = V_{m+1}\underline{H_m}$ with $\boldsymbol{v}_1 = \boldsymbol{v}/\|\boldsymbol{v}\|_2$ and poles $\{\xi_j\}_{j=1}^m$.
8. Compute $\boldsymbol{c} = V_{m+1}^*F\boldsymbol{v}$.

approximation problem: find a unit 2-norm vector $\breve{\boldsymbol{v}} \in \mathcal{Q}_{m+1} = \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$ as

$$\breve{\boldsymbol{v}} = \underset{\substack{\boldsymbol{y}\in\mathcal{Q}_{m+1}\\ \|\breve{\boldsymbol{v}}\|_2=1}}{\operatorname{argmin}} \min_{R_m\in\mathcal{P}_m/q_m} \|F\boldsymbol{y} - R_m(A)\boldsymbol{v}\|_2. \tag{5.2}$$

This means that $F\breve{\boldsymbol{v}}$ is best approximated by an element of $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$. Problem (5.2) is easy to solve. Let $V_{m+1} \in \mathbb{C}^{N\times(m+1)}$ be an orthonormal basis of $\mathcal{Q}_{m+1}$ and write $\boldsymbol{y} = V_{m+1}\boldsymbol{c}$ with $\boldsymbol{c} \in \mathbb{C}^{m+1}$ and $\|\boldsymbol{c}\|_2 = 1$. Then the inner minimum in (5.2) is a linear least squares problem whose solution $R_m(A)\boldsymbol{v}$ is given by orthogonal projection of $F\boldsymbol{y}$ onto $\mathcal{Q}_{m+1}$, i.e., $R_m(A)\boldsymbol{v} = V_{m+1}V_{m+1}^*F\boldsymbol{y}$ minimizes $\|F\boldsymbol{y} - R_m(A)\boldsymbol{v}\|_2$. Hence (5.2) reduces to

$$\begin{aligned}\breve{\boldsymbol{v}} &= \underset{\substack{\boldsymbol{y}=V_{m+1}\boldsymbol{c}\\ \|\boldsymbol{c}\|_2=1}}{\operatorname{argmin}} \|(I - V_{m+1}V_{m+1}^*)F\boldsymbol{y}\|_2\\ &= \underset{\|\boldsymbol{c}\|_2=1}{\operatorname{argmin}} \|(I - V_{m+1}V_{m+1}^*)FV_{m+1}\boldsymbol{c}\|_2.\end{aligned}$$

A minimizer $\boldsymbol{c}$ can be obtained as a right singular vector of $(I - V_{m+1}V_{m+1}^*)FV_{m+1}$ corresponding to a smallest singular value $\sigma_{\min}$. We now exploit that by Theorem 4.4 we can associate with $\breve{\boldsymbol{v}}$ a "new" rational Krylov space $\mathcal{Q}_{m+1}(A, \breve{\boldsymbol{v}}, \breve{q}_m) = \mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$, where the roots of $\breve{q}_m$ (the "new" poles) can be computed from $\boldsymbol{c}$ using Alg. 4.1. The new vector–pole pair $(\breve{\boldsymbol{v}}, \breve{q}_m)$ is optimal in the sense that $\|F\breve{\boldsymbol{v}} - \breve{R}_m(A)\breve{\boldsymbol{v}}\|_2$ is minimal (and equal to $\sigma_{\min}$) among all $\breve{R}_m \in \mathcal{P}_m/\breve{q}_m$, and there is no better vector–pole pair associated with $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$. Replacing $\breve{\boldsymbol{v}}$ back to $\boldsymbol{v}$, we hope that the new rational Krylov space $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, \breve{q}_m)$ contains a better approximation to $F\boldsymbol{v}$ than $\mathcal{Q}_{m+1}(A, \boldsymbol{v}, q_m)$. In this case we have found an improved denominator $\breve{q}_m$ for the rational function R_m in (5.1).

The procedure described is iterated, computing a new rational Krylov space at each iteration and changing the poles by modifying the starting vector. The complete procedure is given in Algorithm 5.1 under the name RKFIT, which stands for *Rational Krylov Fitting*. A MATLAB implementation of RKFIT is available in [2].

