

Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

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

Preprint BUW-IMACM 16/08

Marcel Schweitzer

A two-sided short-recurrence extended Krylov subspace method for nonsymmetric matrices and its relation to rational moment matching

March 6, 2016

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

A two-sided short-recurrence extended Krylov subspace method for nonsymmetric matrices and its relation to rational moment matching

Marcel Schweitzer

March 6, 2016

Abstract We present an extended Krylov subspace analogue of the two-sided Lanczos method, i.e., a method which, given a nonsingular matrix A and vectors $\boldsymbol{b}, \boldsymbol{c}$ with $\langle \boldsymbol{b}, \boldsymbol{c} \rangle \neq 0$, constructs bi-orthogonal bases of the extended Krylov subspaces $\mathcal{E}_m(A, \boldsymbol{b})$ and $\mathcal{E}_m(A^T, \boldsymbol{c})$ via short recurrences. We investigate the connection of the proposed method to rational moment matching for bilinear forms $\boldsymbol{c}^T f(A) \boldsymbol{b}$, similar to known results connecting the two-sided Lanczos method to moment matching. Numerical experiments demonstrate the quality of the resulting approximations and the stability properties of the new extended Krylov subspace method.

Keywords extended Krylov subspaces \cdot short recurrence methods \cdot two-sided method \cdot rational moment matching \cdot Laurent polynomials \cdot bilinear forms

Mathematics Subject Classification (2000) 65F25, 65F30, 65F60

1 Introduction

Given a matrix $A \in \mathbb{R}^{n \times n}$, a function f such that f(A) is defined and vectors $\boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^n$ an important task in many scientific and engineering applications is the approximation of bilinear forms

$$\boldsymbol{c}^T f(A) \boldsymbol{b}. \tag{1.1}$$

Applications in which these bilinear forms occur include network analysis [2,7], electronic structure calculations [1,3,29] or the solution of partial differential equations [21,22].

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

School of Mathematics and Natural Sciences, Bergische Universität Wuppertal, 42097 Wuppertal, Germany, schweitzer@math.uni-wuppertal.de

 $\mathbf{2}$

The case that $A = A^T$ is symmetric and definite and $\mathbf{b} = \mathbf{c}$ is well studied in the literature, see, e.g., [9–12,32] and the references therein. A widely used approach for approximating (1.1) in this case is performing m iterations of the short-recurrence Lanczos process [23, 24] for A and \mathbf{b} , which gives the decomposition

$$AV_m = V_m T_m + t_{m+1,m} v_{m+1} e_m^T, (1.2)$$

where the columns of $V_m = [v_1, \ldots, v_m]$ form an orthonormal basis of the Krylov subspace $\mathcal{K}_m(A, \mathbf{b})$, the matrix T_m is tridiagonal and $\mathbf{e}_m \in \mathbb{R}^m$ is the *m*th canonical unit vector. One then approximates (1.1) by

$$\boldsymbol{b}^T f(A) \boldsymbol{b} \approx \|\boldsymbol{b}\|_2^2 \boldsymbol{e}_1^T f(T_m) \boldsymbol{e}_1.$$
(1.3)

One can show that the approximation (1.3) matches the first 2m moments

$$\|\boldsymbol{b}\|_{2}^{2}\boldsymbol{e}_{1}^{T}T_{m}^{j}\boldsymbol{e}_{1} = \boldsymbol{b}^{T}A^{j}\boldsymbol{b} \text{ for } j = 0, 1, \dots, 2m-1,$$
(1.4)

i.e., it is exact for $f \in \Pi_{2m-1}$, where Π_{2m-1} is the space of all polynomials of degree at most 2m - 1; see, e.g., [32]. When A is nonsymmetric, the orthonormal basis of $\mathcal{K}_m(A, \mathbf{b})$ cannot be generated by short recurrences. A straightforward modification in this case is replacing the Lanczos process by the Arnoldi process and approximating

$$\boldsymbol{b}^{T} f(A) \boldsymbol{b} \approx \|\boldsymbol{b}\|_{2}^{2} \boldsymbol{e}_{1}^{T} f(H_{m}) \boldsymbol{e}_{1}, \qquad (1.5)$$

with H_m the upper Hessenberg matrix containing the corresponding orthogonalization coefficients. However, (1.5) only matches the first m + 1 moments.

Another alternative for the nonsymmetric case is to use the two-sided Lanczos process [28, Section 7.1] (sometimes also called nonsymmetric Lanczos). This method simultaneously generates bases $V_m = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m], W_m = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]$ for the Krylov subspaces $\mathcal{K}_m(A, \boldsymbol{b})$ and $\mathcal{K}_m(A^T, \boldsymbol{c})$, respectively, provided that $\langle \boldsymbol{b}, \boldsymbol{c} \rangle \neq 0$, where $\langle \cdot, \cdot \rangle$ denotes the usual Euclidean inner product. The sequences $\boldsymbol{v}_i, \boldsymbol{w}_i$ are bi-orthonormal, i.e., they satisfy

$$\langle \boldsymbol{v}_i, \boldsymbol{w}_j \rangle = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{else,} \end{cases}$$
(1.6)

and they can be computed using a three-term recurrence similar to the one of the Lanczos process (but requiring not only a multiplication with A but also a multiplication with A^T in each step). The matrix $\hat{T}_m = W_m^T A V_m$ is tridiagonal and fulfills the relations

$$AV_{m} = V_{m}\widehat{T}_{m} + t_{m+1,m}\boldsymbol{v}_{m+1}\boldsymbol{e}_{m}^{T},$$

$$A^{T}W_{m} = W_{m}\widehat{T}_{m}^{T} + t_{m,m+1}\boldsymbol{w}_{m+1}\boldsymbol{e}_{m}^{T}.$$
(1.7)

In [32], it is shown that the approximation

$$\boldsymbol{c}^{T} f(A) \boldsymbol{b} \approx \boldsymbol{e}_{1}^{T} f(\widehat{T}_{m}) \boldsymbol{e}_{1}$$
(1.8)

again matches the first 2m moments, i.e., it has the same degree of exactness as (1.3) in the Hermitian case; see also [13] for related results and [33] for numerical considerations on this topic.

In recent years, extended and rational Krylov subspaces have been heavily investigated as an alternative to standard polynomial Krylov subspaces $\mathcal{K}_m(A, \boldsymbol{b})$, especially in the context of approximating $f(A)\boldsymbol{b}$, the action of a matrix function on a vector, see, e.g., [5, 14-16, 20] and the references therein. Therefore, it is natural to also investigate these methods in the context of approximating bilinear forms (1.1).

In this paper, we concentrate on extended Krylov subspaces, which correspond not only to positive, but also to negative powers of the matrix A, as these also give rise to a short recurrence for computing an orthonormal basis in the Hermitian case. The paper is organized as follows. In Section 2, we briefly review the definition and basic facts about extended Krylov subspaces and their relation to Laurent polynomials. An algorithm for computing bi-orthonormal bases for extended Krylov subspaces corresponding to A and A^T via short recurrences is presented in Section 3. In Section 4, we investigate the properties of the two-sided extended Krylov subspace method in the context of rational moment matching, i.e., considering also negative powers of A in (1.4). Section 5 deals with topics concerning the implementation of the method, e.g., the possibility of early breakdown (similar to what can happen in the two-sided Lanczos process) and how to efficiently deal with the need for solving two linear systems in each iteration of the method. Numerical experiments illustrating the quality of the obtained approximations for (1.1) are reported in Section 6. Concluding remarks and topics for future research are given in Section 7.

Throughout this paper, we assume exact arithmetic unless explicitly stated otherwise.

2 Extended Krylov subspaces

Extended Krylov subspaces have first been introduced in [5] for approximating matrix functions and have been further investigated in, e.g., [17-20, 25, 30, 31]. They are built by not only applying powers of A, but also powers of A^{-1} to extend the basis and they are closely related to the set of Laurent polynomials.

Definition 2.1 Let k, m be two nonnegative integers. Then the set of *Laurent* polynomials of numerator degree at most m and denominator degree at most k is defined as

$$\mathcal{L}_{k}^{m} = \operatorname{span}\{x^{-k}, x^{-(k-1)}, \dots, x^{-1}, 1, x, x^{2}, \dots, x^{m}\}.$$
(2.1)

Obviously, it holds that $\mathcal{L}_0^m = \Pi_m$. The set of Laurent polynomials (2.1) allows to elegantly define extended Krylov subspaces, see also [17–19].

Algorithm 1: Block-wise extended Arnoldi method.

Definition 2.2 Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and $b \in \mathbb{R}^n$. The *m*th extended Krylov subspace corresponding to A and **b** is defined as

$$\mathcal{E}_m(A, \boldsymbol{b}) = \left\{ \phi(A)\boldsymbol{b} : \phi \in \mathcal{L}_m^{m-1} \right\}.$$
(2.2)

It is also possible to define Krylov subspaces corresponding to Laurent polynomials where the numerator and denominator differ by more than one (typically then, $m = \alpha \cdot k + 1$ for some fixed value α), see, e.g., [18], but we will not consider this here for sake of notational simplicity.

Similar to polynomial Krylov subspaces, one can iteratively compute a nested orthonormal basis $\mathcal{V}_m \in \mathbb{R}^{n \times 2m}$ for $\mathcal{E}_m(A, \mathbf{b})$ via the *extended Arnoldi* method given in Algorithm 1.

Defining the matrices $\mathcal{H}_m = \mathcal{V}_m^T A \mathcal{V}_m$ and

$$\eta_{m+1} = \begin{bmatrix} \boldsymbol{v}_{2m+1}^T \\ \boldsymbol{v}_{2m+2}^T \end{bmatrix} A \begin{bmatrix} \boldsymbol{v}_{2m-1}, \boldsymbol{v}_{2m} \end{bmatrix} \in \mathbb{R}^{2 \times 2},$$

one finds the following extended Arnoldi relation

$$A\mathcal{V}_m = \mathcal{V}_m \mathcal{H}_m + \left[\boldsymbol{v}_{2m+1}, \boldsymbol{v}_{2m+2} \right] \eta_{m+1} \begin{bmatrix} \boldsymbol{e}_{2m-1}^T \\ \boldsymbol{e}_{2m}^T \end{bmatrix}, \qquad (2.3)$$

see, e.g., [19, 30]. In case that A is Hermitian, the recurrence for the basis vectors in Algorithm 1 becomes a five-term recurrence, and the matrix \mathcal{H}_m becomes pentadiagonal, see, e.g. [17, 30].

