RATIONAL GAUSS QUADRATURE RULES FOR THE APPROXIMATION OF MATRIX FUNCTIONALS INVOLVING STIELTJES FUNCTIONS

J. ALAHMADI*, M. PRANIĆ†, AND L. REICHEL‡

Abstract. This paper is concerned with computing approximations of matrix functionals of the form $F(A) := \boldsymbol{v}^T f(A)\boldsymbol{v}$, where A is a large symmetric positive definite matrix, $\boldsymbol{v}$ is a vector, and f is a Stieltjes function. We approximate $F(A)$ with the aid of rational Gauss quadrature rules. Associated rational Gauss–Radau and rational anti-Gauss rules are developed. Pairs of rational Gauss and rational Gauss–Radau quadrature rules, or pairs of rational Gauss and rational anti-Gauss quadrature rules, can be used to determine upper and lower bounds, or approximate upper and lower bounds, for $F(A)$. The application of rational Gauss rules, instead of standard Gauss rules, is beneficial when the function f has singularities close to the spectrum of A.

AMS (MOS) subject classification: 65D15, 65D32, 65F60

Keywords: Stieltjes function, matrix function, rational Gauss quadrature, rational Gauss–Radau rule, rational anti-Gauss rule

1. Introduction. We are interested in computing approximations of matrix functionals of the form

$$F(A) := \boldsymbol{v}^T f(A)\boldsymbol{v} \tag{1.1}$$

by quadrature rules, where $A \in \mathbb{R}^{N\times N}$ is a large symmetric positive definite matrix, the vector $\boldsymbol{v} \in \mathbb{R}^N\backslash\{\mathbf{0}\}$ is of unit Euclidean norm, the superscript ^{T} denotes transposition, and f is a Stieltjes function, i.e., f has the representation

$$f(z) = \int_0^\infty \frac{1}{t+z}\, d\mu(t), \quad z \in \mathbb{C} \setminus (-\infty, 0], \tag{1.2}$$

where the nonnegative measure $d\mu$ is such $f(z)$ is well defined; see, e.g., [4, 6, 11, 12, 16, 20, 22] for discussions and illustrations of Stieltjes functions. Examples include

$$f(z) = z^{-a} = \frac{\sin(a\pi)}{\pi}\int_0^\infty \frac{1}{t+z}\, d\mu(t), \quad \text{with} \quad d\mu(t) = t^{-a}dt, \quad a \in (0,1), \tag{1.3}$$

$$f(z) = \frac{\log(1+z)}{z} = \int_0^\infty \frac{1}{t+z}\, d\mu(t), \quad \text{with} \quad d\mu(t) = \begin{cases} 0, & 0 \le t \le 1, \\ t^{-1}dt, & t > 1. \end{cases}$$

Let the matrix A have the spectral factorization

$$A = U\Lambda U^T, \quad \Lambda = \operatorname{diag}[\lambda_1, \lambda_2, \ldots, \lambda_N], \tag{1.4}$$

where the matrix $U \in \mathbb{R}^{N\times N}$ is orthogonal and the eigenvalues λ_i of A are ordered according to $0 < \lambda_1 \le \cdots \le \lambda_N$. The spectral factorization is used in the derivation of the quadrature

*Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: jalahma1@kent.edu

†Department of Mathematics and Informatics, University of Banja Luka, Faculty of Science, M. Stojanovića 2, 51000 Banja Luka, Bosnia and Herzegovina. E-mail: miroslav.pranic@pmf.unibl.org

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

rules of this paper, but application of these rules does not require the computation of this factorization.

Define the vector $[\nu_1, \nu_2, \ldots, \nu_N] := \boldsymbol{v}^T U$. Using the factorization (1.4), the expression (1.1) with f defined by (1.2) can be written as

$$\begin{aligned}\boldsymbol{v}^T f(A)\boldsymbol{v} = \boldsymbol{v}^T U f(\Lambda) U^T \boldsymbol{v} &= \int_0^\infty \boldsymbol{v}^T U (tI + \Lambda)^{-1} U^T \boldsymbol{v}\, d\mu(t) \\ &= \int_0^\infty \sum_{i=1}^N (t + \lambda_i)^{-1} \nu_i^2 d\mu(t) \\ &= \int_0^\infty \int_0^\infty (t + y)^{-1} d\nu(y) d\mu(t) \\ &= \int_0^\infty f(y) d\nu(y) =: \mathcal{I}(f), \end{aligned} \tag{1.5}$$

where the nonnegative measure $d\nu(y)$ has support at the eigenvalues λ_i of A. The associated distribution function $\nu(y)$ can be chosen to be a nondecreasing piece-wise constant function with jumps of height ν_i^2 at the eigenvalues λ_i of A. We discuss the use of rational Gauss-type quadrature rules to approximate the integral (1.5). This is useful when the matrix A is large and the distribution function $\nu(y)$ has many points of increase.

The approximation of expressions $f(A)\boldsymbol{v}$ when f is a Stieltjes function, as well as the computation of error bounds or estimates, has received considerable attention; see, e.g., [4, 11, 12, 22] and references therein. Frommer and Schweitzer [11, 12] approximate $f(A)$ by a polynomial, while Massei and Robol [22] use a rational function. A difficulty with polynomial approximants is that their degree may have to be large to yield the desired accuracy of the computed approximation of (1.1). This is illustrated in Section 5.

Our contribution differs from the works [11, 12, 22] in that we use rational Gauss quadrature rules for the approximation of (1.1). These rules were introduced in [25] and are based on orthogonal rational functions that satisfy short recurrence relations and have prescribed poles. Rational Gauss quadrature rules associated with the measure $d\nu$ in (1.5) can be computed by applying a few steps of a rational Lanczos process to the matrix A with initial vector $\boldsymbol{v}$. These rules are exact for certain rational functions with prescribed poles. This paper defines rational Gauss–Radau quadrature rules, and shows how pairs of rational Gauss and Gauss–Radau rules can be applied to determine upper and lower bounds for the functional (1.1). These bounds are analogues of bounds provided by pairs of standard Gauss and Gauss–Radau rules that have been described by Golub and Meurant [14]. The evaluation of these bounds requires that bounds for extreme eigenvalues of A be available.

We also define rational anti-Gauss quadrature rules and simplified rational anti-Gauss rules. Pairs of rational Gauss and rational anti-Gauss rules or pairs of rational Gauss and simplified rational anti-Gauss rules also provide upper and lower bounds for the functional (1.1) under suitable conditions, and do not require knowledge of extreme eigenvalues of A. Rational anti-Gauss rules provide an extension to rational Gauss quadrature of the anti-Gauss rules associated with (standard) Gauss quadrature rules that were introduced by Laurie [21]. Other modifications and extensions of the anti-Gauss rules by Laurie [21] are discussed in [1, 2, 7, 10, 23].

This paper is organized as follows. Section 2 reviews how (standard) Gauss quadrature rules for the approximation of (1.5) can be determined by carrying out a few steps of the Lanczos process applied to the symmetric matrix A with initial vector $\boldsymbol{v}$. Section 3 discusses

available results on recursion formulas for orthonormal bases for rational Krylov subspaces associated with orthogonal rational functions with several fixed poles. The recursion formulas are determined by the vector $\boldsymbol{v}$, the symmetric matrix A, and by shifted matrices $(A-\alpha_i I)^{-1}$. In our computations, the α_i are negative numbers, which are poles of the rational functions that determine the rational Gauss, rational Gauss–Radau, rational anti-Gauss, and simplified rational anti-Gauss quadrature rules used to approximate (1.1). It also is possible to let some poles appear in complex conjugate pairs. Recursion relations for rational orthogonal functions with poles α_i are reviewed in Section 3, and their application to rational Gauss, rational Gauss–Radau, rational anti-Gauss, and simplified rational anti-Gauss quadrature rules is discussed in Section 4. A nice introduction to rational Gauss rules and their properties is provided by Gautschi [13, Section 3.1.4]. Section 5 presents a few computed examples, and concluding remarks can be found in Section 6.

