

THE EXTENDED GLOBAL LANCZOS METHOD FOR MATRIX FUNCTION APPROXIMATION*

A. H. BENTBIB[†], M. EL GHOMARI[†], C. JAGELS[‡], K. JBILOU[§], AND L. REICHEL[¶]

Dedicated to Walter Gautschi on the occasion of his 90th birthday

Abstract. The need to compute the trace of a large matrix that is not explicitly known, such as the matrix $\exp(A)$, where A is a large symmetric matrix, arises in various applications including in network analysis. The global Lanczos method is a block method that can be applied to compute an approximation of the trace. When the block size is one, this method simplifies to the standard Lanczos method. It is known that for some matrix functions and matrices, the extended Lanczos method, which uses subspaces with both positive and negative powers of A , can give faster convergence than the standard Lanczos method, which uses subspaces with nonnegative powers of A only. This suggests that it may be beneficial to use an extended global Lanczos method instead of the (standard) global Lanczos method. This paper describes an extended global Lanczos method and discusses properties of the associated Gauss-Laurent quadrature rules. Computed examples that illustrate the performance of the extended global Lanczos method are presented.

Key words. extended Krylov subspace, extended moment matching, Laurent polynomial, global Lanczos method, matrix function, Gauss quadrature rule

AMS subject classifications. 65F25, 65F30, 65F60, 33C47

1. Introduction. Let $A \in \mathbb{R}^{n \times n}$ be a large, symmetric matrix, and let $V \in \mathbb{R}^{n \times s}$ be a block vector with $1 \leq s \ll n$. We are interested in approximating expressions of the form

$$(1.1) \quad \mathcal{I}(f) := \text{trace}(V^T f(A)V),$$

where f is a function that is defined on the convex hull of the spectrum of A . The need to evaluate bilinear forms (1.1) arises in various applications such as in network analysis (when $f(t) := \exp(t)$ or $f(t) := t^3$) [3, 9], machine learning [17], electronic structure computation [2, 5, 19], and the solution of ill-posed problems [10, 11].

Bellalij et al. [3] observed that the right-hand side of (1.1) can be expressed as a Stieltjes integral as follows. Introduce the spectral factorization

$$(1.2) \quad A = U\Lambda U^T, \quad \Lambda = \text{diag}[\lambda_1, \dots, \lambda_n],$$

where $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of A and the matrix $U \in \mathbb{R}^{n \times n}$ of eigenvectors is orthogonal. Substituting the spectral factorization (1.2) into the expression (1.1) and setting $\widehat{V} = U^T V$ gives

$$(1.3) \quad \mathcal{I}(f) = \text{trace}(V^T f(A)V) = \sum_{i=1}^n f(\lambda_i) \text{trace}(\widehat{V}_i^T \widehat{V}_i) = \int f(\lambda) d\alpha(\lambda),$$

where $\alpha(\lambda)$ is a nondecreasing real-valued piecewise constant distribution function with possible discontinuities at the eigenvalues λ_i of A ; $d\alpha(\lambda)$ is the associated measure.

*Received November 24, 2018. Accepted November 28, 2018. Published online on January 14, 2019. Recommended by G. Milovanovic.

[†]Laboratory LAMAI, Faculty of Sciences and Technologies, Cadi Ayyad University, Marrakech, Morocco ({a.bentbib@uca.ac.ma, m.elghomari10@gmail.com}).

[‡]Department of Mathematics, Hanover College, Hanover, IN 47243, USA (jagels@hanover.edu).

[§]Laboratory LMPA, 50 Rue F. Buisson, ULCO Calais cedex, France (Khalide.Jbilou@univ-littoral.fr).

[¶]Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA (reichel@math.kent.edu).

Consider the global Krylov subspace

$$(1.4) \quad \mathbb{K}_g^m(A, V) = \text{span}\{V, AV, \dots, A^{m-1}V\} = \{p(A)V : p \in \Pi_{m-1}\},$$

where Π_{m-1} denotes the set of polynomials of degree at most $m - 1$. We tacitly assume here that the block vectors $V, AV, \dots, A^{m-1}V$ are linearly independent. This is the generic situation. An application of m steps of the global Lanczos method to A with initial block vector V then determines the decomposition

$$(1.5) \quad A\mathbb{V}_m = \mathbb{V}_m(T_m \otimes I_s) + \beta_{m+1}V_{m+1}E_m^T,$$

where the $n \times s$ block columns of the matrix $\mathbb{V}_m = [V_1, \dots, V_m] \in \mathbb{R}^{n \times ms}$ form an F -orthonormal basis for the global Krylov subspace (1.4), the initial block vector V_1 is proportional to V , and we assume that $ms \ll n$. The block columns V_1, \dots, V_{m+1} are said to be F -orthonormal if

$$(1.6) \quad (V_j, V_k)_F := \text{trace}(V_j^T V_k) = \begin{cases} 1 & j = k, \\ 0 & j \neq k. \end{cases}$$

The inner product $(\cdot, \cdot)_F$ is associated with the Frobenius norm; we have $\|M\|_F := (M, M)_F^{1/2}$ for any matrix M . Moreover, the matrix $T_m \in \mathbb{R}^{m \times m}$ in (1.5) is symmetric and tridiagonal, $I_s \in \mathbb{R}^{s \times s}$ denotes the identity matrix, $\beta_{m+1} \geq 0$, and $E_m \in \mathbb{R}^{n \times s}$ is the m th block axis vector, i.e., E_m is made up of the columns $(m - 1)s + 1, \dots, ms$ of I_n . The superscript T denotes transposition, and \otimes stands for the Kronecker product. We refer to [8, 15] for derivations of the global Lanczos decomposition (1.5). When the matrix A is large, the dominating cost of computing the decomposition (1.5) is the evaluation of $m - 1$ matrix-block-vector products with the matrix A .

It is attractive to approximate the expression (1.1) by

$$(1.7) \quad \mathcal{G}_m(f) := \|V\|_F^2 e_1^T f(T_m) e_1,$$

because this approximation can be interpreted as an m -point Gauss quadrature rule for the Stieltjes integral (1.3). We have

$$(1.8) \quad \mathcal{I}(p) = \mathcal{G}_m(p) \quad \forall p \in \Pi_{2m-1};$$

see Bellalij et al. [3] for details. The right-hand side of (1.7) can be written as a sum of m terms by a substitution of the spectral factorization of T_m . This substitution shows that the Gaussian nodes are the eigenvalues of T_m and the Gaussian weights are the squares of the first component of the normalized eigenvectors times the square Frobenius norm of V ; see [3] for details. We remark that the main reason for using the approximation (1.7) of (1.1) is that it obviates the need to evaluate $f(A)$, which may be prohibitively expensive when A is large. Instead, only the evaluation of $f(T_m)$, where T_m typically is a fairly small matrix, is required. Numerical methods for the accurate evaluation of f for a small matrix are described by Higham [13].

In the context of approximating expressions of the form $f(A)b$ for some vector $b \in \mathbb{R}^n$, Druskin, Knizhnerman, and Simoncini [7, 16] have shown that it may be possible to approximate such an expression more accurately when using an approximant from an extended Krylov subspace

$$\mathbb{K}^{m+1, m+1}(A, b) = \text{span}\{A^{-m}b, \dots, A^{-1}b, b, Ab, \dots, A^m b\}$$

than when using an approximant in the standard Krylov subspace

$$\mathbb{K}^{2m+1}(A, b) = \text{span}\{b, Ab, \dots, A^{2m}b\}$$

of the same dimension $2m + 1$. We are interested in exploring whether approximations of (1.1) from an extended global Krylov subspace

$$(1.9) \quad \begin{aligned} \mathbb{K}_g^{m+1, m+1}(A, V) &= \text{span}\{V, A^{-1}V, AV, \dots, A^{-m}V, A^mV\} \\ &= \{q(A)V, q = p_1/p_2, p_1, p_2 \in \Pi_m\} \end{aligned}$$

give a higher accuracy than an approximant from the standard global Krylov subspace $\mathbb{K}_g^{2m+1}(A, V)$ of the same dimension.

A possible drawback of the extended global Lanczos method is that the generation of an F -orthonormal basis for the subspace (1.9) requires the solution of m linear systems of equations with the matrix A and right-hand sides that are block vectors; see Algorithm 1 for details. When the matrix A is positive definite and banded with a small bandwidth or when the systems of equations can be solved efficiently by a multigrid method, the generation of an orthonormal basis for the subspace (1.9) may not be much more demanding than computing an F -orthonormal basis for the (standard) global Krylov subspace $\mathbb{K}_g^{2m+1}(A, V)$ with the (standard) global Lanczos method; the latter method does not require the solution of linear systems of equations with the matrix A ; see [3, 8, 15]. We illustrate in Section 4 that the extended global Lanczos method may give approximations of (1.1) of higher accuracy than the (standard) global Lanczos method when these methods use subspaces of the same dimension.

This paper is organized as follows. Section 2 introduces the extended global Lanczos process. The orthonormal basis generated by this method is expressed with the aid of orthonormal Laurent polynomials. Section 3 describes the application of the extended global Lanczos process to the approximation of matrix functions (1.1). We also discuss the analogue of property (1.8) for the approximation by Laurent polynomials. Section 4 presents a few numerical experiments that illustrate the quality of the computed approximations of (1.1). Concluding remarks can be found in Section 5.