Remark 2.1 We just briefly mention that it is also possible to compute a basis for $\mathcal{E}_m(A, \mathbf{b})$ in a slightly different fashion than the one presented in

4

Two-sided extended Krylov subspace methods

Algorithm 1. Instead of generating the basis vectors in a "block-wise fashion" (two at a time), an approach first considered in [30], one can also alternatingly apply multiplications with A and A^{-1} to the last basis vector, resulting in a slightly modified extended Arnoldi decomposition, see, e.g., [17–19].

Comparing the relation (2.3) (for Hermitian A) to (1.2), we find the following main differences, which will also be present in a similar form when comparing the two-sided Lanczos decomposition (1.7) to the two-sided extended Lanczos decomposition to be introduced in Section 3: The matrix \mathcal{H}_m has two additional nonzero off-diagonals compared to T_m , the entries of \mathcal{H}_m are *not* the orthogonalization coefficients from the extended Arnoldi process (but they can be computed cheaply from them, see, e.g., [30]) and $A\mathcal{V}_m$ and $\mathcal{V}_m\mathcal{H}_m$ differ in *two* columns instead of one.

We end this section by collecting some basic but useful properties of extended Krylov spaces in the following proposition.

Proposition 2.1 Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and $\mathbf{b} \in \mathbb{R}^n$. Then

(i) $\mathcal{E}_m(A, \boldsymbol{b}) = \mathcal{K}_{2m}(A, A^{-m}\boldsymbol{b}),$ (ii) $\mathcal{E}_m(A, \boldsymbol{b}) \subseteq \mathcal{K}_{2m+1}(A, A^{-m}\boldsymbol{b}) \subseteq \mathcal{E}_{m+1}(A, \boldsymbol{b}),$ (iii) $A\mathcal{E}_m(A, \boldsymbol{b}) \subseteq \mathcal{K}_{2m+1}(A, A^{-m}\boldsymbol{b}) \subseteq \mathcal{E}_{m+1}(A, \boldsymbol{b}).$

3 The two-sided extended Lanczos process

Short recurrences for computing an orthonormal basis of the extended Krylov subspace (2.2) are only available for Hermitian A. However, bi-orthonormal bases for $\mathcal{E}_m(A, \mathbf{b})$ and $\mathcal{E}_m(A^T, \mathbf{c})$, with $\langle \mathbf{b}, \mathbf{c} \rangle \neq 0$, can again be computed by a five-term recurrence, as we will derive in this section. We begin by presenting an algorithm which computes bi-orthonormal bases by explicit orthogonalization against all previous vectors and then identify which simplifications are possible due to the properties of extended Krylov subspaces. The method with explicit orthogonalization is given as Algorithm 2.

We note that Algorithm 2 may break down in case that one δ_i becomes zero (i.e., when the vectors v_i , w_i are orthogonal). For the following analysis we will always assume that such a breakdown does not occur and postpone the discussion of this topic to Section 5. The results of Algorithm 2 are summarized in the following lemma.

Lemma 3.1 Let $A \in \mathbb{R}^{n \times n}$ be nonsingular and $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ such that $\langle \mathbf{b}, \mathbf{c} \rangle = 1$. Then, after *m* steps of Algorithm 2,

- (i) v_1, \ldots, v_{2m} form a basis of $\mathcal{E}_m(A, \boldsymbol{b})$,
- (ii) $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_{2m}$ form a basis of $\mathcal{E}_m(A^T, \boldsymbol{c})$,
- (iii) the sequences $\boldsymbol{v}_i, \boldsymbol{w}_i, i = 1, \dots, 2m$ satisfy (1.6).

Moreover, defining

$$\widehat{\mathcal{T}}_m = \mathcal{W}_m^T A \mathcal{V}_m, \tag{3.1}$$

Algorithm 2: Construction of bi-orthogonal bases for
$$\mathcal{E}_{m}(A, b)$$
 and
 $\mathcal{E}_{m}(A^{T}, c)$.
Input: $m \in \mathbb{N}, A \in \mathbb{R}^{n \times n}$ nonsingular, $b, c \in \mathbb{R}^{n}$ such that $\langle b, c \rangle = 1$
Output: Bi-orthonormal bases $\mathcal{V}_{m} = [v_{1}, \dots, v_{2m}], \mathcal{W}_{m} = [w_{1}, \dots, w_{2m}]$ of
 $\mathcal{E}_{m}(A, b)$ and $\mathcal{E}_{m}(A^{T}, c)$
 $v_{1} \leftarrow b; w_{1} \leftarrow c;$
 $v_{2} \leftarrow A^{-1}b; w_{2} \leftarrow A^{-T}b;$
 $\alpha_{0} \leftarrow \langle v_{2}, w_{1} \rangle; \beta_{0} \leftarrow \langle w_{2}, v_{1} \rangle;$
 $v_{2} \leftarrow v_{2} - \alpha_{0}v_{1}; w_{2} \leftarrow w_{2} - \beta_{0}w_{1};$
 $\delta_{0} \leftarrow \sqrt{|\langle v_{2}, w_{2} \rangle|; \gamma_{0} \leftarrow \langle v_{2}, w_{2} \rangle/\delta_{0};}$
 $v_{2} \leftarrow v_{2} - \alpha_{0}v_{1}; w_{2} \leftarrow w_{2} - \beta_{0}w_{1};$
 $\delta_{0} \leftarrow \sqrt{|\langle v_{2}, w_{2} \rangle]; \gamma_{0} \leftarrow \langle v_{2}, w_{2} \rangle/\delta_{0};}$
 $v_{2} \leftarrow v_{2} - \alpha_{0}v_{1}; w_{2} \leftarrow w_{2} - \beta_{0}w_{1};$
 $\delta_{0} \leftarrow \sqrt{|\langle v_{2}, w_{2} \rangle]; \gamma_{0} \leftarrow \langle v_{2}, w_{2} \rangle/\delta_{0};}$
 $v_{2} \leftarrow v_{2} - \alpha_{0}v_{1}; w_{2} \leftarrow w_{2} - \beta_{0}w_{1};$
 $\delta_{0} \leftarrow \sqrt{|\langle v_{2}, w_{2} \rangle; |; \gamma_{0} \leftarrow \langle w_{2}, w_{2} \rangle/\delta_{0};}$
 $v_{2} \leftarrow v_{2} - \alpha_{0}v_{1}; w_{2} \leftarrow w_{2} - \beta_{0}w_{1};$
 $\delta_{0} \leftarrow \sqrt{|\langle v_{2}, w_{2} \rangle; |; \gamma_{0} \leftarrow \langle w_{2}, w_{2} \rangle/\delta_{0};}$
 $v_{2} \leftarrow 1/\gamma_{0}v_{2}; w_{2} \leftarrow 1/\delta_{0}w_{2};$
for $i = 1, \dots, 2j$ do
 $\left| \begin{array}{c} \alpha_{i,2j} - i \leftarrow \langle v_{2j+1}, w_{i} \rangle; \\ \beta_{i,2j} - i \leftarrow \langle v_{2j+1}, w_{2j+1} \rangle |; \gamma_{2j-1} \leftarrow \langle v_{2j+1}, w_{2j+1} \rangle/\delta_{2j-1}; \\ w_{2j+1} \leftarrow 1/\delta_{2j+1}w_{2j+1} \rangle; \\ w_{2j+2} \leftarrow A^{-1}w_{2j}; \\ w_{2j+2} \leftarrow A^{-1}w_{2j}; \\ \text{for } i = 1, \dots, 2j + 1$ do
 $\left| \begin{array}{c} \alpha_{i,2j} \leftarrow \langle v_{2j+2}, w_{i} \rangle; \\ \beta_{i,2j} \leftarrow \langle w_{2j+2}, w_{i} \rangle; \\ \beta_{i,2j} \leftarrow \langle w_{2j+2}, w_{i} \rangle; \\ \beta_{i,2j} \leftarrow \langle w_{2j+2}, w_{i} \rangle; \\ \psi_{2j+2} \leftarrow 1/\gamma_{2j+1}w_{2j+2} \rangle; \\ \gamma_{2j} \leftarrow \sqrt{|\langle w_{2j+2}, w_{2j+2} \rangle}; \langle v_{2j} + w_{2j+2} \rangle/\delta_{2j}; \\ w_{2j+2} \leftarrow 1/\delta_{2j+2}w_{2j+2}; \\ w$

and

6

$$\tau_{m+1} = \begin{bmatrix} \boldsymbol{w}_{2m+1}^T \\ \boldsymbol{w}_{2m+2}^T \end{bmatrix} A \begin{bmatrix} \boldsymbol{v}_{2m-1}, \boldsymbol{v}_{2m} \end{bmatrix} \in \mathbb{R}^{2 \times 2},$$

the following relations hold

$$A\mathcal{V}_m = \mathcal{V}_m \widehat{\mathcal{T}}_m + \left[\boldsymbol{v}_{2m+1}, \boldsymbol{v}_{2m+2} \right] \tau_{m+1} \begin{bmatrix} \boldsymbol{e}_{2m-1}^T \\ \boldsymbol{e}_{2m}^T \end{bmatrix}$$
(3.2)

$$A^{T} \mathcal{W}_{m} = \mathcal{W}_{m} \widehat{\mathcal{T}}_{m}^{T} + \left[\boldsymbol{w}_{2m+1}, \boldsymbol{w}_{2m+2} \right] \boldsymbol{\tau}_{m+1}^{T} \begin{bmatrix} \boldsymbol{e}_{2m-1}^{T} \\ \boldsymbol{e}_{2m}^{T} \end{bmatrix}.$$
(3.3)

Proof Assertions (i) and (ii) follow in a straightforward way from Algorithm 2 by standard arguments, in the same way as, e.g., when proving the correctness of the Arnoldi or the two-sided Lanczos method, see, e.g., [28, Sections 6.3 and 7.1]. The bi-orthonormality of the basis vectors, assertion (iii), will be shown by induction. The bi-orthogonality of v_1, v_2 and w_1, w_2 is directly

clear from Algorithm 2. Now assume that v_1, \ldots, v_{2j} with $j \leq m-1$ are bi-orthonormal. Then for $k \leq 2j$,

$$\begin{split} \langle \boldsymbol{w}_k, \boldsymbol{v}_{2j+1} \rangle &= \left\langle \boldsymbol{w}_k, \frac{1}{\gamma_{2j-1}} \left(A \boldsymbol{v}_{2j-1} - \sum_{i=1}^{2j} \alpha_{i,2j-1} \boldsymbol{v}_i \right) \right\rangle \\ &= \frac{1}{\gamma_{2j-1}} \left(\langle \boldsymbol{w}_k, A \boldsymbol{v}_{2j-1} \rangle - \alpha_{k,2j-1} \right) \\ &= \frac{1}{\gamma_{2j-1}} \left(\langle \boldsymbol{w}_k, A \boldsymbol{v}_{2j-1} \rangle - \langle \boldsymbol{w}_k, A \boldsymbol{v}_{2j-1} \rangle \right) = 0. \end{split}$$