Discussion. In this and the following subsections we briefly discuss Alg. 5.1 in a list of comments and some numerical experiments. A more detailed analysis will be given in a separate publication [3].

1. If $A = \mathrm{diag}(\lambda_j)$ and $F = \mathrm{diag}(\varphi_j)$ are diagonal matrices, and $\boldsymbol{v} = [v_1, \ldots, v_N]^T$, then (5.1) corresponds to a rational weighted least squares problem

$$\|F\boldsymbol{v} - R_m(A)\boldsymbol{v}\|_2^2 = \sum_{j=1}^{N} |v_j|^2 \cdot |\varphi_j - R_m(\lambda_j)|^2 \to \min .$$

General rational optimization problems of this type are nonconvex and hence no iterative method for their solution comes with a guarantee to find a global minimum. One popular approach for solving such problems is known as *vector fitting* [20, 19], which similarly to Alg. 5.1 is based on iterative relocation of the poles of rational least squares approximants in partial fraction form.

2. The use of partial fractions as basis functions in [20] may result in poorly conditioned linear algebra problems to be solved, and *orthonormal vector fitting* [10] tries to overcome this problem by using instead an expansion of R_m into orthonormal rational functions; orthonormal with respect to a measure supported on the imaginary axis. The orthonormal rational functions used in [10] are computed by a Gram–Schmidt procedure applied with a quadrature approximation for the inner products. The implications of quadrature-based vector fitting on optimal $\mathcal{H}_2$ approximation of dynamical systems are discussed in [11].

3. In the case that A is a normal matrix, RKFIT can be interpreted as an orthonormal vector fitting algorithm where two rational functions r_m and $\breve{r}_m$ with common denominator q_m are orthonormal with respect to a discrete inner product defined as

$$\langle r_m, \breve{r}_m \rangle := (\breve{r}_m(A)\boldsymbol{v})^*(r_m(A)\boldsymbol{v})/\|\boldsymbol{v}\|_2^2.$$

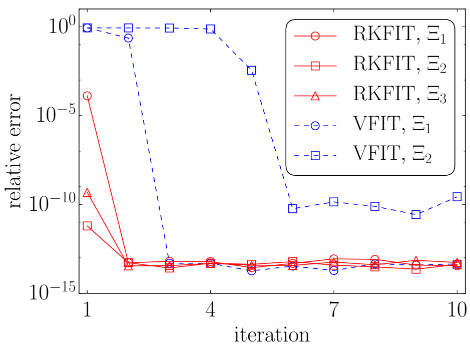
(See [5, 9] for the theory of orthogonal rational functions and their relation to rational Krylov spaces.) RKFIT is different from orthonormal vector fitting in that it uses a discrete inner product defined by a measure not necessarily supported on the imaginary axis. The orthonormal rational functions are computed by the rational Krylov method without the need for explicit quadrature. In [6] it is advocated to use orthogonal rational basis functions with fixed poles for least squares fitting. This leads to a linear least squares problem but does not resolve the problem of pole relocation.

4. If A has Jordan blocks of size 2 or greater then also derivatives of $R_m(z)$ at (some of) the eigenvalues of A are fitted. Consider, for example,

$$A = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix}, \text{ and } R_m(A) = \begin{bmatrix} R_m(\lambda) & R_m'(\lambda) & \cdots & \frac{R_m^{(N-1)}(\lambda)}{(N-1)!} \\ & R_m(\lambda) & \ddots & \vdots \\ & & \ddots & R_m'(\lambda) \\ & & & R_m(\lambda) \end{bmatrix}.$$

Then each component of $R_m(A)\boldsymbol{v}$ is a weighted sum of derivatives $R_m^{(j)}(\lambda)$ and one can use $\boldsymbol{v}$ to choose the weights as required. This generalizes naturally to matrices A with more than one Jordan block.