2. The extended global Lanczos method. We describe the extended global Lanczos method for generating an orthonormal basis for the extended global Krylov subspace (1.9). We assume that the symmetric matrix $A \in \mathbb{R}^{n \times n}$ and the block vector $V \in \mathbb{R}^{n \times s}$ are such that the subspace (1.9) is of full dimension $2m + 1$. The extended global Lanczos method uses short recursion relations for generating an orthonormal basis for (1.9). The existence of short recursion relations follows from the fact that orthogonal Laurent polynomials with respect to an inner product determined by a nonnegative measure with support on (part of) the real axis satisfy short recursion relations. The latter was first observed by Njåstad and Thron [18]. Applications in linear algebra have been described by Simoncini [20] and in [14]. Both the latter references are concerned with the situation when V in (1.9) is a single vector. Simoncini [20] also discusses the generation of an orthonormal basis for an extended Krylov subspace when the matrix A is not symmetric and V is a single vector. Then an orthonormal basis cannot be generated with short recurrence relations.

Heyouni [12] derives an extended global Arnoldi method for generating an orthonormal basis for an extended global Krylov subspace (1.9) when $A \in \mathbb{R}^{n \times n}$ is nonsymmetric and applies this method to solve continuous-time algebraic Riccati equations. The generation of the basis requires “long” recursion relations. We first describe properties and recursion formulas for this method and subsequently explore simplifications that arise when A is symmetric.

Following Heyouni [12], the recursion formulas for the extended global Arnoldi method can be written as

$$(2.1) \quad w_{1,1}V_1 = V, \quad \widehat{V}_2 = A^{-1}V - w_{1,2}V_1, \quad w_{2,2}V_2 = \widehat{V}_2,$$

where $w_{1,1} = \|V\|_F$, $w_{1,2} = (A^{-1}V, V_1)_F$, $w_{2,2} = \|\widehat{V}_2\|_F$, and

$$(2.2) \quad \begin{aligned} h_{2j+1,2j-1}V_{2j+1} &= AV_{2j-1} - \sum_{i=1}^{2j} h_{i,2j-1}V_i, \\ h_{2j+2,2j}V_{2j+2} &= A^{-1}V_{2j} - \sum_{i=1}^{2j+1} h_{i,2j}V_i, \end{aligned} \quad j = 1, 2, \dots, m.$$

The coefficients $h_{i,j}$ are determined so that the block vectors $V_1, V_2, \dots, V_{2m+2}$ are F -orthonormal.

Before describing our proposed algorithm based on pairs of three-term recurrence relations, we provide some useful properties of the extended global Krylov subspaces generated. We will tacitly assume that a breakdown does not occur.

PROPERTIES 1. *Let the matrix $A \in \mathbb{R}^{n \times n}$ be nonsingular, and let $V \in \mathbb{R}^{n \times s}$. Then*

$$\begin{aligned} \mathbb{K}_g^{m+1,m}(A, V) &= \mathbb{K}_{2m}^g(A, A^{-m}V), \\ \mathbb{K}_g^{m+1,m}(A, V) &\subseteq \mathbb{K}_{2m}^g(A, A^{-m+1}V) \subseteq \mathbb{K}_g^{m+2,m+1}(A, V), \\ A^{-1}\mathbb{K}_g^{m+1,m}(A, V) &\subseteq \mathbb{K}_{2m}^g(A, A^{-m-1}V) \subseteq \mathbb{K}_g^{m+2,m+1}(A, V). \end{aligned}$$

Proof. The properties follow straightforwardly from the definition of the global Krylov subspace (1.4) and the extended global Krylov subspace (1.9). \square

PROPERTIES 2. *Let the matrix $A \in \mathbb{R}^{n \times n}$ be nonsingular, and let $V \in \mathbb{R}^{n \times s}$. The F -orthonormal basis $V_1, V_2, \dots, V_{2m+2}$ determined by the recursion formulas (2.1) and (2.2) satisfy, for $j = 1, 2, \dots, m$,*

$$(2.3) \quad AV_{2j-1} \in \text{span}\{V_1, \dots, V_{2j+1}\},$$

$$(2.4) \quad AV_{2j} \in \text{span}\{V_1, \dots, V_{2j+1}\},$$

$$A^{-1}V_{2j-1} \in \text{span}\{V_1, \dots, V_{2j}\},$$

$$A^{-1}V_{2j} \in \text{span}\{V_1, \dots, V_{2j+2}\}.$$

Proof. These properties follow from equations (2.1), (2.2), and the formulas of Properties 1. \square

The existence of three-term recurrence relations for the block vectors V_1, \dots, V_{2m} is a consequence of the following properties.

PROPERTIES 3. *Let $A \in \mathbb{R}^{n \times n}$ be a nonsingular symmetric matrix, and let $V \in \mathbb{R}^{n \times s}$. Then, for $j = 1, 2, \dots, m$,*

$$(2.5) \quad AV_{2j-1} \in \text{span}\{V_{2j-3}, V_{2j-2}, V_{2j-1}, V_{2j}, V_{2j+1}\},$$

$$(2.6) \quad AV_{2j} \in \text{span}\{V_{2j-1}, V_{2j}, V_{2j+1}\},$$

$$(2.7) \quad A^{-1}V_{2j-1} \in \text{span}\{V_{2j-2}, V_{2j-1}, V_{2j}\},$$

$$(2.8) \quad A^{-1}V_{2j} \in \text{span}\{V_{2j-2}, V_{2j-1}, V_{2j}, V_{2j+1}, V_{2j+2}\}.$$

Proof. The assertions require A to be symmetric, i.e., that $(AV, W)_F = (AW, V)_F$ with $V, W \in \mathbb{R}^{n \times s}$. We first show (2.6). It follows from (2.4) that AV_{2j} can be written as

$$AV_{2j} = \sum_{i=1}^j (\alpha_{2i,2j}V_{2i} + \alpha_{2i-1,2j}V_{2i-1}) + \alpha_{2j+1,2j}V_{2j+1},$$

where $\alpha_{1,2j}, \alpha_{2,2j}, \dots, \alpha_{2j+1,2j}$ are scalars. It suffices to show that the coefficients $\alpha_{1,2j}, \alpha_{2,2j}, \dots, \alpha_{2j-2,2j}$ vanish. Since A is a symmetric matrix, we have

$$\begin{aligned}\alpha_{2i-1,2j} &= (V_{2i-1}, AV_{2j})_F = (AV_{2i-1}, V_{2j})_F, \\ \alpha_{2i,2j} &= (V_{2i}, AV_{2j})_F = (AV_{2i}, V_{2j})_F,\end{aligned}$$

and using (2.3) and (2.4), we obtain

$$AV_{2i-1} \in \text{span}\{V_1, \dots, V_{2i+1}\} \quad \text{and} \quad AV_{2i} \in \text{span}\{V_1, \dots, V_{2i+1}\}.$$

Due to the F -orthonormality of V_1, \dots, V_{2j} , the coefficients $\alpha_{1,2j}, \dots, \alpha_{2j-2,2j}$ vanish. It follows that

$$AV_{2j} \in \text{span}\{V_{2j-1}, V_{2j}, V_{2j+1}\}.$$

The properties (2.5), (2.7), and (2.8) can be shown in a similar fashion. \square

We are now in a position to describe the extended global Lanczos method. Notice that the algorithm that implements this method is based on Properties 2. To link the method to a sequence of orthonormal Laurent polynomials, we use (2.7) and (2.6). The properties (2.5) and (2.8) will be used implicitly in Propositions 3.1 and 3.2 below to determine the projected tridiagonal matrices T_{2m} and S_{2m} . The algorithm for the proposed extended global Lanczos method computes an F -orthonormal basis $\{V_1, \dots, V_{2m}\}$ by pairs of three-term recurrence relations. In order to be able to solve all required linear systems of equations with the matrix A in the algorithm, we require that A satisfies

$$\text{trace}(V^T AV) \neq 0 \quad \forall V \in \mathbb{R}^{n \times s}.$$

A sufficient condition for this to hold is that A is symmetric positive definite. We will assume this to be the case.

The extended global Lanczos algorithm can be expressed as follows. After initialization

$$V_1 = V/\delta_1 \quad \text{with} \quad \delta_1 = \|V\|_F, \quad \beta_0 V_0 = 0,$$

the algorithm is determined by the recursion relations which follow from (2.7) and (2.6), for $j = 1, \dots, m$:

$$(2.9) \quad \begin{aligned}\delta_{2j} V_{2j} &= (A^{-1} - \beta_{2j-1} I_n) V_{2j-1} - \beta_{2j-2} V_{2j-2}, \\ \delta_{2j+1} V_{2j+1} &= (A - \alpha_{2j} I_n) V_{2j} - \alpha_{2j-1} V_{2j-1},\end{aligned}$$

with

$$\begin{aligned}\beta_i &= (A^{-1} V_{2j-1}, V_i)_F, & \text{for } i \in \{2j-2, 2j-1\}, \\ \alpha_i &= (AV_{2j}, V_i)_F, & \text{for } i \in \{2j-1, 2j\}.\end{aligned}$$

The $\delta_i > 0$ are normalization coefficients such that $\|V_i\|_F = 1$ for $i \in \{2j, 2j+1\}$. Algorithm 1 summarizes the computations.

We note that Algorithm 1 may break down when some δ_i becomes (numerically) zero because then the next block vector V_i cannot be computed. Breakdown is very rare, and we will assume that all δ_i determined by the algorithm are numerically positive. Then Algorithm 1 computes an F -orthonormal basis V_1, \dots, V_{2m+2} of block vectors $V_j \in \mathbb{R}^{n \times s}$ of the extended global Krylov subspace $\mathbb{K}_g^{m+1, m+1}$.

Algorithm 1 The extended global Lanczos algorithm.

