SHIFTED EXTENDED GLOBAL LANCZOS PROCESSES FOR TRACE ESTIMATION WITH APPLICATION TO NETWORK ANALYSIS

A. H. BENTBIB*, M. EL GHOMARI†, K. JBILOU‡, AND L. REICHEL§

In memory of Mohammed Bellalij.

Abstract. The need to estimate upper and lower bounds for matrix functions of the form trace$(W^T f(A)V)$, where the matrix $A \in \mathbb{R}^{n\times n}$ is large and sparse, $V, W \in \mathbb{R}^{n\times s}$ are block vectors with $1 \leq s \ll n$ columns, and f is a function arises in many applications, including network analysis and machine learning. This paper describes the shifted extended global symmetric and nonsymmetric Lanczos processes and how they can be applied to approximate the trace. These processes compute approximations in the union of Krylov subspaces determined by positive powers of A and negative powers of $A - \sigma I_n$, where the shift σ is a user-chosen parameter. When A is nonsymmetric, transposes of these powers also are used. When A is symmetric and $W = V$, we describe how error bounds or estimates of bounds for the trace can be computed by pairs of Gauss and Gauss-Radau quadrature rules, or by pairs of Gauss and anti-Gauss quadrature rules. These Gauss-type quadrature rules are defined by recursion coefficients for the shifted extended global Lanczos processes. Gauss and anti-Gauss quadrature rules also can be applied to give estimates of error bounds for the trace when A is nonsymmetric and $W \neq V$. Applications to the computation of the Estrada index for networks and to the nuclear norm of a large matrix are presented. Computed examples show the shifted extended symmetric and nonsymmetric Lanczos processes to produce accurate approximations in fewer steps than the standard symmetric and nonsymmetric global Lanczos processes, respectively.

Key words. Shifted extended Krylov subspace, block Lanczos process, extended global symmetric Lanczos process, extended global nonsymmetric Lanczos process, Gauss quadrature, anti-Gauss quadrature, network analysis, Estrada index, trace estimation, nuclear norm estimation.

1. Introduction.

We introduce shifted extended symmetric and nonsymmetric block Lanczos processes. Applications include the analysis of large networks. We first describe this application before discussing the details of the Lanczos processes.

The analysis of networks finds applications in a number of disciplines including social science, engineering, molecular biology, and traffic planning; see [11, 12, 16, 17, 31]. Typically, one is interested in determining the most important vertices of a given network, or to identify global properties of a network.

A network is represented by a graph $G = \{V, E\}$, which consists of a set of *nodes* or *vertices* $V = \{v_i\}_{i=1}^n$ and a set of edges $E = \{\widehat{e}_k = \{v_i, v_j\} : v_i, v_j \in V\}_{k=1}^m$ that connect the vertices. In an *undirected graph* each edge is a "two-way street", while in a *directed graph* at least one edge is a "one-way street". In a *weighted graph*, each edge is assigned a scalar value, which is the weight of the edge; in an *unweighted graph*, all weights are unity.

A *walk* of length k in a graph G is defined as a sequence of vertices and edges such that

$$v_{\ell_0} \overset{\widehat{e}_{j_1}}{\to} v_{\ell_1} \overset{\widehat{e}_{j_2}}{\to} v_{\ell_2} \overset{\widehat{e}_{j_3}}{\to} \dots \overset{\widehat{e}_{j_{k-1}}}{\to} v_{\ell_{k-1}} \overset{\widehat{e}_{j_k}}{\to} v_{\ell_k},$$

*Laboratory LAMAI, Faculty of Sciences and Technologies, Cadi Ayyad University, Marrakech, Morocco. E-mail: a.bentbib@uca.ac.ma

†Laboratory LAMAI, Faculty of Sciences and Technologies, Cadi Ayyad University, Marrakech, Morocco. E-mail: m.elghomari10@gmail.com

‡Laboratory LMPA, 50 Rue F. Buisson, ULCO Calais cedex, France, and Laboratory CSEHS, University UM6P, Benguérir, Morocco. E-mail: Khalide.Jbilou@univ-littoral.fr

§Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: reichel@math.kent.edu

where the edge $\widehat{e}_{j_i} = \{v_{\ell_{i-1}}, v_{\ell_i}\}$ emerges from vertex $v_{\ell_{i-1}}$ and points to vertex v_{ℓ_i} (for a directed graph) or is between the vertices $v_{\ell_{i-1}}$ and v_{ℓ_i} (for an undirected graph). A graph G is connected when there is a walk from any vertex to any other vertex in G. We are concerned with unweighted connected graphs without self-loops and multiple edges. However, the methods described also can be applied to weighted graphs.

We can associate an adjacency matrix $A = [a_{i,j}] \in \mathbb{R}^{n\times n}$ to a graph G. The entries of the adjacency matrix for an unweighted connected graph G without self-loops and multiple edges, with n vertices, are given by

$$a_{i,j} = \begin{cases} 1 & \text{if there is an edge from vertex } v_i \text{ to vertex } v_j \text{ (when } G \text{ is directed)} \\ & \text{or between the vertices } v_i \text{ and } v_j \text{ (when } G \text{ is undirected)}, \\ 0 & \text{otherwise.} \end{cases}$$

The adjacency matrix is symmetric if the graph G is undirected and nonsymmetric otherwise.

The importance of a vertex v_i in a graph G is commonly referred to as its *centrality*. There are several ways to measure centrality. Recently, matrix functions, and in particular the matrix exponential, have received considerable attention for measuring centrality. The number of walks from vertex v_i to vertex v_j of length k are given by $[A^k]_{i,j}$. The *subgraph centrality* of vertex v_i determined by the matrix exponential is defined as

$$[\exp(A)]_{i,i} = 1 + [A]_{i,i} + \frac{[A^2]_{i,i}}{2!} + \frac{[A^3]_{i,i}}{3!} + \dots \ ;$$

see [16, 19, 20]. Thus, the subgraph centrality of the vertex v_i is a weighted average of all walks from v_i back to itself. Longer walks receive a smaller weight than shorter walks. This corresponds to the common modeling assumption that short walks are more important than long ones. The sum of all subgraph centralities of a graph is commonly referred to as the *Estrada index* of G,

$$EE(G) := \sum_{i=1}^{n} [\exp(A)]_{i,i} = \sum_{i=1}^{n} [\exp(\lambda_i)] = \text{trace}(\exp(A)), \tag{1.1}$$

where the λ_i, $i = 1, 2, \dots n$, denote the eigenvalues of the matrix A. The normalized subgraph centrality of the vertex v_i is given by

$$p_i = [\exp(A)]_{i,i}/EE(G), \quad i = 1, 2, \dots, n,$$

and is used to determine the relative importance of vertices: the vertex v_i is important when p_i is relatively large; see [16, Chapter 5]. For some networks, $\ln(EE(G))$ is desired; see [16, p. 99]. We remark that the Estrada index also is a useful measure in statistical thermodynamics [18] and in the investigation of the folding of long-chain molecules [15]. The Estrada index is expensive to compute when the graph G is large.

It is the purpose of this paper to introduce new methods to determine approximations of the Estrada index for large graphs. We will describe novel ways to compute upper and lower bounds for the Estrada index (1.1), or estimates of such bounds, for symmetric and nonsymmetric adjacency matrices for a graph. The methods described also can be applied in machine learning, when $f(t) = \ln(t)$ (see [5, 25, 32]), and in quantum chromodynamics when computing Schatten p-norms, when $f(t) = t^{p/2}$, for some $0 < p \leq 1$; see [6, 34, 36].

Our method for computing upper and lower bounds (or estimates of such bounds) is based on determining upper and lower bounds (or estimates thereof) for expressions of the form

$$\mathcal{I}(f) := \text{trace}(V^T f(A) V), \tag{1.2}$$

where $A \in \mathbb{R}^{n\times n}$ is a large, sparse matrix and $V \in \mathbb{R}^{n\times s}$ is a block vector with $1 \leq s \ll n$ orthonormal columns. Assume for now that the matrix A is symmetric and the function f is analytic on the convex hull of the spectrum of A. Introduce the spectral factorization

$$A = U\Lambda U^T, \qquad \Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n],$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues of A and the matrix $U \in \mathbb{R}^{n\times n}$ of eigenvectors of A is orthogonal. Here and throughout this paper the superscript ^{T} denotes transposition. Then

$$f(A) = Uf(\Lambda)U^T, \qquad f(\Lambda) = \mathrm{diag}[f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)].$$

Bellalij et al. [7] observed that the expression (1.2) can be written as a Stieltjes integral. We have

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

where $\widehat{V} = U^T V$, $\widehat{V}_i$ denotes the ith row of the matrix $\widehat{V}$, and $\alpha(\lambda)$ is a nondecreasing real-valued piecewise constant function with possible discontinuities at the eigenvalues λ_i of A; $d\alpha(\lambda)$ is the associated measure.

Bellalij et al. [7] applied pairs of Gauss and Gauss-Radau quadrature rules to compute upper and lower bounds for expressions of the form (1.2) for certain functions f, including the exponential. The m-point Gauss rule is determined by applying m steps of the global symmetric Lanczos process, which generates an orthonormal basis for the global Krylov subspace

$$\mathbb{K}_m(A, V) := \mathrm{span}\{V, AV, \ldots, A^{m-1}V\} = \{p(A)V : p \in \Pi_{m-1}\}, \tag{1.4}$$

where Π_{m-1} denotes the set of polynomials of degree at most $m-1$. Global Krylov subspace techniques were first proposed by Jbilou et al. [28, 29] for solving linear systems of equations with multiple right-hand sides.

Application of m steps of the global symmetric Lanczos process [29] to A with initial block vector $V \in \mathbb{R}^{n\times s}$ gives the decomposition

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

where $\otimes$ denotes the Kronecker product. The $n \times s$ block columns of the matrix $\mathbb{V}_m = [V_1, V_2, \ldots, V_m] \in \mathbb{R}^{n\times ms}$, where $V_1 = V/\|V\|$ and $\|\cdot\|$ denotes the Frobenius matrix norm, form a basis for the subspace (1.4) that is orthonormal with respect to the inner product

$$\langle X, Y\rangle = \mathrm{trace}(X^T Y), \tag{1.6}$$

i.e.,

$$\langle V_j, V_k\rangle = \mathrm{trace}(V_j^T V_k) = \begin{cases} 1 & j = k, \\ 0 & j \neq k, \end{cases}$$

where we recall that $\langle V_j, V_j\rangle = \|V_j\|^2$. 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 $E_m \in \mathbb{R}^{ms\times s}$ is made up of the columns $(m-1)s+1, (m-1)s+2, \ldots, ms$ of the identity matrix I_{ms}. We tacitly assume that m is small enough so that the decomposition (1.5) with the stated properties exists. This is the generic situation.

The m-point Gauss quadrature rule for the Stieltjes integral (1.3) is given by

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

where $e_1 = [1, 0, \ldots, 0]^T$ denotes the first canonical basis vector. The Gauss rule satisfies

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

see [7] for details.

The associated $(m+1)$-point Gauss-Radau quadrature rule with a specified node ξ can be expressed as

$$\mathcal{R}_{m+1}^{\xi}(f) = \|V\|^2 e_1^T f(T_{m+1,\xi}) e_1, \tag{1.8}$$

where $T_{m+1,\xi} \in \mathbb{R}^{(m+1)\times(m+1)}$ is determined by modifying the last diagonal entry of the matrix T_{m+1} so that $T_{m+1,\xi}$ has an eigenvalue at ξ. Here T_{m+1} is the matrix associated with the $(m+1)$-point Gauss quadrature rule. The Gauss-Radau quadrature rule satisfies

$$\mathcal{I}(p) = \mathcal{R}_{m+1}^{\xi}(p), \quad \forall p \in \Pi_{2m};$$

see [7].

When the integrand f in (1.3) is $2m+1$ times continuously differentiable and the derivatives $f^{(2m)}$ and $f^{(2m+1)}$ do not change sign in the convex hull of the spectrum of A, the Radau point ξ can be chosen to be one of the endpoints of the convex hull so that the quadrature rules $\mathcal{G}_m(f)$ and $\mathcal{R}_{m+1}^{\xi}(f)$ bracket $\mathcal{I}(f)$; see [7] for details. This result follows from the seminal work by Golub and Meurant [23].

When $f^{(2m)}$ or $f^{(2m+1)}$ change sign in the convex hull of the spectrum of A, pairs of Gauss and Gauss-Radau quadrature rule are not guaranteed to bracket $\mathcal{I}(f)$. In this situation it may be attractive to compute estimates of upper and lower bounds for $\mathcal{I}(f)$ by evaluating pairs of Gauss and anti-Gauss quadrature rules. Anti-Gauss rules were introduced by Laurie [30] to estimate the quadrature error of Gauss rules applied to the approximation of integrals of a real-valued function with respect to a nonnegative real-valued measure. Recent applications to the approximation of matrix functions are described in [1, 2, 8].

If the function f cannot be approximated accurately by a polynomial of low to moderate degree, then Gauss-type quadrature rules (1.7) and (1.8) typically will not furnish accurate approximations of the expression (1.2). This situation occurs, for instance, when the function f or one of its low-order derivatives has a singularity at or close to some eigenvalue of A. Then it may be beneficial to approximate f by a rational function with a pole at or close to a singularity of f or of one of its derivatives. In fact, Druskin and Knizhnerman [14] have shown that it also may be beneficial to approximate entire functions f by rational functions with a pole in the finite complex plane, compared to polynomial approximations. Therefore, Druskin and Knizhnerman [14] suggested the application of extended Krylov subspaces when the matrix A is nonsingular. These subspaces are determined by both positive and negative powers of A.

The shifted extended global Krylov subspaces used in this paper generalize the extended Krylov subspaces applied by Druskin and Knizhnerman [14] by allowing a real or complex shift σ. Thus, we consider approximation methods for (1.2) that use shifted extended Krylov subspaces of the form

$$\mathbb{K}_m^{\sigma}(A, V) := \text{span}\left\{V, AV, A^2V, \ldots, A^{m-1}V, (A-\sigma I_n)^{-1}V, (A-\sigma I_n)^{-2}V, \ldots, (A-\sigma I_n)^{-m}V\right\} \subset \mathbb{R}^{n\times s}, \tag{1.9}$$

where the shift σ is distinct from the eigenvalues of A. We assume that m is small enough so that the block vectors in the right-hand side of (1.9) are linearly independent.

This paper presents Gauss, Gauss-Radau, and anti-Gauss quadrature rules associated with shifted extended Krylov subspaces (1.9). These rules are used to approximate (1.2). Numerical examples in Section 6 illustrate that these quadrature rules may yield significantly more accurate approximations than Gauss and Gauss-Radau rules that are based on the "standard" global Krylov subspaces (1.4).

We also are interested in determining upper and lower bounds for expressions of the form

$$\mathcal{I}(f) := \operatorname{trace}(W^T f(A)V), \tag{1.10}$$

when the matrix A is nonsymmetric and $W, V \in \mathbb{R}^{n\times s}$, $1 \leq s \ll n$, are block vectors that might be distinct. Our analysis assumes that the matrix A is diagonalizable and has the spectral factorization

$$A = PDP^{-1}, \quad \Lambda = \operatorname{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n] \in \mathbb{C}^{n\times n}, \tag{1.11}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues of A and $P \in \mathbb{C}^{n\times n}$ is a nonsingular matrix of unit eigenvectors. However, the application of the numerical methods described does not require existence of the factorization (1.11). A discussion on the situation when the factorization (1.11) does not exist is provided by Pozza et al. [33].