5. If $F = f(A)$, each iteration of Alg. 5.1 requires the computation of $f(A)V_{m+1}$. If one is interested in scalar rational approximation problems where A is a diagonal matrix (see point 1), or a Jordan block matrix (see point 4), then $f(A)$ is easy to compute. Otherwise rational Krylov techniques can be used to approximate $f(A)V_{m+1}$ directly (see, e.g., the review [22]).



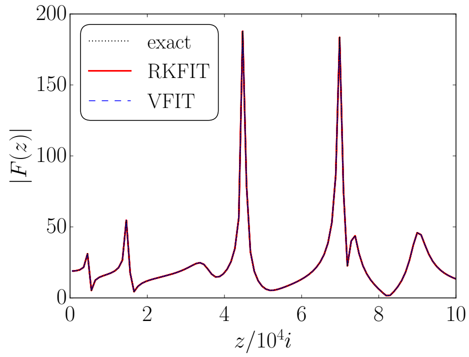


Fig. 5.1: Least-squares approximation of a rational function $F(z)$ of type $(19, 18)$ using RKFIT and the vector fitting code VFIT. Left: Relative error $\|F(A)\boldsymbol{v} - R_m(A)\boldsymbol{v}\|_2/\|F(A)\boldsymbol{v}\|_2$ after each iteration of RKFIT (solid red) and VFIT (dashed blue). The two convergence curves for each method correspond to different choices for the initial poles Ξ_1 (circles), Ξ_2 (squares), and Ξ_3 (triangles), respectively. Right: Plot of $|F(z)|$ over an interval on the imaginary axis overlaid with the approximants $|R_m(z)|$ obtained after 10 iterations of RKFIT and VFIT with initial poles Ξ_1 (the curves are visually indistinguishable).

6. If A, F, and $\boldsymbol{v}$ are real-valued and the initial poles $\{\xi_j\}_{j=1}^m$ appear in complex conjugate pairs, it is natural to enforce real arithmetic in all operations of Alg. 5.1. This can be achieved by using the real form of the rational Krylov method [30] instead of Alg. 3.1, and a generalized real Schur form (see, e.g., [36, § 3.1]) in Step 2 of Alg. 4.1. Our RKFIT implementation [2] provides a `'real'` option for this purpose.

7. The output vector $\boldsymbol{c} \in \mathbb{C}^{m+1}$ returned by Alg. 5.1 collects the coefficients of the approximant $R_m(A)\boldsymbol{v}$ in the rational Krylov basis V_{m+1}, i.e., $R_m(A)\boldsymbol{v} = V_{m+1}\boldsymbol{c}$. Using Theorem 3.2 and Remark 3.3 we find that $R_m(z)$ can be evaluated for any point $z \in \mathbb{C}$ by computing a full QR factorization of $z\underline{K_m} - \underline{H_m}$ and forming an inner product of $\boldsymbol{c}$ with the last column $\boldsymbol{q}_{m+1}^{(z)}$ of the Q factor scaled by its first entry, i.e.,

$$R_m(z) = \frac{(\boldsymbol{q}_{m+1}^{(z)})^* \boldsymbol{c}}{(\boldsymbol{q}_{m+1}^{(z)})^* \boldsymbol{e}_1}.$$

In the following we discuss three experiments with the aim of providing further insight and showing the applicability of RKFIT. Accompanying MATLAB scripts to reproduce these experiments are available as part of [2]. All computations were performed with MATLAB version R2013a on an Intel Core i5-3570 processor running Scientific Linux, Release 6.4 (Carbon).

5.1. Experiment 1: Fitting an artificial frequency response. We first consider a diagonal matrix $A \in \mathbb{C}^{N \times N}$ with $N = 200$ logarithmically spaced eigenvalues in the interval $[10^{-5}i, 10^5 i]$. The matrix $F = F(A)$ is a rational matrix function of type $(19, 18)$ given in partial fraction form in [20, subsection 4.1], and $\boldsymbol{v} = [1, 1, \ldots, 1]^T$. We compare RKFIT to the vector fitting code VFIT [20, 19] which is available online.[3]

We consider three different sets of starting poles, namely

[3] `http://www.sintef.no/Projectweb/VECTFIT/Downloads/VFUT3/` as of November 2014.