Input: symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$, initial block vector $V \in \mathbb{R}^{n \times s}$

Output: F-orthonormal basis $V_1, \dots, V_{2(m+1)}$ of $\mathbb{K}_g^{m+1, m+1}(A, V)$

- 1: $\delta_1 = \|V\|_F$; $V_1 = V/\delta_1$; $V_0 = 0$;
 - 2: **for** $j = 1$ **to** m
 - 3: $\widehat{V} = A^{-1}V_{2j-1}$;
 - 4: $\beta_{2j-2} = (\widehat{V}, V_{2j-2})_F$; $\widehat{V} = \widehat{V} - \beta_{2j-2}V_{2j-2}$;
 - 5: $\beta_{2j-1} = (\widehat{V}, V_{2j-1})_F$; $\widehat{V} = \widehat{V} - \beta_{2j-1}V_{2j-1}$;
 - 6: $\delta_{2j} = \|\widehat{V}\|_F$; $V_{2j} = \widehat{V}/\delta_{2j}$;
 - 7: $\widehat{V} = AV_{2j}$;
 - 8: $\alpha_{2j-1} = (\widehat{V}, V_{2j-1})_F$; $\widehat{V} = \widehat{V} - \alpha_{2j-1}V_{2j-1}$;
 - 9: $\alpha_{2j} = (\widehat{V}, V_{2j})_F$; $\widehat{V} = \widehat{V} - \alpha_{2j}V_{2j}$;
 - 10: $\delta_{2j+1} = \|\widehat{V}\|_F$; $V_{2j+1} = \widehat{V}/\delta_{2j+1}$;
 - 11: **end for**
 - 12: $\widehat{V} = A^{-1}V_{2m+1}$;
 - 13: $\beta_{2m} = (\widehat{V}, V_{2m})_F$; $\widehat{V} = \widehat{V} - \beta_{2m}V_{2m}$;
 - 14: $\beta_{2m+1} = (\widehat{V}, V_{2m+1})_F$; $\widehat{V} = \widehat{V} - \beta_{2m+1}V_{2m+1}$;
 - 15: $\delta_{2m+2} = \|\widehat{V}\|_F$; $V_{2m+2} = \widehat{V}/\delta_{2m+2}$;
-

Below we show some properties of the basis V_1, \dots, V_{2m+2} determined by Algorithm 1. In particular, we show how the block vectors V_j are related to orthogonal Laurent polynomials. The following spaces of Laurent polynomials of degree $2m - 1$ and $2m$ will be used:

$$\begin{aligned}\Delta_{-m, m-1} &= \text{span}\{1, x^{-1}, x, \dots, x^{m-1}, x^{-m}\}, \\ \Delta_{-m, m} &= \text{span}\{1, x^{-1}, x, \dots, x^{m-1}, x^{-m}, x^m\}.\end{aligned}$$

Elements of these spaces are known as Laurent polynomials.

PROPOSITION 2.1. *Let $V_1, V_2, \dots, V_{2m+1}, V_{2m+2}$ be the F-orthonormal basis of $\mathbb{K}_g^{m+2, m+1}(A, V)$ determined by Algorithm 1. Then*

$$\begin{aligned}V_{2j-1} &= p_{j-1}(A)V + q_{j-1}(A^{-1})V = R_{2j-2}(A)V, \\ V_{2j} &= r_{j-1}(A^{-1})A^{-1}V + s_{j-1}(A)V = R_{2j-1}(A)V,\end{aligned}$$

where $\deg(p_{j-1}) = \deg(r_{j-1}) = j-1$, $\deg(q_{j-1}) \leq j-1$, and $\deg(s_{j-1}) \leq j-1$. Moreover, the Laurent polynomials R_{2j-1} and R_{2j-2} live in the spaces $\Delta_{-j, j-1}$ and $\Delta_{-(j-1), j-1}$, respectively.

Proof. The result follows from the recursion formulas of Algorithm 1. \square

The following proposition shows that the Laurent polynomials of Proposition 2.1 satisfy pairs of three-term recursion relations.

PROPOSITION 2.2. *Let A be a symmetric positive definite matrix. Then there is a sequence of Laurent polynomials R_0, R_1, \dots, R_{2m} that are orthonormal with respect to the bilinear form*

$$(2.10) \quad (P, Q)_F = \text{trace}(P(A)V, Q(A)V) = \int P(\lambda)Q(\lambda)d\alpha(\lambda),$$

where $d\alpha$ is the measure in (1.3). Its support lives on the positive real axis. Moreover, these Laurent polynomials satisfy a pair of three-term recurrence relations of the form

$$(2.11) \quad \begin{aligned} \delta_{2j} R_{2j-1}(x) &= (x^{-1} - \beta_{2j-1}) R_{2j-2}(x) - \beta_{2j-2} R_{2j-3}(x), \\ \delta_{2j+1} R_{2j}(x) &= (x - \alpha_{2j}) R_{2j-1}(x) - \alpha_{2j-1} R_{2j-2}(x), \end{aligned} \quad j = 1, \dots, m,$$

where $R_0 = 1/\delta_1$ and $R_{-1} = 0$.

Proof. According to Proposition 2.1, there exists a sequence of Laurent polynomials R_0, R_1, \dots, R_{2m} such that $V_i = R_{i-1}(A)V$, for $i = 1, \dots, 2m + 2$. Let $R_0 = 1/\delta_1$. It follows from the spectral factorization (1.2) that

$$(R_{i-1}, R_{j-1})_F = \text{trace}(V_i^T V_j) = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Thus, the Laurent polynomials R_j are orthonormal with respect to the bilinear form (2.10). Substituting $V_i = R_{i-1}(A)V$ into (2.9) gives (2.11). \square

Next we express the Laurent polynomials R_j in terms of the power basis,

$$(2.12) \quad \begin{aligned} R_{2j-1}(x) &= r_{2j-1,-j} x^{-j} + r_{2j-1,j-1} x^{j-1} + \sum_{i=-(j-1)}^{j-2} r_{2j-1,i} x^i, \\ R_{2j}(x) &= r_{2j,j} x^j + r_{2j,-j} x^{-j} + \sum_{i=-(j-1)}^{j-1} r_{2j,i} x^i. \end{aligned}$$

The coefficients $r_{2j-1,j-1}$ and $r_{2j,j}$ are the leading coefficient of R_{2j-1} and R_{2j} , respectively, and $r_{2j-1,-j}$ and $r_{2j,-j}$ are the trailing coefficient of R_{2j-1} and R_{2j} , respectively. The sequence R_0, R_1, \dots, R_{2m} of Laurent polynomials is said to be regular if the leading and trailing coefficients of each Laurent polynomial are nonvanishing. The following result shows the relation between leading and trailing coefficients and the coefficients in the recursion formulas of Proposition 2.2.

PROPOSITION 2.3. *Let the sequence of orthonormal Laurent polynomials be defined by the recursion formulas of Proposition 2.2, and let the α_j , β_j , and δ_j be the recursion coefficients. Then the leading and trailing coefficients of the Laurent polynomials can be computed recursively. Let $\alpha_0 = 1$ and $r_{0,0} = 1/\delta_1$. Then the leading and trailing coefficients are given by, for $j = 1, 2, \dots, m$,*

$$(2.13) \quad r_{2j-1,-j} = \frac{r_{2j-2,-(j-1)}}{\delta_{2j}} = (-1)^{j-1} \frac{\alpha_0 \alpha_2 \dots \alpha_{2j-2}}{\delta_1 \delta_2 \dots \delta_{2j}},$$

$$(2.14) \quad r_{2j-1,(j-1)} = -\frac{\beta_{2j-1}}{\delta_{2j}} r_{2j-2,(j-1)} = (-1)^j \frac{\beta_1 \beta_3 \dots \beta_{2j-1}}{\delta_1 \delta_2 \dots \delta_{2j}},$$

$$(2.15) \quad r_{2j,j} = \frac{r_{2j-1,j-1}}{\delta_{2j+1}} = (-1)^j \frac{\beta_1 \beta_3 \dots \beta_{2j-1}}{\delta_1 \delta_2 \dots \delta_{2j+1}},$$

$$(2.16) \quad r_{2j,-j} = -\frac{\alpha_{2j}}{\delta_{2j+1}} r_{2j-1,-j} = (-1)^j \frac{\alpha_2 \alpha_4 \dots \alpha_{2j}}{\delta_1 \delta_2 \dots \delta_{2j+1}}.$$

Proof. Equation (2.11) with $j = 1$ and the fact that R_0 is a constant yields that the coefficient of the trailing reciprocal power x^{-1} of R_1 is $r_{1,-1} = 1/\delta_1 \delta_2$ and its constant term is $r_{1,0} = -\beta_1/\delta_1 \delta_2$. The result follows by observing how the leading positive powers and trailing reciprocal powers are incremented in the recursion formulas. The coefficient of the

leading term of R_{2j} , $r_{2j,2j-1}$, is determined by the coefficient of the leading term of R_{2j-1} divided by δ_{2j+1} ; the coefficient of its trailing term is the coefficient of the trailing term of R_{2j-1} multiplied by $-\alpha_{2j}$ and divided by δ_{2j+1} . The leading and trailing coefficients of R_{2j+1} are found in a similar way. The right-hand sides of (2.13)–(2.16) follow by substitution. \square

PROPOSITION 2.4. *Let A be a symmetric positive definite matrix. Then the Laurent polynomials defined by (2.12) form a regular sequence.*

Proof. The bilinear form $(U, AW)_F$ defines an inner product for a normed vector space when A is a symmetric positive definite matrix. Hence, the coefficients $\alpha_{2i} = (V_{2i}, AV_{2i})_F$ and $\beta_{2i-1} = (V_{2i-1}, A^{-1}V_{2i-1})_F$ are nonvanishing for $i = 1, 2, \dots, m$. An application of Proposition 2.3 shows that the sequence of Laurent polynomials $\{R_i\}_{i=0}^{2m}$ is regular. \square

The following proposition allows us to express the leading and trailing coefficients in terms of inner products involving Laurent polynomials.