The function $f(A)$ can be defined by $f(A) = Pf(\Lambda)P^{-1}$. Substituting (1.11) into (1.10), and setting $\widehat{W} = P^T W \in \mathbb{C}^{n\times s}$ and $\widehat{V} = P^{-1}V \in \mathbb{C}^{n\times s}$, yield

$$\mathcal{I}(f) = \operatorname{trace}(W^T f(A)V) = \sum_{i=1}^{n} f(\lambda_i)\operatorname{trace}(\widehat{W}_i^T \widehat{V}_i) = \int f(\lambda)d\mu(\lambda), \tag{1.12}$$

where $\widehat{W}_i$ and $\widehat{V}_i$ denote the ith rows of the matrices $\widehat{W}$ and $\widehat{V}$, respectively. Further, $d\mu(\lambda)$ is a measure with support at the eigenvalues of A.

We will present the shifted extended global nonsymmetric Lanczos process for generating biorthogonal bases $\{V_i\}_{i=1}^{2m}$ and $\{W_i\}_{i=1}^{2m}$ for the shifted extended global Krylov subspaces $\mathbb{K}_m^\sigma(A, V)$ and $\mathbb{K}_m^\sigma(A^T, W)$, respectively. Thus,

$$\langle V_j, W_k\rangle = \operatorname{trace}(W_k^T V_j) = \begin{cases} 1 & j = k, \\ 0 & j \neq k. \end{cases}$$

We describe Gauss and anti-Gauss quadrature rules for estimating upper and lower error bounds for the computed approximation of (1.10). This work extends the quadrature rules discussed by Fenu et al. [21] to shifted extended Krylov subspaces, and it extends the recursion relations for the extended global symmetric Lanczos process described in [9, 26] to the shifted symmetric and nonsymmetric Lanczos processes.

This paper is organized as follows. Section 2 reviews results in [8] on the extended global Lanczos process applied to a symmetric matrix. We also discuss the connection between the shifted extended global symmetric Lanczos process and Gauss-type quadrature rules. Section 3 is concerned with the computation of anti-Gauss quadrature rules associated with the subspace (1.9) when A is symmetric. The shifted extended global nonsymmetric Lanczos process for generating biorthogonal bases for the spaces $\mathbb{K}_m^\sigma(A, V)$ and $\mathbb{K}_m^\sigma(A^T, W)$ is described in Section 4. These bases are expressed with the aid of shifted orthogonal Laurent polynomials. Section 5 discusses the computation of anti-Gauss-Laurent quadrature rules. Numerical experiments with applications to network analysis and Schatten p-norm computations are presented in Section 6 to illustrate the quality of the computed approximations. Section 7 contains concluding remarks.

2. The shifted extended global symmetric Lanczos process. This section discusses the shifted extended global symmetric Lanczos process and its relation to shifted Gauss-Laurent and Gauss-Laurent-Radau quadrature rules. While introducing a shift $\sigma \neq 0$ is straightforward theoretically, it is important in applications. The shift affects the coefficients in the recursion relation satisfied by the orthonormal block vectors that make up a basis for the shifted extended Krylov subspaces (1.9). We require σ to be real and outside the convex hull of the spectrum of A.

2.1. Preliminaries and notation. We begin by recalling some definitions and notation that will be used throughout this paper. The Kronecker product of two matrices $A = [a_{i,j}]$ and $B = [b_{i,j}]$ is defined by $A \otimes B = [a_{i,j}B]$ and satisfies the following properties:

1. $(A \otimes B)(C \otimes D) = AC \otimes BD$,
2. $(A \otimes B)^T = A^T \otimes B^T$.

Definition 2.1. *[10] Let the matrices $M = [M_1, M_2, \ldots, M_s] \in \mathbb{R}^{n\times sp}$ and $N = [N_1, N_2, \ldots, N_\ell] \in \mathbb{R}^{n\times \ell p}$ be partitioned into block columns M_i and N_i of size $n \times p$, respectively. Then the $\diamond$-product of the matrices M and N is given by*

$$M^T \diamond N = [\langle N_j, M_i\rangle]_{i=1,2,\ldots,s}^{j=1,2,\ldots,\ell} \in \mathbb{R}^{s\times \ell}.$$

The following proposition gives some properties of the $\diamond$-product. We refer to this product as the "diamond product".

Proposition 2.2. *[10] Let $A, B, C \in \mathbb{R}^{n\times ps}$, $D \in \mathbb{R}^{n\times n}$, $L \in \mathbb{R}^{p\times p}$, and $\alpha \in \mathbb{R}$. Then*

1. $(A + B)^T \diamond C = A^T \diamond C + B^T \diamond C$,
2. $A^T \diamond (B + C) = A^T \diamond B + A^T \diamond C$,
3. $(\alpha A)^T \diamond C = \alpha(A^T \diamond C)$,
4. $(A^T \diamond B)^T = B^T \diamond A$,
5. $(DA)^T \diamond B = A^T \diamond (D^T B)$,
6. $A^T \diamond (B(L \otimes I_p)) = (A^T \diamond B)L$.

2.2. The shifted extended global symmetric Lanczos process. This subsection describes the application of the extended global symmetric Lanczos process to the shifted matrix $A - \sigma I_n$ to generate an orthonormal basis $\{V_j\}_{j=1}^{2m}$ of block vectors $V_j \in \mathbb{R}^{n\times s}$ for the shifted extended global Krylov subspace (1.9). This basis is computed by short recurrence formulas as follows:

$$\begin{aligned} V_1 &= \frac{V}{\alpha_{1,1}}, \\ V_2 &= \frac{\widetilde{V}_2}{\alpha_{2,2}}, \quad \widetilde{V}_2 = (A - \sigma I_n)^{-1} V_1 - \alpha_{1,2} V_1, \end{aligned} \tag{2.1}$$

where $\alpha_{1,1} = \|V\|$, $\alpha_{1,2} = \langle (A - \sigma I_n)^{-1} V_1, V_1\rangle$, $\alpha_{2,2} = \|\widetilde{V}_2\|$. For $j = 1, 2, \ldots, m$, we have

$$\begin{aligned} h_{2j+1,2j-1} V_{2j+1} &= \widetilde{V}_{2j+1} = A V_{2j-1} - \sum_{i=2j-3}^{2j} h_{i,2j-1} V_i, \\ h_{2j+2,2j} V_{2j+2} &= \widetilde{V}_{2j+2} = (A - \sigma I_n)^{-1} V_{2j} - \sum_{i=2j-2}^{2j+1} h_{i,2j} V_i. \end{aligned} \tag{2.2}$$

The coefficients $h_{i,j}$ are determined so that the block vectors $V_1, V_2, \ldots, V_{2m+2}$ are orthonormal. This leads to the expressions

$$\begin{array}{rclcrcl} h_{i,2j-1} & = & \langle AV_{2j-1}, V_i\rangle, & \qquad & h_{2j+1,2j-1} & = & \|\widetilde{V}_{2j+1}\|, \\ h_{i,2j} & = & \langle (A-\sigma I_n)^{-1}V_{2j}, V_i\rangle, & & h_{2j+2,2j} & = & \|\widetilde{V}_{2j+2}\|. \end{array} \tag{2.3}$$

We provide recursion relations for computing the coefficients $h_{i,j}$ below.

PROPOSITION 2.3. *Let the coefficients $\alpha_{i,j}$ and $h_{i,j}$ be defined by* (2.1), (2.2), *and* (2.3). *They can be computed as follows:*

$$\begin{aligned} h_{1,2} &= \alpha_{2,2}, \\ h_{2,1} &= \frac{1}{\alpha_{2,2}}[1 - \alpha_{1,2}h_{1,1} + \sigma\alpha_{1,2}], \\ h_{2,3} &= -\frac{\alpha_{1,2}h_{1,3}}{\alpha_{2,2}}. \end{aligned}$$

For $j = 2, 3, \ldots, m$, we have

$$\begin{aligned} h_{2j-3,2j-1} &= h_{2j-1,2j-3}, \\ h_{2j-2,2j-1} &= \frac{-1}{h_{2j-2,2j-4}}h_{2j-3,2j-4}h_{2j-3,2j-1}, \\ h_{2j,2j-1} &= \frac{1}{h_{2j,2j-2}}\left[\sigma h_{2j-1,2j-2} - \sum_{i=2j-3}^{2j-1} h_{i,2j-2}h_{i,2j-1}\right], \\ h_{2j-2,2j} &= h_{2j,2j-2}, \\ h_{2j-1,2j} &= \frac{-1}{h_{2j-1,2j-3}}h_{2j-2,2j-3}h_{2j-2,2j}, \\ h_{2j+1,2j} &= \frac{1}{h_{2j+1,2j-1}}\left[\sigma h_{2j-1,2j} - \sum_{i=2j-2}^{2j} h_{i,2j-1}h_{i,2j}\right]. \end{aligned}$$

Proof. From the relations (2.1) and (2.3), and due to the orthonormality of the block vectors $\{V_i\}_{i=1}^{2m+2}$, we get

$$\begin{aligned} h_{1,2} &= \langle (A-\sigma I_n)^{-1}V_2, V_1\rangle = \langle (A-\sigma I_n)^{-1}V_1, V_2\rangle \\ &= \langle \alpha_{2,2}V_2 + \alpha_{1,2}V_1, V_2\rangle = \alpha_{2,2}. \end{aligned}$$

The second equation in (2.1) yields

$$\alpha_{2,2}AV_2 = V_1 + \sigma\alpha_{2,2}V_2 - \alpha_{1,2}AV_1 + \alpha_{1,2}\sigma V_1,$$

and it follows that

$$\begin{aligned} h_{2,1} &= \langle V_1, AV_2\rangle \\ &= \frac{1}{\alpha_{2,2}}\langle V_1, V_1 + \sigma\alpha_{2,2}V_2 - \alpha_{1,2}AV_1 + \alpha_{1,2}\sigma V_1\rangle \\ &= \frac{1}{\alpha_{2,2}}[1 - \alpha_{1,2}h_{1,1} + \sigma\alpha_{1,2}] \end{aligned}$$

and

$$h_{2,3} = \langle V_3, AV_2 \rangle = \frac{-1}{\alpha_{2,2}} \langle V_3, \alpha_{1,2} AV_1 \rangle = -\frac{\alpha_{1,2} h_{1,3}}{\alpha_{2,2}}.$$

We also have

$$\begin{aligned} h_{2j-3,2j-1} &= \langle AV_{2j-1}, V_{2j-3} \rangle = \langle AV_{2j-3}, V_{2j-1} \rangle \\ &= \langle h_{2j-1,2j-3} V_{2j-1} + \sum_{i=2j-5}^{2j-2} h_{i,2j-3} V_i, V_{2j-1} \rangle = h_{2j-1,2j-3}, \\ h_{2j-2,2j-1} &= \langle AV_{2j-1}, V_{2j-2} \rangle = \langle AV_{2j-2}, V_{2j-1} \rangle. \end{aligned}$$

Multiplying the second equation in (2.2) by $(A - \sigma I_n)$ from the left gives

$$h_{2j-2,2j-4} AV_{2j-2} = V_{2j-4} - \sum_{i=2j-6}^{2j-3} h_{i,2j-4} (A - \sigma I_n) V_i + \sigma h_{2j-2,2j-4} V_{2j-2},$$

which implies that

$$h_{2j-2,2j-1} = -\frac{1}{h_{2j-2,2j-4}} \left[\sum_{i=2j-6}^{2j-3} h_{i,2j-4} h_{i,2j-1} \right].$$

Since, $h_{i,2j-1} = 0$ for $i = 1, 2, \ldots, 2j-4$, it follows that

$$h_{2j-2,2j-1} = -\frac{1}{h_{2j-2,2j-4}} h_{2j-3,2j-4} h_{2j-3,2j-1}.$$

For the coefficient $h_{2j,2j-1}$, we have

$$h_{2j,2j-1} = \langle AV_{2j-1}, V_{2j} \rangle = \langle AV_{2j}, V_{2j-1} \rangle,$$

and using the expression for AV_{2j} and similar manipulations as above give

$$\begin{aligned} h_{2j,2j-1} &= \frac{1}{h_{2j,2j-2}} \left[\sigma h_{2j-1,2j-2} - \sum_{i=2j-3}^{2j-1} h_{i,2j-2} h_{i,2j-1} \right], \\ h_{2j-2,2j} &= \langle (A - \sigma I_n)^{-1} V_{2j}, V_{2j-2} \rangle = \langle (A - \sigma I_n)^{-1} V_{2j-2}, V_{2j} \rangle \\ &= \langle h_{2j,2j} V_{2j} + \sum_{i=2j-4}^{2j-1} h_{i,2j-2} V_i, V_{2j} \rangle = h_{2j,2j-2}, \\ h_{2j-1,2j} &= \langle (A - \sigma I_n)^{-1} V_{2j}, V_{2j-1} \rangle = \langle (A - \sigma I_n)^{-1} V_{2j-1}, V_{2j} \rangle. \end{aligned}$$

According to the first relation in (2.2), we can express $(A - \sigma I_n)^{-1} V_{2j-1}$ as

$$(A - \sigma I_n)^{-1} V_{2j-1} = \frac{1}{h_{2j-1,2j-3}} \left[V_{2j-3} - \sum_{i=2j-5}^{2j-2} h_{i,2j-3} (A - \sigma I_n)^{-1} V_i + \sigma (A - \sigma I_n)^{-1} V_{2j-3} \right].$$

The orthogonality condition and fact that $h_{i,2j} = 0$ for $i < 2j-2$ lead to

$$\begin{aligned}
h_{2j-1,2j} &= \frac{-1}{h_{2j-1,2j-3}} h_{2j-2,2j-3} h_{2j-2,2j}, \\
h_{2j+1,2j} &= \langle (A-\sigma I_n)^{-1} V_{2j}, V_{2j+1} \rangle = \langle (A-\sigma I_n)^{-1} V_{2j+1}, V_{2j} \rangle \\
&= \left\langle \frac{1}{h_{2j+1,2j-1}} \left[V_{2j-1} - \sum_{i=2j-3}^{2j} h_{i,2j-1} (A-\sigma I_n)^{-1} V_i + \sigma (A-\sigma I_n)^{-1} V_{2j-1} \right], V_{2j} \right\rangle \\
&= \frac{1}{h_{2j+1,2j-1}} \left[\sigma h_{2j-1,2j} - \sum_{i=2j-2}^{2j} h_{i,2j-1} h_{i,2j} \right].
\end{aligned}$$

This completes the proof. $\square$

Example: Let $m = 3$. Then the matrix $H = [h_{i,j}] \in \mathbb{R}^{8\times 6}$ is of the form

$$H = \begin{bmatrix}
h_{1,1} & h_{1,2} & h_{1,3} & 0 & 0 & 0 \\
h_{2,1} & h_{2,2} & h_{2,3} & h_{2,4} & 0 & 0 \\
h_{3,1} & h_{3,2} & h_{3,3} & h_{3,4} & h_{3,5} & 0 \\
0 & h_{4,2} & h_{4,3} & h_{4,4} & h_{4,5} & h_{4,6} \\
0 & 0 & h_{5,3} & h_{5,4} & h_{5,5} & h_{5,6} \\
0 & 0 & 0 & h_{6,4} & h_{6,5} & h_{6,6} \\
0 & 0 & 0 & 0 & h_{7,5} & h_{7,6} \\
0 & 0 & 0 & 0 & 0 & h_{8,6}
\end{bmatrix}$$

All entries denoted by $h_{i,j}$ in the matrix may be nonvanishing. Moreover, entries on the second superdiagonal equal entries on the second subdiagonal, i.e., $h_{i,i+2} = h_{i+2,i}$ for $i = 1, 2, 3, 4$. $\square$

We next discuss some useful properties of the shifted extended global Krylov subspaces. Here and below we will tacitly assume that the number of steps of the shifted extended global symmetric Lanczos process is small enough to avoid breakdown. This is the generic situation; breakdown is very rare.