- Ξ_1: 9 log-spaced poles in $[10^3 i, 10^5 i]$ and their complex conjugates;
- Ξ_2: 12 log-spaced poles in $[10^6 i, 10^9 i]$ and their complex conjugates;
- Ξ_3: 18 infinite poles (applicable to RKFIT only);

and run 10 iterations of RKFIT and VFIT, respectively.

The numerical results are shown in Figure 5.1. On the left we see the relative error $\|F(A)\boldsymbol{v} - R_m(A)\boldsymbol{v}\|_2/\|F(A)\boldsymbol{v}\|_2$ after each iteration. We observe that RKFIT converges within the first 2 iterations for all three sets of initial poles Ξ_1, Ξ_2, and Ξ_3. VFIT requires 3 iterations starting with Ξ_1 and it fails to converge within 10 iterations when being initialized with the poles Ξ_2. In the later case MATLAB warnings about solves of close-to-singular linear systems seem to indicate that the partial fraction basis used in VFIT is ill-conditioned. RKFIT, on the other hand, always uses discrete orthonormal rational bases and performs robustly with respect to changes in the initial poles. The choice of infinite initial poles Ξ_3 is interesting in that it requires no a-priori knowledge of the pole location (choosing all poles to be infinite is not possible in the available VFIT code). On the right of Figure 5.1 we show a plot $|F(z)|$ over an interval on the imaginary axis together with the RKFIT and VFIT approximants $|R_m(z)|$. This plot essentially coincides with [20, Figure 1] (it does not exactly coincide as apparently the figure in that paper has been produced with a smaller number of sampling points, causing some "spikes" to be missed or reduced).

5.2. Experiment 2: Square root of a symmetric matrix. We consider the approximation of $F\boldsymbol{v}$ with the matrix square root $F = A^{1/2}$, $A = \mathrm{tridiag}(-1, 2, -1) \in \mathbb{R}^{100\times 100}$, and $\boldsymbol{v} = [1, 0, \ldots, 0]^T$. Again, we test different sets of initial poles, namely

- Ξ_1: 16 log-spaced poles in $[-10^8, -10^{-8}]$;
- Ξ_2: 16 linearly spaced poles in $[0, 4]$;
- Ξ_3: 16 infinite poles (applicable to RKFIT only).

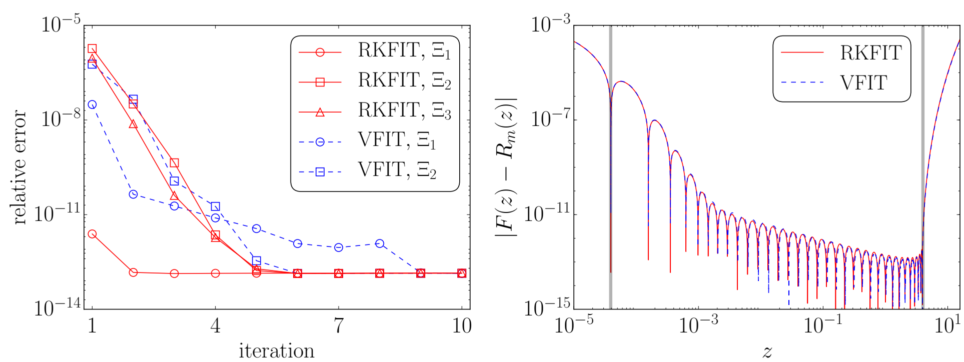


Fig. 5.2: Least-squares approximation of the function $F(z) = z^{1/2}$ using RKFIT and the vector fitting code VFIT. Left: This plot shows the relative approximation error $\|F(A)\boldsymbol{v} - R_m(A)\boldsymbol{v}\|_2/\|F(A)\boldsymbol{v}\|_2$ after each iteration of RKFIT (solid red) and VFIT (dashed blue). The convergence curves for each method correspond to different choices for the initial poles Ξ_1 (circles), Ξ_2 (squares), and Ξ_3 (triangles), respectively. Right: This is the plot of $|F(z) - R_m(z)|$ over an interval on the positive real axis obtained after 10 iterations of RKFIT and VFIT with initial poles Ξ_1. The vertical lines indicate the spectral interval of A.