PROPOSITION 2.5. *The leading and trailing coefficients of the Laurent polynomials R_0, R_1, \dots, R_{2m} , defined by (2.12), can be expressed as*

$$(2.17) \quad r_{2j-1,-j} = \frac{1}{(R_{2j-1}, x^{-j})_F}, \quad (2.19) \quad r_{2j,j} = \frac{1}{(R_{2j}, x^j)_F},$$

$$(2.18) \quad r_{2j-1,j-1} = \frac{\alpha_{2j}}{(R_{2j-1}, x^j)_F}, \quad (2.20) \quad r_{2j,-j} = \frac{\beta_{2j+1}}{(R_{2j}, x^{-(j+1)})_F}.$$

Due to the F -orthogonality of R_0, R_1, \dots, R_{2m} , we have $(R_i, x^{-(\lfloor i/2 \rfloor + 1)})_F \neq 0$ and $(R_i, x^{\lfloor (i+1)/2 \rfloor})_F \neq 0$, for all $i = 0, 1, \dots, 2m$, where $\lfloor t \rfloor$ denotes the integer part of $t \geq 0$. This secures that the denominators in the expressions (2.17)–(2.20) are nonvanishing.

Proof. It follows from (2.12) that

$$R_{2j-1}(x) = r_{2j-1,-j}x^{-j} + r_{2j-1,j-1}x^{j-1} + \sum_{i=-(j-1)}^{j-2} r_{2j-1,i}x^i,$$

$$\in r_{2j-1,-j}x^{-j} + s(x), \quad \text{with } s \in \text{span}\{R_0, \dots, R_{2j-2}\}.$$

The orthonormality of the Laurent polynomials R_0, \dots, R_{2m} yields

$$(R_{2j-1}, R_{2j-1})_F = r_{2j-1,-j}(R_{2j-1}, x^{-j})_F,$$

which shows (2.17). To show (2.18), we multiply the first equation in (2.12) by x . Then

$$xR_{2j-1}(x) = r_{2j-1,j-1}x^j + r_{2j-1,-j}x^{-j+1} + \sum_{i=-(j-1)}^{j-2} r_{2j-1,i}x^{i+1}$$

$$\in r_{2j-1,j-1}x^j + s(x), \quad \text{with } s \in \text{span}\{R_0, \dots, R_{2j-2}\}.$$

Again using the orthonormality of the R_i , we obtain

$$(xR_{2j-1}, R_{2j-1})_F = r_{2j-1,j-1}(R_{2j-1}, x^j)_F,$$

which gives (2.18). The relations (2.19) and (2.20) can be shown similarly. \square

PROPOSITION 2.6. *Let the sequence of orthonormal Laurent polynomials be determined by Proposition 2.2. The recursion coefficients $\{\alpha_j\}_{j=1}^{2m}$, $\{\beta_j\}_{j=1}^{2m+1}$, and $\{\delta_j\}_{j=2}^{2m+2}$ can be*

expressed as

$$\begin{aligned} \beta_1 &= \frac{(R_0, x^{-1})_F}{(R_0, x^0)_F}, \\ \beta_{2j} &= \frac{(R_{2j}, x^{-(j+1)})_F}{(R_{2j-1}, x^{-j})_F}, & \beta_{2j+1} &= -\beta_{2j} \frac{(R_{2j-1}, x^j)_F}{(R_{2j}, x^j)_F}, & j &\geq 1, \\ \alpha_{2j-1} &= \frac{(R_{2j-1}, x^j)_F}{(R_{2j-2}, x^{j-1})_F}, & \alpha_{2j} &= -\alpha_{2j-1} \frac{(R_{2j-2}, x^{-j})_F}{(R_{2j-1}, x^{-j})_F}, & j &\geq 1, \\ \delta_{2j} &= -\frac{(R_{2j-3}, x^{(j-1)})_F (R_{2j-1}, x^{-j})_F}{(R_{2j-2}, x^{j-1})_F (R_{2j-3}, x^{-(j-1)})_F}, & j &\geq 1, \\ \delta_{2j+1} &= -\frac{(R_{2j}, x^j)_F (R_{2j-2}, x^{-j})_F}{(R_{2j-2}, x^{j-1})_F (R_{2j-1}, x^{-j})_F}, \end{aligned}$$

Proof. First consider the coefficient α_{2j-1} . We have

$$\alpha_{2j-1} = (xR_{2j-1}, R_{2j-2})_F = (R_{2j-1}, xR_{2j-2})_F.$$

It follows from (2.12) that

$$\begin{aligned} R_{2j-2}(x) &= r_{2j-2, j-1}x^{j-1} + r_{2j-2, -j+1}x^{-j+1} + \sum_{i=-(j-2)}^{j-2} r_{2j-2, i}x^i, \\ xR_{2j-2}(x) &\in r_{2j-2, j-1}x^j + s(x), \quad \text{with } s \in \text{span}\{R_0, \dots, R_{2j-2}\}. \end{aligned}$$

The orthonormality of R_0, \dots, R_{2m} now yields

$$\alpha_{2j-1} = r_{2j-2, j-1}(x^j, R_{2j-1})_F = \frac{(R_{2j-1}, x^j)_F}{(R_{2j-2}, x^{j-1})_F}.$$

The expression for β_{2j} can be shown similarly. It follows from the fact that

$$\beta_{2j} = (x^{-1}R_{2j}, R_{2j-1})_F = (R_{2j}, x^{-1}R_{2j-1})_F.$$

We turn to α_{2j} . It follows from (2.14) that

$$r_{2j-1, j-1} = -\frac{\beta_{2j-1}}{\delta_{2j}} r_{2j-2, j-1}.$$

Now using (2.18) and (2.19) yields

$$(2.21) \quad \alpha_{2j} = -\frac{\beta_{2j-1}}{\delta_{2j}} \frac{(R_{2j-1}, x^j)_F}{(R_{2j-2}, x^{j-1})_F}.$$

On the other hand, we have

$$r_{2j-1, -j} = \frac{r_{2j-2, -(j-1)}}{\delta_{2j}},$$

and using (2.17) and (2.20), we obtain

$$(2.22) \quad \frac{1}{(R_{2j-1}, x^{-j})_F} = -\frac{\beta_{2j-1}}{\delta_{2j}} \frac{1}{(x^{-j}, R_{2j-2})_F}.$$

Combining (2.21) and (2.22) gives α_{2j} . To show the expression for β_1 , we use the fact that $R_1(x) = r_{1,-1}x^{-1} + r_{1,0}$ and the expressions for $r_{1,0}$ and $r_{1,-1}$ in (2.13) and (2.14). This yields

$$-\frac{r_{1,0}}{r_{1,-1}} = \frac{(x^{-1}, R_0)_F}{(x^0, R_0)_F} = \beta_1.$$

The expression for β_{2j+1} follows from applications of Propositions 2.3 and 2.5. We obtain

$$(2.23) \quad r_{2j,j} = \frac{r_{2j-1,j-1}}{\delta_{2j+1}},$$

$$\frac{1}{(R_{2j}, x^j)_F} = \frac{\alpha_{2j}}{(R_{2j-1}, x^j)_F} \frac{1}{\delta_{2j+1}},$$

$$(2.24) \quad r_{2j,-j} = \frac{\alpha_{2j}}{\delta_{2j+1}} r_{2j-1,-j},$$

$$\frac{\beta_{2j+1}}{(R_{2j}, x^{-j-1})_F} = -\frac{\alpha_{2j}}{\delta_{2j+1}} \frac{1}{(R_{2j-1}, x^{-j})_F}.$$

Combining (2.23) and (2.24) gives the expression for β_{2j+1} . Finally, the expressions for δ_{2j} and δ_{2j+1} can be shown by using (2.21) and (2.23). \square

We summarize the above results in the following theorem.

THEOREM 2.7. *Let A be a symmetric positive definite matrix, let $V \in \mathbb{R}^{n \times s}$ be a block vector, and let $d\alpha$ denote the measure defined in (1.3). Then there is a sequence of regular Laurent polynomials R_0, R_1, \dots, R_{2m} that are orthonormal with respect to the bilinear form (2.10). The Laurent polynomials can be computed recursively by pairs of three-term recurrence relations of the form*

$$(2.25) \quad \begin{aligned} \delta_{2j} R_{2j-1}(x) &= (x^{-1} - \beta_{2j-1}) R_{2j-2}(x) - \beta_{2j-2} R_{2j-3}(x), \\ \delta_{2j+1} R_{2j}(x) &= (x - \alpha_{2j}) R_{2j-1}(x) - \alpha_{2j-1} R_{2j-2}(x), \end{aligned} \quad j = 1, 2, \dots, m,$$

with $R_0 = 1/\mu_0$, $R_{-1} = 0$, and $\mu_0 = \|V\|_F$, where

$$\begin{aligned} \beta_1 &= \frac{\mu_{-1}}{\mu_0}, \\ \alpha_{2j-1} &= \frac{\rho_{2j-1}}{\rho_{2j-2}}, & \alpha_{2j} &= -\alpha_{2j-1} \frac{\sigma_{2j-2}}{\sigma_{2j-1}}, & \delta_{2j+1} &= -\frac{\sigma_{2j-2}}{\sigma_{2j-1}} \frac{\rho_{2j}}{\rho_{2j-2}}, \\ \beta_{2j-2} &= \frac{\sigma_{2j-2}}{\sigma_{2j-3}}, & \beta_{2j-1} &= -\beta_{2j-2} \frac{\rho_{2j-3}}{\rho_{2j-2}}, & \delta_{2j} &= -\frac{\rho_{2j-3}}{\rho_{2j-2}} \frac{\sigma_{2j-1}}{\sigma_{2j-3}}. \end{aligned}$$

Moreover,

$$\mu_n = \int x^n d\alpha(x), \quad \rho_n = (R_n, x^{\lfloor (n+1)/2 \rfloor})_F, \quad \sigma_n = (R_n, x^{-\lfloor (n/2) + 1 \rfloor})_F.$$

We assume that $ms \ll n$ is sufficiently small to avoid breakdown of the recursion relations.