Application of m steps of the shifted extended global symmetric Lanczos process to the matrix A with initial block vector V_1 of unit norm yields the decomposition

$$\begin{aligned}
A\mathbb{V}_{2m} &= \mathbb{V}_{2m+1}(\widetilde{\mathbb{T}}_{2m} \otimes I_s) \\
&= \mathbb{V}_{2m}(\mathbb{T}_{2m} \otimes I_s) + V_{2m+1}(\left[t_{2m+1,2m-1},\ t_{2m+1,2m}\right] E_m^T \otimes I_s),
\end{aligned}$$

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

$$\mathbb{T}_{2m} = [t_{i,j}] = \mathbb{V}_{2m}^T \diamond A\mathbb{V}_{2m} \in \mathbb{R}^{2m\times 2m},$$

with $t_{i,j} = \langle AV_j, V_i \rangle$, $i, j = 1, 2, \ldots, 2m$. The matrices

$$\mathbb{V}_{2m} = [V_1, V_2, \ldots, V_{2m}], \quad \mathbb{V}_{2m+1} = [V_1, V_2, \ldots, V_{2m+1}]$$

are made up of orthonormal block vectors $V_j \in \mathbb{R}^{n\times s}$. Also introduce the matrix

$$\widetilde{\mathbb{T}}_{2m} = \mathbb{V}_{2m+1}^T \diamond A\mathbb{V}_{2m} \in \mathbb{R}^{(2m+1)\times 2m}. \tag{2.4}$$

The entries of $\mathbb{T}_{2m}$ and $\widetilde{\mathbb{T}}_{2m}$ can be expressed in terms of recursion coefficients for the shifted extended global symmetric Lanczos process as shown below. This makes them easy to compute.

Proposition 2.4. *Let the coefficients* $h_{i,j}$ *and* $\alpha_{i,j}$ *be defined by* (2.1) *and* (2.2). *The nontrivial entries of the matrices* $\mathbb{T}_{2m} = [t_{i,j}]$ *and* $\widetilde{\mathbb{T}}_{2m} = [t_{i,j}]$ *can be expressed as*

$$\begin{aligned} t_{i,2j-1} &= h_{i,2j-1}, \quad \textit{for } i = 2j-3, 2j-2, \ldots, 2j+1, \quad j = 1, 2, \ldots, m, \\ t_{1,2} &= t_{2,1}, \\ t_{2,2} &= \sigma - \frac{\alpha_{1,2}}{\alpha_{2,2}} t_{2,1}, \\ t_{3,2} &= -\frac{\alpha_{1,2}}{\alpha_{2,2}} t_{3,1}. \end{aligned}$$

Moreover, for $j = 1, 2, \ldots, m-1$, *we have*

$$\begin{aligned} t_{2j+1,2j+2} &= t_{2j+2,2j+1}, \\ t_{2j+2,2j+2} &= \sigma - \frac{h_{2j+1,2j}}{h_{2j+2,2j}} t_{2j+2,2j+1}, \\ t_{2j+3,2j+2} &= -\frac{h_{2j+1,2j}}{h_{2j+2,2j}} t_{2j+3,2j+1}. \end{aligned}$$

Proof. The proof is similar to that of [9, Proposition 3.1]. □

The orthonormal basis $\{V_j\}_{j=1}^{2m}$ for (1.9) can be expressed with the aid of orthogonal shifted Laurent polynomials, i.e.,

$$V_{2j-1} = R_{2j-2}(A)V \quad \text{and} \quad V_{2j} = R_{2j-1}(A)V, \quad j = 1, 2, \ldots, m, \tag{2.5}$$

where R_{2j-1} and R_{2j} are shifted Laurent polynomials that live in the spaces

$$\begin{aligned} R_{2j-1}(x) \in \Delta_{-j,j-1} &:= \operatorname{span}\{1, (x-\sigma)^{-1}, x, \ldots, x^{j-1}, (x-\sigma)^{-j}\}, \\ R_{2j}(x) \in \Delta_{-j,j} &:= \operatorname{span}\{1, (x-\sigma)^{-1}, x, \ldots, (x-\sigma)^{-j}, x^{j}\}. \end{aligned}$$

Proposition 2.5. *Let* A *be a symmetric matrix and let the coefficients* $\alpha_{i,j}$ *and* $h_{i,j}$ *be given by* (2.1) *and* (2.2). *Then the sequence of shifted Laurent polynomials* $R_0, R_1, \ldots, R_{2m}$, *determined by* (2.5), *are orthonormal with respect to the bilinear form*

$$\langle P, Q \rangle = \operatorname{trace}((P(A)V)^T Q(A)V) = \int P(\lambda)Q(\lambda)d\alpha(\lambda),$$

where $d\alpha$ *is the measure defined in* (1.3). *These shifted Laurent polynomials satisfy a pair of five-term recurrence relations of the form*

$$\begin{aligned} h_{2j+1,2j-1}R_{2j}(x) &= xR_{2j-2}(x) - \sum_{i=2j-3}^{2j} h_{i,2j-1}R_{i-1}(x), \\ h_{2j+2,2j}R_{2j+1}(x) &= (x-\sigma)^{-1}R_{2j-1}(x) - \sum_{i=2j-2}^{2j+1} h_{i,2j}R_{i-1}(x), \end{aligned}$$

where $R_1(x) = (1/\alpha_{2,2})[(x-\sigma)^{-1}R_0(x) - \alpha_{1,2}R_0(x)]$, $R_0(x) = 1/\alpha_{1,1}$, *and* $R_{-2} = R_{-1} = 0$.

Proof. A similar result is shown in [9, Theorem 2.7]. The proposition can be shown by modifying the proof presented there. □

2.3. Shifted Gauss-Laurent quadrature rules. The shifted extended global symmetric Lanczos approximation of the Stieltjes integral (1.2) is given by

$$\mathcal{G}_{2m}^{\sigma}(f) = \|V\|^2 e_1^T f(\mathbb{T}_{2m}) e_1 = \sum_{i=1}^{2m} f(\theta_i) w_i, \tag{2.6}$$

where θ_i denotes the ith eigenvalue of $\mathbb{T}_{2m}$ and $w_i = \|V\|^2 u_{i,1}^2$. Here $u_{i,1}$ is the first component of the normalized eigenvector u_i of $\mathbb{T}_{2m}$ associated with the eigenvalue θ_i. Using the same techniques as in [9, Proposition 3.4], we can show that the zeros of R_{2m} are the eigenvalues of $\mathbb{T}_{2m}$. Then, applying (2.6), we find that

$$\mathcal{G}_{2m}^{\sigma}(R_{2m}) = 0. \tag{2.7}$$

We will show below that (2.6) is a shifted Gauss-Laurent quadrature rule. The following properties help us to establish this fact.

PROPOSITION 2.6. *Let the matrix pairs $\{\mathbb{V}_{2m}, \mathbb{T}_{2m}\}$ and $\{\mathbb{W}_{2m}, \mathbb{H}_{2m}\}$, where $\mathbb{V}_{2m} = [V_1, V_2, \ldots, V_{2m}]$ and $\mathbb{W}_{2m} = [W_1, W_2, \ldots, W_{2m}]$, be associated to the shifted extended global Krylov subspace $\mathbb{K}_m^{\sigma}(A, V)$ and the global Krylov subspace $\mathbb{K}_{2m}(A, (A-\sigma I_n)^{-m} V)$, respectively, and let $P_{2m} = \mathbb{W}_{2m}^T \diamond \mathbb{V}_{2m} \in \mathbb{R}^{2m\times 2m}$. Then the matrices $\mathbb{T}_{2m}$ and $\mathbb{H}_{2m}$ are similar, i.e., $\mathbb{H}_{2m} = P_{2m}\mathbb{T}_{2m}P_{2m}^T$, where $P_{2m}^T P_{2m} = I_{2m}$. Moreover, the matrices $\mathbb{T}_{2m}$ and $\mathbb{H}_{2m}$ satisfy the properties:*

1. $(\mathbb{T}_{2m} - \sigma I_{2m})^m (\mathbb{V}_{2m}^T \diamond (A - \sigma I_n)^{-m} V_1) = e_1$,
2. $(\mathbb{H}_{2m} - \sigma I_{2m})^m e_1 = P_{2m}(\mathbb{T}_{2m} - \sigma I_{2m})^m P_{2m}^T e_1$,
3. $\|(A - \sigma I_n)^{-m} V_1\|^2 e_1^T f(\mathbb{H}_{2m})(\mathbb{H}_{2m} - \sigma I_{2m})^{2m} e_1 = e_1^T f(\mathbb{T}_{2m}) e_1$.

Further, we have

$$\|(A - \sigma I_n)^{-m} V_1\|^2 = \frac{1}{e_1^T(\mathbb{H}_{2m} - \sigma I_{2m})^{2m} e_1}.$$

Proof. We first show that $P_{2m}^T P_{2m} = I_{2m}$. By using the properties of the $\diamond$-product, we obtain

$$P_{2m}^T P_{2m} = (\mathbb{W}_{2m}^T \diamond \mathbb{V}_{2m})(\mathbb{V}_{2m}^T \diamond \mathbb{W}_{2m}) = \mathbb{W}_{2m}^T \diamond (\mathbb{V}_{2m}([\mathbb{V}_{2m}^T \diamond \mathbb{W}_{2m}] \otimes I_s)).$$

Since the shifted extended global subspace can be regarded as a global Krylov subspace, i.e.,

$$\mathbb{K}_m^{\sigma}(A, V) = \mathbb{K}_{2m}(A, (A - \sigma I_n)^{-m} V),$$

it follows that the columns of the matrix $\mathbb{W}_{2m}$ belong to $\mathbb{K}_m^{\sigma}(A, V)$. Therefore,

$$P_{2m}^T P_{2m} = \mathbb{W}_{2m}^T \diamond \mathbb{W}_{2m} = I_{2m}.$$

Using the definition of $\mathbb{H}_{2m}$, we obtain

$$\begin{aligned}\mathbb{H}_{2m} &= \mathbb{W}_{2m}^T \diamond A\mathbb{W}_{2m} = \mathbb{W}_{2m}^T \diamond [A\mathbb{V}_{2m}(\mathbb{V}_{2m}^T \diamond \mathbb{W}_{2m} \otimes I_s)] \\ &= (\mathbb{W}_{2m}^T \diamond A\mathbb{V}_{2m})(\mathbb{V}_{2m}^T \diamond \mathbb{W}_{2m}) \\ &= (\mathbb{W}_{2m}^T \diamond \mathbb{V}_{2m})(\mathbb{V}_{2m}^T \diamond A\mathbb{V}_{2m})(\mathbb{V}_{2m}^T \diamond \mathbb{W}_{2m}) = P_{2m}\mathbb{T}_{2m}P_{2m}^T.\end{aligned}$$

Thus, the matrices $\mathbb{T}_{2m}$ and $\mathbb{H}_{2m}$ are similar.

An application of a slightly modified form of [9, Lemma 3.8] gives

$$(A - \sigma I_n)^{-m} V_1 = \mathbb{V}_{2m}[(\mathbb{T}_{2m} - \sigma I_{2m})^{-m} e_1 \otimes I_s].$$

Multiplying the last equation by $\mathbb{V}_{2m}^T$ from the left and using properties of the $\diamond$-product, we obtain

$$\mathbb{V}_{2m}^T \diamond (A - \sigma I_n)^{-m} V_1 = (\mathbb{T}_{2m} - \sigma I_{2m})^{-m} e_1,$$

which shows the first assertion.

The second assertion follows by using the fact that $\mathbb{H}_{2m} = P_{2m}\mathbb{T}_{2m}P_{2m}^T$ and the orthogonality of the matrix P_{2m}.

For the third assertion, we have

$$e_1^T (f(\mathbb{H}_{2m})(\mathbb{H}_{2m} - \sigma I_{2m})^{2m} e_1 = e_1^T (\mathbb{H}_{2m} - \sigma I_{2m})^m (f(\mathbb{H}_{2m})(\mathbb{H}_{2m} - \sigma I_{2m})^m e_1.$$

An application of the second assertion shows that the above expression is equal to

$$e_1^T P_{2m}(\mathbb{T}_{2m} - \sigma I_{2m})^m f(\mathbb{T}_{2m})(\mathbb{T}_{2m} - \sigma I_{2m})^m P_{2m}^T e_1.$$

On the other hand, we have

$$\begin{aligned}(\mathbb{T}_{2m} - \sigma I_{2m})^m P_{2m}^T e_1 &= \frac{1}{\|(A - \sigma I_n)^{-m} V_1\|} (\mathbb{T}_{2m} - \sigma I_{2m})^m [\mathbb{V}_{2m}^T \diamond (A - \sigma I_n)^{-m} V_1] \\ &= \frac{e_1}{\|(A - \sigma I_n)^{-m} V_1\|},\end{aligned}$$

where the last equality follows from the first assertion. This concludes the proof of the last assertion. ◻

Theorem 2.7. *Let A be a symmetric matrix. Apply m steps of the shifted extended global Lanczos process with the initial block vector $V \in \mathbb{R}^{n\times s}$ to A to evaluate the expression* (2.6). *Then this expression is a $2m$-point shifted Gauss-Laurent quadrature rule associated with the measure $d\alpha$ in* (1.3)*, i.e.,*

$$\mathcal{G}_{2m}^{\sigma}(f) = \mathcal{I}(f) \quad \forall f \in \Delta_{-2m,2m-1}.$$

Moreover, if the function f is $4m$ times continuously differentiable in the convex hull of the spectrum of A, then the reminder term for this rule is given by

$$\mathcal{E}_{2m}(f) := \mathcal{I}(f) - \mathcal{G}_{2m}^{\sigma}(f) = \frac{d^{4m}}{dx^{4m}}(fw)_{x=\widetilde{\theta}} \frac{1}{(4m)!} \int \prod_{j=1}^{2m} (x - \theta_j)^2 w(x) d\alpha(x), \tag{2.8}$$

where

$$w(x) := \frac{1}{e_1^T (\mathbb{H}_{2m} - \sigma I_{2m})^{2m} e_1} (x - \sigma)^{2m}$$

and the scalar $\widetilde{\theta}$ lives in the largest open interval contained in the convex hull of the spectrum of A.