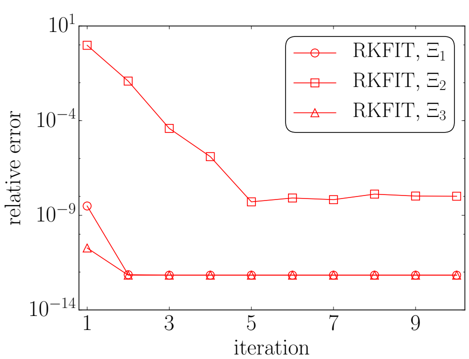
Note that the initial poles Ξ_1 are placed on the branch cut of $z^{1/2}$, which is a reasonable initial guess for the poles of R_m. Some of the poles Ξ_2 are located very close to the eigenvalues of A whose spectral interval is approximately $[0, 4]$. The convergence of the relative error per iteration of RKFIT and VFIT is shown on the left of Figure 5.2. In order to use VFIT for this problem we have diagonalized A and provided the code with weights corresponding to the components of $\boldsymbol{v}$ in the eigenvector basis of A. All tests converge within at most 9 iterations, with the fastest convergence achieved by RKFIT with initial guess Ξ_1. On the right of Figure 5.2 we show the error $|z^{1/2} - R_m(z)|$ over an interval containing the spectrum of A.

5.3. Experiment 3: Exponential of a nonnormal matrix. We consider the approximation of $F\boldsymbol{v}$ with the matrix exponential $F = \exp(A)$ of a Grcar matrix A of size $N = 100$ generated in MATLAB via `A = -5*gallery('grcar',N,3)`. The eigenvalues and 10^{-6}-pseudospectrum of A are shown on the right of Figure 5.3. The vector is $\boldsymbol{v} = [1, 1, \ldots, 1]^T$ and we consider different sets of initial poles for RKFIT,

- Ξ_1: 16 poles equal to 0;
- Ξ_2: 16 poles equal to -10;
- Ξ_3: 16 infinite poles.

Note that A is not diagonalizable and therefore VFIT cannot be applied as in the previous two experiments. On the left of Figure 5.3 we observe excellent convergence of RKFIT within 2 iterations starting with the initial poles Ξ_1 and Ξ_3.

With the initial poles Ξ_2 the error stagnates on a higher level, possibly trapped nearby a non-global minimum. As is the case with any nonlinear iteration, RKFIT is not guaranteed to converge to a global minimum (if it even exists). We currently do not have a good explanation why the initial guess Ξ_2 is bad, but we have verified that $\xi = -10$ lies in the 10^{-6}-pseudospectrum of A and hence the initial rational Krylov space may have too large components in just a few eigendirections of A.



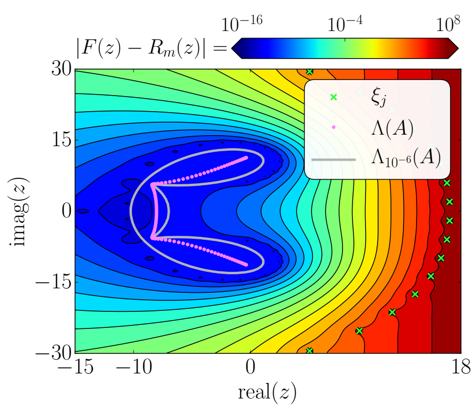


Fig. 5.3: Least-squares approximation of the function $F(z) = \exp(z)$ using RKFIT. Left: This plot shows the relative approximation error $\|F(A)\boldsymbol{v} - R_m(A)\boldsymbol{v}\|_2/\|F(A)\boldsymbol{v}\|_2$ after each iteration of RKFIT (solid red) for different choices of initial poles Ξ_1, Ξ_2, and Ξ_3, respectively. Right: A plot of $|F(z) - R_m(z)|$ over a region in the complex plane together with the poles of R_m (green crosses), where R_m is the rational least squares approximant obtained after 10 iterations of RKFIT with initial poles Ξ_1. The eigenvalues of the Grcar matrix and its 10^{-6}-pseudospectrum are also shown.