2. Gauss quadrature rules. This section reviews the application of the symmetric Lanczos process to the symmetric positive definite matrix $A \in \mathbb{R}^{N\times N}$ to evaluate Gauss quadrature rules for the approximation of the functional (1.1); see, e.g., Golub and Meurant [14] for further details. The application of $1 \leq m \ll N$ steps of the symmetric Lanczos process to A with initial unit vector $\boldsymbol{v}$ gives the Lanczos decomposition

$$AV_m = V_m T_m + t_{m-1,m}\boldsymbol{v}_m \boldsymbol{e}_m^T, \tag{2.1}$$

where the matrix $V_m = [\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m-1}] \in \mathbb{R}^{N\times m}$, with $\boldsymbol{v}_0 := \boldsymbol{v}$, has orthonormal columns that form a basis for the Krylov subspace

$$\mathbb{K}_m(A, \boldsymbol{v}) = \text{span}\{\boldsymbol{v}, A\boldsymbol{v}, \ldots, A^{m-1}\boldsymbol{v}\}.$$

Moreover, the unit vector $\boldsymbol{v}_m \in \mathbb{R}^N$ satisfies $V_m^T \boldsymbol{v}_m = \boldsymbol{0}$, and $T_m = [t_{ij}]_{i,j=0}^{m-1} \in \mathbb{R}^{m\times m}$ is a symmetric positive definite tridiagonal matrix; the scalar $t_{m-1,m}$ is nonnegative. Here and below $\boldsymbol{e}_j = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ denotes the jth axis vector. We assume that $1 \leq m \ll N$ is chosen small enough so that the decomposition (2.1) with the stated properties exists. This is the generic situation.

It follows from the recursion relation (2.1) for the columns of V_m that the jth column can be expressed as

$$\boldsymbol{v}_{j-1} = p_{j-1}(A)\boldsymbol{v}, \quad j = 1, 2, \ldots, m,$$

where p_{j-1} is a polynomial of degree $j-1$. Golub and Meurant [14] show that the expression

$$\mathcal{G}_m(f) := \boldsymbol{e}_1^T f(T_m)\boldsymbol{e}_1 \tag{2.2}$$

is the m-point Gauss quadrature rule for the approximation of (1.1), i.e.,

$$\mathcal{G}_m(f) = \boldsymbol{v}^T f(A)\boldsymbol{v}, \qquad \forall f \in \mathbb{P}_{2m},$$

where $\mathbb{P}_{2m}$ denotes the space of all polynomials of degree at most $2m-1$, i.e., a space of dimension $2m$. This Gauss rule is associated with the bilinear form

$$\langle f, g\rangle := (f(A)\boldsymbol{v})^T (g(A)\boldsymbol{v}), \tag{2.3}$$

which is an inner product for polynomials f and g of sufficiently low degree. This can be seen by substituting the spectral factorization (1.4) into (2.3). Substituting the spectral

factorization of the matrix T_m into the right-hand side of (2.2) shows that $\mathcal{G}_m(f)$ indeed is a quadrature rule with m nodes.

When the integrand f has one or several singularities close to the support of the measure $d\nu$ in (1.5), Gauss rules (2.2) with a moderate number of nodes, m, may yield poor approximations of the functional (1.1). This difficulty can be remedied by using rational Gauss rules. They were first discussed by Gonchar and López Lagomasino [15], and have subsequently received considerable attention; see, e.g., [5, 13, 17, 19].

We will use the rational Gauss quadrature rules described in [25] and define associated rational Gauss–Radau rules, as well as rational anti-Gauss and simplified anti-Gauss rules. The computation of the rational Gauss rules described in [25] is based on the observation in [24] that a sequence of certain orthogonal rational functions satisfy short recursion relations, i.e., the number of terms in the recursion relations can be bounded independently of the number of orthogonal rational functions in the sequence. We note that different sequences of orthogonal rational functions that satisfy a three-term recursion relation have been developed by Deckers and Bultheel [8]. These sequences also can be used to construct rational Gauss rules; see [9]. We use the rational Gauss quadrature rules described in [25], because their computation requires fewer linear systems of equations with matrices that are determined from A to be solved than the approach in [9].

3. Recursion relations for rational Krylov subspaces. The first part of this section reviews results in [24] on recursion relations for certain orthogonal rational functions. The number of terms in the recursion relations depends on the number of distinct poles of the rational functions and on the ordering of certain elementary rational basis functions. The recursion relations are applied in a rational Lanczos process, which is described in Subsection 3.1.

Introduce linear spaces of rational functions with finite poles,

$$\mathbb{Q}_{i,k_i} = \text{span}\left\{ \frac{1}{(y-\alpha_i)^j} : \ j = 1, 2, \ldots, k_i \right\}, \quad i = 1, 2, \ldots, \ell, \tag{3.1}$$

where the α_i are real distinct poles of multiplicity $k_i, i = 1, 2, \ldots, \ell$. They are assumed to lie outside the convex hull of the support of the measure $d\nu$. In the application of this paper, they will lie on the negative real axis.

Now let $\alpha_{\ell+2i}$, $i = 1, 2, \ldots, \hat{\ell}$, denote distinct complex conjugate poles with nonvanishing imaginary part, and assume that each pole and its complex conjugate are adjacent, i.e., $\alpha_{\ell+2i} = \bar{\alpha}_{\ell+2i-1}$, where the bar denotes complex conjugation. Since we are interested in integrating functions f, such that $f(y) \in \mathbb{R}$ for $y > 0$, we replace each pair of rational functions

$$\frac{1}{(y-\alpha_{\ell+2i-1})^j}, \quad \frac{1}{(y-\bar{\alpha}_{\ell+2i-1})^j}$$

by a pair

$$\frac{1}{(y^2+\beta_i y+\gamma_i)^j}, \quad \frac{y}{(y^2+\beta_i y+\gamma_i)^j},$$

where the coefficients $\beta_i, \gamma_i \in \mathbb{R}$ are defined by $y^2+\beta_i y+\gamma_i = (y-\alpha_{\ell+2i-1})(y-\bar{\alpha}_{\ell+2i-1})$. Analogously to (3.1), we define the spaces

$$\mathbb{W}_{i,2s_i} = \text{span}\left\{ \frac{1}{(y^2+\beta_i y+\gamma_i)^j}, \frac{y}{(y^2+\beta_i y+\gamma_i)^j} : \ j = 1, 2, \ldots, s_i \right\}, \quad i = 1, 2, \ldots, \hat{\ell},$$

where s_i is the multiplicity of the complex conjugate poles $\alpha_{\ell+2i-1}$ and $\alpha_{\ell+2i}$.