The same line of argument shows that $\langle \boldsymbol{v}_k, \boldsymbol{w}_{2j+1} \rangle = 0$ for $k \leq 2j$. We further have

$$egin{aligned} &\langle m{w}_{2j+1}, m{v}_{2j+1}
angle &= \left\langle m{w}_{2j+1}, rac{1}{\gamma_{2j-1}} \left(A m{v}_{2j-1} - \sum_{i=1}^{2j} lpha_{i,2j-1} m{v}_i
ight)
ight
angle \ &= rac{1}{\gamma_{2j-1}} \left\langle m{w}_{2j+1}, A m{v}_{2j-1}
ight
angle = 1. \end{aligned}$$

Similar arguments can now be made to show that v_{2j+2} and w_{2j+2} are orthogonal to w_1, \ldots, w_{2j+1} and v_1, \ldots, v_{2j+1} respectively, and that $\langle v_{2j+2}, w_{2j+2} \rangle =$ 1. We conclude the proof of the lemma by proving the matrix relations (3.2) and (3.3). Let $k \leq 2m - 2$, so that $v_k \in \mathcal{E}_{m-1}(A, \mathbf{b})$ and thus, according to Proposition 2.1, $Av_k \in \mathcal{E}_m(A, \mathbf{b})$. Due to the bi-orthonormality of the sequences v_i, w_i and the fact that v_1, \ldots, v_{2m} is a basis of $\mathcal{E}_m(A, \mathbf{b})$, we can write

$$A\boldsymbol{v}_{k} = \sum_{i=1}^{2m} \langle \boldsymbol{w}_{i}, A\boldsymbol{v}_{k} \rangle \, \boldsymbol{v}_{i}.$$

$$(3.4)$$

In the same way, for $2m - 1 \le k \le 2m$, we have $Av_k \in \mathcal{E}_{m+1}(A, b)$, so that we can decompose Av_k in terms of the basis v_1, \ldots, v_{2m+2} as

$$A\boldsymbol{v}_{k} = \sum_{i=1}^{2m+2} \langle \boldsymbol{w}_{i}, A\boldsymbol{v}_{k} \rangle \boldsymbol{v}_{i}.$$
(3.5)

Recasting the relations (3.4) and (3.5) into matrix form proves (3.2). The proof of (3.3) works along the same lines. \Box

By Lemma 3.1, Algorithm 2 works correctly (assuming no breakdown occurs) and can thus be used as a starting point for deriving a short recurrence for computing the bi-orthonormal bases $v_i, w_i, i = 1, ..., 2m$, by showing that most of the orthogonalization coefficients $\alpha_{i,j}, \beta_{i,j}$ must be zero. The precise result is stated in the following lemma.

Lemma 3.2 Let $A \in \mathbb{R}^{n \times n}$ be nonsingular, let $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ such that $\langle \mathbf{b}, \mathbf{c} \rangle = 1$, and let $\alpha_{i,j}, \beta_{i,j}, i, j = 1, \ldots, 2m$ be the orthogonalization coefficients from Algorithm 2. Then, for all $j = 1, \ldots, m$,

$$\alpha_{i,2j-1} = \beta_{i,2j-1} = 0 \text{ for } i < 2j - 3,$$

$$\alpha_{i,2j} = \beta_{i,2j} = 0 \text{ for } i < 2j - 2.$$

Proof According to Algorithm 2, we have

$$\alpha_{i,2j-1} = \langle A \boldsymbol{v}_{2j-1}, \boldsymbol{w}_i \rangle = \langle \boldsymbol{v}_{2j-1}, A^T \boldsymbol{w}_i \rangle.$$

Now $A^T \boldsymbol{w}_i \in \mathcal{E}_{j-1}(A^T, \boldsymbol{c})$ if i < 2j - 3, and \boldsymbol{v}_{2j-1} is orthogonal to this space. In the same way

$$\alpha_{i,2j} = \left\langle A^{-1} \boldsymbol{v}_{2j}, \boldsymbol{w}_i \right\rangle = \left\langle \boldsymbol{v}_{2j}, A^{-T} \boldsymbol{w}_i \right\rangle$$

and $A^{-T} \boldsymbol{w}_i \in \mathcal{E}_{j-1}(A^T, \boldsymbol{c})$, to which \boldsymbol{v}_{2j} is orthogonal, if i < 2j - 2. The proof for $\beta_{i,2j-1}$ and $\beta_{i,2j}$ is completely analogous. \Box

According to Lemma 3.2, each of the basis vectors v_i, w_i in Algorithm 2 has to be orthogonalized against *at most four* previous basis vectors, thus proving that a short recurrence for the bases exists. A further optimization of Algorithm 2 can be obtained by reducing the number of inner products that need to be evaluated. This is indeed possible, as the coefficients $\alpha_{i,j}, \beta_{i,j}$ obey certain recursion relations, as stated in the next lemma.

Lemma 3.3 Let $A \in \mathbb{R}^{n \times n}$ be nonsingular, let $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ such that $\langle \mathbf{b}, \mathbf{c} \rangle = 1$, and let $\alpha_{i,j}, \beta_{i,j}, \gamma_j, \delta_j, i, j = 1, \ldots, 2m$ be the orthogonalization coefficients from Algorithm 2. Then for all $j = 2, \ldots, m$ the following recursive relations hold

$$\alpha_{2j-3,2j-1} = \delta_{2j-3},\tag{3.6}$$

$$\alpha_{2j-2,2j-1} = -\frac{1}{\delta_{2j-4}} \beta_{2j-3,2j-4} \alpha_{2j-3,2j-1}, \qquad (3.7)$$

$$\alpha_{2j,2j-1} = -\frac{1}{\delta_{2j-2}} \sum_{i=2j-3}^{2j-1} \beta_{i,2j-2} \alpha_{i,2j-1}, \qquad (3.8)$$

$$\alpha_{2j-2,2j} = \delta_{2j-2},\tag{3.9}$$

$$\alpha_{2j-1,2j} = -\frac{1}{\delta_{2j-3}} \beta_{2j-2,2j-3} \alpha_{2j-2,2j}, \qquad (3.10)$$

$$\alpha_{2j+1,2j} = -\frac{1}{\delta_{2j-1}} \sum_{i=2j-2}^{2j} \beta_{i,2j-1} \alpha_{i,2j}, \qquad (3.11)$$

$$\beta_{2j-1,2j-1} = \alpha_{2j-1,2j-1},\tag{3.12}$$

$$\beta_{2j,2j} = \alpha_{2j-1,2j-1},\tag{3.13}$$

Relations (3.6)–(3.11) also hold when the roles of α and β are exchanged and δ is replaced by γ .

Proof From Algorithm 2 and Lemma 3.2, we find

$$\begin{aligned} \alpha_{2j-3,2j-1} &= \langle A \boldsymbol{v}_{2j-1}, \boldsymbol{w}_{2j-3} \rangle = \langle \boldsymbol{v}_{2j-1}, A^T \boldsymbol{w}_{2j-3} \rangle \\ &= \left\langle \boldsymbol{v}_{2j-1}, \delta_{2j-3} \boldsymbol{w}_{2j-1} + \sum_{i=2j-5}^{2j-2} \beta_{i,2j-1} \boldsymbol{w}_i \right\rangle = \delta_{2j-3}, \end{aligned}$$

where the last equality follows from the bi-orthonormality of the v_i, w_i , proving (3.6). Furthermore, it holds

$$\begin{aligned} \alpha_{2j-2,2j-1} &= \left\langle A \boldsymbol{v}_{2j-1}, \frac{1}{\delta_{2j-4}} \left(A^{-T} \boldsymbol{w}_{2j-4} - \sum_{i=2j-6}^{2j-3} \beta_{i,2j-4} \boldsymbol{w}_{i} \right) \right\rangle \\ &= -\frac{1}{\delta_{2j-4}} \sum_{i=2j-6}^{2j-3} \beta_{i,2j-4} \left\langle A \boldsymbol{v}_{2j-1}, \boldsymbol{w}_{i} \right\rangle \\ &= -\frac{1}{\delta_{2j-4}} \beta_{2j-3,2j-4} \alpha_{2j-3,2j-1}, \end{aligned}$$

where we used that $\alpha_{i,2j-1} = 0$ for i < 2j-3. This proves (3.7). To show (3.8), consider

$$\begin{aligned} \alpha_{2j,2j-1} &= \langle A \boldsymbol{v}_{2j-1}, \boldsymbol{w}_{2j} \rangle \\ &= \left\langle A \boldsymbol{v}_{2j-1}, \frac{1}{\delta_{2j-2}} \left(A^{-T} \boldsymbol{w}_{2j-2} - \sum_{i=2j-4}^{2j-1} \beta_{i,2j-2} \boldsymbol{w}_{i} \right) \right\rangle \\ &= -\frac{1}{\delta_{2j-2}} \sum_{i=2j-4}^{2j-1} \beta_{i,2j-2} \langle A \boldsymbol{v}_{2j-1}, \boldsymbol{w}_{i} \rangle \\ &= -\frac{1}{\delta_{2j-2}} \sum_{i=2j-3}^{2j-1} \beta_{i,2j-2} \alpha_{i,2j-1}. \end{aligned}$$

Recursion formula (3.9) follows from

$$\alpha_{2j-2,2j} = \left\langle A^{-1} \boldsymbol{v}_{2j}, \boldsymbol{w}_{2j-2} \right\rangle = \left\langle \boldsymbol{v}_{2j}, A^{-T}, \boldsymbol{w}_{2j-2} \right\rangle$$
$$= \left\langle \boldsymbol{v}_{2j}, \delta_{2j-2} \boldsymbol{w}_{2j} + \sum_{i=2j-4}^{2j-1} \beta_{i,2j} \boldsymbol{w}_i \right\rangle = \delta_{2j-2}.$$

Using the same techniques as before, we find 1

1.