Proof. According to Proposition 2.6, we have

$$\mathcal{G}_{2m}^{\sigma}(f) = \|V\|^2 e_1^T f(\mathbb{H}_{2m}) w(\mathbb{H}_{2m}) e_1.$$

Therefore, $\mathcal{G}_{2m}^{\sigma}$ is a $2m$-point rational Gauss quadrature rule; see [22]. The remainder term for this rule is given by

$$\mathcal{I}(f) - \mathcal{G}_{2m}^{\sigma}(f) = \frac{d^{4m}}{dx^{4m}}(fw)_{x=\widetilde{\theta}} \frac{1}{(4m)!} \int \prod_{j=1}^{2m} (x - \alpha_j)^2 w(x) d\alpha(x),$$

where α_j is the jth eigenvalues of $\mathbb{H}_{2m}$. The proof is completed since the matrices $\mathbb{H}_{2m}$ and $\mathbb{T}_{2m}$ are similar. More details are provided in the proof of [27, Corollary 5.5], which has to be applied to the matrix $A - \sigma I_n$ with the inner product (1.6). ◻

2.4. Shifted Gauss-Laurent-Radau quadrature rules. A $(2m+1)$-point shifted Gauss-Laurent-Radau quadrature rule is obtained by assigning one of the quadrature nodes, denoted by ξ, and determining the remaining $2m$ quadrature nodes and the $2m+1$ weights so that the resulting quadrature rule is exact for all shifted Laurent polynomials of as high an order as possible. Application of m steps of the shifted extended global Lanczos process to the matrix A with initial block vector $V \in \mathbb{R}^{n\times s}$ determines the matrix $\widetilde{\mathbb{T}}_{2m} \in \mathbb{R}^{(2m+1)\times 2m}$ defined by (2.4). The matrix $\mathbb{T}_{2m}$ is a leading principal submatrix. We introduce the $(2m+1)$-point shifted Gauss-Laurent-Radau rule

$$\mathcal{R}^{\sigma,\xi}_{2m+1}(f) = \|V\|^2 e_1^T f(\mathbb{T}_{2m+1,\xi}) e_1,$$

where

$$\mathbb{T}_{2m+1,\xi} := \begin{bmatrix} \mathbb{T}_{2m} & \tau \\ \tau^T & \widetilde{\alpha}_\xi \end{bmatrix} \in \mathbb{R}^{(2m+1)\times(2m+1)},$$

and $\tau \in \mathbb{R}^{2m}$ contains the first $2m$ entries of the last column of $\mathbb{T}_{2m+1}$. The entry $\widetilde{\alpha}_\xi$ is determined so that the matrix $\mathbb{T}_{2m+1,\xi}$ has an eigenvalue at ξ, where ξ is a chosen node outside the convex hull of the spectrum of $\mathbb{T}_{2m}$. The parameter $\widetilde{\alpha}_\xi$ is computed similarly as described in [24, p. 561]. Thus, we solve the equation

$$\mathbb{T}_{2m+1,\xi} \begin{bmatrix} x \\ -1 \end{bmatrix} = \xi \begin{bmatrix} x \\ -1 \end{bmatrix}, \quad x \in \mathbb{R}^{2m},$$

which can be written as

$$\begin{cases} \mathbb{T}_{2m} x - \tau = \xi x, \\ \tau^T x - \widetilde{\alpha}_\xi = -\xi, \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = (\mathbb{T}_{2m} - \xi I_{2m})^{-1}\tau, \\ \widetilde{\alpha}_\xi = \xi + \tau^T x. \end{cases}$$

Therefore, $\widetilde{\alpha}_\xi$ can be expressed in terms of ξ as follows

$$\widetilde{\alpha}_\xi = \xi + \tau^T (\mathbb{T}_{2m} - \xi I_{2m})^{-1}\tau.$$

THEOREM 2.8. *The $(2m+1)$-point shifted Gauss-Laurent-Radau rule associated with the measure $d\alpha$ satisfies*

$$\mathcal{R}^{\sigma,\xi}_{2m+1}(f) = \mathcal{I}(f), \quad \forall f \in \Delta_{-2m,2m}.$$

Furthermore, if the function f is $4m+1$ times continuously differentiable in the convex hull of the spectrum of A, then the remainder term for this rule is given by

$$\mathcal{I}(f) - \mathcal{R}^{\sigma,\xi}_{2m+1}(f) = \frac{d^{4m+1}}{dx^{4m+1}}(fw)_{x=\widetilde{\theta}_\xi} \frac{1}{(4m+1)!} \int (x-\xi) \prod_{j=1}^{2m} (x-\theta_{\xi,j})^2 w(x) d\alpha(x), \tag{2.9}$$

where $\widetilde{\theta}_\xi$ lies in the largest open interval contained in the convex hull of the spectrum of A and ξ. The scalars $\theta_{\xi,1}, \theta_{\xi,2}, \ldots, \theta_{\xi,2m}, \xi$ denote the eigenvalues of $\mathbb{T}_{2m+1,\xi}$, and the function w is defined in Theorem 2.7.

Proof. The proof is analogous to the proof of Theorem 2.7 and [27, Theorem 7.1]. The latter proof has to be applied to the matrix $M = A - \sigma I_n$ with the inner product (1.6). ◻

The remainder terms for the shifted Gauss-Laurent rule (2.8) and the shifted Gauss-Laurent-Radau rule (2.9) allow us to bound the expression $\mathcal{I}(f)$ in (1.2) from above and below if the derivatives $(f\cdot w)^{(4m)}$

and $(f\cdot w)^{(4m+1)}$ are of constant sign in the convex hull of the spectrum of A and ξ. For instance, when the derivatives $(f\cdot w)^{(4m)}$ and $(f\cdot w)^{(4m+1)}$ are nonnegative on this interval and ξ is larger than or equal to the largest eigenvalue of A, the expression $\mathcal{I}(f)$ can be bounded according to

$$\mathcal{G}_{2m}^{\sigma}(f) \leq \mathcal{I}(f) \leq \mathcal{R}_{2m+1}^{\sigma,\xi}(f).$$

However, when at least one of the derivatives $(f\cdot w)^{(4m)}$ or $(f\cdot w)^{(4m+1)}$ changes sign on the convex hull of the spectrum of A and ξ, pairs of shifted Gauss-Laurent and shifted Gauss-Laurent-Radau rules are not guaranteed to produce upper and lower bounds for $\mathcal{I}(f)$.

3. Shifted anti-Gauss-Laurent quadrature rules. Laurie [30] introduced anti-Gauss rules for the estimation of the quadrature error for Gauss rules applied to the integration of real-valued functions on a real interval. Extensions to (standard) block Krylov subspace methods are described in [1, 2, 21]. This section introduces the $(2m+1)$-point shifted anti-Gauss-Laurent quadrature rule, denoted by $\mathcal{A}_{2m+1}^{\sigma}$, associated with the shifted Gauss-Laurent rule (2.8). It is characterized by

$$(\mathcal{I} - \mathcal{A}_{2m+1}^{\sigma})(f) = -\mathcal{E}_{2m}(f), \quad \forall f \in \Delta_{-2m,2m+1},$$

where $\mathcal{E}_{2m}$ is the error of the shifted Gauss-Laurent quadrature rule defined in (2.8). This is equivalent to

$$\mathcal{A}_{2m+1}^{\sigma}(f) = (2\mathcal{I} - \mathcal{G}_{2m}^{\sigma})(f), \quad \forall f \in \Delta_{-2m,2m+1}.$$

Therefore, $\mathcal{A}_{2m+1}^{\sigma}(f)$ may be considered a $(2m+1)$-point shifted Gauss-Laurent quadrature rule with respect to the bilinear form $\langle \cdot, \cdot \rangle_A$ determined by the functional $(2\mathcal{I} - \mathcal{G}_{2m}^{\sigma})(f)$ and given by

$$\langle P, Q \rangle_A := (2\mathcal{I} - \mathcal{G}_{2m}^{\sigma})(PQ) = 2\langle P, Q\rangle - \|V\|^2 e_1^T P(\mathbb{T}_{2m}) Q(\mathbb{T}_{2m}) e_1. \tag{3.1}$$

Let $\widetilde{R}_0, \widetilde{R}_1, \ldots, \widetilde{R}_{2m}$ be the first $2m+1$ shifted orthonormal Laurent polynomials with respect to the bilinear form (3.1). They satisfy a pair of five-term recurrence relations of the form

$$\begin{aligned}
\widetilde{h}_{2j+1,2j-1}\widetilde{R}_{2j}(x) &= x\widetilde{R}_{2j-2}(x) - \sum_{i=2j-3}^{2j} \widetilde{h}_{i,2j-1}\widetilde{R}_{i-1}(x),\\
\widetilde{h}_{2j+2,2j}\widetilde{R}_{2j+1}(x) &= (x-\sigma)^{-1}\widetilde{R}_{2j-1}(x) - \sum_{i=2j-2}^{2j+1} \widetilde{h}_{i,2j}\widetilde{R}_{i-1}(x)
\end{aligned}$$

for $j = 1, 2, \ldots$, where $\widetilde{R}_{-2} = \widetilde{R}_{-1} = 0$ and

$$\widetilde{R}_0(x) = 1/\widetilde{\alpha}_{1,1}, \quad \widetilde{R}_1(x) = (1/\widetilde{\alpha}_{2,2})[(x-\sigma)^{-1}\widetilde{R}_0(x) - \widetilde{\alpha}_{1,2}\widetilde{R}_0(x)].$$

Furthermore,

$$\widetilde{h}_{i,2j-1} = \langle x\widetilde{R}_{2j-2}, \widetilde{R}_{i-1}\rangle_A, \quad \widetilde{h}_{i,2j} = \langle (x-\sigma)^{-1}\widetilde{R}_{2j-1}, \widetilde{R}_{i-1}\rangle_A. \tag{3.2}$$

The coefficients $\widetilde{h}_{2j+1,2j-1}$ and $\widetilde{h}_{2j+2,2j}$ are determined so that $\langle \widetilde{R}_{2j}, \widetilde{R}_{2j}\rangle_A = 1$ and $\langle \widetilde{R}_{2j+1}, \widetilde{R}_{2j+1}\rangle_A = 1$. Due to Theorem 2.7, shifted anti-Gauss-Laurent quadrature rules yield the same result as shifted Gauss-Laurent quadrature rules for all shifted Laurent polynomials in $\Delta_{-2m,2m-1}$, i.e.,

$$\text{if } P \text{ and } Q \text{ are shifted Laurent polynomials such that } PQ \in \Delta_{-2m,2m-1}, \text{ then } \langle P, Q\rangle_A = \langle P, Q\rangle. \tag{3.3}$$

Using this property in (3.2) gives

$$\begin{aligned}
\widetilde{h}_{i,j} &= h_{i,j}, \quad i,j = 1,2,\ldots,2m-1,\\
\widetilde{h}_{2m,i} &= h_{2m,i}, \quad i \in \{2m-2, 2m-1\},\\
\widetilde{h}_{i,2m} &= h_{i,2m}, \quad i \in \{2m-2, 2m-1\}.
\end{aligned}$$

This shows that $\widetilde{R}_j(x) = R_j(x)$ for $j = 0,1,\ldots,2m-1$. In addition, we have

$$\widetilde{h}_{2m+1,2m-1}\widetilde{R}_{2m}(x) = h_{2m+1,2m-1}R_{2m}(x). \tag{3.4}$$

Using the properties $\langle \widetilde{R}_{2m}, \widetilde{R}_{2m}\rangle_A = \langle R_{2m}, R_{2m}\rangle = 1$ and (2.7) gives

$$\widetilde{h}^2_{2m+1,2m-1} = 2h^2_{2m+1,2m-1}.$$

Thus, we can choose $\widetilde{h}_{2m+1,2m-1} = \sqrt{2}h_{2m+1,2m-1}$. Substituting this expression into (3.4) yields

$$\widetilde{R}_{2m}(x) = \frac{1}{\sqrt{2}}R_{2m}(x).$$

We turn to the determination of the entries of the symmetric pentadiagonal matrix $\mathbb{T}^a_{2m+1} = [\widetilde{t}_{i,j}] \in \mathbb{R}^{(2m+1)\times(2m+1)}$ associated with the $(2m+1)$-point shifted anti-Gauss-Laurent rule,

$$\mathcal{A}^{\sigma}_{2m+1}(f) = \|V\|^2 e_1^T f(\mathbb{T}^a_{2m+1})e_1, \tag{3.5}$$

where $\widetilde{t}_{i,j} = \langle x\widetilde{R}_{i-1}, \widetilde{R}_{j-1}\rangle_A$ for $i,j = 1,2,\ldots,2m+1$. Recall that $t_{i,j} = \langle xR_{i-1}, R_{j-1}\rangle$. Using (3.3), we find

$$\widetilde{t}_{i,j} = t_{i,j}, \text{ for } i,j = 1,2,\ldots,2m.$$

We obtain from (2.7) that

$$\mathcal{G}^{\sigma}_{2m}(R_{2m}R_{2m-2}) = \mathcal{G}^{\sigma}_{2m}(R_{2m}R_{2m-1}) = \mathcal{G}^{\sigma}_{2m}(R_{2m}R_{2m}) = 0.$$

Therefore,

$$\begin{aligned}
\widetilde{t}_{2m+1,2m-1} = \langle x\widetilde{R}_{2m}, \widetilde{R}_{2m-2}\rangle_A &= \frac{1}{\sqrt{2}}\langle xR_{2m}, R_{2m-2}\rangle_A\\
&= \frac{1}{\sqrt{2}}[2\langle xR_{2m}, R_{2m-2}\rangle - \mathcal{G}^{\sigma}_{2m}(R_{2m}R_{2m-2})]\\
&= \sqrt{2}\, t_{2m+1,2m-1}.
\end{aligned}$$

In the same manner, we get the remaining entries of $\mathbb{T}^a_{2m+1}$,

$$\widetilde{t}_{2m+1,2m} = \sqrt{2}t_{2m+1,2m} \quad \text{and} \quad \widetilde{t}_{2m+1,2m+1} = t_{2m+1,2m+1}.$$

In conclusion, the symmetric pentadiagonal matrix $\mathbb{T}^a_{2m+1}$ associated with the $(2m+1)$-point shifted anti-Gauss-Laurent rule (3.5) can be obtained from the matrix $\mathbb{T}_{2m+1} = \mathbb{V}^T_{2m+1} \diamond A\mathbb{V}_{2m+1}$ associated with the $(2m+1)$-point shifted Gauss-Laurent rule by multiplying the entries $t_{2m+1,2m-1}$ and $t_{2m+1,2m}$ by $\sqrt{2}$, i.e.,

$$\mathbb{T}^a_{2m+1} = \begin{bmatrix} \mathbb{T}_{2m} & \Psi_{2m} \\ \Psi^T_{2m} & t_{2m+1,2m+1} \end{bmatrix},$$

where $\Psi_{2m} = [0,\ldots,0,\sqrt{2}t_{2m+1,2m-1}, \sqrt{2}t_{2m+1,2m}]^T \in \mathbb{R}^{2m}$. Algorithm 1 describes how an approximation of (1.3) and an error estimate can be computed by a pair of shifted Gauss-Laurent and anti-Gauss-Laurent quadrature rules when the matrix A is symmetric. In the spirit of Laurie [30], we approximate (1.3) by $U_{\text{app}}(f) = (\mathcal{G}^{\sigma}_{2m}(f) + \mathcal{A}^{\sigma}_{2m+1}(f))/2$ and estimate the error in $U_{\text{app}}(f)$ by the difference $|\mathcal{G}^{\sigma}_{2m}(f) - \mathcal{A}^{\sigma}_{2m+1}(f)|/|\mathcal{G}^{\sigma}_{2m}(f)|$.