Proof. The result follows from Propositions 2.2, 2.4, and 2.6. \square

3. Application to the approximation of matrix functions. We begin by introducing some notation. Partition the matrices $M = [M_1, \dots, M_p] \in \mathbb{R}^{n \times ps}$ and $N = [N_1, \dots, N_l] \in \mathbb{R}^{n \times ls}$ into block columns $M_i, N_j \in \mathbb{R}^{n \times s}$, and define the \diamond -product of the matrices M and N as

$$M^T \diamond N = [\text{trace}(M_i^T N_j)]_{i=1, \dots, p}^{j=1, \dots, l} \in \mathbb{R}^{p \times l}.$$

The matrix M is said to be F -orthonormal if $M^T \diamond M = I_p$. The \diamond -product satisfies

$$M^T \diamond (B(L \otimes I_s)) = (M^T \diamond B)L,$$

where $L \in \mathbb{R}^{l \times l}$; see [1, 3, 4] for further details.

Let the block columns of the matrix $\mathbb{V}_{2m} = [V_1, V_2, \dots, V_{2m}] \in \mathbb{R}^{n \times 2ms}$ form an orthonormal basis for the extended Krylov subspace $\mathbb{K}_g^{m+1, m}$. Thus, the block columns satisfy (1.6). Introduce the matrices

$$(3.1) \quad T_{2m} = \mathbb{V}_{2m}^T \diamond A \mathbb{V}_{2m} \in \mathbb{R}^{2m \times 2m}, \quad S_{2m} = \mathbb{V}_{2m}^T \diamond A^{-1} \mathbb{V}_{2m} \in \mathbb{R}^{2m \times 2m}.$$

The expression (1.1) can be approximated by

$$(3.2) \quad G_{2m}^e(f) = \|V\|_F^2 e_1^T f(T_{2m}) e_1.$$

We will see that this approximation is analogous to (1.7). Define the spectral factorization

$$(3.3) \quad T_{2m} = U_{2m} \Theta_{2m} U_{2m}^T,$$

where

$$U_{2m} = [u_1, \dots, u_{2m}] \in \mathbb{R}^{2m \times 2m}, \quad U_{2m}^T U_{2m} = I_{2m}, \quad \Theta_{2m} = \text{diag}[\theta_1, \dots, \theta_{2m}].$$

Substituting (3.3) into (3.2) yields

$$(3.4) \quad G_{2m}^e(f) = \sum_{i=1}^{2m} f(\theta_i) w_i,$$

where $w_i = \|V\|_F^2 u_{i,1}^2$. Here $u_{i,1}$ denotes the first component of the normalized eigenvector u_i of T_{2m} . We will see below that (3.4) is a Gauss–Laurent quadrature rule. An adaptation of the Christoffel–Darboux formula to the extended global Lanczos method shows that the weights w_i satisfy

$$w_i = \frac{1}{\sum_{k=0}^{2m-1} R_k^2(\theta_i)}, \quad i = 1, 2, \dots, 2m;$$

see, e.g., Bultheel et al. [6] for details. The next two propositions express the entries of T_{2m} and S_{2m} in terms of the recursion coefficients. This will allow us to compute the entries quite efficiently.

PROPOSITION 3.1. *Let the matrix A be symmetric positive definite, and let $\{\alpha_j\}_{j=1}^{2m}$, $\{\beta_j\}_{j=1}^{2m-1}$, and $\{\delta_j\}_{j=2}^{2m+2}$ be the recursion coefficients for the orthonormal Laurent polynomials; cf. Proposition 2.2. The matrix $T_{2m} = [t_{i,j}]$ in (3.1) is symmetric and pentadiagonal with the nontrivial entries, for $j = 1, 2, \dots, m$,*

$$(3.5) \quad t_{i,2j} = \alpha_i \quad \text{for } i \in \{2j-1, 2j\},$$

$$(3.6) \quad t_{2j+1,2j} = \delta_{2j+1},$$

$$(3.7) \quad t_{2j-3,2j-1} = \frac{-\delta_{2j-2} \delta_{2j-1}}{\beta_{2j-3}},$$

$$(3.8) \quad t_{2j-2,2j-1} = \delta_{2j-1},$$

$$(3.9) \quad t_{2j-1,2j-1} = \frac{1 - \beta_{2j-2} \delta_{2j-1} - \delta_{2j} \alpha_{2j-1}}{\beta_{2j-1}},$$

$$(3.10) \quad t_{2j,2j-1} = \alpha_{2j-1},$$

$$(3.11) \quad t_{2j+1,2j-1} = \frac{-\delta_{2j} \delta_{2j+1}}{\beta_{2j-1}}.$$

Proof. We have $t_{i,j} = \text{trace}(V_i^T AV_j)$. Therefore, (3.5) and (3.6) follow from the definitions of α_{2j-1} , α_{2j} , and δ_{2j+1} in (2.9). Using the symmetry of T_{2m} and the definitions of α_{2j-1} and δ_{2j-1} , the formulas (3.8) and (3.10) are obtained from

$$\begin{aligned} t_{2j-2,2j-1} &= t_{2j-1,2j-2} = \text{trace}(V_{2j-1}^T AV_{2j-2}) = \delta_{2j-1}, \\ t_{2j,2j-1} &= t_{2j-1,2j} = \text{trace}(V_{2j-1}^T AV_{2j}) = \alpha_{2j-1}. \end{aligned}$$

The formulas (3.7), (3.9), and (3.11) are obtained from the expressions for AV_{2j-1} for $j = 1, 2, \dots, m$. Thus, multiplying the first equality in (2.9) by A from the left gives

$$\beta_{2j-1} AV_{2j-1} = V_{2j-1} - \beta_{2j-2} AV_{2j-2} - \delta_{2j} AV_{2j},$$

and using the second equation in (2.9), we see that

$$\begin{aligned} \beta_{2j-1} AV_{2j-1} &= -\beta_{2j-2} \alpha_{2j-3} V_{2j-3} - \beta_{2j-2} \alpha_{2j-2} V_{2j-2} \\ (3.12) \quad &+ (1 - \beta_{2j-2} \delta_{2j-1} - \delta_{2j} \alpha_{2j-1}) V_{2j-1} \\ &- \delta_{2j} \alpha_{2j} V_{2j} - \delta_{2j} \delta_{2j+1} V_{2j+1}. \end{aligned}$$

In view of that $t_{2j-1,2j-1} = \text{trace}(V_{2j-1}^T AV_{2j-1})$ and $t_{2j+1,2j-1} = \text{trace}(V_{2j+1}^T AV_{2j-1})$ and due to the orthonormality of the set of block vectors $\{V_i\}_{i=1}^{2(m+1)}$, we obtain (3.9) and (3.11). Finally,

$$t_{2j-3,2j-1} = t_{2j-1,2j-3} = \frac{-\delta_{2j-2} \delta_{2j-1}}{\beta_{2j-3}},$$

which shows (3.7). This concludes the proof of the proposition. \square

PROPOSITION 3.2. *Let the matrix A be symmetric positive definite, and let $\{\alpha_j\}_{j=1}^{2m}$, $\{\beta_j\}_{j=1}^{2m-1}$, and $\{\delta_j\}_{j=2}^{2m+2}$ be recursion coefficients for the orthonormal Laurent polynomials; cf. (2.2). The matrix $S_{2m} = [s_{i,j}]$ in (3.1) is symmetric and pentadiagonal with the nontrivial entries, for $j = 1, 2, \dots, m$,*

$$\begin{aligned} s_{i,2j-1} &= \beta_i && \text{for } i \in \{2j-2, 2j-1\}, \\ s_{2j,2j-1} &= \delta_{2j}, \\ s_{2j-2,2j} &= \frac{-\delta_{2j-1} \delta_{2j}}{\alpha_{2j-2}}, \\ s_{2j-1,2j} &= \delta_{2j}, \\ s_{2j,2j} &= \frac{1 - \alpha_{2j-1} \delta_{2j} - \delta_{2j+1} \beta_{2j}}{\alpha_{2j}}, \\ s_{2j+1,2j} &= \beta_{2j}, \\ s_{2j+2,2j} &= \frac{-\delta_{2j+1} \delta_{2j+2}}{\alpha_{2j}}. \end{aligned}$$

Proof. The proof is analogous to the proof of Proposition 3.1. The relation

$$\begin{aligned} \alpha_{2j} A^{-1} V_{2j} &= -\alpha_{2j-1} \beta_{2j-2} V_{2j-2} - \alpha_{2j-1} \beta_{2j-1} V_{2j-1} \\ (3.13) \quad &+ (1 - \alpha_{2j-1} \delta_{2j} - \delta_{2j+1} \beta_{2j}) V_{2j} \\ &- \delta_{2j+1} \beta_{2j+1} V_{2j+1} - \delta_{2j+1} \delta_{2j+2} V_{2j+2} \end{aligned}$$

is used. \square

The following relations can be shown by expressing equations (2.9), (3.12), and (3.13) in matrix form:

$$(3.14) \quad A\mathbb{V}_{2m} = \mathbb{V}_{2m}(T_{2m} \otimes I_s) + V_{2m+1} \begin{bmatrix} t_{2m+1,2m-1} & t_{2m+1,2m} \end{bmatrix} E_{2m-1}^T \otimes I_s,$$

$$(3.15) \quad A^{-1}\mathbb{V}_{2m} = \mathbb{V}_{2m}(S_{2m} \otimes I_s) + [V_{2m+1}, V_{2m+2}] \begin{bmatrix} 0 & s_{2m+1,2m} \\ 0 & s_{2m+2,2m} \end{bmatrix} E_{2m-1}^T \otimes I_s,$$

where the matrix $E_{2m-1} = [e_{2m-1}, e_{2m}] \in \mathbb{R}^{2m \times 2}$ is made up of the last two columns of the identity matrix $I_{2m} \in \mathbb{R}^{2m \times 2m}$.