$$\begin{aligned} \alpha_{2j-1,2j} &= \langle A^{-1} \boldsymbol{v}_{2j}, \boldsymbol{w}_{2j-1} \rangle \\ &= \left\langle A^{-1} \boldsymbol{v}_{2j}, \frac{1}{\delta_{2j-3}} \left(A^T \boldsymbol{w}_{2j-3} - \sum_{i=2j-5}^{2j-2} \beta_{i,2j-3} \boldsymbol{w}_i \right) \right\rangle \\ &= -\frac{1}{\delta_{2j-3}} \sum_{i=2j-5}^{2j-2} \beta_{i,2j-3} \left\langle A^{-1} \boldsymbol{v}_{2j}, \boldsymbol{w}_i \right\rangle \\ &= -\frac{1}{\delta_{2j-3}} \beta_{2j-2,2j-3} \alpha_{2j-2,2j}, \end{aligned}$$

thus proving (3.10). Equation (3.11) follows due to

$$\begin{aligned} \alpha_{2j+1,2j} &= \left\langle A^{-1} \boldsymbol{v}_{2j}, \boldsymbol{w}_{2j+1} \right\rangle \\ &= \left\langle A^{-1} \boldsymbol{v}_{2j}, \frac{1}{\delta_{2j-1}} \left(A^T \boldsymbol{w}_{2j-1} - \sum_{i=2j-3}^{2j} \beta_{i,2j-1} \boldsymbol{w}_i \right) \right\rangle \\ &= -\frac{1}{\delta_{2j-1}} \sum_{i=2j-3}^{2j} \beta_{i,2j-1} \left\langle A^{-1} \boldsymbol{v}_{2j}, \boldsymbol{w}_i \right\rangle \\ &= -\frac{1}{\delta_{2j-2}} \sum_{i=2j-2}^{2j} \beta_{i,2j-1} \alpha_{i,2j}. \end{aligned}$$

Equations (3.12) and (3.13) follow directly from

$$\beta_{2j-1,2j-1} = \langle A^T w_{2j-1}, w_{2j-1} \rangle = \langle w_{2j-1}, A v_{2j-1} \rangle = \alpha_{2j-1,2j-1}$$

and

10

$$\beta_{2j,2j} = \left\langle A^T \boldsymbol{w}_{2j}, \boldsymbol{w}_{2j} \right\rangle = \left\langle \boldsymbol{w}_{2j}, A \boldsymbol{v}_{2j} \right\rangle = \alpha_{2j,2j}$$

The analogues of (3.6)–(3.11) for the β -values can be shown in exactly the same way, by systematically switching the roles of α and β and replacing δ by γ . \Box

Using the recursion relations from Lemma 3.3, only *four* inner products have to be computed in each iteration of Algorithm 2, although orthogonalization against *eight* vectors is necessary and an *additional four* vectors have to be normalized.

The matrix $\widehat{\mathcal{T}}_m$ from (3.1) will play a crucial role for computing an approximation for (1.1) similar to (1.8). As the entries of $\widehat{\mathcal{T}}_m$ are, in contrast to the entries of the matrix \widehat{T}_m from the polynomial two-sided Lanczos process, different from the orthogonalization coefficients and thus not directly available, we next derive recursion formulas for them. This allows to compute them more efficiently, without explicitly performing the matrix-matrix products $\mathcal{W}_m^T A \mathcal{V}_m$, but more importantly it allows to retrieve the matrix $\widehat{\mathcal{T}}_m$ without the need for storing the "old" basis vectors v_i, w_i .

Lemma 3.4 Let $A \in \mathbb{R}^{n \times n}$ be nonsingular, let $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ such that $\langle \mathbf{b}, \mathbf{c} \rangle = 1$, let $\alpha_{i,j}, \beta_{i,j}, \gamma_j, \delta_j, i, j = 1, \ldots, 2m$ be the orthogonalization coefficients from Algorithm 2 and let $\widehat{\mathcal{T}}_m = (t_{i,j})_{i,j=1,\ldots,2m}$ be defined by (3.1). Then, for all $j = 2, \ldots, m$, the following recursive relations hold

$$t_{i,2j-1} = \alpha_{i,2j-1} \text{ for } i = 2j-3, \dots, 2j+1,$$
 (3.14)

$$t_{2j+1,2j-1} = \gamma_{2j-1}, \tag{3.15}$$

$$t_{2j-1,2j} = -\frac{1}{\gamma_{2j-2}} \sum_{i=2j-3}^{2j-1} \alpha_{i,2j-2} t_{2j-1,i}, \qquad (3.16)$$

$$t_{2j,2j} = -\frac{1}{\gamma_{2j-2}} \alpha_{2j-1,2j-2} t_{2j,2j-1}, \qquad (3.17)$$

$$t_{2j+1,2j} = -\frac{1}{\gamma_{2j-2}} \alpha_{2j-1,2j-2} t_{2j+1,2j-1}.$$
(3.18)

All other entries of $\widehat{\mathcal{T}}_m$ are zero.

Proof Due to $\widehat{\mathcal{T}}_m = \mathcal{W}_m^T A \mathcal{V}_m$, we have $t_{i,j} = \langle \boldsymbol{w}_i, A \boldsymbol{v}_j \rangle$. Therefore, $t_{i,2j-1} = \alpha_{i,2j-1}$ for all $i = 1, \ldots, 2m$ from the definition of α . Together with Lemma 3.2, this shows (3.14) and (3.15) as well as that all other entries of $\widehat{\mathcal{T}}_m$ in the odd-numbered columns are zero.

We proceed by showing that all entries $t_{i,2j}$ for $i \notin \{2j-1,2j,2j+1\}$ are zero. We have $t_{i,2j} = \langle \boldsymbol{w}_i, A\boldsymbol{v}_{2j} \rangle$. Using Proposition 2.1, we find $A\boldsymbol{v}_{2j} \in A\mathcal{E}_j(A, \boldsymbol{b}) \subseteq \mathcal{K}_{2j+1}(A, A^{-j}\boldsymbol{b}) = \operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{2j+1}\}$. For i > 2j+1, the vector \boldsymbol{w}_i is orthogonal to this space, so that $t_{i,2j} = 0$ in this case. On the other hand $t_{i,2j} = \langle \boldsymbol{w}_i, A \boldsymbol{v}_{2j} \rangle = \langle A^T \boldsymbol{w}_i, \boldsymbol{v}_{2j} \rangle$ and using a similar argument as above, we find that \boldsymbol{v}_{2j} is orthogonal to $A^T \boldsymbol{w}_i$ when i < 2j - 1. It remains to show that the recursion formulas (3.16)–(3.18) hold for the three nonzero entries of the even-numbered columns of $\widehat{\mathcal{T}}_m$. We have

$$\langle \boldsymbol{w}_{2j-1}, A \boldsymbol{v}_{2j} \rangle = \left\langle \boldsymbol{w}_{2j-1}, \frac{1}{\gamma_{2j-2}} \left(\boldsymbol{v}_{2j-2} - \sum_{i=2j-4}^{2j-1} \alpha_{i,2j-2} A \boldsymbol{v}_i \right) \right\rangle$$

$$= -\frac{1}{\gamma_{2j-2}} \sum_{i=2j-4}^{2j-1} \left\langle \boldsymbol{w}_{2j-1}, \alpha_{i,2j-2} A \boldsymbol{v}_i \right\rangle$$

$$= -\frac{1}{\gamma_{2j-2}} \sum_{i=2j-3}^{2j-1} \alpha_{i,2j-2} t_{2j-1,i},$$

where we used the definition of $t_{2j-1,i}$ and the fact that $t_{2j-1,2j-4} = 0$. Next, consider

$$\langle \boldsymbol{w}_{2j}, A \boldsymbol{v}_{2j} \rangle = \left\langle \boldsymbol{w}_{2j}, \frac{1}{\gamma_{2j-2}} \left(\boldsymbol{v}_{2j-2} - \sum_{i=2j-4}^{2j-1} \alpha_{i,2j-2} A \boldsymbol{v}_i \right) \right\rangle$$

= $-\frac{1}{\gamma_{2j-2}} \alpha_{2j-1,2j-2} t_{2j,2j-1}$

using the already proven nonzero structure of $\widehat{\mathcal{T}}_m$. Finally, equation (3.18) follows from

$$\langle \boldsymbol{w}_{2j+1}, A \boldsymbol{v}_{2j} \rangle = \left\langle \boldsymbol{w}_{2j+1}, \frac{1}{\gamma_{2j-2}} \left(\boldsymbol{v}_{2j-2} - \sum_{i=2j-4}^{2j-1} \alpha_{i,2j-2} A \boldsymbol{v}_i \right) \right\rangle$$

= $-\frac{1}{\gamma_{2j-2}} \alpha_{2j-1,2j-2} t_{2j+1,2j-1}.$

This concludes the proof of the lemma. \Box

Using the results of Lemma 3.2–3.4, we are now in a position to formulate a more efficient, short-recurrence method which is equivalent to Algorithm 2 in exact arithmetic. This method, the *two-sided extended Lanczos process*, is given in Algorithm 3.

When aiming to estimate a bilinear form (1.1), one can define an approximation based on Algorithm 3 in the usual way as

$$\boldsymbol{c}^{T} f(A) \boldsymbol{b} \approx \boldsymbol{e}_{1}^{T} f(\widehat{\mathcal{T}}_{m}) \boldsymbol{e}_{1}.$$
(3.19)

In the next section, we show that the approximation (3.19) is exact for Laurent polynomials of numerator degree at most 2m - 1 and denominator degree at most 2m, and that this is a higher degree of exactness than what is obtained by using the extended Arnoldi method in general. Numerical experiments comparing the accuracy of the approximation (3.19) to other approaches are given in Section 6.

Algorithm 3: Two-sided extended Lanczos process.