Let

$$k = \sum_{i=1}^{\ell} k_i, \quad s = \sum_{i=1}^{\hat{\ell}} s_i, \tag{3.2}$$

and introduce the $(m+1)$-dimensional linear space

$$\mathbb{S}_{m+1} := \mathbb{P}_{m+1-k-2s} \oplus \mathbb{Q}_{1,k_1} \oplus \cdots \oplus \mathbb{Q}_{\ell,k_\ell} \oplus \mathbb{W}_{1,2s_1} \oplus \cdots \oplus \mathbb{W}_{\hat{\ell},2s_{\hat{\ell}}}, \tag{3.3}$$

where we assume that the k_i and s_i are chosen so that $0 \leq k + 2s \leq m - 1$. Then the space (3.3) contains linear functions. Let

$$\Psi_{m+1} = \{\psi_0, \psi_1, \ldots, \psi_m\} \tag{3.4}$$

denote an elementary basis for the space $\mathbb{S}_{m+1}$, i.e., $\psi_0(y) = 1$ and each basis function $\psi_i(y)$, for $i = 1, 2, \ldots, m$, is one of the functions

$$y^j, \ \frac{1}{(y-\alpha_i)^j}, \ \frac{1}{(y^2+\beta_i y+\gamma_i)^j}, \ \frac{y}{(y^2+\beta_i y+\gamma_i)^j} \tag{3.5}$$

for some positive integers i and j. The notation $\psi_s \prec \psi_t$ indicates that the basis function ψ_s comes before ψ_t. We say that the ordering of the basis functions (3.4) is *natural* if $\psi_0(y) = 1$ and the remaining functions ψ_j, $j = 1, 2, \ldots, m$, satisfy:

1. $y^p \prec y^{p+1}$ for all integers $p > 0$,
2. $\frac{1}{(y-\alpha_i)^p} \prec \frac{1}{(y-\alpha_i)^{p+1}}$ for all integers $p > 0$ and every real pole α_i,
3. $\frac{1}{(y^2+\beta_i y+\gamma_i)^p} \prec \frac{y}{(y^2+\beta_i y+\gamma_i)^p} \prec \frac{1}{(y^2+\beta_i y+\gamma_i)^{p+1}}$ for all positive integers p and every pair $\{\beta_i, \gamma_i\}$,
4. if $\psi_j(y) = \frac{1}{(y^2+\beta_i y+\gamma_i)^p}$, then $\psi_{j+1}(y) = \frac{y}{(y^2+\beta_i y+\gamma_i)^p}$.

In particular, if $\psi_j = y^p$ for some $p > 1$, then some preceding elementary basis function ψ_i with $i < j$ equals y^{p-1}. An analogous statement also holds for negative powers $(y - \alpha_i)^{-p}$ and $(y^2 + \beta_i y + \gamma_i)^{-p}$ for $p > 1$.

It is shown in [24] that a Stieltjes procedure for orthogonalizing elementary basis functions with respect to some inner product only requires short recursion relations when the basis functions are in natural order. The same holds if we orthogonalize the elementary basis functions with respect to a bilinear form $[f, g] = \mathcal{L}(fg)$, where $\mathcal{L}$ is a linear functional on $\mathbb{S}_{m+1}$, as long as $\mathcal{L}(\phi_j^2) \neq 0$. In this section, we orthogonalize the elementary basis functions with respect to the bilinear form (2.3); in Section 4 also other bilinear forms will be used. The bilinear form (2.3) is an inner product for all elementary basis functions of low enough order.

A recursion relation is said to be *short* if the number of terms in the relation can be bounded independently of the number of orthogonal functions generated. A well-known special case is furnished by the situation when all elementary basis functions are monomials, i.e., $\psi_j(y) = y^j$, $j = 0, 1, 2, \ldots$. Then $\phi_0, \phi_1, \phi_2, \ldots$ form a family of orthonormal polynomials with ϕ_j of degree j for all j. These polynomials satisfy a three-term recurrence relation. When the elementary basis functions are allowed to be of the form (3.5), the number of terms required in the recurrence relation depends on how often powers of the same function appear in the sequence (3.4). This is explained below.

We mention three types of recursion relations from [24] that are needed in this paper and start with the recursion relation for $y\phi_r$:

$$y\phi_r(y) = \sum_{j=-n_1}^{n_2} c_{r,r+j}\phi_{r+j}(y), \quad r = 0, 1, \dots , \tag{3.6}$$

where $r - n_1$ is the largest integer smaller than r such that ψ_{r-n_1} is a monomial if there is such a monomial (otherwise $r - n_1 = 0$), and $r + n_2$ is the smallest integer larger than r such that ψ_{r+n_2} is a monomial.

In order to introduce a real pole α_i we need the following recursions:

$$\frac{1}{y - \alpha_i}\phi_r(y) = \sum_{j=-n_3}^{n_4} c^{(i)}_{r,r+j}\phi_{r+j}(y), \quad r = 0, 1, 2, \dots , \tag{3.7}$$

where $r - n_3$ is the largest integer smaller than r such that ψ_{r-n_3} is a rational function with a pole at α_i if there is such a rational function (otherwise $r - n_3 = 0$), and $r + n_4$ is the smallest integer larger than r such that ψ_{r+n_4} is a rational function with a pole at α_i.

To introduce a pair of complex conjugate poles, we use the following formulas with $p \in \{0, 1\}$:

$$\frac{y^p}{y^2 + \beta_j y + \gamma_j}\phi_r(y) = \sum_{i=-n_5}^{n_6} c^{(j)}_{r,r+i}\phi_{r+i}(y), \quad r = 0, 1, 2, \dots , \tag{3.8}$$

where $r - n_5$ is the largest integer smaller than r such that $\psi_{r-n_5}(y) = (y^2 + \beta_j y + \gamma_j)^s$ if there is such a rational function (otherwise $r - n_5 = 0$), and $r + n_6$ is the smallest integer larger than r such that $\psi_{r+n_6}(y) = y(y^2 + \beta_j y + \gamma_j)^s$. Note that the first orthonormal rational functions ϕ_i and the first elementary basis functions ψ_i span the same space.

Define the vector of orthonormal rational functions,

$$\Phi_m(y) := [\phi_0(y), \phi_1(y), \dots , \phi_{m-1}(y)].$$

If ψ_m is a monomial and $m - d$ is the largest integer smaller than m such that ψ_{m-d} is a monomial, then the recursion formulas (3.6) can be written in the matrix form

$$y\Phi_m(y) = H_m\Phi_m(y) + \sum_{j=1}^{d} h_{m-j,m}\phi_m(y)\boldsymbol{e}_{m+1-j}. \tag{3.9}$$

The matrix H_m has the following block-diagonal structure: it has $m - k - 2s - 1$ square blocks along the diagonal such that any two consecutive blocks overlap in one diagonal element. The jth block of H_m is of dimension $r \times r$, where $r - 2$ is the number of rational functions between consecutive monomials y^{j-1} and y^j. More precisely, the jth block of H_m is the submatrix $H_m(r_1 : r_2, r_1 : r_2)$ (in MATLAB notation) with the entries h_{ij} for $r_1 \le i, j \le r_2$, where $\psi_{r_1}(y) = y^{j-1}$ and $\psi_{r_2}(y) = y^j$. Also, $r = r_2 - r_1 + 1$. The non-zero entries of $H_m = [h_{ij}]_{i,j=0}^{m-1}$ are recursion coefficients for the functions $\phi_i, i = 0, 1, \dots , m - 1$. They satisfy

$$h_{i,j} = \langle y\phi_i(y), \phi_j(y)\rangle = \langle y\phi_j(y), \phi_i(y)\rangle = h_{j,i}, \tag{3.10}$$

which shows that the matrix H_m depends only on the first m elementary basis functions ψ_j, $j = 0, 1, \ldots, m-1$, and does not depend on ψ_m. In this paper we will always assume that ψ_m (or ψ_{m+1} if we deal with H_{m+1}) is a monomial for only one reason: then the zeros of ϕ_m are eigenvalues of H_m.