Algorithm 1 Approximation of trace$(V^T f(A)V)$ by pairs of shifted Gauss-Laurent and shifted anti-Gauss Laurent quadrature rules for a symmetric matrix A.

Input: Symmetric matrix A, initial block V, parameter σ, and function f.

1. Choose tolerance $\epsilon > 0$ and the maximum number of iterations I_{max}.
2. $\alpha_{1,1} = \|V\|$; $V_1 = V/\alpha_{1,1}$;
3. $\alpha_{1,2} = \langle (A - \sigma I_n)^{-1} V_1, V_1 \rangle$; $\widetilde{V}_2 = (A - \sigma I_n)^{-1} V_1 - \alpha_{1,2} V_1$;
4. $\alpha_{2,2} = \|\widetilde{V}_2\|$; $V_2 = \widetilde{V}_2/\alpha_{2,2}$; $\widetilde{V}_3 = AV_1$; $h_{1,1} = \langle \widetilde{V}_3, V_1 \rangle$;
5. for $j = 1 : I_{max}$
   (a) $\widetilde{V}_{2j+2} = (A - \sigma I_n)^{-1} V_{2j}$; $h_{2j,2j} = \langle \widetilde{V}_{2j+2}, V_{2j} \rangle$;
   (b) Compute $h_{i,2j-1}$ and $h_{i,2j}$ from recursion relations given by Proposition 2.3.
   (c) $\widetilde{V}_{2j+1} = \widetilde{V}_{2j+1} - \sum\limits_{i=2j-3}^{2j} h_{i,2j-1} V_i$; $h_{2j+1,2j-1} = \|\widetilde{V}_{2j+1}\|$;
   (d) if $j = 1$
       $t_{1:2,1} = h_{1:2,1}$; $t_{1,2} = t_{2,1}$; $t_{2,2} = \sigma - \alpha_{1,2} t_{2,1}/\alpha_{2,2}$;
   else
       $t_{2j-3:2j,2j-1} = h_{2j-3:2j,2j-1}$; $t_{2j-1,2j} = t_{2j,2j-1}$;
       $t_{2j,2j} = \sigma - h_{2j-1,2j-2} t_{2j,2j-1}/h_{2j,2j-2}$;
   end
   (e) $G^{\sigma}_{2j}(f) = e_1^T f(\mathbb{T}_{2j}) e_1$;
   (f) $V_{2j+1} = \widetilde{V}_{2j+1}/h_{2j+1,2j-1}$;
   (g) $\widetilde{V}_{2j+2} = \widetilde{V}_{2j+2} - \sum\limits_{i=2j-2}^{2j+1} h_{i,2j} V_i$; $h_{2j+2,2j} = \|\widetilde{V}_{2j+2}\|$;
   (h) if $j = 1$, $\quad t_{3,1} = h_{3,1}$; $t_{3,2} = -\alpha_{1,2} t_{3,1}/\alpha_{2,2}$;
   else
       $t_{2j+1,2j-1} = h_{2j+1,2j-1}$; $t_{2j+1,2j} = -h_{2j-1,2j-2} t_{2j+1,2j-1}/h_{2j,2j-2}$;
   end
   (i) $\Psi_{2j} = \sqrt{2}\,[0, \ldots, 0, t_{2j+1,2j-1}, t_{2j+1,2j}]^T$;
   (j) $\widetilde{V}_{2j+3} = AV_{2j+1}$; $h_{2j+1,2j+1} = \langle V_{2j+1}, \widetilde{V}_{2j+3} \rangle$;
   (k) Compute $\mathbb{T}^a_{2j+1} = \begin{bmatrix} \mathcal{T}_{2j} & \Psi_{2j} \\ \Psi^T_{2j} & h_{2j+1,2j+1} \end{bmatrix}$ and $A^{\sigma}_{2j+1}(f) = e_1^T f(\mathbb{T}^a_{2j+1}) e_1$;
   (l) if $|G^{\sigma}_{2j}(f) - A^{\sigma}_{2j+1}(f)|/|G^{\sigma}_{2j}(f)| < \epsilon$
       $U_{\text{app}}(f) = \alpha^2_{1,1}[G^{\sigma}_{2j}(f) + A^{\sigma}_{2j+1}(f)]/2$; Break;
   end
   (m) $V_{2j+2} = \widetilde{V}_{2j+2}/h_{2j+2,2j}$;
   (n) end

Output: Approximation $U_{\text{app}}(f)$ of trace$(V^T f(A)V)$.

4. The shifted extended global nonsymmetric Lanczos process. This section describes the recursion relations for the shifted extended global nonsymmetric Lanczos process. This process generates two biorthogonal bases of block vectors $\{V_j\}_{j=1}^{2m}$ and $\{W_j\}_{j=1}^{2m}$ for the shifted extended global Krylov subspaces $\mathbb{K}^{\sigma}_m(A, V)$ and $\mathbb{K}^{\sigma}_m(A^T, W)$. These bases can be computed with short recurrence formulas. We have

$$\begin{array}{rclrcl} \alpha_{1,1} &=& |\langle W, V \rangle|^{1/2}, \quad V_1 = V/\alpha_{1,1}, & \beta_{1,1} &=& \langle W, V \rangle/\alpha_{1,1}, \quad W_1 = W/\beta_{1,1}, \\ \widehat{V}_2 &=& (A - \sigma I_n)^{-1} V_1 - \alpha_{1,2} V_1, & \widehat{W}_2 &=& (A^T - \sigma I_n)^{-1} W_1 - \beta_{1,2} W_1, \\ \alpha_{1,2} &=& \langle W_1, (A - \sigma I_n)^{-1} V_1 \rangle, & \beta_{1,2} &=& \langle V_1, (A^T - \sigma I_n)^{-1} W_1 \rangle, \\ \alpha_{2,2} &=& |\langle \widehat{W}_2, \widehat{V}_2 \rangle|^{1/2}, \quad V_2 = \widehat{V}_2/\alpha_{2,2}, & \beta_{2,2} &=& \langle \widehat{W}_2, \widehat{V}_2 \rangle/\alpha_{2,2}, \quad W_2 = \widehat{W}_2/\beta_{2,2}, \end{array} \tag{4.1}$$

and for $j = 1, 2, \ldots, m$,

$$\begin{array}{rcl} \widehat{V}_{2j+1} & = & h_{2j+1,2j-1}V_{2j+1} = AV_{2j-1} - \displaystyle\sum_{i=2j-3}^{2j} h_{i,2j-1}V_i, \\ \widehat{W}_{2j+1} & = & g_{2j+1,2j-1}W_{2j+1} = A^T W_{2j-1} - \displaystyle\sum_{i=2j-3}^{2j} g_{i,2j-1}W_i, \\ \widehat{V}_{2j+2} & = & h_{2j+2,2j}V_{2j+2} = (A - \sigma I_n)^{-1}V_{2j} - \displaystyle\sum_{i=2j-2}^{2j+1} h_{i,2j}V_i, \\ \widehat{W}_{2j+2} & = & g_{2j+2,2j}W_{2j+2} = (A^T - \sigma I_n)^{-1}W_{2j} - \displaystyle\sum_{i=2j-2}^{2j+1} g_{i,2j}W_i, \end{array} \tag{4.2}$$

where

$$\begin{array}{rclrcl} h_{i,2j-1} & = & \langle AV_{2j-1}, W_i \rangle, & g_{i,2j-1} & = & \langle A^T W_{2j-1}, V_i \rangle, \\ h_{i,2j} & = & \langle (A - \sigma I_n)^{-1} V_{2j}, W_i \rangle, & g_{i,2j} & = & \langle (A^T - \sigma I_n)^{-1} W_{2j}, V_i \rangle, \end{array}$$

and

$$\begin{array}{rclrcll} h_{2j+1,2j-1} & = & |\alpha_{2j+1}|^{1/2}, & g_{2j+1,2j-1} & = & \alpha_{2j+1}/h_{2j+1,2j-1}, & \text{with } \ \alpha_{2j+1} = \langle \widehat{W}_{2j+1}, \widehat{W}_{2j+1} \rangle, \\ h_{2j+2,2j} & = & |\alpha_{2j+2}|^{1/2}, & g_{2j+2,2j} & = & \alpha_{2j+2}/h_{2j+2,2j}, & \text{with } \ \alpha_{2j+2} = \langle \widehat{V}_{2j+2}, \widehat{W}_{2j+2} \rangle. \end{array}$$

Similarly as for the shifted extended global symmetric Lanczos process, the coefficients $h_{i,j}$ and $g_{i,j}$ can be computed recursively.

Proposition 4.1. *The coefficients $h_{i,j}$, $g_{i,j}, \alpha_{i,j}$, and $\beta_{i,j}$, defined by the relations* (4.1) *and* (4.2), *can be computed recursively as follows:*

$$\begin{aligned} h_{1,2} &= \alpha_{2,2}, \\ h_{2,1} &= \frac{1}{\beta_{2,2}}[1 - \beta_{1,2}h_{1,1} + \sigma\beta_{1,2}], \\ h_{2,3} &= -\frac{\beta_{1,2}h_{1,3}}{\beta_{2,2}}. \end{aligned}$$

For $j = 2, 3, \ldots, m$, we have

$$\begin{aligned} h_{2j-3,2j-1} &= g_{2j-1,2j-3}, \\ h_{2j-2,2j-1} &= \frac{-1}{g_{2j-2,2j-4}} g_{2j-3,2j-4} h_{2j-3,2j-1}, \\ h_{2j,2j-1} &= \frac{1}{g_{2j,2j-2}}\left[\sigma g_{2j-1,2j-2} - \sum_{i=2j-3}^{2j-1} g_{i,2j-2}h_{i,2j-1}\right], \\ h_{2j-2,2j} &= g_{2j,2j-2}, \\ h_{2j-1,2j} &= \frac{-1}{g_{2j-1,2j-3}} g_{2j-2,2j-3} h_{2j-2,2j}, \\ h_{2j+1,2j} &= \frac{1}{g_{2j+1,2j-1}}\left[\sigma h_{2j-1,2j} - \sum_{i=2j-2}^{2j} g_{i,2j-1}h_{i,2j}\right], \\ g_{2j-1,2j-1} &= h_{2j-1,2j-1}, \\ g_{2j,2j} &= h_{2j,2j}. \end{aligned}$$

These relations also hold when the $h_{i,j}$ and $\alpha_{i,j}$ are replaced by $g_{i,j}$ and $\beta_{i,j}$, respectively, and vice versa.

Proof. The relations can be shown similarly as those in Proposition 2.3. ∎

Introduce the pentadiagonal matrix

$$\widehat{\mathbb{T}}_{2m} = [t_{i,j}] = \widehat{\mathbb{W}}_{2m}^T \diamond A\widehat{\mathbb{V}}_{2m} \in \mathbb{R}^{2m\times 2m}. \tag{4.3}$$

Here

$$t_{i,j} = \langle AV_j, W_i\rangle,\ i,j = 1,2,\ldots,2m,$$

and the matrices $\widehat{\mathbb{V}}_{2m} = [V_1, V_2, \ldots, V_{2m}]$ and $\widehat{\mathbb{W}}_{2m} = [W_1, W_2, \ldots, W_{2m}]$ are defined by the recursion relations for the shifted extended global nonsymmetric Lanczos process,

$$\begin{array}{rcl} A\widehat{\mathbb{V}}_{2m} & = & \widehat{\mathbb{V}}_{2m}(\mathbb{T}_{2m}\otimes I_s) + V_{2m+1}(\tau_m E_m^T \otimes I_s), \\ A^T\widehat{\mathbb{W}}_{2m} & = & \widehat{\mathbb{W}}_{2m}(\widehat{\mathbb{T}}_{2m}^T\otimes I_s) + W_{2m+1}(\widehat{\tau}_m E_m^T \otimes I_s), \end{array} \tag{4.4}$$

with $E_m = [e_{2m-1}, e_{2m}] \in \mathbb{R}^{2m\times 2}$ and

$$\begin{array}{rcl} \widehat{\tau}_m & = & [\widehat{t}_{2m+1,2m-1}, \widehat{t}_{2m+1,2m}] := [\langle A^T W_{2m-1}, V_{2m+1}\rangle, \langle A^T W_{2m}, V_{2m+1}\rangle], \\ \tau_m & = & [t_{2m+1,2m-1}, t_{2m+1,2m}]. \end{array}$$

The entries of $\widehat{\mathbb{T}}_{2m}$, τ_m, and $\widehat{\tau}_m$ are computed recursively as shown below.

Proposition 4.2. *Let the coefficients $h_{i,j}$, $g_{i,j}$, $\alpha_{i,j}$, and $\beta_{i,j}$ be defined by* (4.1) *and* (4.2). *The matrix $\widehat{\mathbb{T}}_{2m} = [t_{i,j}]$ in* (4.3) *and the coefficients τ_m and $\widehat{\tau}_m$ in* (4.4) *can be computed as follows:*

$$\begin{aligned} t_{i,2j-1} &= h_{i,2j-1}, \quad \text{for } i \in \{2j-3, 2j-2, \ldots, 2j+1\},\ j = 1,2,\ldots,m, \\ t_{1,2} &= \frac{1}{\alpha_{2,2}}\Big[1 - \alpha_{1,2}(t_{1,1} - \sigma)\Big], \\ t_{2,2} &= \sigma - \frac{\alpha_{1,2}}{\alpha_{2,2}} t_{2,1}, \\ t_{3,2} &= -\frac{\alpha_{1,2}}{\alpha_{2,2}} t_{3,1}, \\ \widehat{t}_{3,1} &= g_{3,1}, \\ \widehat{t}_{3,2} &= -\frac{\beta_{1,2}}{\beta_{2,2}} \widehat{t}_{3,1}, \end{aligned}$$

For $j = 1,2,\ldots,m-1$, we have

$$\begin{aligned} t_{2j+1,2j+2} &= \frac{-1}{h_{2j+2,2j}}\left[\sum_{i=2j-1}^{2j+1} h_{i,2j} t_{2j+1,i} - \sigma h_{2j+1,2j}\right], \\ t_{2j+2,2j+2} &= \sigma - \frac{h_{2j+1,2j}}{h_{2j+2,2j}} t_{2j+2,2j+1}, \\ t_{2j+3,2j+2} &= -\frac{h_{2j+1,2j}}{h_{2j+2,2j}} t_{2j+3,2j+1}, \\ \widehat{t}_{2j+3,2j+1} &= g_{2j+3,2j+1}, \\ \widehat{t}_{2j+3,2j+1} &= -\frac{g_{2j+1,2j}}{g_{2j+2,2j}} \widehat{t}_{2j+3,2j+1}. \end{aligned}$$

Proof. The recursions can be shown in a similar manner as [9, Proposition 3.1]. However, some adjustments have to be made, because here we apply a nonsymmetric Lanczos process instead of a symmetric Lanczos process, and the matrix is $A - \sigma I_n$ instead of A. □

There are two sequences of shifted Laurent polynomials $p_0, p_1, \ldots, p_m$ and $q_0, q_1, \ldots, q_m$ that are biorthogonal with respect to the bilinear form

$$\langle P, Q\rangle = \text{trace}(P(A)V, Q(A)W) = \int P(\lambda)Q(\lambda)d\mu(\lambda),$$

where $d\mu$ is the measure in (1.12). These shifted Laurent polynomials satisfy a pair of five-term recurrence relations of the form

$$\begin{aligned}
h_{2j+1,2j-1}p_{2j}(x) &= xp_{2j-2}(x) - \sum_{i=2j-3}^{2j} h_{i,2j-1}p_{i-1}(x),\\
g_{2j+1,2j-1}q_{2j}(x) &= xq_{2j-2}(x) - \sum_{i=2j-3}^{2j} g_{i,2j-1}q_{i-1}(x),\\
h_{2j+2,2j}p_{2j+1}(x) &= (x-\sigma)^{-1}p_{2j-1}(x) - \sum_{i=2j-2}^{2j+1} h_{i,2j}p_{i-1}(x),\\
g_{2j+2,2j}q_{2j+1}(x) &= (x-\sigma)^{-1}q_{2j-1}(x) - \sum_{i=2j-2}^{2j+1} g_{i,2j}q_{i-1}(x),
\end{aligned} \tag{4.5}$$

where $p_1(x) = (1/\alpha_{2,2})[(x-\sigma)^{-1}p_0(x) - \alpha_{1,2}p_0(x)]$, $p_0(x) = 1/\alpha_{1,1}$, and
$q_1(x) = (1/\beta_{2,2})[(x-\sigma)^{-1}q_0(x) - \beta_{1,2}q_0(x)]$, $q_0(x) = 1/\beta_{1,1}$, $p_{-2} = p_{-1} = q_{-2} = q_{-1} = 0$.