Input: $m \in \mathbb{N}$, $A \in \mathbb{R}^{n \times n}$ nonsingular, $b, c \in \mathbb{R}^n$ such that $\langle b, c \rangle = 1$ **Output**: Bi-orthonormal bases $\mathcal{V}_m = [v_1, \ldots, v_{2m}], \mathcal{W}_m = [w_1, \ldots, w_{2m}]$ of $\mathcal{E}_m(A, b)$ and $\mathcal{E}_m(A^T, c)$, compressed matrix $\widehat{\mathcal{T}}_m = \mathcal{W}_m^T A \mathcal{V}_m$ $b; w_1 \leftarrow c;$ v_1 $\boldsymbol{v}_2 \leftarrow A^{-1}\boldsymbol{b}; \boldsymbol{w}_2 \leftarrow A^{-T}\boldsymbol{b};$ $\alpha_0 \leftarrow \langle \boldsymbol{v}_2, \boldsymbol{w}_1 \rangle; \beta_0 \leftarrow \langle \boldsymbol{w}_2, \boldsymbol{v}_1 \rangle;$ $\begin{array}{l} \mathbf{v}_{2} \leftarrow \mathbf{v}_{2} - \alpha_{0} \mathbf{v}_{1}; \mathbf{w}_{2} \leftarrow \mathbf{w}_{2} - \beta_{0} \mathbf{w}_{1}; \\ \delta_{0} \leftarrow \sqrt{|\langle \mathbf{v}_{2}, \mathbf{w}_{2} \rangle|}; \gamma_{0} \leftarrow \langle \mathbf{v}_{2}, \mathbf{w}_{2} \rangle / \delta_{0}; \end{array}$ $\boldsymbol{v}_2 \leftarrow 1/\gamma_0 \boldsymbol{v}_2; \boldsymbol{w}_2 \leftarrow 1/\delta_0 \boldsymbol{w}_2;$ $\begin{array}{c|c} \mathbf{for} & j = 1, 2, \dots, m \ \mathbf{do} \\ & \mathbf{v}_{2j+1} \leftarrow A \mathbf{v}_{2j-1}; \\ & \mathbf{v}_{T} \end{array}$ $\boldsymbol{w}_{2j+1} \leftarrow A^T \boldsymbol{w}_{2j-1};$ $v_{2j+2} \leftarrow A^{-1} v_{2j}; \\ w_{2j+2} \leftarrow A^{-T} w_{2j};$ Compute $\alpha_{i,2j-1}, \alpha_{i,2j}, \beta_{i,2j-1}, \beta_{i,2j}$ via (3.6)–(3.13) $\begin{aligned} & v_{2j+1} \leftarrow v_{2j+1} - \sum_{i=2j-3}^{2j} \alpha_{i,2j-1} v_i; \\ & w_{2j+1} \leftarrow w_{2j+1} - \sum_{i=2j-3}^{2j} \beta_{i,2j-1} v_i; \\ & w_{2j+1} \leftarrow w_{2j+1} - \sum_{i=2j-3}^{2j} \beta_{i,2j-1} w_i; \\ & \delta_{2j-1} \leftarrow \sqrt{|\langle v_{2j+1}, w_{2j+1} \rangle|}; \\ & \gamma_{2j-1} \leftarrow \langle v_{2j+1}, w_{2j+1} \rangle| \rangle \langle \delta_{2j-1}; \end{aligned}$ $v_{2j+1} \leftarrow 1/\gamma_{2j+1} v_{2j+1};$ $\begin{array}{c} v_{2j+1} \leftarrow 1/|z_{j+1} < z_{j+1}, \\ w_{2j+1} \leftarrow 1/\delta_{2j+1} w_{2j+1}, \\ v_{2j+2} \leftarrow v_{2j+2} - \sum_{i=2j-2}^{2j+1} \alpha_{i,2j} v_i; \\ w_{2j+2} \leftarrow w_{2j+2} - \sum_{i=2j-2}^{2j+1} \beta_{i,2j} w_i; \\ \end{array}$ $\delta_{2j} \leftarrow \sqrt{\left|\left\langle \boldsymbol{v}_{2j+2}, \boldsymbol{w}_{2j+2} \right\rangle\right|}; \gamma_{2j} \leftarrow \left\langle \boldsymbol{v}_{2j+2}, \boldsymbol{w}_{2j+2} \right\rangle / \delta_{2j};$ $\boldsymbol{v}_{2j+2} \leftarrow 1/\gamma_{2j+2} \boldsymbol{v}_{2j+2};$ $w_{2j+2} \leftarrow 1/\delta_{2j+2} w_{2j+2};$ Compute $t_{i,2j-1}, t_{i,2j}$ via (3.14)–(3.18)

4 Rational moment matching

The standard two-sided Lanczos algorithm is related to (polynomial) moment matching via

$$\boldsymbol{c}^{T} p(A) \boldsymbol{b} = \boldsymbol{e}_{1}^{T} p(\widehat{T}_{m}) \boldsymbol{e}_{1}$$
 for all $p \in \Pi_{2m-1}$

see, e.g., [32]. In this section, we prove that an analogous relation holds for the two-sided extended Lanczos algorithm and Laurent polynomials. We begin by stating the precise result in the following theorem.

Theorem 4.1 Let $A \in \mathbb{R}^{n \times n}$ be nonsingular, let $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ such that $\langle \mathbf{b}, \mathbf{c} \rangle = 1$ and let $\widehat{\mathcal{T}}_m$ be defined as in (3.1), with $\mathcal{W}_m, \mathcal{V}_m$ computed by Algorithm 3. Then, if $\widehat{\mathcal{T}}_m$ is nonsingular,

$$\boldsymbol{c}^{T}\phi(A)\boldsymbol{b} = \boldsymbol{e}_{1}^{T}\phi(\widehat{\mathcal{T}}_{m})\boldsymbol{e}_{1}$$
 for all $\phi \in \mathcal{L}_{2m}^{2m-1}$

To be able to prove Theorem 4.1, we need some further auxiliary results on properties of the matrix $\hat{\mathcal{T}}_m$ and the *inverse projection matrix*

$$\widehat{\mathcal{S}}_m := \mathcal{W}_m^T A^{-1} \mathcal{V}_m. \tag{4.1}$$

The matrix $\widehat{\mathcal{S}}_m$ from (4.1) satisfies the identities

$$A^{-1}\mathcal{V}_{m} = \mathcal{V}_{m}\widehat{\mathcal{S}}_{m} + [\mathbf{v}_{2m+1}, \mathbf{v}_{2m+2}] \sigma_{m+1} \begin{bmatrix} \mathbf{e}_{2m-1}^{T} \\ \mathbf{e}_{2m}^{T} \end{bmatrix}, \qquad (4.2)$$
$$A^{-T}\mathcal{W}_{m} = \mathcal{W}_{m}\widehat{\mathcal{S}}_{m}^{T} + [\mathbf{w}_{2m+1}, \mathbf{w}_{2m+2}] \sigma_{m+1}^{T} \begin{bmatrix} \mathbf{e}_{2m-1}^{T} \\ \mathbf{e}_{2m}^{T} \end{bmatrix},$$

where

$$\sigma_{m+1} = \begin{bmatrix} \boldsymbol{w}_{2m+1}^T \\ \boldsymbol{w}_{2m+2}^T \end{bmatrix} A^{-1} \begin{bmatrix} \boldsymbol{v}_{2m-1}, \boldsymbol{v}_{2m} \end{bmatrix} \in \mathbb{R}^{2 \times 2}.$$

We omit the proof of these relations, as it is completely analogous to how the relations for $\widehat{\mathcal{T}}_m$ are proven in Lemma 3.1. The nonzero structure of $\widehat{\mathcal{S}}_m$ is very similar, but not completely identical to that of $\widehat{\mathcal{T}}_m$.

Proposition 4.1 Let $\widehat{S}_m = (s_{k,\ell})_{k,\ell=1,\ldots,2m}$ be defined by (4.1). Then for all $j = 1, \ldots, m$

$$s_{i,2j-1} = 0 \text{ for } i < 2j-2 \text{ or } i > 2j$$

 $s_{i,2j} = 0 \text{ for } i < 2j-2 \text{ or } i > 2j+2$

Proof The result can be proven analogous to the result on the nonzero structure of $\widehat{\mathcal{T}}_m$. \Box

Exploiting the nonzero structure of $\widehat{\mathcal{T}}_m$ and $\widehat{\mathcal{S}}_m$ allows us to prove the following results on exact approximation of certain polynomials in A and A^{-1} .

Lemma 4.1 Let $A \in \mathbb{R}^{n \times n}$ be nonsingular, let $\mathbf{b}, \mathbf{c} \in \mathbb{R}^n$ such that $\langle \mathbf{b}, \mathbf{c} \rangle = 1$ and let $\widehat{\mathcal{T}}_m$ and $\widehat{\mathcal{S}}_m$ be defined as in (3.1) and (4.1), respectively, with $\mathcal{W}_m, \mathcal{V}_m$ computed by Algorithm 3. Then

$$\boldsymbol{c}^{T} \boldsymbol{p}(A) \boldsymbol{b} = \boldsymbol{e}_{1}^{T} \boldsymbol{p}(\widehat{\mathcal{T}}_{m}) \boldsymbol{e}_{1} \text{ for } \boldsymbol{p} \in \boldsymbol{\Pi}_{2m-1}$$

$$(4.3)$$

$$\boldsymbol{c}^{T} \boldsymbol{p}(A^{-1})\boldsymbol{b} = \boldsymbol{e}_{1}^{T} \boldsymbol{p}(\widehat{\mathcal{S}}_{m})\boldsymbol{e}_{1} \text{ for } \boldsymbol{p} \in \Pi_{2m+1}$$

$$(4.4)$$

Proof We begin by proving relation (4.3). It suffices to prove the statement for monomials $p(z) = z^j$ for j = 1, ..., 2m - 1. For j = 1 and $m \ge 2$, from (3.2) and using $\mathcal{V}_m \mathbf{e}_1 = \mathbf{b}$ we have

$$A \boldsymbol{b} = A \mathcal{V}_m \boldsymbol{e}_1 = \mathcal{V}_m \mathcal{T}_m \boldsymbol{e}_1$$

where the second term vanishes because $e_{2m-1}^T e_1 = e_{2m}^T e_1 = 0$. Now, for $j \leq m-1$, we inductively find

$$A^{j}\boldsymbol{b} = \mathcal{V}_{m}\widehat{\mathcal{T}}_{m}^{j-1}\boldsymbol{e}_{1} + [\boldsymbol{v}_{2m+1}, \boldsymbol{v}_{2m+2}]\tau_{m+1} \begin{bmatrix} \boldsymbol{e}_{2m-1}^{T} \\ \boldsymbol{e}_{2m}^{T} \end{bmatrix} \widehat{\mathcal{T}}_{m}^{j-1}\boldsymbol{e}_{1}.$$
(4.5)

Due to the nonzero structure of $\widehat{\mathcal{T}}_m$, the vector $\widehat{\mathcal{T}}_m^{j-1} e_1$ can only have nonzeros in its first 2j-1 entries. Thus, for $j \leq m-1$, the second term on the right-hand side of (4.5) again vanishes and we have

$$A^{j}\boldsymbol{b} = \mathcal{V}_{m}\mathcal{T}_{m}^{j}\boldsymbol{e}_{1} \text{ for } j \leq m-1$$

$$(4.6)$$

Repeating this line of argument starting from the decomposition (3.3) yields

$$(A^T)^j \boldsymbol{c} = \mathcal{W}_m \left(\widehat{\mathcal{T}}_m^T\right)^j \boldsymbol{e}_1 \text{ for } j \le m-1.$$
 (4.7)

Combining (4.6) and (4.7) gives

$$\boldsymbol{c}^{T} A^{2m-1} \boldsymbol{b} = ((A^{T})^{m-1} \boldsymbol{c})^{T} A(A^{m-1} \boldsymbol{b}) = (\mathcal{W}_{m} \left(\widehat{\mathcal{T}}_{m}^{T}\right)^{m-1} \boldsymbol{e}_{1})^{T} A(\mathcal{V}_{m} \widehat{\mathcal{T}}_{m}^{m-1} \boldsymbol{e}_{1})$$
$$= \boldsymbol{e}_{1}^{T} \widehat{\mathcal{T}}_{m}^{m-1} \mathcal{W}_{m}^{T} A \mathcal{V}_{m} \widehat{\mathcal{T}}_{m}^{m-1} \boldsymbol{e}_{1} = \boldsymbol{e}_{1}^{T} \widehat{\mathcal{T}}_{m}^{2m-1} \boldsymbol{e}_{1}.$$
(4.8)

A similar argument obviously holds for lower powers of A, so that this establishes (4.3).