6. Summary and future work. We introduced the notion of generalized rational Krylov decompositions and studied their connections with rational Krylov spaces. We generalized the implicit Q theorem to the rational case and have provided some insight for the continuation combination proposed by Ruhe [32] for building rational Krylov spaces. Algorithms for transforming generalized RKDs and thereby changing the poles and starting vector of the associated spaces were presented. These algorithms, in particular Alg. 4.2, can also be employed for implicit restarting in polynomial and rational Krylov decompositions and even eigenvalue assignment, cf. [4]. A comparison with existing algorithms for the same purpose might be interesting.

We introduced the RKFIT method for rational least squares fitting. A detailed study of the convergence properties of RKFIT will be subject of future work. We will extend the MATLAB code in [2] to return the computed rational approximant in partial fraction form, although this conversion itself may be ill-conditioned in particular when the poles of the approximant are far away from the eigenvalues of A. A further extension of RKFIT will handle "lucky breakdowns" in the case when the rational Krylov space becomes (nearly) A-invariant. It should be possible to robustify RKFIT as it was done for linearized least squares and Padé approximation in [17, 16], where close-to-zero singular values lead to a reduction of the approximation degree.

REFERENCES

[1] 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.

[2] M. Berljafa and S. Güttel, *A Rational Krylov Toolbox for MATLAB*, MIMS EPrint 2014.56, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2014. Available for download at `http://guettel.com/rktoolbox/`.

[3] M. Berljafa and S. Güttel, *RKFIT: A robust algorithm for rational least squares approximation*, MIMS EPrint, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2014. In prepartaion.

[4] Z. Bujanović and Z. Drmač, *A new framework for implicit restarting of the Krylov–Schur algorithm*, to appear in Numer. Linear Algebra Appl., (2014).

[5] A. Bultheel, P. Gonzalez-Vera, E. Hendriksen, and O. Njastad, *Orthogonal Rational Functions*, Cambridge University Press, Cambridge, UK, 1999.

[6] A. Bultheel, M. Van Barel, and P. Van gucht, *Orthogonal basis functions in discrete least-squares rational approximation*, J. Comput. Appl. Math., 164-165 (2004), pp. 175–194.

[7] G. De Samblanx and A. Bultheel, *Using implicitly filtered RKS for generalised eigenvalue problems*, J. Comput. Appl. Math., 107 (1999), pp. 195–218.

[8] G. De Samblanx, K. Meerbergen, and A. Bultheel, *The implicit application of a rational filter in the RKS method*, BIT, 37 (1997), pp. 925–947.

[9] K. Deckers and A. Bultheel, *Rational Krylov sequences and orthogonal rational functions*, Tech. Report TW499, Department of Computer Science, KU Leuven, Leuven, Belgium, 2007.

[10] D. Deschrijver, B. Haegeman, and T. Dhaene, *Orthonormal vector fitting: A robust macro-modeling tool for rational approximation of frequency domain responses*, IEEE Transactions on Advanced Packaging, 30 (2007), pp. 216–225.

[11] Z. Drmač, S. Gugercin, and C. Beattie, *Quadrature-based vector fitting: Implications for H2 system approximation*, arXiv preprint arXiv:1403.4655, (2014).

[12] 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.

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

[14] K. Gallivan, E. Grimme, and P. Van Dooren, *A rational Lanczos algorithm for model reduction*, Numer. Algorithms, 12 (1996), pp. 33–63.

[15] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, USA, 4th ed., 2013.

[16] P. Gonnet, S. Güttel, and L. N. Trefethen, *Robust Padé approximation via SVD*, SIAM Rev., 55 (2013), pp. 101–117.

[17] P. Gonnet, R. Pachón, and L. N. Trefethen, *Robust rational interpolation and least-squares*, Electron. Trans. Numer. Anal., 38 (2011), pp. 146–167.

[18] S. Gugercin, A. Antoulas, and C. Beattie, *A rational Krylov iteration for optimal H2 model reduction*, in Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems, Kyoto, Japan, 2006, pp. 1665–1667.