The associated $2m$-point shifted Gauss-Laurent quadrature rule is given by

$$\mathcal{G}_{2m}^{\sigma}(f) = \langle V, W\rangle e_1^T f(\widehat{\mathbb{T}}_{2m})e_1 = \sum_{i=1}^{2m} f(\mu_i)w_i, \tag{4.6}$$

where μ_i denotes the ith eigenvalue of $\widehat{\mathbb{T}}_{2m}$ and $w_i = \langle V, W\rangle u_{i,1}^2$. Here $u_{i,1}$ is the first component of the normalized eigenvector u_i of $\widehat{\mathbb{T}}_{2m}$. This quadrature rule satisfies

$$\mathcal{G}_{2m}^{\sigma}(f) = \mathcal{I}(f), \quad \forall f \in \Delta_{-2m,2m-1}.$$

This can be shown similarly as related results in [27] or in Section 3.

Lemma 4.3. *Let the shifted Laurent polynomials p_{2m} and q_{2m} be determined by the recursion relations* (4.5). *Then*

$$\mathcal{G}_{2m}^{\sigma}(p_{2m}) = \mathcal{G}_{2m}^{\sigma}(q_{2m}) = 0.$$

Proof. Consider the vectors of shifted Laurent polynomials

$$\vec{P}_{2m}(x) = [p_0(x), p_1(x), \ldots, p_{2m-1}(x)], \quad \vec{Q}_{2m}(x) = [q_0(x), q_1(x), \ldots, q_{2m-1}(x)].$$

Using (4.4), these vectors can be expressed as

$$\begin{aligned}
x\vec{P}_{2m}(x) &= \vec{P}_{2m}(x)\widehat{\mathbb{T}}_{2m} + p_{2m}(x)\tau_m E_m^T,\\
x\vec{Q}_{2m}(x) &= \vec{Q}_{2m}(x)\widehat{\mathbb{T}}_{2m}^T + q_{2m}(x)\widehat{\tau}_m E_m^T.
\end{aligned}$$

It follows that the $2m$ zeros of the Laurent polynomials p_{2m} and q_{2m} are the eigenvalues of the matrix $\widehat{\mathbb{T}}_{2m}$. On the other hand, we have in view of (4.6) that $\mathcal{G}^{\sigma}_{2m}(p_{2m}) = \mathcal{G}^{\sigma}_{2m}(q_{2m}) = 0$. This completes the proof. □

5. Shifted anti-Gauss-Laurent quadrature rules for nonsymmetric matrices. This section extends the construction of the shifted anti-Gauss-Laurent quadrature rules of Section 3 to the situation when the matrix A is nonsymmetric. Introduce the $(2m+1)$-point shifted anti-Gauss-Laurent quadrature rule

$$\mathcal{A}^{\sigma}_{2m+1}(f) = \langle W, V\rangle \sum_{i=1}^{2m+1} e_1^T f(\widehat{\mathbb{T}}^a_{2m+1}) e_1,$$

where the matrix $\widehat{\mathbb{T}}^a_{2m+1} = [\widetilde{t}_{i,j}] \in \mathbb{R}^{(2m+1)\times(2m+1)}$ is such that

$$\mathcal{A}^{\sigma}_{2m+1}(f) = (2\mathcal{I} - \mathcal{G}^{\sigma}_{2m})(f), \quad \forall f \in \Delta_{-2m,2m+1}.$$

The entries of $\widehat{\mathbb{T}}^a_{2m+1}$ are given by $\widetilde{t}_{i,j} = \langle x\widetilde{p}_{j-1}, \widetilde{q}_{i-1}\rangle_a$, where

$$\langle P, Q\rangle_a := (2\mathcal{I} - \mathcal{G}^{\sigma}_{2m})(PQ) = 2\langle P, Q\rangle - \langle W, V\rangle e_1^T P(\widehat{\mathbb{T}}_{2m}) Q(\widehat{\mathbb{T}}_{2m}) e_1. \tag{5.1}$$

Let $\widetilde{p}_0, \widetilde{p}_1, \ldots, \widetilde{p}_{2m}$ and $\widetilde{q}_0, \widetilde{q}_1, \ldots, \widetilde{q}_{2m}$ be the first biorthogonal shifted Laurent polynomials with respect to the bilinear form (5.1). These shifted Laurent polynomials satisfy a pair of five-term recurrence relations of the form

$$\begin{aligned}
&\widetilde{p}_0(x) = 1/\widetilde{\alpha}_{1,1}, \quad \widetilde{p}_1(x) = (1/\widetilde{\alpha}_{2,2})[(x-\sigma)^{-1}\widetilde{p}_0(x) - \widetilde{\alpha}_{1,2}\widetilde{p}_0(x)],\\
&\widetilde{q}_0(x) = 1/\widetilde{\beta}_{1,1}, \quad \widetilde{q}_1(x) = (1/\widetilde{\beta}_{2,2})[(x-\sigma)^{-1}\widetilde{q}_0(x) - \widetilde{\beta}_{1,2}\widetilde{q}_0(x)],
\end{aligned}$$

and

$$\begin{aligned}
\widetilde{h}_{2j+1,2j-1}\widetilde{p}_{2j}(x) &= x\widetilde{p}_{2j-2}(x) - \sum_{i=2j-3}^{2j} \widetilde{h}_{i,2j-1}\widetilde{p}_{i-1}(x),\\
\widetilde{h}_{2j+2,2j}\widetilde{p}_{2j+1}(x) &= (x-\sigma)^{-1}\widetilde{p}_{2j-1}(x) - \sum_{i=2j-2}^{2j+1} \widetilde{h}_{i,2j}\widetilde{p}_{i-1}(x),\\
\widetilde{g}_{2j+1,2j-1}\widetilde{q}_{2j}(x) &= x\widetilde{q}_{2j-2}(x) - \sum_{i=2j-3}^{2j} \widetilde{g}_{i,2j-1}\widetilde{q}_{i-1}(x),\\
\widetilde{g}_{2j+2,2j}\widetilde{q}_{2j+1}(x) &= (x-\sigma)^{-1}\widetilde{q}_{2j-1}(x) - \sum_{i=2j-2}^{2j+1} \widetilde{g}_{i,2j}\widetilde{q}_{i-1}(x),
\end{aligned}$$

where $\widetilde{p}_{-2} = \widetilde{p}_{-1} = \widetilde{q}_{-2} = \widetilde{q}_{-1} = 0$ and

$$\begin{aligned}
&\widetilde{h}_{i,2j-1} = \langle x\widetilde{q}_{2j-2}, \widetilde{q}_{i-1}\rangle_a, \quad \widetilde{h}_{i,2j} = \langle (x-\sigma)^{-1}\widetilde{p}_{2j-1}, \widetilde{q}_{i-1}\rangle_a,\\
&\widetilde{g}_{i,2j-1} = \langle x\widetilde{p}_{2j-2}, \widetilde{q}_{i-1}\rangle_a, \quad \widetilde{g}_{i,2j} = \langle (x-\sigma)^{-1}\widetilde{p}_{2j-1}, \widetilde{q}_{i-1}\rangle_a.
\end{aligned} \tag{5.2}$$

The coefficients $\widetilde{h}_{2j+1,2j-1}$, $\widetilde{h}_{2j+2,2j}$, $\widetilde{g}_{2j+1,2j-1}$, and $\widetilde{g}_{2j+2,2j}$ are determined so that

$$\langle \widetilde{p}_{2j}, \widetilde{q}_{2j}\rangle_a = 1 \text{ and } \langle \widetilde{p}_{2j+1}, \widetilde{q}_{2j+1}\rangle_a = 1.$$

Using the fact that for two shifted Laurent polynomials P and Q such that $PQ \in \Delta_{-2m,2m-1}$, we have $\langle P, Q\rangle_a = \langle P, Q\rangle$, and applying the relations (5.2), we obtain

$$\begin{aligned}
&\widetilde{p}_j = p_j, \quad \widetilde{q}_j = q_j, \qquad \text{for } j = 0, 1, \ldots, 2m-1,\\
&\widetilde{p}_{2m} = \frac{1}{\sqrt{2}} p_{2m}, \quad \widetilde{q}_{2m} = \frac{1}{\sqrt{2}} q_{2m}.
\end{aligned}$$

These formulas show that the entries of the pentadiagonal matrix $\widehat{\mathbb{T}}^a_{2m+1}$ associated to the $(2m+1)$-point shifted anti-Gauss-Laurent quadrature rule are

$$\widetilde{t}_{i,j} = t_{i,j}, \quad \text{for} \quad i,j = 1,2,\ldots,2m$$

and

$$\begin{array}{rclcrcl} \widetilde{t}_{2m+1,2m-1} & = & \sqrt{2}t_{2m+1,2m-1}, & & \widetilde{t}_{2m-1,2m+1} & = & \sqrt{2}\,\widehat{t}_{2m+1,2m-1}, \\ \widetilde{t}_{2m+1,2m} & = & \sqrt{2}t_{2m+1,2m}, & & \widetilde{t}_{2m,2m+1} & = & \sqrt{2}\,\widehat{t}_{2m+1,2m}, \\ \widetilde{t}_{2m+1,2m+1} & = & t_{2m+1,2m+1}. & & & & \end{array}$$

In summary, the pentadiagonal matrix $\widehat{\mathbb{T}}^a_{2m+1}$ is obtained from the matrix $\widehat{\mathbb{T}}_{2m+1} = \widehat{\mathbb{W}}^T_{2m+1} \diamond A\widehat{\mathbb{V}}_{2m+1}$ associated with the $(2m+1)$-point shifted Gauss-Laurent rule as follows

$$\widehat{\mathbb{T}}^a_{2m+1} = \begin{bmatrix} \widehat{\mathbb{T}}_{2m} & \Phi_{2m} \\ \Psi^T_{2m} & t_{2m+1,2m+1} \end{bmatrix} \in \mathbb{R}^{(2m+1)\times(2m+1)},$$

where $\Phi_{2m} = [0,\ldots,0,\sqrt{2}\,\widehat{t}_{2m+1,2m-1},\sqrt{2}\,\widehat{t}_{2m+1,2m}]^T$ and $\Psi_{2m} = [0,\ldots,0,\sqrt{2}t_{2m+1,2m-1},\sqrt{2}t_{2m+1,2m}]^T$.

Algorithm 2 describes how an approximation of (1.3) and an error estimate can be computed by a pair of shifted Gauss-Laurent and shifted anti-Gauss-Laurent quadrature rules when the matrix A is nonsymmetric. The computed approximation, $U_{\rm app}f$, is the average of approximations of (1.3) determined by shifted Gauss-Laurent and anti-Gauss-Laurent quadrature rules. Similarly as in Algorithm 1, we approximate (1.10) by $U_{\rm app}(f) = (\mathcal{G}^\sigma_{2m}(f) + \mathcal{A}^\sigma_{2m+1}(f))/2$ and estimate the error in $U_{\rm app}(f)$ by the difference $|\mathcal{G}^\sigma_{2m}(f) - \mathcal{A}^\sigma_{2m+1}(f)|/|\mathcal{G}^\sigma_{2m}(f)|$.

6. Numerical experiments. This section presents some numerical results that illustrate the performance of the shifted extended Gauss-Laurent-type quadrature rules based on the global shifted extended symmetric or nonsymmetric Lanczos processes. All experiments were carried out in MATLAB R2015a on a computer with an Intel Core i-3 processor and 3.89 GB of RAM. The computations were done with about 15 significant decimal digits.

Upper and lower bounds for trace$(f(A))$ can be determined as the sum of upper and lower bounds for trace$(E_j^T f(A)E_j)$, $j = 1,2,\ldots,n_s$, with $n_s = \lfloor (n+s-1)/s \rfloor$, with the initial block vectors $E_j = [e_{s(j-1)+1},\ldots,e_{\min\{sj,n\}}] \in \mathbb{R}^{n\times s}$, provided that the integrand f is such that pairs of shifted Gauss-Laurent and Gauss-Laurent-Radau rules yield upper and lower bounds. This is described for (standard) Gauss and Gauss-Radau quadrature rules in [7]. We instead apply pairs of shifted Gauss-Laurent and anti-Gauss-Laurent quadrature rules to determine approximations of upper and lower bounds. Denote the computed approximations of trace$(E_j^T f(A)E_j)$ by $U_{{\rm lower},j}(A)$ and $U_{{\rm upper},j}(A)$, respectively. Then

$$U_{\rm lower}(A) := \sum_{j=1}^{n_s} U_{{\rm lower},j}(A), \quad U_{\rm upper}(A) := \sum_{j=1}^{n_s} U_{{\rm upper},j}(A)$$

provide approximations for upper and lower bounds for trace(A).

We report the magnitude of the estimated relative error

$$RelErr(A) = \frac{|U_{\rm upper}(A) - U_{\rm lower}(A)|}{|U_{\rm upper}(A)|}. \tag{6.1}$$

The (standard) Gauss and Gauss-Radau rules described in [7] provide upper and lower bounds for trace$(E_j^T f(A)E_j)$ when $f(t) = \exp(t)$. In this situation, $U_{{\rm upper},j}(A)$ and $U_{{\rm lower},j}(A)$ denote these bounds.

Algorithm 2 Approximation of trace($W^T f(A)V$) by pairs of shifted Gauss-Laurent and shifted anti-Gauss-Laurent quadrature rules for a nonsymmetric matrix A.

Input: Nonsymmetric matrix A, initial block vectors V, W, parameter σ and function f.