Algorithm 1 breaks down when $\delta_j = 0$. The next result addresses this situation.

PROPOSITION 3.3. *Assume that the coefficients δ_j generated by Algorithm 1 satisfy $\delta_j > 0$, for $2 \leq j \leq 2m$, and $\delta_{2m+1} = 0$. Then*

$$\text{trace}(V^T f(A)V) = \|V\|_F^2 e_1^T f(T_{2m})e_1$$

for any function f that is continuous on the convex hull of the spectrum of A .

Proof. Under the conditions of the proposition, the decomposition (3.14) simplifies to

$$A\mathbb{V}_{2m} = \mathbb{V}_{2m}(T_{2m} \otimes I_s), \quad \mathbb{V}_{2m} = [V_1, \dots, V_{2m}].$$

Then an application of [3, Proposition 2.1] gives

$$f(A)\mathbb{V}_{2m} = \mathbb{V}_{2m}f(T_{2m} \otimes I_s) = \mathbb{V}_{2m}f(T_{2m}) \otimes I_s.$$

Multiplying the above expression by $E_1 = e_1 \otimes I_s$ from the right-hand side yields

$$f(A)V_1 = \mathbb{V}_{2m}(f(T_{2m}) \otimes I_s)E_1 = \mathbb{V}_{2m}(f(T_{2m})e_1 \otimes I_s)$$

and

$$\begin{aligned} \text{trace}(V_1^T f(A)V_1) &= V_1^T \diamond [\mathbb{V}_{2m}(f(T_{2m})e_1 \otimes I_s)] \\ &= (V_1^T \diamond \mathbb{V}_{2m})(f(T_{2m})e_1) = e_1^T f(T_{2m})e_1. \end{aligned}$$

Substituting $V = \|V\|_F V_1$ into $\text{trace}(V_1^T f(A)V_1)$ completes the proof. \square

PROPOSITION 3.4. *Let the Laurent polynomials R_0, R_1, \dots, R_{2m} be determined by the recursion relations (2.2). Then the $2m$ zeros of the Laurent polynomial R_{2m} are the eigenvalues of the matrix T_{2m} in (3.1).*

Proof. Consider the vector of Laurent polynomials

$$\vec{R}_{2m}(x) = [R_0(x), R_1(x), \dots, R_{2m-1}(x)]^T.$$

Then the recurrence relation (2.25) can be expressed as

$$x\vec{R}_{2m}(x) = \vec{R}_{2m}(x)T_{2m} + R_{2m}(x)[t_{2m+1,2m-1}, t_{2m+1,2m}]E_{2m-1}^T,$$

where E_{2m-1} is defined below of (3.15). Let x_j be a zero of R_{2m} . Then

$$T_{2m}\vec{R}_{2m}(x_j) = x_j\vec{R}_{2m}(x_j).$$

This completes the proof. \square

Proposition 3.3 shows that in case of a breakdown of Algorithm 1, the trace (1.1) can be computed exactly after m steps of the algorithm. However, breakdown is rare. We are

therefore interested in how well (1.1) is approximated when no breakdown occurs. It is shown in [3] that when the (standard) global Lanczos method is applied, we have

$$\text{trace}(V^T p(A)V) = \|V\|_F^2 e_1^T p(T_m) e_1, \quad \forall p \in \Pi_{2m-1}.$$

The following result is analogous for Laurent polynomials. We show that the approximation (3.2) corresponds to a $2m$ -point Gauss-Laurent quadrature rule, i.e., the rule is exact for Laurent polynomials of positive degree at most $2m - 1$ and negative degree at most $2m$.

THEOREM 3.5. *Let A be a symmetric positive definite matrix. Carry out m steps of Algorithm 1 with an initial block vector $V \in \mathbb{R}^{n \times s}$. Assume that no breakdown occurs. The algorithm then determines the entries of the symmetric pentadiagonal matrix T_{2m} defined by (3.1), and*

$$\text{trace}(V^T p(A)V) = \|V\|_F^2 e_1^T p(T_{2m}) e_1, \quad \forall p \in \Delta_{-2m, 2m-1}.$$

To show this theorem, we require some auxiliary results on the properties of the matrices T_{2m} and S_{2m} in (3.1).

LEMMA 3.6. *Let the matrices T_{2m} and S_{2m} be defined by (3.1), and let $E_{2m-1} = [e_{2m-1}, e_{2m}] \in \mathbb{R}^{2m \times 2}$. Then*

$$\begin{aligned} E_{2m-1}^T T_{2m}^j e_1 &= [0 \ 0]^T, & j = 0, 1, \dots, m-2, \\ E_{2m-1}^T S_{2m}^j e_1 &= [0 \ 0]^T, & j = 0, 1, \dots, m-1. \end{aligned}$$

Proof. The result follows by exploiting the structure of the symmetric pentadiagonal matrices T_{2m} and S_{2m} ; see Propositions 3.1 and 3.2. \square

The following result relates positive powers of S_{2m} to negative powers of T_{2m} .

LEMMA 3.7. *Let T_{2m} and S_{2m} be given by (3.1), and let $e_1 = [1, 0, \dots, 0]^T \in \mathbb{R}^{2m}$. Then*

$$(3.16) \quad S_{2m}^j e_1 = T_{2m}^{-j} e_1 \quad \text{for } j = 1, 2, \dots, m.$$

Proof. Using induction, we begin by showing that

$$\begin{aligned} T_{2m}^j S_{2m}^j &= T_{2m}^{j-1} S_{2m}^{j-1} - T_{2m}^{j-1} (\mathbb{V}_{2m}^T \diamond A[V_{2m+1}, V_{2m+2}]) \tilde{S}_{m+1, m} E_{2m-1}^T S_{2m}^{j-1}, \\ & \quad j = 1, \dots, m, \end{aligned}$$

with $\tilde{S}_{m+1, m} = \begin{bmatrix} 0 & s_{2m+1, 2m} \\ 0 & s_{2m+2, 2m} \end{bmatrix}$. Using the decomposition (3.14), (3.15), and the properties of the Kronecker product, we obtain

$$I_{2m} = T_{2m} S_{2m} + (\mathbb{V}_{2m}^T \diamond A[V_{2m+1}, V_{2m+2}]) \tilde{S}_{m+1, m} E_{2m-1}^T.$$

Let $j = 2, 3, \dots, m$, and assume that

$$\begin{aligned} T_{2m}^k S_{2m}^k &= T_{2m}^{k-1} S_{2m}^{k-1} - T_{2m}^{k-1} (\mathbb{V}_{2m}^T \diamond A[V_{2m+1}, V_{2m+2}]) \tilde{S}_{m+1, m} E_{2m-1}^T S_{2m}^{k-1}, \\ & \quad k = 1, \dots, j-1. \end{aligned}$$

By induction, we have

$$\begin{aligned} T_{2m}^j S_{2m}^j &= T_{2m} T_{2m}^{j-1} S_{2m}^{j-1} S_{2m} \\ &= T_{2m}^{j-1} S_{2m}^{j-1} - T_{2m}^{j-1} (\mathbb{V}_{2m}^T \diamond A[V_{2m+1}, V_{2m+2}]) \tilde{S}_{m+1, m} E_{2m-1}^T S_{2m}^{j-1}. \end{aligned}$$

Multiplying this equation by e_1 from the right gives

$$T_{2m}^j S_{2m}^j e_1 = T_{2m}^{j-1} S_{2m}^{j-1} e_1 - T_{2m}^{j-1} (\mathbb{V}_{2m}^T \diamond A[V_{2m+1}, V_{2m+2}]) \tilde{S}_{m+1, m} E_{2m-1}^T S_{2m}^{j-1} e_1.$$

An application of Lemma 3.6 yields

$$T_{2m}^j S_{2m}^j e_1 = T_{2m}^{j-1} S_{2m}^{j-1} e_1, \quad j = 1, 2, \dots, m.$$

This completes the proof. \square

LEMMA 3.8. *Let the matrix A be symmetric positive definite, and let the matrices T_{2m} and S_{2m} be defined by (3.1). Let $\mathbb{V}_{2m} = [V_1, V_2, \dots, V_{2m}]$ be the matrix in (3.14), i.e., the block columns $V_j \in \mathbb{R}^{n \times s}$ are F -orthonormal, and the first block column is given by $V_1 = V/\|V\|_F$. Then*

$$(3.17) \quad A^j V_1 = \mathbb{V}_{2m} (T_{2m}^j e_1 \otimes I_s), \quad j = 0, 1, \dots, m-1,$$

$$(3.18) \quad A^{-j} V_1 = \mathbb{V}_{2m} (S_{2m}^j e_1 \otimes I_s), \quad j = 0, 1, \dots, m,$$

$$(3.19) \quad A^{-j} V_1 = \mathbb{V}_{2m} (T_{2m}^{-j} e_1 \otimes I_s), \quad j = 0, 1, \dots, m.$$

Proof. We obtain from (3.14) that

$$A \mathbb{V}_{2m} = \mathbb{V}_{2m} (T_{2m} \otimes I_s) + V_{2m+1} ([t_{2m+1, 2m-1}, t_{2m+1, 2m}] E_{2m-1}^T \otimes I_s).$$