[19] B. Gustavsen, *Improving the pole relocating properties of vector fitting*, IEEE Trans. Power Del., 21 (2006), pp. 1587–1592.

[20] B. Gustavsen and A. Semlyen, *Rational approximation of frequency domain responses by vector fitting*, IEEE Trans. Power Del., 14 (1999), pp. 1052–1061.

[21] S. Güttel, *Rational Krylov Methods for Operator Functions*, PhD thesis, Institut für Numerische Mathematik und Optimierung der Technischen Universität Bergakademie Freiberg, Freiberg, Germany, 2010. Available online at `http://nbn-resolving.de/urn:nbn:de:bsz:105-qucosa-27645`.

[22] S. Güttel, *Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection*, GAMM Mitteilungen, 36 (2013), pp. 8–31.

[23] S. Güttel, R. V. Beeumen, K. Meerbergen, and W. Michiels, *NLEIGS: A class of fully rational Krylov methods for nonlinear eigenvalue problems*, to appear in SIAM J. Sci. Comput., (2014).

[24] E. Jarlebring and H. Voss, *Rational Krylov for nonlinear eigenproblems, an iterative projection method*, Appl. Math., 50 (2005), pp. 543–554.

[25] B. Kågström and P. Poromaa, *Computing eigenspaces with specified eigenvalues of a regular matrix pair (A, B) and condition estimation: theory, algorithms and software*, Numer. Algorithms, 12 (1996), pp. 369–407.

[26] D. Kressner, *Block algorithms for reordering standard and generalized Schur forms*, ACM Trans. Math. Software, 32 (2006), pp. 521–532.

[27] T. Mach, M. Pranić, and R. Vandebril, *Computing approximate (block) rational Krylov subspaces without explicit inversion with extensions to symmetric matrices*, Electron. Trans. Numer. Anal., 43 (2014), pp. 100–124.

[28] K. Meerbergen, *Changing poles in the rational Lanczos method for the Hermitian eigenvalue problem*, Numer. Linear Algebra Appl., 8 (2001), pp. 33–52.

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

[30] A. Ruhe, *The rational Krylov algorithm for nonsymmetric eigenvalue problems. III: Complex shifts for real matrices*, BIT, 34 (1994), pp. 165–176.

[31] A. Ruhe, *Rational Krylov algorithms for nonsymmetric eigenvalue problems. II. Matrix pairs*, Linear Algebra Appl., 198 (1994), pp. 283–295.

[32] A. Ruhe, *Rational Krylov: A practical algorithm for large sparse nonsymmetric matrix pencils*, SIAM J. Sci. Comput., 19 (1998), pp. 1535–1551.

[33] A. Ruhe, *The rational Krylov algorithm for nonlinear matrix eigenvalue problems*, J. Math. Sci. (N. Y.), 114 (2003), pp. 1854–1856.

[34] A. Ruhe and D. Skoogh, *Rational Krylov algorithms for eigenvalue computation and model reduction*, in Applied Parallel Computing Large Scale Scientific and Industrial Problems, B. Kågström, J. Dongarra, E. Elmroth, and J. Waśniewski, eds., vol. 1541 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 1998, pp. 491–502.

[35] G. W. Stewart, *A Krylov–Schur algorithm for large eigenproblems*, SIAM J. Matrix Anal. Appl., 23 (2001), pp. 601–614.

[36] G. W. Stewart, *Matrix Algorithms, Volume II: Eigensystems*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001.

[37] G. W. Stewart, *Addendum to "A Krylov–Schur algorithm for large eigenproblems"*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 599–601.

[38] M. Van Barel, D. Fasino, L. Gemignani, and N. Mastronardi, *Orthogonal rational functions and structured matrices*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 810–829.

[39] R. Van Beeumen, K. Meerbergen, and W. Michiels, *A rational Krylov method based on Hermite interpolation for nonlinear eigenvalue problems*, SIAM J. Sci. Comput., 35 (2013), pp. A327–A350.

[40] R. Vandebril and D. S. Watkins, *A generalization of the multishift QR-algorithm*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 759–779.