1. Choose tolerance $\epsilon > 0$ and the maximum number of iterations I_{max}.
2. $\alpha_{1,1} = |\langle W, V\rangle|^{1/2}$; $V_1 = V/\alpha_{1,1}$; $\beta_{1,1} = \langle W, V\rangle/\alpha_{1,1}$; $W_1 = W/\beta_{1,1}$;
3. $\alpha_{1,2} = \langle W_1, (A - \sigma I_n)^{-1}V_1\rangle$; $\widetilde{V}_2 = (A - \sigma I_n)^{-1}V_1 - \alpha_{1,2}V_1$;
4. $\beta_{1,2} = \langle V_1, (A^T - \sigma I_n)^{-1}W_1\rangle$; $\widetilde{W}_2 = (A^T - \sigma I_n)^{-1}W_1 - \beta_{1,2}W_1$;
5. $\alpha_{2,2} = |\langle \widetilde{W}_2, \widetilde{V}_2\rangle|^{1/2}$; $V_2 = \widetilde{V}_2/\alpha_{2,2}$; $\beta_{2,2} = |\langle \widetilde{W}_2, \widetilde{V}_2\rangle/\alpha_{2,2}$; $W_1 = \widetilde{W}_2/\beta_{2,2}$;
6. $\widetilde{V}_3 = AV_1$; $\widetilde{W}_3 = A^T W_1$; $h_{1,1} = \langle \widetilde{W}_3, V\rangle$; $g_{1,1} = h_{1,1}$;
7. For $j = 1 : I_{max}$
   (a) $\widetilde{V}_{2j+2} = (A - \sigma I_n)^{-1}V_{2j}$; $\widetilde{W}_{2j+2} = (A^T - \sigma I_n)^{-1}W_{2j}$;
   (b) $h_{2j,2j} = \langle \widetilde{W}_{2j+2}, V_{2j}\rangle$; $g_{2j,2j} = h_{2j,2j}$;
   (c) Compute $h_{i,2j-1}, g_{i,2j-1}, h_{i,2j}$; $g_{i,2j}$ from recursion relations given by
   (d) $\widetilde{V}_{2j+1} = \widetilde{V}_{2j+1} - \sum_{i=2j-3}^{2j} h_{i,2j-1}V_i$; $\widetilde{W}_{2j+1} = \widetilde{W}_{2j+1} - \sum_{i=2j-3}^{2j} g_{i,2j-1}W_i$;
   (e) $h_{2j+1,2j-1} = |\langle \widetilde{W}_{2j+1}, \widetilde{V}_{2j+1}\rangle|^{1/2}$; $g_{2j+1,2j-1} = \langle \widetilde{W}_{2j+1}, \widetilde{V}_{2j+1}\rangle/h_{2j+1,2j-1}$;
   (f) if $j = 1$
   $t_{1:2,1} = h_{1:2,1}$; $t_{1,2} = [1 - \alpha_{1,2}(t_{1,1} - \sigma)]/\alpha_{2,2}$; $t_{2,2} = \sigma - \alpha_{1,2}t_{2,1}/\alpha_{2,2}$;
   else
   $t_{2j-3:2j,2j-1} = h_{2j-3:2j,2j-1}$; $t_{2j-1,2j} = -[\sum_{i=2j-3}^{2j-1} h_{i,2j-2}t_{2j-1,i} - \sigma h_{2j-1,2j-2}]/h_{2j,2j-2}$;
   $t_{2j,2j} = \sigma - h_{2j-1,2j-2}t_{2j,2j-1}/h_{2j,2j-2}$;
   end
   (g) $G^{\sigma}_{2j}(f) = e_1^T f(\widehat{\mathbb{T}}_{2j})e_1$;
   (h) $V_{2j+1} = \widetilde{V}_{2j+1}/h_{2j+1,2j-1}$; $W_{2j+1} = \widetilde{W}_{2j+1}/g_{2j+1,2j-1}$;
   (i) $\widetilde{V}_{2j+2} = \widetilde{V}_{2j+2} - \sum_{i=2j-2}^{2j+1} h_{i,2j}V_i$; $\widetilde{W}_{2j+2} = \widetilde{W}_{2j+2} - \sum_{i=2j-2}^{2j+1} g_{i,2j}W_i$;
   (j) $h_{2j+2,2j} = \left|\langle \widetilde{W}_{2j+2}, \widetilde{V}_{2j+2}\rangle\right|^{1/2}$; $g_{2j+2,2j} = \langle \widetilde{W}_{2j+2}, \widetilde{V}_{2j+2}\rangle/h_{2j+2,2j}$;
   (k) if $j = 1$, $t_{3,1} = h_{3,1}$; $t_{3,2} = -\alpha_{1,2}t_{3,1}/\alpha_{2,2}$; else
   $t_{2j+1,2j-1} = h_{2j+1,2j-1}$; $t_{2j+1,2j} = -h_{2j-1,2j-2}t_{2j+1,2j-1}/h_{2j,2j-2}$;
   end
   (l) $\Psi_{2j} = \sqrt{2}\,[0, \ldots, 0, t_{2j+1,2j-1}, t_{2j+1,2j}]$; $\Phi_{2j} = \sqrt{2}\,[0, \ldots, 0, t_{2j-1,2j+1}, t_{2j,2j+1}]^T$;
   (m) $\widetilde{V}_{2j+3} = AV_{2j+1}$; $\widetilde{W}_{2j+3} = A^T W_{2j+1}$; $h_{2j+1,2j+1} = \langle \widetilde{W}_{2j+3}, V_{2j+1}\rangle$; $g_{2j+1,2j+1} = h_{2j+1,2j+1}$;
   (n) Compute $\widehat{\mathbb{T}}^a_{2j+1} = \begin{bmatrix} \widehat{\mathbb{T}}_{2j} & \Phi_{2j} \\ \Psi^T_{2j} & h_{2j+1,2j+1} \end{bmatrix}$ and $A^{\sigma}_{2j+1}(f) = e_1^T f(\widehat{\mathbb{T}}^a_{2j+1})e_1$;
   (o) if $|G^{\sigma}_{2j}(f) - A^{\sigma}_{2j+1}(f)|/|G^{\sigma}_{2j}(f)| < \epsilon$
   $U_{\text{app}}(f) = \langle W, V\rangle[G^{\sigma}_{2j}(f) + A^{\sigma}_{2j+1}(f)]/2$; Break;
   end
   (p) $V_{2j+2} = \widetilde{V}_{2j+2}/h_{2j+2,2j}$; $W_{2j+2} = \widetilde{W}_{2j+2}/g_{2j+2,2j}$;
   (q) end

Output: Approximation $U_{\text{app}}(f)$ of trace($W^T f(A)V$).

The first two subsections compare the performance of pairs of shifted Gauss-Laurent and anti-Gauss-Laurent quadrature rules for symmetric matrices A, as implemented by Algorithm 1, to the performance of (standard) Gauss and Gauss-Radau quadrature (GQ) rules based on the global Lanczos algorithm described in [7, Algorithm 2]. In the third subsection, we compare the application of shifted Gauss-

Laurent and anti-Gauss-Laurent quadrature rules, as implemented by Algorithm 2, to the MATLAB function expm. The block size of E_j is set to $s = 60$ and the stopping tolerance ϵ in Algorithms 1 and 2 is set to $2 \cdot 10^{-3}$. The GQ method is terminated analogously.

The shift parameter is set to $\sigma = 1.01\lambda_{\max}$ or $\sigma = 1.01\lambda_{\min}$, where $\lambda_{\max}$ and $\lambda_{\min}$ are estimates of the largest and smallest eigenvalues of A; we assume here that $\lambda_{\max}$ is positive and $\lambda_{\min}$ is negative. Several techniques can be applied to determine such estimates, including using Gershgorin's disks [37], the irbleigs method [3, 4] for symmetric matrices A, and the MATLAB command eigs, which implements an implicitly restarted Krylov method [35] and can be applied for symmetric and nonsymmetric matrices, or the power method. We use the latter method with initial vector $v = [1, 1, \ldots, 1]^T$.

The systems of equations with the matrix $A - \sigma I_n$ in Algorithm 1 and the systems of equations with the matrices $A - \sigma I_n$ and $A^T - \sigma I_n$ in Algorithm 2 are solved by using the backslash operator \ of MATLAB. This operator computes an LU or Cholesky factorization of $A - \sigma I_n$[1].

6.1. Application to undirected graphs in network analysis. We compute approximations of the Estrada index $EE(A)$ for some undirected networks using the shifted Gauss-Laurent-type rules determined by Algorithm 1. These rules are compared to the Gauss-type quadrature rules based on the global Lanczos algorithm. These rules are denoted by GQ in the tables and described in [7]. We choose the prescribed eigenvalue $\xi = \lambda_{\max}$ for GQ. Then we have the bounds

$$\mathcal{G}_m(f) \leq \operatorname{trace}(E_j^T \exp(A) E_j) \leq \mathcal{R}_{m+1}^{\xi}(f), \qquad f(t) = \exp(t).$$

for every m, where $\mathcal{G}_m(f)$ and $\mathcal{R}_{m+1}^{\xi}(f)$ are defined by (1.7) and (1.8), respectively; see [7] for details. We consider six real-world undirected networks, which can be found in the SuiteSparse Matrix Collection [13]. Some details on these matrices are presented in Table 6.1, including the sparsity of each adjacency matrix, i.e., the ratio between the number of nonzero elements and the total number of elements, n^2. Table 6.2 reports the required CPU time (Time) in seconds, the total number of matrix-vector product (MVP) evaluations, and the relative error (6.1) achieved with these methods. We also report the total number of linear systems solved (LSS) in Algorithm 1. The results show Algorithm 1 to be faster and require a smaller number of MVP evaluations than the GQ algorithm to estimate the Estrada index $EE(G)$. To illustrate the quality of the computed bounds of $\operatorname{trace}(V^T \exp(A) V)$ determined by the shifted anti-Gauss-Laurent quadrature rules, we consider the networks as-22july06 and Erdos972. We choose block size $s = 60$ and $V = E_1$. Figure 6.1 displays the computed approximations of upper and lower bounds for for $EE(G)$ for these networks versus the number of iterations. As can be observed, standard Gauss-type quadrature rules based on the global Lanczos method require many more steps to bracket $\operatorname{trace}(V^T \exp(A) V)$ tightly.

6.2. Application to computing the nuclear norm. The nuclear norm of a general matrix $X \in \mathbb{R}^{m \times n}$ is defined as

$$\|X\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i,$$

where the σ_i are singular values of X. It is impractical or unfeasible to use the singular value decomposition of X to compute the nuclear norm of a large matrix. Computation of the nuclear norm can be

[1] If the matrix $A \in \mathbb{R}^{n \times n}$ is symmetric the operator \ first seeks to compute the Cholesky factorization of $A - \sigma I_n$. If this is not possible, because $A - \sigma I_n$ is not positive definite, then an LU factorization is determined by Gaussian elimination with partial pivoting. The computed factorization is used to solve the linear system of equations with the matrix $A - \sigma I_n$.

Table 6.1
Adjacency matrix properties.

Matrix	# Nodes	# Edges	λ_{max}	Sparsity	Application
Undirected graphs:					
Arenas/email	1133	10902	20.74	$8.5 \cdot 10^{-3}$	interchange network
Pajek/Erdos972	5488	7085	14.45	$4.70 \cdot 10^{-4}$	collaboration network
SNAP/as-735	7716	13895	46.89	$4.45 \cdot 10^{-4}$	computer server network
SNAP/Oregon-1	11492	23409	60.33	$3.54 \cdot 10^{-4}$	road network
Newman/as-22july06	22963	96872	71.61	$1.83 \cdot 10^{-4}$	structure of internet routers
Newman/cond-mat-2005	40421	351384	47.63	$2.15 \cdot 10^{-4}$	collaboration network
Directed graphs:					
SNAP/p2p-Gnutella08	6301	20777	5.12	$5.23 \cdot 10^{-4}$	peer to peer network
Pajek/EVA	8497	6726	1.85	$9.32 \cdot 10^{-5}$	corporate inter-relationships
Pajek/California	9664	16150	7.41	$1.73 \cdot 10^{-4}$	web search
SNAP/p2p-Gnutella04	10879	39994	4.45	$3.38 \cdot 10^{-4}$	peer to peer network

Table 6.2
CPU time in seconds, RelErr, number of matrix-vector product evaluations (MVP) and number of linear system solves (LSS) for computing the Estrada index for several undirected networks.

Matrix	GQ [7]			Algorithm 1			
	Time	RelErr	MVP	Time	RelErr	MVP	LSS
email	2.52	$7.36 \cdot 10^{-4}$	9830	2.24	$2.21 \cdot 10^{-4}$	4585	4585
Erdos972	38.79	$6.63 \cdot 10^{-4}$	47384	30.97	$1.84 \cdot 10^{-4}$	26360	26360
as-735	59.63	$9.30 \cdot 10^{-4}$	67416	40.92	$6.75 \cdot 10^{-6}$	23040	23040
Oregon-1	406.56	$9.18 \cdot 10^{-4}$	110180	71.98	$2.56 \cdot 10^{-5}$	34476	34476
as-22july06	1856	$9.06 \cdot 10^{-4}$	233770	309	$4.43 \cdot 10^{-5}$	68889	68889
cond-mat-2005	9459	$9.40 \cdot 10^{-4}$	666397	3160	$5.78 \cdot 10^{-5}$	206185	206185

considered a trace estimation problems for the symmetric positive semidefinite matrices $A = X^TX$ or $A = XX^T$ [36]. The nuclear norm of X can be expressed as