The proof of (4.4) proceeds along the same lines, the difference being that we have

$$A^{-j}\boldsymbol{b} = \mathcal{V}_m \widehat{\mathcal{T}}_m^{-j}\boldsymbol{e}_1 \text{ and } (A^{-T})^j \boldsymbol{c} = \mathcal{W}_m \left(\widehat{\mathcal{S}}_m^T\right)^j \boldsymbol{e}_1 \text{ for } j \leq m$$

so that we obtain exactness up to a polynomial degree of 2m + 1 in this case, as stated in (4.4). \Box

Next, we need to relate powers of $\widehat{\mathcal{S}}_m$ to negative powers of $\widehat{\mathcal{T}}_m$.

Lemma 4.2 Let the assumptions of Lemma 4.1 hold and further let $\widehat{\mathcal{T}}_m$ be nonsingular. Then

$$\boldsymbol{e}_1^T p(\widehat{\mathcal{S}}_m) \boldsymbol{e}_1 = \boldsymbol{e}_1^T p(\widehat{\mathcal{T}}_m^{-1}) \boldsymbol{e}_1 \text{ for } p \in \Pi_{2m}$$

Proof It again suffices to prove the statement of the lemma for monomials. Multiplying (4.2) by $\mathcal{W}_m^T A$ from the left gives

$$I = \widehat{\mathcal{T}}_m \widehat{\mathcal{S}}_m + \mathcal{W}_m^T A\left[\boldsymbol{v}_{2m+1}, \boldsymbol{v}_{2m+2}\right] \sigma_{m+1} \begin{bmatrix} \boldsymbol{e}_{2m-1}^T \\ \boldsymbol{e}_{2m}^T \end{bmatrix}$$
(4.9)

which directly implies

$$\widehat{\mathcal{T}}_m \widehat{\mathcal{S}}_m \boldsymbol{e}_1 = \boldsymbol{e}_1.$$

In addition, (4.9) can be rearranged to yield

$$\widehat{\mathcal{T}}_{m}\widehat{\mathcal{S}}_{m} = I - \mathcal{W}_{m}^{T}A\left[\boldsymbol{v}_{2m+1}, \boldsymbol{v}_{2m+2}\right]\sigma_{m+1} \begin{bmatrix} \boldsymbol{e}_{2m-1}^{T} \\ \boldsymbol{e}_{2m}^{T} \end{bmatrix}.$$
(4.10)

Using (4.10), we find

$$\widehat{\mathcal{T}}_{m}^{j}\widehat{\mathcal{S}}_{m}^{j}\boldsymbol{e}_{1} = \widehat{\mathcal{T}}_{m}^{j-1}\left(I - \mathcal{W}_{m}^{T}A\left[\boldsymbol{v}_{2m+1}, \boldsymbol{v}_{2m+2}\right]\sigma_{m+1}\left[\boldsymbol{e}_{2m-1}^{T}\right]\right)\widehat{\mathcal{S}}_{m}^{j-1}\boldsymbol{e}_{1}.$$
 (4.11)

By induction, and noting that only the first 2j - 2 entries of $\widehat{\mathcal{S}}_m^{j-1} e_1$ may be nonzero, (4.11) gives

$$\widehat{\mathcal{T}}_m^j \widehat{\mathcal{S}}_m^j \pmb{e}_1 = \widehat{\mathcal{T}}_m^{j-1} \widehat{\mathcal{S}}_m^{j-1} \pmb{e}_1 = \pmb{e}_1$$

Preprint – Preprint – Preprint – Preprint – Preprin

14

for $j \leq m$. Multiplying by $\widehat{\mathcal{T}}_m^{-j}$ from the left gives

$$\widehat{\mathcal{S}}_{m}^{j} \boldsymbol{e}_{1} = \widehat{\mathcal{T}}_{m}^{-j} \boldsymbol{e}_{1} \text{ for } j \leq m.$$

$$(4.12)$$

Analogous to (4.12), we can derive

$$\left(\widehat{\mathcal{S}}_{m}^{T}\right)^{j} \boldsymbol{e}_{1} = \left(\widehat{\mathcal{T}}_{m}^{T}\right)^{-j} \boldsymbol{e}_{1} \text{ for } j \leq m.$$

$$(4.13)$$

Using (4.12) and (4.13) together gives

$$\boldsymbol{e}_{1}^{T}\widehat{\boldsymbol{\mathcal{S}}}_{m}^{2m}\boldsymbol{e}_{1} = \left(\left(\widehat{\boldsymbol{\mathcal{S}}}_{m}^{T}\right)^{m}\boldsymbol{e}_{1}\right)^{T}\left(\widehat{\boldsymbol{\mathcal{S}}}_{m}^{m}\boldsymbol{e}_{1}\right) = \left(\left(\widehat{\boldsymbol{\mathcal{T}}}_{m}^{T}\right)^{-m}\boldsymbol{e}_{1}\right)^{T}\left(\widehat{\boldsymbol{\mathcal{T}}}_{m}^{-m}\boldsymbol{e}_{1}\right) = \boldsymbol{e}_{1}^{T}\widehat{\boldsymbol{\mathcal{T}}}_{m}^{-2m}\boldsymbol{e}_{1}.$$

$$(4.14)$$

As (4.14) obviously also holds for lower powers of \widehat{S}_m , this concludes the proof of the lemma. \Box

Using the results of Lemma 4.1 and 4.2, we are in a position to prove Theorem 4.1.

Proof of Theorem 4.1 Let $\phi \in \mathcal{L}_{2m}^{2m-1}$. Then, $\phi(A) = p(A) + q(A^{-1})$ with $p \in \Pi_{2m-1}$ and $q \in \Pi_{2m}$. According to Lemma 4.1, this yields

$$\boldsymbol{c}^{T}\phi(A)\boldsymbol{b} = \boldsymbol{c}^{T}p(A)\boldsymbol{b} + \boldsymbol{c}^{T}q(A^{-1})\boldsymbol{b} = \boldsymbol{e}_{1}^{T}p(\widehat{\mathcal{T}}_{m})\boldsymbol{e}_{1} + \boldsymbol{e}_{1}^{T}q(\widehat{\mathcal{S}}_{m})\boldsymbol{e}_{1}.$$
 (4.15)

Further using Lemma 4.2, we can rewrite the right-hand side of (4.15) to give

$$\boldsymbol{c}^{T}\boldsymbol{\phi}(A)\boldsymbol{b} = \boldsymbol{e}_{1}^{T}\boldsymbol{p}(\widehat{\mathcal{T}}_{m})\boldsymbol{e}_{1} + \boldsymbol{e}_{1}^{T}\boldsymbol{q}(\widehat{\mathcal{T}}_{m}^{-1})\boldsymbol{e}_{1}.$$
(4.16)

Noting that the right-hand side of (4.16) is exactly $\boldsymbol{e}_1^T \phi(\widehat{\mathcal{T}}_m) \boldsymbol{e}_1$ completes the proof of the theorem.

Next, we compare the result of Theorem 4.1 to the situation one faces when using the standard extended Arnoldi method, Algorithm 1. This is very similar to the comparison of two-sided Lanczos and Arnoldi in the polynomial case in [32]. For Hermitian A, it is shown in [19] that 4m (rational) moments are matched by the matrix \mathcal{H}_m from the extended Arnoldi method, i.e., the same number as what was proven in Theorem 4.1 for the extended two-sided Lanczos method. In the non-Hermitian case, however, the situation is different. We will only briefly state the most important results in the following and refrain from giving all proofs in full detail, as most of them are almost identical to those presented before for the two-sided extended Lanczos method, with obvious modifications.

Performing *m* steps of the extended Arnoldi method for *A* and **b** produces an orthonormal basis \mathcal{V}_m of $\mathcal{E}_m(A, \mathbf{b})$ and a block upper Hessenberg matrix \mathcal{H}_m (with (2×2) -blocks) which fulfills the relation (2.3) and an inverse projection matrix $\mathcal{G}_m = \mathcal{V}_m^T A^{-1} \mathcal{V}_m$ which satisfies

$$A^{-1}\mathcal{V}_m = \mathcal{V}_m\mathcal{G}_m + \left[\boldsymbol{v}_{2m+1}, \boldsymbol{v}_{2m+2}\right]\theta_{m+1} \begin{bmatrix} \boldsymbol{e}_{2m-1}^T \\ \boldsymbol{e}_{2m}^T \end{bmatrix},$$

where

16

$$\theta_{m+1} = \begin{bmatrix} \boldsymbol{w}_{2m+1}^T \\ \boldsymbol{w}_{2m+2}^T \end{bmatrix} A^{-1} \begin{bmatrix} \boldsymbol{v}_{2m-1}, \boldsymbol{v}_{2m} \end{bmatrix} \in \mathbb{R}^{2 \times 2}.$$

The nonzero structure of the matrices \mathcal{H}_m and \mathcal{G}_m below the diagonal is exactly the same as that of $\widehat{\mathcal{T}}_m$ and $\widehat{\mathcal{S}}_m$, respectively. Thus, by exploiting this structure, one can, analogously to the proof of Lemma 4.1, show that $p(A)\mathbf{b} =$ $\|\mathbf{b}\|_2 \mathcal{V}_m p(\mathcal{H}_m) \mathbf{e}_1$ and $q(A^{-1})\mathbf{b} = \|\mathbf{b}\|_2 \mathcal{V}_m q(\mathcal{G}_m) \mathbf{e}_1$ for $p \in \Pi_{m-1}, q \in \Pi_m$. Similarly, $q(\mathcal{H}_m^{-1}) \mathbf{e}_1 = q(\mathcal{G}_m) \mathbf{e}_1$ for $q \in \Pi_m$ can be shown.