Multiplying this equation by $E_1 = e_1 \otimes I_s$ from the right gives

$$A V_1 = A \mathbb{V}_{2m} E_1 = \mathbb{V}_{2m} (T_{2m} \otimes I_s) E_1 + V_{2m+1} ([t_{2m+1, 2m-1}, t_{2m+1, 2m}] E_{2m-1}^T \otimes I_s) E_1.$$

Using properties of the Kronecker product and the fact that $E_{2m-1}^T e_1 = [0, 0]$, we obtain

$$A V_1 = \mathbb{V}_{2m} (T_{2m} e_1 \otimes I_s).$$

Let $j = 2, 3, \dots, m-1$, and assume that

$$A^k V_1 = \mathbb{V}_{2m} (T_{2m}^k e_1 \otimes I_s), \quad k = 0, 1, \dots, j-1.$$

We will show the identity

$$A^j V_1 = \mathbb{V}_{2m} (T_{2m}^j e_1 \otimes I_s)$$

by induction. We have

$$A^j V_1 = A \cdot A^{j-1} V_1 = A \mathbb{V}_{2m} (T_{2m}^{j-1} e_1 \otimes I_s).$$

Using the decomposition (3.14), we obtain

$$\begin{aligned} A^j V_1 &= [\mathbb{V}_{2m} (T_{2m} \otimes I_s) + V_{2m+1} ([t_{2m+1, 2m-1}, t_{2m+1, 2m}] E_{2m-1}^T \otimes I_s)] (T_{2m}^{j-1} e_1 \otimes I_s) \\ &= \mathbb{V}_{2m} (T_{2m}^j e_1 \otimes I_s) + V_{2m+1} ([t_{2m+1, 2m-1}, t_{2m+1, 2m}] E_{2m-1}^T T_{2m}^{j-1} e_1 \otimes I_s). \end{aligned}$$

An application of Lemma 3.6 gives

$$A^j V_1 = \mathbb{V}_{2m} (T_{2m}^j e_1 \otimes I_s).$$

According to (3.11) and Lemma 3.6 and by using the same techniques as above, we find (3.18). Finally, (3.19) follows from (3.17) and Lemma 3.7. \square

LEMMA 3.9. *Let the matrix A be symmetric positive definite, and let the matrices T_{2m} and S_{2m} be defined by (3.1). Let V_1 be the initial block vector computed in Algorithm 1. Then*

$$(3.20) \quad \text{trace}(V_1^T A^j V_1) = e_1^T T_{2m}^j e_1, \quad \text{for } j = 0, 1, \dots, 2m - 1,$$

$$(3.21) \quad \text{trace}(V_1^T A^{-j} V_1) = e_1^T S_{2m}^j e_1, \quad \text{for } j = 0, 1, \dots, 2m + 1,$$

$$(3.22) \quad \text{trace}(V_1^T A^{-j} V_1) = e_1^T T_{2m}^{-j} e_1, \quad \text{for } j = 0, 1, \dots, 2m.$$

Proof. The proof is based on an application of Lemma 3.8. Let $j = j_1 + j_2$ with $0 \leq j_1, j_2 < m$. Consider

$$\begin{aligned} \text{trace}(V_1^T A^j V_1) &= \text{trace}(V_1^T A^{j_1} A^{j_2} V_1) \\ &= \text{trace}(V_1^T A^{j_1} \mathbb{V}_{2m}(T_{2m}^{j_2} e_1 \otimes I_s)) \\ &= (V_1^T A^{j_1}) \diamond \mathbb{V}_{2m}(T_{2m}^{j_2} e_1 \otimes I_s) \\ &= (V_1^T A^{j_1} \diamond \mathbb{V}_{2m}) T_{2m}^{j_2} e_1 = e_1^T T_{2m}^{j_1} T_{2m}^{j_2} e_1 \\ &= e_1^T T_{2m}^j e_1. \end{aligned}$$

For the power $2m - 1$ of A , we have

$$\begin{aligned} \text{trace}(V_1^T A^{2m-1} V_1) &= \text{trace}(V_1^T A^m A^{m-1} V_1) \\ &= V_1^T A^m \diamond \mathbb{V}_{2m}(T_{2m}^{m-1} e_1 \otimes I_s) \\ &= (V_1^T A^{m-1} A \diamond \mathbb{V}_{2m}) T_{2m}^{m-1} e_1 \\ &= [(e_1^T T_{2m}^{m-1} \otimes I_s) \mathbb{V}_{2m}^T A \diamond \mathbb{V}_{2m}] T_{2m}^{m-1} e_1 \\ &= e_1^T T_{2m}^{m-1} (\mathbb{V}_{2m}^T A \diamond \mathbb{V}_{2m}) T_{2m}^{m-1} e_1 \\ &= e_1^T T_{2m}^{2m-1} e_1. \end{aligned}$$

This shows (3.20). The same techniques can be used to show (3.21). Finally, (3.22) can be established by substituting (3.16) into (3.21). \square

We are in a position to prove Theorem 3.5.

Proof of Theorem 3.5. Let $p \in \Delta_{-2m, 2m-1}$. Then $p(x) = \sum_{i=-2m}^{2m-1} c_i x^i$ for certain coefficients c_i . Let $V = \|V\|_F V_1$. Then

$$\text{trace}(V^T p(A) V) = \|V\|_F^2 \text{trace}(V_1^T p(A) V_1).$$

It suffices to show that

$$\text{trace}(V_1^T p(A) V_1) = e_1^T p(T_{2m}) e_1.$$

We have

$$\text{trace}(V_1^T p(A) V_1) = \sum_{i=-2m}^{2m-1} c_i \text{trace}(V_1^T A^i V_1).$$

Then Lemma 3.9 gives

$$\text{trace}(V_1^T p(A) V_1) = \sum_{i=-2m}^{2m-1} c_i e_1^T T_{2m}^i e_1 = e_1^T p(T_{2m}) e_1.$$

This concludes the proof of Theorem 3.5. \square

Algorithm 2 describes how approximations of $\text{trace}(V^T f(A) V)$ are computed by the extended global Lanczos method.

Algorithm 2 Approximation of $\text{trace}(V^T f(A)V)$ by the extended global Lanczos method.

Input: symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$, function f , initial block vector $V \in \mathbb{R}^{n \times s}$

Output: approximation G_{2m} of $\text{trace}(V^T f(A)V)$

- 1: Choose tolerance $\epsilon > 0$ and the maximal number of iterations I_{max} . $G_0 = 1$.
- 2: **for** $m = 1$ **to** I_{max}
- 3: Compute block vectors $[V_{2m-1}, V_{2m}]$ to update \mathbb{V}_{2m} by Algorithm 1.
- 4: Compute matrix $T_{2m} \in \mathbb{R}^{2m \times 2m}$ using Proposition 3.1.
- 5: Compute $G_{2m} := G_{2m}^e(f) = \sum_{i=1}^{2m} f(\theta_i)w_i$ given by (3.4).
- 6:

$$\epsilon_{m,m-1} := \begin{cases} \frac{|G_{2m} - G_{2m-2}|}{|G_{2m}|} & \text{if } G_{2m} \neq 0 \\ |G_{2m} - G_{2m-2}| & \text{if } G_{2m} = 0 \end{cases}$$

- 7: **if** $\epsilon_{m,m-1} < \epsilon$ **then exit**
- 8: **end for**

4. Numerical experiments. We illustrate the performance of the extended global Lanczos method when applied to the approximation of expressions of the form (1.1). All experiments are carried out in MATLAB R2015a on a computer with an Intel Core i-3 processor and 3.89 GBytes of RAM. The computations are done with about 15 significant decimal digits. In all the experiments, the initial block vector V is generated randomly with uniformly distributed entries in the interval $[0, 1]$. The parameter ϵ in Algorithm 2 is set to 10^{-7} unless specified otherwise.

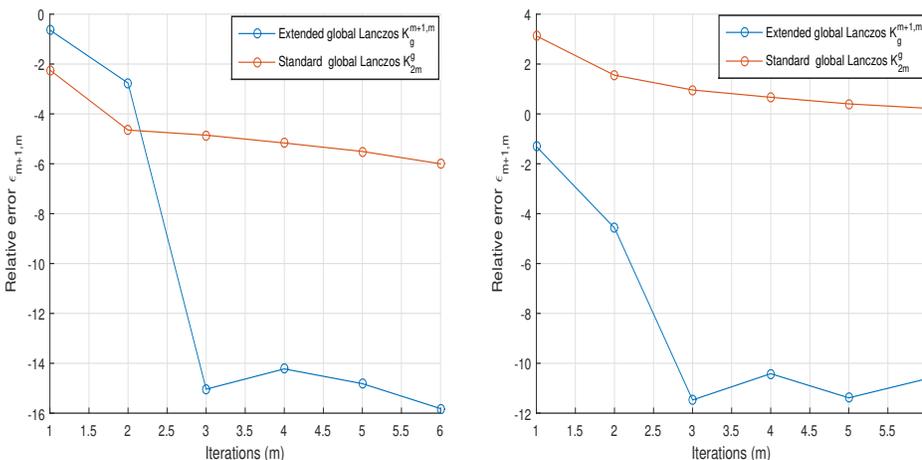


FIG. 4.1. Convergence of the extended global Lanczos method and the (standard) global Lanczos when approximating $\text{trace}(V^T (A^{-6} + A^5)V)$ (right plot) and $\text{trace}(V^T (B^{-6} + B^5)V)$ (left plot).