The following example illustrates the structure of the matrix H_m.

Example 3.1. Let $\alpha_i, \beta_j, \gamma_j$ be defined as described in the beginning of this section and consider the elementary basis

$$\left\{1, \frac{1}{y-\alpha_1}, y, \frac{1}{(y-\alpha_1)^2}, y^2, \ldots, \frac{1}{y-\alpha_\ell}, y^{k-k_\ell+1}, \ldots, \frac{1}{(y-\alpha_\ell)^{k_\ell}}, y^k, \frac{1}{y^2+\beta_1 y+\gamma_1},\right.$$
$$\frac{y}{y^2+\beta_1 y+\gamma_1}, y^{k+1}, \ldots, \frac{1}{(y^2+\beta_1 y+\gamma_1)^{s_1}}, \frac{y}{(y^2+\beta_1 y+\gamma_1)^{s_1}}, y^{k+s_1}, \frac{1}{y^2+\beta_2 y+\gamma_2},$$
$$\left.\frac{y}{y^2+\beta_2 y+\gamma_2}, y^{k+s_1+1}, \ldots, \frac{1}{(y^2+\beta_{\hat\ell} y+\gamma_{\hat\ell})^{s_{\hat\ell}}}, \frac{y}{(y^2+\beta_{\hat\ell} y+\gamma_{\hat\ell})^{s_{\hat\ell}}}, y^{k+s}\right\}.$$

This basis together with the function $\psi_m(y) = y^{k+s+1}$, where k and s are defined by (3.2), and $m = 1 + 2k + 3s$, satisfy the requirements of natural ordering and make up the space $\mathbb{S}_{m+1}$. The matrix H_m in this case has k 3×3 blocks and s trailing 4×4 blocks along the diagonal:

$$H_m = \begin{bmatrix} * & * & * & & & & & & \\ * & * & * & & & & & & \\ * & * & * & * & * & & & & \\ & & * & * & * & & & & \\ & & * & * & * & & & & \\ & & & & & \ddots & & & \\ & & & & & & * & * & * & * \\ & & & & & & * & * & * & * \\ & & & & & & * & * & * & * \\ & & & & & & * & * & * & * & * & * & * \\ & & & & & & & & * & * & * & * \\ & & & & & & & & * & * & * & * \\ & & & & & & & & * & * & * & * \end{bmatrix}.$$

Matrix entries that may be nonvanishing are marked by $*$. If we place one more rational function between two consecutive monomials y^{j-1} and y^j in the elementary basis, then the size of the jth block of H_m increases by one. Similarly, removing one rational function between two consecutive monomials decreases the size of the corresponding block by one. □

3.1. The rational Lanczos process. In the rest of the paper we focus on functionals $\mathcal{I}$ defined by (1.1) when A is a symmetric positive definite matrix, and $\boldsymbol{v}$ is a normalized vector. If the functions ϕ_j, $j = 0, 1, \ldots, m$, are orthonormal with respect to the bilinear form (2.3), then the vectors

$$\{\boldsymbol{v}_0 = \phi_0(A)\boldsymbol{v}, \ldots, \boldsymbol{v}_j = \phi_j(A)\boldsymbol{v}\} \tag{3.11}$$

form an orthonormal basis for the rational Krylov subspace

$$\mathbb{K}_{j+1}(A, \boldsymbol{v}) = \operatorname{span}\{\psi_0(A)\boldsymbol{v}, \psi_1(A)\boldsymbol{v}, \ldots, \psi_j(A)\boldsymbol{v}\}, \tag{3.12}$$

for $j = 0, 1, \ldots, m$. Indeed,

$$\boldsymbol{v}_j^T \boldsymbol{v}_i = (\phi_j(A)\boldsymbol{v})^T(\phi_i(A)\boldsymbol{v}) = \langle \phi_j, \phi_i \rangle.$$

The vectors $\boldsymbol{v}_j$ satisfy the same recursion relations as rational functions ϕ_j and can be constructed by the rational Lanczos process. This process is analogous to the Stieltjes-type procedure [24, Algorithm 3.1] for computing an orthonormal basis for the space $\mathbb{S}_{m+1}$ which is based on the recursion relations (3.6), (3.7) and (3.8).

Algorithm 1 The rational Lanczos process.

Input: $\boldsymbol{v} \in \mathbb{R}^N \backslash \{\boldsymbol{0}\}$, a sequence of matrix-valued elementary basis functions $\psi_0, \psi_1, \ldots, \psi_m$, and functions for evaluating matrix-vector products with A and for solving linear systems of equations with matrices of the form $A - \alpha_\ell I$ and $A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I$. Thus, we do not explicitly form the elementary basis functions A^j, $(A - \alpha_\ell I)^{-j}$, and $(A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I)^{-j}$. The given sequence of elementary basis functions implicitly defines the integers j, ℓ, and $\hat{\ell}$ used below.
Output: Orthonormal basis $\{\boldsymbol{v}_r\}_{r=0}^m$.
Initialization: $\boldsymbol{v}_0 := \boldsymbol{v}/\|\boldsymbol{v}\|; r := 1;$
while $r \leq m$ **do**
  if $\psi_r = A^j$ for some $j \in \mathbb{N}$ **then**
    $\boldsymbol{u} := A\boldsymbol{v}_{r-1};$
    for $i = \hat{r} : r-1$ **do**
      $c_{r-1,i} := \boldsymbol{v}_i^T \boldsymbol{u}; \ \boldsymbol{u} := \boldsymbol{u} - c_{r-1,i}\boldsymbol{v}_i;$
    end for
    $\delta_r := \|\boldsymbol{u}\|; \ \boldsymbol{v}_r := \boldsymbol{u}/\delta_r;$
    $r = r + 1$
  else if $\psi_r = (A - \alpha_\ell I)^{-j}$ for some $j, \ell \in \mathbb{N}$ **then**
    $\boldsymbol{u} := (A - \alpha_\ell I)^{-1}\boldsymbol{v}_{r-1};$
    for $i = \hat{r} : r-1$ **do**
      $c_{r-1,i} := \boldsymbol{v}_i^T \boldsymbol{u}; \ \boldsymbol{u} := \boldsymbol{u} - c_{r-1,i}\boldsymbol{v}_i;$
    end for
    $\delta_r := \|\boldsymbol{u}\|; \ \boldsymbol{v}_r := \boldsymbol{u}/\delta_r;$
    $r = r + 1$
  else if $\psi_r = (A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I)^{-j}$ for some $j, \hat{\ell} \in \mathbb{N}$ **then**
    $\boldsymbol{u} := (A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I)^{-1}\boldsymbol{v}_{r-1};$
    $\boldsymbol{w} := A(A^2 + \beta_{\hat{\ell}} A + \gamma_{\hat{\ell}} I)^{-1}\boldsymbol{v}_{r-1};$
    for $i = \hat{r} : r-1$ **do**
      $c_{r-1,i} := \boldsymbol{v}_i^T \boldsymbol{u}; \ \boldsymbol{u} := \boldsymbol{u} - c_{r-1,i}\boldsymbol{v}_i;$
      $d_{r-1,i} := \boldsymbol{v}_i^T \boldsymbol{w}; \ \boldsymbol{w} := \boldsymbol{w} - d_{r-1,i}\boldsymbol{v}_i;$
    end for
    $\delta_r := \|\boldsymbol{u}\|; \ \boldsymbol{v}_r := \boldsymbol{u}/\delta_r;$
    $\eta := \boldsymbol{v}_r^T \boldsymbol{w}; \ \boldsymbol{w} := \boldsymbol{w} - \eta\boldsymbol{v}_r;$
    $\delta_r' := \|\boldsymbol{w}\|; \ \boldsymbol{v}_{r+1} := \boldsymbol{w}/\delta_r';$
    $r = r + 2$
  end if
end while