As the vector \boldsymbol{c} is not part of the iteration in this case, and one has no relation for $A^T \mathcal{V}_m$ in the standard extended Arnoldi method, one cannot use a relation similar to (4.8) to reach a higher degree of exactness, thus finding

$$\boldsymbol{c}^{T}\phi(A)\boldsymbol{b} = \|\boldsymbol{b}\|_{2}\boldsymbol{u}^{T}\phi(\mathcal{H}_{m})\boldsymbol{e}_{1} \text{ for } \phi \in \mathcal{L}_{m+1}^{m}$$
(4.17)

with $\boldsymbol{u} = \mathcal{V}_m^T \boldsymbol{c}$. In case that $\boldsymbol{b} = \boldsymbol{c}$, we can slightly improve the result of (4.17), as explained in the following. One easily proves the relations

$$\|\boldsymbol{b}\|_{2}\mathcal{V}_{m}\mathcal{H}_{m}^{m}\boldsymbol{e}_{1} = \mathcal{V}_{m}\mathcal{V}_{m}^{T}A^{m}\boldsymbol{e}_{1} \text{ and } \|\boldsymbol{b}\|_{2}\mathcal{V}_{m}\mathcal{G}_{m}^{m+1}\boldsymbol{e}_{1} = \mathcal{V}_{m}\mathcal{V}_{m}^{T}A^{-(m+1)}\boldsymbol{e}_{1},$$
(4.18)

which can be used to deduce

$$\|\boldsymbol{b}\|_{2}\boldsymbol{b}^{T}\mathcal{V}_{m}\mathcal{H}_{m}^{m}\boldsymbol{e}_{1} = \boldsymbol{b}^{T}\mathcal{V}_{m}\mathcal{V}_{m}^{T}A^{m}\boldsymbol{e}_{1}.$$
(4.19)

Equation (4.19) is equivalent to

$$\|\boldsymbol{b}\|_2^2 \boldsymbol{e}_1^T \mathcal{H}^m \boldsymbol{e}_1 = (\mathcal{V}_m \mathcal{V}_m^T \boldsymbol{b})^T A^m \boldsymbol{b} = \boldsymbol{b}^T A^m \boldsymbol{b},$$

using $\mathcal{V}_m \mathcal{V}_m^T \boldsymbol{b} = \boldsymbol{b}$, as $\mathcal{V}_m \mathcal{V}_m^T$ is the orthogonal projector onto $\mathcal{E}_m(A, \boldsymbol{b})$. Similarly, one also finds

$$\|\boldsymbol{b}\|_{2}^{2} \boldsymbol{e}_{1}^{T} \mathcal{G}^{m+1} \boldsymbol{e}_{1} = \boldsymbol{b}^{T} A^{-(m+1)} \boldsymbol{b}$$

from (4.18). For $\boldsymbol{b} = \boldsymbol{c}$, we thus have

$$\boldsymbol{b}^T \phi(A) \boldsymbol{b} = \| \boldsymbol{b} \|_2^2 \boldsymbol{e}_1^T \phi(\mathcal{H}_m) \boldsymbol{e}_1 \text{ for } \phi \in \mathcal{L}_{m+1}^m$$

A similar, general result for Laurent polynomials of higher degree is not possible, as one can easily construct examples for which $\|\boldsymbol{b}\|_2^2 \boldsymbol{e}_1^T \phi(\mathcal{H}_m) \boldsymbol{e}_1$ is inexact both for $\phi \in \mathcal{L}_{m+2}^m$ and \mathcal{L}_{m+1}^{m+1} .

5 Implementation issues

In this section we briefly comment on some topics concerning the implementation of Algorithm 3.

5.1 Early breakdown

Until now, we assumed that no breakdown occurs during the execution of the method. It can, however, happen that $\langle v_{2j+1}, w_{2j+1} \rangle$ or $\langle v_{2j+2}, w_{2j+2} \rangle$ vanishes and the method cannot be continued in its present form. As for the polynomial two-sided Lanczos methods, we have to distinguish two cases.

- (i) If the inner product vanishes because one of the involved vectors is the zero vector, this means that either $\mathcal{E}_m(A, \mathbf{b})$ or $\mathcal{E}_m(A^T, \mathbf{c})$ is invariant (depending on whether one of the \mathbf{v}_i or one of the \mathbf{w}_i is zero) and we are in the presence of a *lucky breakdown*. In this case, $f(A)\mathbf{b} \in \mathcal{E}_m(A, \mathbf{b})$ or $f(A)^T \mathbf{c} \in \mathcal{E}_m(A^T, \mathbf{c})$, and we retrieve the exact value of $\mathbf{c}^T f(A)\mathbf{b}$.
- (ii) If the inner product vanishes, but both involved vectors are nonzero, the method suffers from a *serious breakdown*. In this case, we do not know anything about the approximation properties of the computed subspaces. In order to be able to continue the iteration, one may, e.g., apply *look-ahead techniques* similar to those used for the polynomial two-sided Lanczos method (if possible), see, e.g., [8]. As this is a rather technical and extensive topic, it is well beyond the scope of this paper to go into details on this.

5.2 Dealing with the linear system solves

Another important topic on which we want to comment is the practicability of the proposed method. At first sight, it may seem that investing an additional linear system solve with A^T may not be worth the saved orthogonalization cost due to short recurrences, in contrast to the polynomial case, where one just has to invest an additional matrix-vector product with A^T . We can, however, make the following points for our method:

- (i) If the linear systems with A are solved by a direct method, i.e., by computing an LU-factorization A = LU, one can re-use this factorization also for the systems with A^T , as then $A^T = U^T L^T$ is an LU-factorization of A^T . Thus, if the forward-backward substitution for solving the systems is significantly cheaper than computing the factorization, the additional cost introduced by the second linear system may be acceptable.
- (ii) In case one uses an iterative method for the linear system solves (resulting in a so-called *inner-outer method* as discussed, e.g., in [6,27] in the context of the shift-invert Lanczos method), it is possible to use a solver based on the two-sided Lanczos process, see, e.g. [28, Section 7.2], which can solve both systems simultaneously, so that one obtains the solution of the second linear system essentially for free.
- (iii) When approximating a bilinear form $c^T f(A) b$ for $b \neq c$ and a Hermitian matrix A by the (extended) Lanczos method, one commonly does so by rewriting the bilinear form as

$$\boldsymbol{c}^{T} f(A) \boldsymbol{b} = \frac{1}{4} (\boldsymbol{c} + \boldsymbol{b})^{T} f(A) (\boldsymbol{c} + \boldsymbol{b}) - \frac{1}{4} (\boldsymbol{c} - \boldsymbol{b})^{T} f(A) (\boldsymbol{c} - \boldsymbol{b}), \quad (5.1)$$



Fig. 5.1: Magnitude of the entries of $\mathcal{V}_{20}^T \mathcal{W}_{20}$ (logarithmic scale) when applying (a) Algorithm 2, (b) Algorithm 2 with short recurrences, (c) Algorithm 2 to a three-dimensional convection-diffusion problem.

and then applies the method to both terms on the right-hand side of (5.1) separately, see, e.g. [12, Section 11.1]. In this case, assuming that approximating both terms to the desired accuracy takes about the same number of iterations, the computational work is also doubled in this case, which therefore seems to be quite natural in the presence of a bilinear form involving two different vectors $\mathbf{b} \neq \mathbf{c}$, even in the Hermitian case.

5.3 Stability in floating-point arithmetic

It is known in the polynomial Krylov case that two-sided methods may be prone to numerical instabilities. In particular, the computed bases $\mathcal{V}_m, \mathcal{W}_m$ are often not bi-orthonormal any longer after a rather small number of iterations of the method in finite precision computations, mainly caused by the use of short recurrences. Additional potential for instability lies in the recursion formulas (3.6)-(3.13) for computing the orthogonalization coefficients. While a detailed analysis of the behavior of the method in floating point arithmetic is beyond the scope of this paper, we at least illustrate the behavior one can expect by a small numerical example. We run the same experiment using (i) Algorithm 2, (ii) Algorithm 2 with short recurrences but explicit computation of all coefficients, and (iii) Algorithm 3. We apply the method to a matrix A corresponding to the finite difference discretization of a three-dimensional convection-diffusion equation with N = 32 points in each spatial direction and Péclet numbers $Pe_1 = 40$, $Pe_2 = 20$, resulting in a highly nonsymmetric problem (see also Section 6 for a more detailed description of a similar model problem). Figure 5.1 depicts the magnitude of the entries of $\mathcal{V}_{20}^T \mathcal{W}_{20}$ for the three different versions of the algorithm. As is expected, the orthogonality of the basis vectors is lost in floating point arithmetic after some iterations when using short recurrences, especially in Algorithm 3. In figure 5.2, the convergence behavior of the three methods for approximating $\mathbf{c}^T f(A) \mathbf{b}$ with random vectors \mathbf{b}, \mathbf{c} and $f(z) = z^{-1/2}$ is shown. Despite the severe loss of orthogonality in Algorithm 3, all methods behave almost exactly the same and manage to reach a very high accuracy, indicating that the behavior observed in Figure 5.1



Fig. 5.2: Error norm when using (a) Algorithm 2, (b) Algorithm 2 with short recurrences, (c) Algorithm 2 for approximating $c^T A^{-1/2} b$ for a three-dimensional convection-diffusion problem.

does not necessarily mean that the method will fail to produce results of high accuracy.

6 Numerical experiments

In this Section, we illustrate some properties of the presented method by means of a few numerical examples. All experiments are performed in MATLAB R2015a.

We begin by investigating a small-scale problem which mainly serves the purpose to confirm the result of Theorem 4.1 by numerical evidence. Consider the tridiagonal, nonsymmetric matrix

$$A = \operatorname{tridiag}(1, 2, -1) \in \mathbb{R}^{100 \times 100},$$

b the normalized vector of all ones and **c** the first canonical unit vector, scaled such that $\langle \boldsymbol{b}, \boldsymbol{c} \rangle = 1$. We approximate $\boldsymbol{c}^T f(A) \boldsymbol{b}$ for the Laurent polynomial $f(z) = z^5 + z^{-6}$ by the extended two-sided Lanczos method and the extended Arnoldi method. The convergence history of both methods is depicted in Figure 6.1. As predicted by Theorem 4.1, the extended two-sided Lanczos method finds the exact value of the bilinear form (up to approximately machine precision) after m = 3 iterations, which is not the case for the extended Arnoldi method (which, in addition, shows unstable behavior after reaching an error norm slightly below 10^{-9} and fails to converge closer to the exact value.

For our next experiment, we choose the matrix vanHeukelum/cage11 from the University of Florida Sparse Matrix Collection [4]. The matrix A is the adjacency matrix of a directed, weighted graph with 39,082 nodes. We approximate the upper-left entry of the exponential of A, i.e., $e_1^T \exp(A)e_1$, which



Fig. 6.1: Convergence history of the extended two-sided Lanczos and extended Arnoldi method for approximating $\boldsymbol{c}^T(A^5 + A^{-6})\boldsymbol{b}$ where $A = \operatorname{tridiag}(1, 2, -1) \in \mathbb{R}^{100 \times 100}$.

corresponds to the *subgraph centrality* of node *i* in the graph. Thus, we are in the situation $\mathbf{b} = \mathbf{c}$, in which the extended Arnoldi method can be expected to obtain slightly better results than for $\mathbf{b} \neq \mathbf{c}$, cf. Section 4. As the results in Figure 6.2 show, the two-sided extended Lanczos method again finds are very accurate approximation rapidly, and it again shows superior stability in comparison to the extended Arnoldi method.