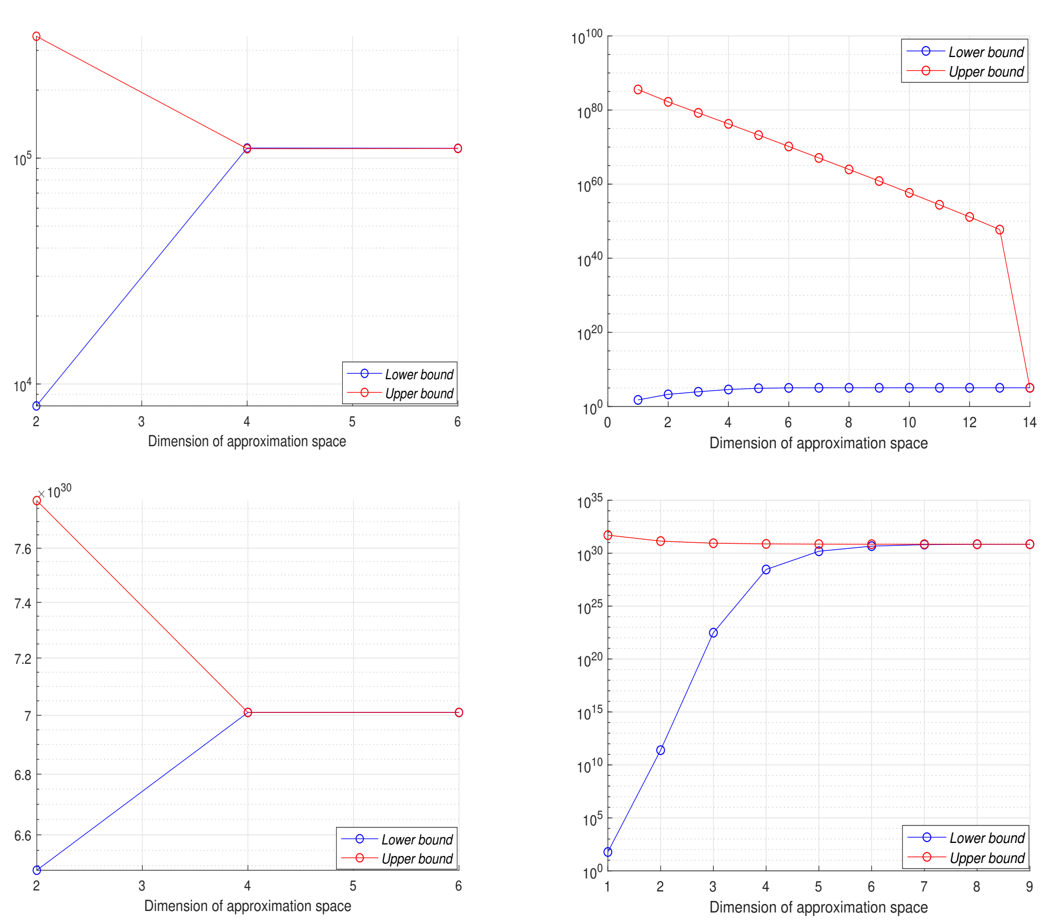
$$\|X\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i = \sum_{i=1}^{\min\{m,n\}} \lambda_i^{1/2} = \mathrm{trace}(A^{1/2}),$$

where the λ_i are the eigenvalues of A. We consider the same data sets as used in [36]. All matrices were obtained from [13]. We compare the performance of Algorithm 1 to the performance of the GQ method [7] when estimating the nuclear norm of the matrices in Table 6.3. The prescribed eigenvalue ξ in the GQ algorithm is chosen to be $\xi = 0$. Let $\xi = \lambda_{\min}$. Then we have the bounds

$$\mathcal{R}_{m+1}^{\xi}(f) \leq \mathrm{trace}(V^T A^{1/2} V) \leq \mathcal{G}_m(f), \qquad f(t) = t^{1/2};$$

see [7] for details. Table 6.3 displays the CPU time (Time) required in seconds, the total number of matrix-vector product (MVP) evaluations, and the relative error (RelErr) in approximations determined by these methods. The symbol ($--$) signifies that the stopping criterion was not satisfied within 3 hours of execution time. The table shows Algorithm 1 to be faster and to require fewer matrix-vector product evaluations than the GQ algorithm.

FIG. 6.1. *Upper and lower bounds or estimates thereof for* $\mathrm{trace}(V^T \exp(A) V)$. *Top row: Erdos972 graph. Bottom row: as-22july06 graph. Left plot: Algorithm 1. Right plot: GQ method [7].*



To illustrate the quality of the computed (approximate) bounds determined by the GQ algorithm and Algorithm 1, we consider $\mathrm{trace}(V^T \sqrt{X_1^T X_1} V)$ and $\mathrm{trace}(V^T \sqrt{X_2^T X_2} V)$, where X_1 and X_2 are the adjacency matrices for the Erdos992 and FA graphs, respectively. The initial block vector V is generated randomly with uniformly distributed entries in the interval $[0, 1]$; the matrix has $s = 60$ columns. Figure 6.2 shows the upper and lower bounds for $\mathrm{trace}(V^T \sqrt{X_1^T X_1} V)$ and $\mathrm{trace}(V^T \sqrt{X_2^T X_2} V)$ produced by the the GQ method and the estimates of upper and lower bounds determined by Algorithm 1 versus the number of iterations. The figure demonstrates the effectiveness of Algorithm 1.

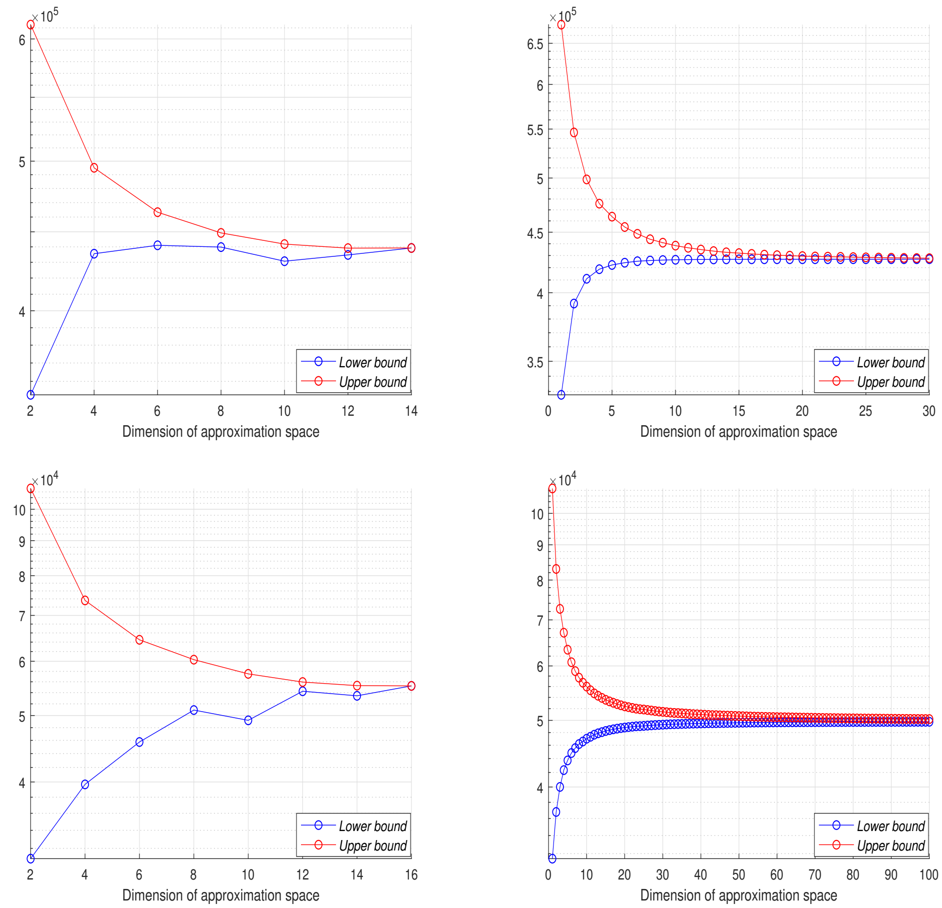
6.3. Application to directed graphs in network analysis. We consider the computation of the Estrada index for some directed graphs that model real-world directed networks. The adjacency matrices are nonsymmetric. These computations illustrate the performance of the shifted Gauss-Laurent-type quadrature rules determined by the nonsymmetric Lanczos process and implemented by Algorithm 2. We use the adjacency matrices p2p-Gnutella08, EVA, California, and p2p-Gnutella04 from [13]. Some properties on these matrices are given in Table 6.1. In Table 6.4, we show the CPU time required by

Table 6.3

CPU time in seconds, RelErr, number of matrix-vector product evaluations (MVP) and number of linear system solves (LSS) for computing the nuclear norm.

Matrix	GQ [7]			Algorithm 1			
	Time	RelErr	MVP	Time	RelErr	MVP	LSS
Erdos992	1400	$1.9 \cdot 10^{-3}$	279060	274.97	$1.33 \cdot 10^{-4}$	44080	44080
deter3	134.16	$1.8 \cdot 10^{-3}$	90396	171.75	$7.39 \cdot 10^{-4}$	37308	37308
California	554.04	$3.6 \cdot 10^{-3}$	901864	321.13	$2.38 \cdot 10^{-4}$	67108	67108
FA	−−	−−	−−	4274	$2.76 \cdot 10^{-4}$	79299	79299

Fig. 6.2. *Upper and lower bounds or estimates thereof for* trace$(V^T\sqrt{A}V)$. *Top row: Erdos992 graph. Bottom row: FA matrix. Left plot: Algorithm 1. Right plot: GQ method [7].*



Algorithm 2 and the MATLAB function `expm`. We also show the approximation of the Estrada index computed by the Algorithm 2. As illustrated by this table, the computational cost for the function `expm` is much higher than for Algorithm 2.

Table 6.4
CPU time in seconds for Algorithm 2 and the MATLAB function expm *for computing the Estrada index for several directed graphs.*

Matrix	Algorithm 2		expm
	Time	Approximation	Time
p2p-Gnutella08	35.11	$6.12 \cdot 10^3$	165.63
EVA	12.82	$8.47 \cdot 10^3$	80.95
California	25.51	$1.14 \cdot 10^4$	130.16
p2p-Gnutella04	219.79	$1.06 \cdot 10^4$	1616.64

7. Conclusion. This paper describes the extended shifted symmetric and nonsymmetric Lanczos processes. These algorithms are used to compute shifted Gauss-Laurent-type quadrature rules. The matrices of recursion coefficients for these Lanczos processes are shown to be pentadiagonal. This results in computations with short recursion formulas. Applications to the determination of estimates of upper and lower bounds for the trace of matrix functions are described. Also applications to the computation of the nuclear norm of a large matrix are described. The computed examples illustrate the effectiveness of the proposed methods.

Acknowledgment. The authors would like to thank Carl Jagels and Giuseppe Rodriguez for carefully reading the manuscript and for comments that lead to an improved presentation. Research by LR was supported in part by NSF grant DMS-1720259.

REFERENCES

[1] H. Alqahtani and L. Reichel, Simplified anti-Gauss quadrature rules with applications in linear algebra, Numer. Algorithms, 77 (2018) 577–602.
[2] H. Alqahtani and L. Reichel, Generalized block anti-Gauss quadrature rules, Numer. Math., 143 (2019) 605–648.
[3] J. Baglama, D. Calvetti, and L. Reichel, IRBL: An implicitly restarted block-Lanczos method for large-scale Hermitian eigenproblems, SIAM J. Sci. Comput., 24 (2003) 1650–1677.
[4] J. Baglama, D. Calvetti, and L. Reichel, Algorithm 827: irbleigs: A MATLAB program for computing a few eigenpairs of a large sparse Hermitian matrix, ACM Trans. Math. Software, 29 (2003) 337–348.
[5] Z. Bai and G. Golub, Bounds for the trace of the inverse and the determinant of symmetric positive definite matrices, Ann. Numer. Math., 4 (1997) 29–38.
[6] 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. 074204.
[7] M. Bellalij, L. Reichel, G. Rodriguez, and H. Sadok, Bounding matrix functionals via partial global block Lanczos decomposition, Appl. Numer. Math., 94 (2015) 127–139.
[8] A. H. Bentbib, M. El Guide, and K. Jbilou, The block Lanczos algorithm for linear ill-posed problems, Calcolo, 54 (2017) 711–732.
[9] A. H. Bentbib, M. El Ghomari, C. Jagels, K. Jbilou, and L. Reichel, The extended global Lanczos method for matrix function approximation, Electron. Trans. Numer. Anal., 50 (2018) 144–163.
[10] R. Bouyouli, K. Jbilou, R. Sadaka, and H. Sadok, Convergence properties of some block Krylov subspace methods for multiple linear systems, J. Comput. Appl. Math., 196 (2006) 498–511.
[11] G. Caldarelli, Scale-Free Networks, Oxford University Press, Oxford, 2007.
[12] O. De la Cruz Cabrera, M. Matar, and L. Reichel, Analysis of directed networks via the matrix exponential, J. Comput. Appl. Math, 355 (2019) 182–192.
[13] T. Davis and Y. Hu, The SuiteSparse Matrix Collection, https://sparse.tamu.edu.
[14] V. Druskin and L. Knizhnerman, Extended Krylov subspace approximations of the matrix square root and related functions, SIAM J. Matrix Anal. Appl., 19 (1998) 755–771.
[15] E. Estrada, Characterization of the folding degree of proteins, Bioinformatics, 18 (2002) 697–704.
[16] E. Estrada, The Structure of Complex Networks: Theory and Applications, Oxford University Press, Oxford, 2011.
[17] E. Estrada, M. Fox, D. Higham, and G. L. Oppo, eds., Network Science, Complexity in Nature and Technology,

Springer, New York, 2010.
[18] E. Estrada and N. Hatano, Statistical-mechanical approach to subgraph centrality in complex networks, Chem. Phys. Lett., 439 (2007) 247–251.
[19] E. Estrada and D. J. Higham, Network properties revealed through matrix functions, SIAM Rev., 52 (2010) 696–714.
[20] E. Estrada and J. A. Rodriguez-Velazquez, Subgraph centrality in complex networks, Phys. Rev. E, 71 (2005), Art. 056103.
[21] C. Fenu, D. Martin, L. Reichel, and G. Rodriguez, Block Gauss and anti-Gauss quadrature with application to networks, SIAM J. Matrix Anal. Appl., 34 (2013) 1655–1684.
[22] W. Gautschi, Orthogonal Polynomials: Computation and Approximation, Oxford University Press, Oxford, 2004.
[23] G. H. Golub, and G. Meurant, Matrices, Moments and Quadrature with Applications, Princeton University Press, Princeton, 2010.
[24] G. H. Golub and C. F. Van Loan, Matrix Computations, 4th ed., Johns Hopkins University Press, Baltimore, 2013.
[25] I. Han, D. Malioutov, and J. Shin, Large-scale log-determinant computation through stochastic Chebyshev expansions, in Proceedings of The 32nd International Conference on Machine Learning, F. Bach and D. Blei, eds., Lille, France, 2015, JMLR Workshop and Conference Proceedings, 37 (2015) 908–917.
[26] C. Jagels, K. Jbilou, and L. Reichel, The extended global Lanczos method, Gauss-Radau quadrature, and matrix function approximation, J. Comput. Appl. Math., 381 (2021) Art. 113027.
[27] C. Jagels and L. Reichel, The structure of matrices in rational Gauss quadrature, Math. Comp. 82 (2013) 2035-2060.
[28] K. Jbilou, A. Messaoudi, and H. Sadok, Global FOM and GMRES algorithms for matrix equations, Appl. Numer. Math., 31 (1999) 49–63.
[29] K. Jbilou, H. Sadok, and A. Tinzefte, Oblique projection methods for multiple linear systems. Electron. Trans. Num. Anal., 20 (2005) 119–138.
[30] D. P. Laurie, Anti-Gaussian quadrature formulas, Math. Comp., 65 (1996) 735–747.
[31] M. E. J. Newman, Networks: An Introduction, Oxford University Press, Oxford, 2010.
[32] T. T. Ngo, M. Bellalij, and Y. Saad, The trace ratio optimization problem, SIAM Rev., 54 (2012) 545–569.
[33] S. Pozza, M. S. Pranić, and Z. Strakoš, The Lanczos algorithm and complex Gauss quadrature, Electron. Trans. Numer. Anal., 50 (2018) 1–19.
[34] Y. Saad, J. Chelikowsky, and S. Shontz, Numerical methods for electronic structure calculations of materials, SIAM Rev., 52 (2010) 3–54.
[35] D. C. Sorensen, Numerical methods for large eigenvalue problems, Acta Numer., 11 (2002) 519–584.
[36] S. Ubaru, J. Chen, and Y. Saad, Fast estimation of $tr(f(A))$ via stochastic Lanczos quadrature, SIAM J. Matrix Anal. Appl., 38 (2017) 1075–1099.
[37] R. S. Varga, Geršgorin and His Circles, Springer, Berlin, 2004.