The implementation of the rational Lanczos process requires the solution of linear systems

of equations with matrices of the forms $A - \alpha_i I$ and $A^2 + \beta_i A + \gamma_i I$, where $\alpha_i, \beta_i, \gamma_i$ are suitable real scalars. Throughout this paper the matrix A is assumed to have a structure that allows fairly efficient computation of $(A - \alpha_i I)^{-1}\boldsymbol{v}_{r-1}$ and $(A^2 + \beta_i A + \gamma_i I)^{-1}\boldsymbol{v}_{r-1}$ by direct or iterative methods. This is the case, for instance, when A is banded with a small bandwidth or when A is obtained by discretizing an elliptic partial differential equation and the required vectors can be computed by a multigrid method. The norm $\|\cdot\|$ in Algorithm 1 denotes the Euclidean vector norm.

We assume in Algorithm 1 that the basis (3.4) satisfies conditions 1-4 of natural ordering. The value of $\hat{r}$ is such that $\psi_{\hat{r}}$ is a basis function having the same pole(s) as ψ_r but of smaller order; otherwise $\hat{r} = 0$. Performing m steps of Algorithm 1 yields an orthonormal basis $\{\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ for the rational Krylov subspace $\mathbb{K}_{m+1}(A, \boldsymbol{v})$. The matrix

$$V_m = [\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m-1}] \in \mathbb{R}^{N\times m}, \quad \text{with } \boldsymbol{v}_0 = \boldsymbol{v},$$

and the symmetric matrix

$$H_m = [h_{i,j}]_{i,j=0}^{m-1} \in \mathbb{R}^{m\times m}, \quad h_{i,j} = \mathcal{I}(y\phi_i\phi_j) = \boldsymbol{v}_i^T A\boldsymbol{v}_j,$$

satisfy the decomposition

$$AV_m = V_m H_m + \sum_{j=1}^{d} h_{m-j,m}\boldsymbol{v}_m \boldsymbol{e}_{m+1-j}^T, \tag{3.13}$$

where $\boldsymbol{v}_m \in \mathbb{R}^N$ is such that $V_m^T\boldsymbol{v}_m = \boldsymbol{0}$. The orthonormal basis V_m for the subspace $\mathbb{K}_{m+1}(A, \boldsymbol{v})$ is nested, i.e., $\mathbb{K}_1 \subseteq \mathbb{K}_2 \subseteq \ldots$, where $\mathbb{K}_i$ is spanned by the first i columns of the matrix V_m. The orthogonal projection of A onto the rational Krylov subspace (3.12) is given by

$$H_m = V_m^T A V_m.$$

We tacitly assume that m is small enough so that the decomposition (3.13) with the stated properties exists. This is the generic situation. Algorithm 1 breaks down before m steps have been carried out if $\delta_r = 0$ or $\delta_r^{'} = 0$. Ramification of breakdown is out of scope of this paper.

4. Application to rational Gauss quadrature. This section discusses rational Gauss quadrature rules for the approximation of functionals (1.1) when f is a Stieltjes function (1.2) and A is a symmetric positive definite matrix. Then

$$F(A) = \boldsymbol{v}^T f(A)\boldsymbol{v} = \int_0^\infty \boldsymbol{v}^T(tI + A)^{-1}\boldsymbol{v}\, d\mu(t) = \mathcal{I}(f). \tag{4.1}$$

It is shown in [25] that the expression

$$\widehat{\mathcal{G}}_m(f) := \boldsymbol{e}_1^T f(H_m)\boldsymbol{e}_1 \tag{4.2}$$

is a rational Gauss quadrature rule for the approximation of the expression (4.1). It is characterized by the property

$$\widehat{\mathcal{G}}_m(f) = \mathcal{I}(f), \quad \forall f \in \mathbb{S}_{2m}, \tag{4.3}$$

where

$$\mathbb{S}_{2m} := \mathbb{P}_{2m-2k-4s} \oplus \mathbb{Q}_{1,2k_1} \oplus \cdots \oplus \mathbb{Q}_{\ell,2k_\ell} \oplus \mathbb{W}_{1,4s_1} \oplus \cdots \oplus \mathbb{W}_{\hat{\ell},4s_{\hat{\ell}}} = \frac{\mathbb{P}_{2m}}{w^2}$$

denotes a $2m$-dimensional linear space of certain rational functions and

$$w(y) = \prod_{i=1}^{\ell} (y - \alpha_i)^{k_i} \prod_{j=1}^{\hat{\ell}} \left(y^2 + \beta_j y + \gamma_j\right)^{s_j} \tag{4.4}$$

is a polynomial of degree $k + 2s$ with k and s defined by (3.2). The scalars $\alpha_i, \beta_j, \gamma_j$ are determined as described in the beginning of Section 3. Thus, the m-point rational Gauss rule (4.2) is exact for all function in the space $\mathbb{S}_{2m}$ of dimension $2m$. This is analogous to the situation for standard Gauss rules: the m-point Gauss rule (2.2) is exact for all polynomials in the space $\mathbb{P}_{2m}$ of dimension $2m$.

In view of (1.5), the rational Gauss rule (4.2) can be written as

$$\widehat{\mathcal{G}}_m(f) = \int_0^\infty \boldsymbol{e}_1^T (tI + H_m)^{-1} \boldsymbol{e}_1 d\mu(t) = \boldsymbol{e}_1^T f(H_m) \boldsymbol{e}_1, \tag{4.5}$$

where we recall that the measure $d\mu$ defines the function f; the rational Gauss rule is with respect to the measure $d\nu$ in (1.5).

The orthonormal rational function ϕ_m has m distinct zeros $\{y_i\}_{i=1}^m$ that lie in the convex hull of the support of the measure $d\nu$, and they are the eigenvalues of H_m in (3.9); see [25, Theorem 2.5]. Recall that we assume ψ_m to be a monomial. Thus the rational function ϕ_m in (3.9) can be represented as

$$\phi_m(y) = c_m \frac{\prod_{i=1}^m (y - y_i)}{w(y)} \in \mathbb{S}_{m+1}, \tag{4.6}$$

where c_m is a constant.