EXAMPLE 4.1. The purpose of this example is to illustrate the degree of exactness of (3.2) ($4m - 1$ moments are matched) shown in Theorem 3.5. Consider the symmetric positive

TABLE 4.1

Example 4.2: The symmetric matrix $A \in \mathbb{R}^{n \times n}$ has eigenvalues distributed in the interval $[10^{-1}, 10^6]$ and a random orthogonal eigenvector matrix. The block size is $s = 6$.

$f(x)$	Extended global Lanczos, subspace $\mathbb{K}_g^{m+1,m}$			Standard global Lanczos \mathbb{K}_g^m		
	$(m+1, m)$	Error	Time (sec)	m	Error	SF
e^{-x}	(25,24)	2.2×10^{-7}	1.4	470	8.1×10^{-5}	17.3
\sqrt{x}	(59,58)	9.2×10^{-7}	3.9	599	9.9×10^{-7}	8.7
$x^{-1/4}$	(49,48)	8.9×10^{-7}	3.0	607	9.1×10^{-5}	19.0
$\ln(x)$	(77,76)	9.5×10^{-7}	7.9	612	1.7×10^{-3}	7.1
$e^{-\sqrt{x}}$	(19,18)	2.3×10^{-7}	0.6	607	5.7×10^{-5}	61.0
x^{-4}	(3,2)	1.6×10^{-11}	0.2	308	7.8×10^{-3}	76.3

TABLE 4.2

Example 4.3: The matrix $A \in \mathbb{R}^{n \times n}$ is a discretized negative Laplacian with $n = 10000$. Block size is $s = 20$.

$f(x)$	Extended global Lanczos $\mathbb{K}_g^{m+1,m}$			Standard global Lanczos \mathbb{K}_g^m		
	$(m+1, m)$	Error	Time (sec)	m	Error	SF
e^{-x}	(5,4)	1.1×10^{-7}	1.7	143	9.4×10^{-7}	38.3
\sqrt{x}	(9,8)	9.4×10^{-7}	4.1	67	8.2×10^{-7}	3.4
$x^{-1/4}$	(9,8)	3.0×10^{-7}	4.1	84	9.5×10^{-7}	6.2
$\ln(x)$	(9,8)	5.6×10^{-7}	3.6	73	8.7×10^{-7}	5.1
$e^{-\sqrt{x}}$	(4,3)	3.0×10^{-7}	1.6	105	9.4×10^{-7}	28.0
x^{-4}	(3,2)	1.5×10^{-13}	0.9	111	9.9×10^{-7}	40.0

definite tridiagonal matrix $A = \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{n \times n}$ and the symmetric positive definite Toeplitz $B = [b_{i,j}]_{i=1}^n$, where $b_{i,j} = 1/(1 + |i - j|)$ with $n = 1000$. We choose a block size $s = 6$ and $f(x)$ to be the Laurent polynomial $p(x) = x^{-6} + x^5$. Figure 4.1 shows that the approximation of $\text{trace}(V^T p(A)V)$ determined by the extended global Lanczos method is the exact value of (1.1) after $m = 3$ iterations, up to round-off errors, in agreement with the theory. The standard global Lanczos requires many more steps to give an accurate approximation of the expression (1.1).

In the following three examples, we compare the performance of Algorithm 2 with the performance of the standard global Lanczos algorithm. The speedup factor shows the ratio of execution times for these methods, i.e., $\text{SF} := \text{time}_{\text{standard}}/\text{time}_{\text{extended}}$.

EXAMPLE 4.2. Let $A \in \mathbb{R}^{n \times n}$ with $n = 1000$ be a symmetric positive definite matrix, whose eigenvalues are log-uniformly distributed in the interval $[10^{-1}, 10^6]$. The block size is $s = 6$. Results for several functions f are reported in Table 4.1.

EXAMPLE 4.3. The matrix A is the discretization of the negative Laplacian $L_A(u) = -u_{xx} - u_{yy}$ on the unit square $[0, 1]^2$ with homogeneous Dirichlet boundary conditions. The number of inner grid points in each coordinate direction is 100. A discretization by the standard 5-point stencil gives a symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$ with $n = 10000$. The block size is chosen to be $s = 20$. Computed results are shown in Table 4.2.

EXAMPLE 4.4. Let $A = n^2 \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{n \times n}$ with $n = 50000$. The block size is set to $s = 50$. Table 4.3 reports the performance of the methods. A dash ‘-’ signifies that the maximum allowed number of iterations I_{max} is reached before the stopping criterion is satisfied.

5. Conclusion. This paper describes an extended global Lanczos method for the approximation of bilinear forms of the form $\text{trace}(V^T f(A)V)$. An F -orthonormal basis of block vectors for extended Krylov subspaces is computed by short recurrence formulas. The

TABLE 4.3

Example 4.4: the matrix $A \in \mathbb{R}^{n \times n}$ is positive definite tridiagonal with $n = 50000$. Block size is $s = 50$.

$f(x)$	Extended global Lanczos $\mathbb{K}_g^{m+1,m}$			Standard global Lanczos \mathbb{K}_g^m		
	$(m+1, m)$	Error	Time (sec)	m	Error	SF
e^{-x}	(4,3)	2.5×10^{-8}	6.0	-	-	-
\sqrt{x}	(9,8)	8.1×10^{-4}	21.9	-	-	-
$x^{-1/4}$	(10,9)	9.1×10^{-5}	24.5	-	-	-
$\ln(x)$	(19,18)	9.9×10^{-5}	65.4	-	-	-
$e^{-\sqrt{x}}$	(4,3)	3.1×10^{-7}	6.1	-	-	-
x^{-4}	(3,2)	4.0×10^{-8}	3.7	-	-	-

numerical results show the extended global Lanczos method to require fewer iterations than the (standard) global Lanczos method. Timings show the former method to be faster than the latter for the functions and matrices considered in Examples 4.2–4.4. The extended global Lanczos method is particularly attractive to use when linear systems of equations with the matrix A can be solved efficiently.

Acknowledgment. Research of LR was supported in part by NSF grants DMS-1729509 and DMS-1720259.

REFERENCES

- [1] O. ABIDI, M. HEYOUNI, AND K. JBILOU, *On some properties of the extended block and global Arnoldi methods with applications to model reduction*, Numer. Algorithms, 75 (2017), pp. 285–304.
- [2] S. BARONI, R. GEBAUER, O. B. MALCIOGLU, Y. SAAD, P. UMARI, AND J. XIAN, *Harnessing molecular excited states with Lanczos chains*, J. Phys. Condens. Mat., 22 (2010), Art. Id. 074204, 8 pages.
- [3] M. BELLALIJ, L. REICHEL, G. RODRIGUEZ, AND H. SADOK, *Bounding matrix functionals via partial global block Lanczos decomposition*, Appl. Numer. Math., 94 (2015), pp. 127–139.
- [4] R. BOUYOULI, K. JBILOU, R. SADAQA, AND H. SADOK, *Convergence properties of some block Krylov subspace methods for multiple linear systems*, J. Comput. Appl. Math., 196 (2006), pp. 498–511.
- [5] D. R. BOWLER AND T. MIYAZAKI, *$O(n)$ methods in electronic structure calculations*, Rep. Prog. Phys., 75 (2012), Art. Id. 036503, 43 pages.
- [6] A. BULTHEEL, P. GONZÁLEZ-VERA, AND R. ORIVE, *Quadrature on the half-line and two point Padé approximants to Stieljes functions. Part I: Algebraic aspects*, J. Comput. Appl. Math., 65 (1995), pp. 57–72.
- [7] 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.
- [8] L. ELBOUYAHYAOU, A. MESSAOUDI, AND H. SADOK, *Algebraic properties of the block GMRES and block Arnoldi methods*, Electron. Trans. Numer. Anal., 33 (2009), pp. 207–220.
<http://etna.ricam.oeaw.ac.at/vol.33.2008-2009/pp207-220.dir/pp207-220.pdf>
- [9] E. ESTRADA, *The Structure of Complex Networks: Theory and Applications*, Oxford University Press, Oxford, 2011.
- [10] C. FENU, L. REICHEL, G. RODRIGUEZ, AND H. SADOK, *GCV for Tikhonov regularization by partial SVD*, BIT, 57 (2017), pp. 1019–1039.
- [11] P. C. HANSEN, *Rank-Deficient and Discrete Ill-Posed Problems*, SIAM, Philadelphia, 1998.
- [12] M. HEYOUNI, *Extended Arnoldi methods for large low-rank Sylvester matrix equations*, Appl. Numer. Math., 60 (2010), pp. 1171–1182.
- [13] N. J. HIGHAM, *Functions of Matrices*, SIAM, Philadelphia, 2008.
- [14] C. JAGELS AND L. REICHEL, *The extended Krylov subspace method and orthogonal Laurent polynomials*, Linear Algebra Appl., 431 (2009), pp. 441–458.
- [15] K. JBILOU, A. MESSAOUDI, AND H. SADOK, *Global FOM and GMRES algorithms for matrix equations*, Appl. Numer. Math., 31 (1999), pp. 49–63.
- [16] 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.

- [17] T. T. NGO, M. BELLALIJ, AND Y. SAAD, *The trace ratio optimization problem*, SIAM Rev., 54 (2012), pp. 545–569.
- [18] O. NJÅSTAD AND W. J. THRON, *The theory of sequences of orthogonal L-polynomials*, in Padé Approximants and Continued Fractions, H. Waadeland and H. Wallin, eds., Det Kongelige Norske Videnskabers Selskab Skrifter 1, DKNVS, Trondheim, 1983, pp. 54–91.
- [19] Y. SAAD, J. CHELIKOWSKY, AND S. SHONTZ, *Numerical methods for electronic structure calculations of materials*, SIAM Rev., 52 (2010), pp. 3–54.
- [20] V. SIMONCINI, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comput., 29 (2007), pp. 1268–1288.