In a last experiment, we compare our two-sided extended Krylov subspace method to the polynomial two-sided Lanczos, in order to illustrate that there are indeed situations in which it might be beneficial to invest the additional work for linear system solves in a practical application. We consider a semidiscretization of the partial differential equation

$$\frac{\partial u}{\partial t} - \Delta u + \tau_1 \frac{\partial u}{\partial x_1} + \tau_2 \frac{\partial u}{\partial x_2} = 0 \quad \text{on } (0,1)^2 \times (0,T),$$
$$u(x,t) = 0 \quad \text{on } \partial(0,1)^2 \text{ for all } t \in [0,T], \quad (6.1)$$
$$u(x,0) = u_0(x) \text{ for all } x \in (0,1)^2.$$

Discretizing the differential operator $-\Delta u + \tau_1 \frac{\partial u}{\partial x_1} + \tau_2 \frac{\partial u}{\partial x_2}$ by central differences with uniform discretization step size $h = \frac{1}{n+1}$ yields the matrix

$$A = -\frac{1}{h^2} \left(I \otimes C_1 + C_2 \otimes I \right) \in \mathbb{R}^{n^2 \times n^2}$$
(6.2)

20



Fig. 6.2: Convergence history of the extended two-sided Lanczos and extended Arnoldi method for approximating $e_1^T \exp(A)e_1$ where $A \in \mathbb{R}^{39,082 \times 39,082}$ is the adjacency matrix of a weighted, directed graph.

Table 6.1: Number of iterations necessary to approximate $\boldsymbol{c}^T \exp(-tA)\boldsymbol{b}$ to an accuracy of 10^{-3} , where A is given by (6.2), for different n by the polynomial and extended two-sided Lanczos method.

	n = 32	n = 64	n = 128	n = 256
polynomial	7	7	26	99
extended	4	6	6	6

with

$$C_{i} = \begin{bmatrix} -2 & 1 - \frac{\tau_{i}h}{2} \\ 1 + \frac{\tau_{i}h}{2} & -2 & 1 - \frac{\tau_{i}h}{2} \\ & 1 + \frac{\tau_{i}h}{2} & \ddots & \ddots \\ & & \ddots & \ddots & 1 - \frac{\tau_{i}h}{2} \\ & & 1 + \frac{\tau_{i}h}{2} & -2 \end{bmatrix} \in \mathbb{R}^{n \times n}, i = 1, 2;$$

see, e.g., [26]. Approximating quantities of the form $\mathbf{c}^T \exp(-tA)\mathbf{b}$, where A is a discretized differential operator, is, e.g., of importance in so-called *Krylov Subspace Spectral methods*, see, e.g., [21, 22]. We consider a discretization of (6.1) with varying number of discretization points, with the convection coefficients τ_i , i = 1, 2 chosen such that the Péclet numbers $\operatorname{Pe}_i = \frac{\tau_i h}{2}$ are equal to $\operatorname{Pe}_1 = .2$ and $\operatorname{Pe}_2 = .1$, respectively, for all discretizations. In Table 6.1, we compare the number of iterations necessary for the extended and polynomial two-sided Lanczos method to approximate $\mathbf{c}^T \exp(-tA)\mathbf{b}$ to an accuracy of 10^{-3} , where the time step is t = .005 and \mathbf{b}, \mathbf{c} are chosen randomly.

We observe that the extended two-sided Lanczos method scales extremely well, with the number of iterations being almost independent of the discretization step size, while the number of iterations in the polynomial two-sided Lanczos method increases rapidly for smaller discretization step sizes. In addition, we observed that the polynomial Lanczos method often fails to reach higher accuracies than 10^{-3} for larger n, while this was not a problem in the extended method.

7 Conclusions

We presented a new algorithm for computing bi-orthonormal bases of extended Krylov subspaces by short (five-term) recurrences, which can be seen as an analogue to the (polynomial) two-sided Lanczos method.

We investigated the relation of this method to rational moment matching and proved that after m steps, bilinear forms $c^T \phi(A) b$, where ϕ is a Laurent polynomial of denominator degree 2m and numerator degree 2m - 1, are evaluated exactly. In contrast, we showed that m steps of the standard extended Arnoldi method only evaluate Laurent polynomials of denominator degree mand numerator degree m - 1 exactly (or m + 1 and m, respectively, if b = c).

We only touched very briefly on the topics of breakdown and performance of the method in finite precision arithmetic. A thorough investigation of these issues was well beyond the scope of this paper, but seems to be an interesting topic for future research. Another direction for further research is the derivation of short recurrences for extended Krylov subspaces where the number of steps taken in the "positive" and "negative direction" is different, thus allowing to apply more matrix-vector products than linear system solves.

Acknowledgment The author would like to thank Michele Benzi for providing a very useful overview of literature on application areas for the proposed method.

References

- S. BARONI, R. GEBAUER, O. B. MALCIOĞLU, Y. SAAD, P. UMARI, AND J. XIAN, Harnessing molecular excited states with Lanczos chains, J. Phys-Condens. Mat., 22 (2010), p. 074204.
- M. BENZI AND P. BOITO, Quadrature rule-based bounds for functions of adjacency matrices, Linear Algebra Appl., 433 (2010), pp. 637–652.
- D. R. BOWLER AND T. MIYAZAKI, O(n) methods in electronic structure calculations, Rep. Prog. Phys., 75 (2012), p. 036503.
- T. A. DAVIS AND Y. HU, The University of Florida Sparse Matrix Collection, ACM T. Math. Software, 38 (2011), pp. 1–25.
- V. DRUSKIN AND L. KNIZHNERMAN, Extended Krylov subspaces: Approximation of the matrix square root and related functions, SIAM J. Matrix Anal. Appl., 19 (1998), pp. 775–771.
- J. VAN DEN ESHOF AND M. HOCHBRUCK, Preconditioning Lanczos approximations to the matrix exponential, SIAM J. Sci. Comput., 27 (2006), pp. 1438–1457.
- E. ESTRADA AND D. HIGHAM, Network properties revealed through matrix functions, SIAM Rev., 52 (2010), pp. 696–714.

- 8. R. W. FREUND, M. H. GUTKNECHT, AND N. M. NACHTIGAL, An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices, SIAM J. Sci. Comput., 14 (1993), pp. 137-158.
- 9. G. H. GOLUB, Bounds for matrix moments, Rocky Mt. J. Math., 4 (1974), pp. 207-212.
- 10. G. H. GOLUB AND G. MEURANT, Matrices, moments and quadrature, in Numerical Analysis 1993, Pitman Research Notes in Mathematics Series, Vol. 303, D. F. Griffiths and G. A. Watson, eds., Essex, England, 1994, pp. 105-156.
- 11. , Matrices, moments and quadrature. II. How to compute the norm of the error in iterative methods, BIT, 37 (1997), pp. 687-705.
- -, Matrices, Moments and Quadrature with Applications, Princeton Series in Ap-12.plied Mathematics, Princeton University Press, Princeton and Oxford, 2010. H. GUO AND R. A. RENAUT, Estimation of $u^T f(A)v$ for large-scale unsymmetric ma-
- 13.trices, Numer. Linear Algebra Appl., 11 (2004), pp. 75-89.
- 14. S. GÜTTEL, Rational Krylov Methods for Operator Functions, PhD thesis, Fakultät für Mathematik und Informatik der Technischen Universität Bergakademie Freiberg, 2010.
- 15. S. GÜTTEL, Rational Krylov approximation of matrix functions: Numerical methods and optimal pole selection, GAMM-Mitteilungen, 36 (2013), pp. 8–31.
- 16. S. GÜTTEL AND L. KNIZHNERMAN, A black-box rational Arnoldi variant for Cauchy-Stieltjes matrix functions, BIT, 53 (2013), pp. 595-616.
- 17.C. JAGELS AND L. REICHEL, The extended Krylov subspace method and orthogonal Laurent polynomials, Linear Algebra Appl., 431 (2009), pp. 441–458.
- 18. -, Recursion relations for the extended Krylov subspace method, Linear Algebra Appl., 434 (2011), pp. 1716–1732.
- 19., The structure of matrices in rational Gauss quadrature, Math. Comp., 82 (2013), pp. 2035–2060.
- 20. 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.
- J. V. LAMBERS, A spectral time-domain method for computational electrodynamics, 21.Adv. Appl. Math. Mech., 1 (2009), pp. 781-798.
- 22., Solution of time-dependent PDE through component-wise approximation of matrix functions, IAENG, Int. J. Appl. Math., 41 (2011), pp. 1-10.
- 23.C. LANCZOS, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Res. Nat. Bur. Stand., 45 (1950), pp. 255-282.
- -, Solutions of systems of linear equations by minimized iterations, J. Res. Nat. 24.Bur. Stand., 49 (1952), pp. 33–53.
- 25.C. MERTENS AND R. VANDEBRIL, Short recurrences for computing extended Krylov bases for Hermitian and unitary matrices, Numer. Math., 131 (2015), pp. 303-328.
- 26.I. MORET AND P. NOVATI, RD-rational approximations of the matrix exponential., BIT, 44 (2004), pp. 595-615.
- 27. M. POPOLIZIO AND V. SIMONCINI, Acceleration techniques for approximating the matrix exponential operator, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 657-683.
- 28.Y. SAAD, Iterative Methods for Sparse Linear Systems, 2nd ed., SIAM, 2003.
- 29.Y. SAAD, J. R. CHELIKOWSKY, AND S. M. SHONTZ, Numerical methods for electronic structure calculations of materials, SIAM Rev., 52 (2010), pp. 3-54.
- 30. V. SIMONCINI, A new iterative method for solving large-scale Lyapunov matrix equations, SIAM J. Sci. Comput., 29 (2007), pp. 1268-1288.
- 31. -, Extended Krylov subspace for parameter dependent systems, Appl. Numer. Math., 60 (2010), pp. 550–560.
- 32. Z. STRAKOŠ, Model reduction using the Vorobyev moment problem, Numer. Algorithms, 51 (2009), pp. 363–379.
- 33. Z. STRAKOŠ AND P. TICHÝ, On efficient numerical approximation of the bilinear form c*A⁻¹b, SIAM J. Sci. Comput., 33 (2011), pp. 565–587.