The remainder term for rational Gauss rules (4.5) can be derived by considering rational Hermite interpolation. The following result is shown similarly as [19, Theorem 5.4]. The Lagrange fundamental functions associated with the function (4.6) are defined by

$$l_i(y) := \frac{\phi_m(y)}{\phi_m'(y_i)(y - y_i)}, \quad i = 1, 2, \ldots, m,$$

and satisfy

$$l_i(y_j) = \begin{cases} 1 & \text{if} \quad j = i, \\ 0 & \text{if} \quad j \neq i. \end{cases}$$

Introduce the rational Hermite interpolation function

$$\hat{L}(y) := \sum_{i=1}^m \left(\hat{l}_i(y) f(y_i) + \tilde{l}_i(y) f'(y_i) \right), \tag{4.7}$$

where

$$\hat{l}_i(y) = \left[1 - 2(y - y_i) l_i'(y_i)\right] l_i^2(y), \quad \tilde{l}_i(y) = (y - y_i) l_i^2(y).$$

Then the function $\hat{L}(y)$ satisfies the interpolation conditions

$$\hat{L}(y_i) = f(y_i), \quad \hat{L}'(y_i) = f'(y_i), \quad i = 1, 2, \ldots, m.$$

Theorem 1. *Let $\hat{L}$ be the rational Hermite function* (4.7) *determined by the interpolation nodes $y_1 < y_2 < \ldots < y_m$, which are the m zeros of $\phi_m \in \mathbb{S}_{m+1}$. Assume that f is $2m$ times continuously differentiable in the interval between the nodes y_1 and y_m. Then for some scalar $c = c(y)$ depending on y in this interval, we have*

$$f(y) = \hat{L}(y) + \frac{d^{2m}}{dt^{2m}} \left(w(t)^2 f(t)\right)_{t=c} \frac{\phi_m^2(y)}{c_m^2 (2m)!}. \tag{4.8}$$

Proof. If $y = y_i$ for some $i = 1, 2, \ldots, m$, then

$$f(y) = \hat{L}(y),$$

in (4.8). Thus, the error vanishes. The result holds for an arbitrary constant c. Now, assume that $y \neq y_i$ for all i. In this case, we consider

$$g(t) = w^2(t)(f(t) - \hat{L}(t)) - w^2(y)(f(y) - \hat{L}(y)) \prod_{i=1}^{m} \frac{(t - y_i)^2}{(y - y_i)^2},$$

where $w^2(x)\hat{L}(x)$ is a polynomial of degree $2m - 1$. The function g is $2m$ times continuously differentiable and has $2m + 1$ zeros. By Rolle's theorem, there exists a scalar $c = c(y)$ depending on y in the interval between the nodes y_1 and y_m, for which

$$\begin{aligned}
0 &= \frac{d^{2m}}{dt^{2m}} (g(t))_{t=c} \\
&= \frac{d^{2m}}{dt^{2m}} \left(w(t)^2 f(t)\right)_{t=c} - w^2(y)(f(y) - \hat{L}(y)) \frac{d^{2m}}{dt^{2m}} \left(\prod_{i=1}^{m} \frac{(t - y_i)^2}{(y - y_i)^2} \right)_{t=c} \\
&= \frac{d^{2m}}{dt^{2m}} \left(w(t)^2 f(t)\right)_{t=c} - w^2(y)(f(y) - \hat{L}(y)) \frac{(2m)!}{\prod_{i=1}^{m} (y - y_i)^2}.
\end{aligned}$$

Rearranging terms, and dividing by w^2 yields (4.8). □

Theorem 1 can be used to construct an error term for the m-point rational Gauss quadrature rule $\widehat{\mathcal{G}}_m$. The following results can be shown similarly as [19, Corollary 5.5].

Corollary 2. *Assume that f is $2m$ times continuously differentiable in the convex hull of the spectrum of A. Then the remainder term for the rational Gauss rule* (4.5) *can be expressed as*

$$\mathcal{E}_m(f) := \mathcal{I}(f) - \widehat{\mathcal{G}}_m(f) = \frac{d^{2m}}{dy^{2m}} (w^2(y) f(y))_{y=c} \cdot \frac{1}{(2m)!} \int_0^\infty \frac{\prod_{j=1}^{m} (y - y_j)^2}{w^2(y)} d\nu(y) \tag{4.9}$$

for some scalar c in an interval that contains the spectrum of A.

Proof. Theorem 1 shows that the rational Hermite interpolation function $\hat{L}$ lies in $\mathbb{S}_{2m}$. It now follows from (4.3) that

$$\widehat{\mathcal{G}}_m(\hat{L}) = \mathcal{I}(\hat{L}),$$

and we obtain

$$(\mathcal{I} - \widehat{\mathcal{G}}_m)(f) = \mathcal{I}(f - \hat{L}) = \frac{d^{2m}}{dt^{2m}} \left(w(t)^2 f(t)\right)_{t=c} \frac{\mathcal{I}(\phi_m^2(y))}{c_m^2 (2m)!}.$$

This shows (4.9). □

Assume that f satisfies the conditions of Corollary 2, and that

$$\frac{d^{2m}}{dy^{2m}}\left(w^2(y)f(y)\right) \geq 0, \tag{4.10}$$

in some open interval containing the spectrum of A. Then Corollary 2 gives a lower bound for $\mathcal{I}(f)$. We have

$$\widehat{\mathcal{G}}_m(f) \leq \mathcal{I}(f).$$

4.1. Rational Gauss–Radau quadrature rules. This subsection discusses the computation of rational Gauss–Radau rules and error bounds that can be determined with these rules. This approach of bracketing (1.1) is a rational analogue of the technique advocated by Golub and Meurant [14] for computing upper and lower bounds for (1.1) by evaluating pairs of (standard) Gauss and Gauss–Radau quadrature rules.

The $(m+1)$-point rational Gauss–Radau quadrature rule with a prescribed node θ can be expressed as

$$\widehat{\mathcal{R}}^{\theta}_{m+1}(f) = \int_0^{\infty} \boldsymbol{e}_1^T (tI + H^{\theta}_{m+1})^{-1} \boldsymbol{e}_1 d\mu(t) = \boldsymbol{e}_1^T f(H^{\theta}_{m+1}) \boldsymbol{e}_1 \tag{4.11}$$

for a suitable matrix $H^{\theta}_{m+1} \in \mathbb{R}^{(m+1)\times(m+1)}$. This quadrature rule is characterized by the property

$$\widehat{\mathcal{R}}^{\theta}_{m+1}(f) = \mathcal{I}(f), \quad \forall f \in \mathbb{S}_{2m+1},$$

where

$$\mathbb{S}_{2m+1} := \mathbb{P}_{2m+1-2k-4s} \oplus \mathbb{Q}_{1,2k_1} \oplus \cdots \oplus \mathbb{Q}_{\ell,2k_\ell} \oplus \mathbb{W}_{1,4s_1} \oplus \cdots \oplus \mathbb{W}_{\hat{\ell},4s_{\hat{\ell}}} = \frac{\mathbb{P}_{2m+1}}{w^2}.$$

The standard $(m+1)$-point Gauss–Radau rule for the approximation of (1.1) can be determined by modifying the tridiagonal matrix T_{m+1} in (2.1). Analogously, we will show that the rational Gauss–Radau rule (4.11) can be determined by modifying the last diagonal entry of the matrix H_{m+1}.

Note that in our definition of the rational $(m+1)$-point Gauss–Radau rule we assume exactness on the space $\frac{\mathbb{P}_{2m+1}}{w^2}$ with the same w as in definition of $\widehat{\mathcal{G}}_m$. In other words, we do not introduce a new finite pole. Thus, we can introduce the finite poles in the same order as in the construction of $\widehat{\mathcal{G}}_m$: take ψ_m to be a monomial, and determine the matrix H_{m+1} such that

$$y\Phi_{m+1}(y) = H_{m+1}\Phi_{m+1}(y) + \varphi_{m+1}(y)\boldsymbol{e}_{m+1}, \tag{4.12}$$

and

$$\boldsymbol{e}_1^T f(H_{m+1})\boldsymbol{e}_1 = \mathcal{I}(f), \quad \forall f \in \frac{\mathbb{P}_{2m+2}}{w^2}.$$

We will show that the rational Gauss–Radau rule (4.11) with a fixed node $\theta \leq \lambda_1$ can be determined by replacing the last diagonal entry $h_{m,m}$ of H_{m+1} by $h^{\theta}_{m,m}$ so that the resulting matrix

$$H^{\theta}_{m+1} = \begin{bmatrix} H_m & \boldsymbol{w}_m \\ \boldsymbol{w}_m^T & h^{\theta}_{m,m} \end{bmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}, \quad \boldsymbol{w}_m = [0,\ldots,0,h_{m,m-d},\ldots,h_{m,m-1}]^T \in \mathbb{R}^m,$$

has an eigenvalue at $\theta \le \lambda_1$. Only d trailing entries of the vector $\boldsymbol{w}_m$ might be nonvanishing, where d is the same as in (3.9). We use $\varphi_{m+1} = h_{m,m+1}\phi_{m+1}$ instead of ϕ_{m+1} for simplicity. The last equation in (4.12) can be written in the form

$$\varphi_{m+1}(y) = (y - h_{m,m})\phi_m(y) - \sum_{i=m-d}^{m-1} h_{m,i}\phi_i(y).$$

First we show that

$$\boldsymbol{e}_1^T f(H_{m+1}^\theta)\boldsymbol{e}_1 = \mathcal{I}(f), \quad \forall f \in \frac{\mathbb{P}_{2m+1}}{w^2}, \tag{4.13}$$

for any $h_{m,m}^\theta$. Replacing $h_{m,m}$ in the matrix H_{m+1} by $h_{m,m} + c$, we obtain the recursion coefficients for the rational functions $\phi_0, \phi_1, \ldots, \phi_m, \tilde{\varphi}_{m+1}$, where

$$\tilde{\varphi}_{m+1} = \varphi_{m+1} - c\phi_m.$$

We see that $\tilde{\varphi}_{m+1}$ is orthogonal to the space $\frac{\mathbb{P}_m}{w} = \operatorname{span}\{\phi_0, \ldots, \phi_{m-1}\}$ with respect to the bilinear form defined by the integral $\mathcal{I}$. Then we can construct the new functional $\tilde{\mathcal{I}}$ on the space $\frac{\mathbb{P}_{2m+2}}{w^2}$ such that $\tilde{\mathcal{I}}(f) = \mathcal{I}(f)$ for $f \in \frac{\mathbb{P}_{2m+1}}{w^2}$ and $\tilde{\mathcal{I}}(\tilde{\varphi}_{m+1}\phi_m) = 0$. Thus, $\phi_0, \phi_1, \ldots, \phi_m, \tilde{\varphi}_{m+1}$ is the sequence of orthogonal rational functions with respect to $\tilde{\mathcal{I}}$, and

$$\boldsymbol{e}_1^T f(H_{m+1}^\theta)\boldsymbol{e}_1 = \tilde{\mathcal{I}}(f), \quad \forall f \in \frac{\mathbb{P}_{2m+2}}{w^2}.$$

The formula (4.13) follows from the fact that $\tilde{\mathcal{I}}(f) = \mathcal{I}(f)$ for $f \in \frac{\mathbb{P}_{2m+1}}{w^2}$.

We finish the construction of the rational Gauss–Radau quadrature by choosing c so that $\tilde{\varphi}_{m+1}(\theta) = 0$. Thus, we get

$$c = \frac{\varphi_{m+1}(\theta)}{\phi_m(\theta)}.$$

The rational function $\tilde{\varphi}_{m+1}$ can be written in the form

$$\tilde{\varphi}_{m+1}(y) = a\frac{(y-\theta)q_m(y)}{w(y)},$$

where a is a constant and q_m is a polynomial of degree m. Since $\tilde{\varphi}_{m+1}$ is orthogonal to the space $\frac{\mathbb{P}_m}{w}$, we see that $\frac{q_m}{w}$ is orthogonal to the space $\frac{\mathbb{P}_m}{w}$ with respect to the integral $I(f) = \mathcal{I}((y-\theta)f)$ with nonnegative measure $(y-\theta)d\nu(y)$. Thus, we conclude that the rational Gauss–Radau quadrature $\widehat{\mathcal{G}}_{m+1}$ has a node at θ and m distinct nodes in the open interval that contains the spectrum of A.

Similarly, we may define rational Gauss–Radau rules with a fixed node $\theta \ge \lambda_N$. The following result is shown in the same manner as [19, Theorem 7.1].

Theorem 3. *Assume that f is $2m+1$ times continuously differentiable in the convex hull of the spectrum of A and θ. Then the remainder term for the rational Gauss–Radau rule* (4.11) *can be written as*

$$\begin{aligned}\mathcal{E}_{m+1}(f) :=& \mathcal{I}(f) - \widehat{\mathcal{R}}_{m+1}^\theta(f) \\ =& \frac{d^{2m+1}}{dy^{2m+1}}(w^2(y)f(y))_{y=c} \cdot \frac{1}{(2m+1)!}\int_0^\infty (y-\theta)\frac{\prod_{j=1}^m (y-y_j)^2}{w^2(y)}d\nu(y),\end{aligned} \tag{4.14}$$

where the scalar c lies in an interval that contains the spectrum of A and θ.

Proof. Let the function ϕ_m in Theorem 1 have an additional node θ and assume that the rational Hermite interpolation function $\hat{L}^\theta \in \mathbb{S}_{2m+1}$ also interpolates f at $y_1, y_2, \ldots, y_m$ and θ. Then by modifying the error term (4.8) and applying Corollary 2, we obtain (4.14). □

Assume that f satisfies the conditions of Theorem 3, and that

$$\frac{d^{2m+1}}{dy^{2m+1}}\left(w^2(y)f(y)\right) \geq 0 \tag{4.15}$$

in some open interval containing the spectrum of A and θ. Since

$$\begin{aligned}\int_0^\infty (y-\lambda_1)\frac{\prod_{j=1}^m (y-y_j)^2}{w^2(y)}d\nu(y) \geq 0,\\ \int_0^\infty (y-\lambda_N)\frac{\prod_{j=1}^m (y-y_j)^2}{w^2(y)}d\nu(y) \leq 0,\end{aligned} \tag{4.16}$$

the remainder terms for the rational Gauss–Radau rule with a prescribed node $\theta \leq \lambda_1$ or $\theta \geq \lambda_N$ are of opposite sign. It follows that

$$\widehat{\mathcal{R}}_{m+1}^{\lambda_1}(f) \leq \mathcal{I}(f) \leq \widehat{\mathcal{R}}_{m+1}^{\lambda_N}(f). \tag{4.17}$$

Analogously to (4.17), when the derivative $(w^2 f)^{(2m)}$ in (4.9) is of constant sign in an interval that contains the spectrum of A, and the derivative $(w^2 f)^{(2m+1)}$ in (4.14) is of constant sign in an interval that contains the spectrum of A and θ, and the Radau point $\theta \leq \lambda_1$ or $\theta \geq \lambda_N$ is suitably chosen, the values $\widehat{\mathcal{G}}_m(f)$ and $\widehat{\mathcal{R}}_{m+1}^\theta(f)$ bracket (1.1). We tacitly assume here that θ is chosen so that $\widehat{\mathcal{R}}_{m+1}^\theta(f)$ can be evaluated.

It is known that for every Stieltjes functions f, we have for all nonnegative integers k and ℓ,

$$(-1)^k \frac{d^{k+\ell}}{dx^{k+\ell}}(x^\ell f(x)) \geq 0, \quad x > 0; \tag{4.18}$$

see, e.g., Sokal [26, Theorem 1] and references therein. Setting $\ell = 0$ shows that Stieltjes functions are completely monotonic. Formula (4.18) demonstrates that for certain simple polynomials (4.4), such as $w(x) = x$, the properties (4.10) and (4.15) hold. Then pairs of rational Gauss and rational Gauss–Radau rules, or pairs of rational Gauss–Radau rules, can be used to bracket (1.1); this is a consequence of Corollary 2 and Theorem 3. However, the bracketing cannot be guaranteed for all polynomials (4.4). Moreover, we recall that the application of a Gauss–Radau rule to bounding the error requires knowledge of the location of the largest or smallest eigenvalues of A in order to allocate the Radau point.

4.2. Rational anti-Gauss quadrature rules. When the derivatives $(w^2 f)^{(2m)}$ or $(w^2 f)^{(2m+1)}$ change sign on the convex hull of the spectrum of A, pairs of rational Gauss and rational Gauss–Radau quadrature rules are not guaranteed to bracket (1.1). In this case, estimates of upper and lower bounds for (1.1) can be determined by evaluating appropriate pairs of rational Gauss and anti-Gauss quadrature rules. An advantage of this approach is that the sign of derivatives of $(w^2 f)$ are allowed to change in an interval that contains the spectrum of A. Moreover, knowledge of the location of the largest or smallest eigenvalues of A is not required.

Laurie [21] introduced the (standard) $(m+1)$-point anti-Gauss quadrature rule that gives an error of the same magnitude and of opposite sign as the (standard) m-point Gauss

quadrature rule when applied to polynomials in $\mathbb{P}_{2m+2}$. The evaluation of the standard $(m+1)$-point anti-Gauss quadrature rule requires the computation of $m+1$ steps of the (standard) Lanczos process; see, e.g., [7] for details. In this subsection, we will show that rational anti-Gauss rules can be computed in an analogous fashion.

The $(m+1)$-point rational anti-Gauss rule $\widetilde{\mathcal{G}}_{m+1}$ associated with the functional $\mathcal{I}$ defined by (1.5) is determined by the requirement that

$$(\mathcal{I} - \widetilde{\mathcal{G}}_{m+1})(f) = -(\mathcal{I} - \widehat{\mathcal{G}}_m)(f), \quad \forall f \in \mathbb{S}_{2m+2}, \tag{4.19}$$

where

$$\mathbb{S}_{2m+2} := \mathbb{P}_{2m+2-2k-4s} \oplus \mathbb{Q}_{1,2k_1} \oplus \cdots \oplus \mathbb{Q}_{\ell,2k_\ell} \oplus \mathbb{W}_{1,4s_1} \oplus \cdots \oplus \mathbb{W}_{\hat{\ell},4s_{\hat{\ell}}} = \frac{\mathbb{P}_{2m+2}}{w^2}.$$

We will show that $\widetilde{\mathcal{G}}_{m+1}(f)$ can be expressed as

$$\widetilde{\mathcal{G}}_{m+1}(f) = \int_0^\infty \boldsymbol{e}_1^T (tI + \widetilde{H}_{m+1})^{-1} \boldsymbol{e}_1 d\mu(t) = \boldsymbol{e}_1^T f(\widetilde{H}_{m+1}) \boldsymbol{e}_1 \tag{4.20}$$

for a suitable matrix $\widetilde{H}_{m+1} \in \mathbb{R}^{(m+1)\times(m+1)}$.

The relation (4.19) shows that $\widetilde{\mathcal{G}}_{m+1}$ is the $(m+1)$-point rational Gauss rule for the functional

$$\mathcal{J}(f) := (2\mathcal{I} - \widehat{\mathcal{G}}_m)(f).$$

Define, analogously to (3.11), the vectors

$$\tilde{\boldsymbol{v}}_j = \tilde{\phi}_j(A)\boldsymbol{v}, \quad j = 0, 1, \ldots, m+1,$$

where the $\tilde{\phi}_j$ are orthonormal rational functions with respect to the bilinear form

$$\{f, g\} := \mathcal{J}(fg), \tag{4.21}$$

i.e, $\{\tilde{\phi}_i, \tilde{\phi}_j\} = 0$ for $i \neq j$, and $\{\tilde{\phi}_j, \tilde{\phi}_j\} = 1$ for all j. These orthonormal functions satisfy recurrence relations of the form

$$y\tilde{\Phi}_{m+1}(y) = \widetilde{H}_{m+1}\tilde{\Phi}_{m+1}(y) + \tilde{\varphi}_{m+1}(y)\boldsymbol{e}_{m+1}, \tag{4.22}$$

where $\tilde{\Phi}_{m+1}(y) = [\tilde{\phi}_0, \tilde{\phi}_1, \ldots, \tilde{\phi}_m]^T$, and the remainder term $\tilde{\varphi}_{m+1}(y)\boldsymbol{e}_{m+1}$ in (4.22) is of the form $\tilde{h}_{m,m+1}\tilde{\phi}_{m+1}(y)\boldsymbol{e}_{m+1}$ since the elementary basis functions ψ_m and ψ_{m+1} are monomials. Analogously to (3.10), the matrix $\widetilde{H}_{m+1} = [\tilde{h}_{ij}]_{i,j=0}^m$ is determined by the coefficients of the recursion relation that express the bilinear form (4.21),

$$\tilde{h}_{i,j} = \{y\tilde{\phi}_i, \tilde{\phi}_j\} = \{y\tilde{\phi}_j, \tilde{\phi}_i\} = \tilde{h}_{j,i}. \tag{4.23}$$

Introducing the elementary basis functions ψ_i in the same way as in construction of the rational Gauss–Radau rule gives the matrix $\widetilde{H}_{m+1}$ with the same block-diagonal structure as H_{m+1}. It follows from (4.3) and (4.19) that for rational functions f and g such that $fg \in \mathbb{S}_{2m}$, we have

$$\{f, g\} = \langle f, g\rangle = \mathcal{I}(fg).$$

These equalities show that

$$\tilde{h}_{i,j} = h_{i,j}, \quad i,j = 0,1,\ldots,m-1.$$

Therefore, $\tilde{\phi}_j = \phi_j$ for $0 \le j < m$, and

$$\tilde{H}_{m+1} = \begin{bmatrix} H_m & \tilde{\boldsymbol{w}}_m \\ \tilde{\boldsymbol{w}}_m^T & \tilde{h}_{m,m} \end{bmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}, \quad \tilde{\boldsymbol{w}}_m = \begin{bmatrix} 0,\ldots,0,\tilde{h}_{m-d,m},\ldots,\tilde{h}_{m-1,m} \end{bmatrix}^T \in \mathbb{R}^m.$$

Further, we get

$$h_{m-1,m}\phi_m(y) = (y - h_{m-1,m-1})\phi_{m-1}(y) - \sum_{j=m-d}^{m-2} h_{m-1,j}\phi_j(y) = \tilde{h}_{m-1,m}\tilde{\phi}_m(y).$$

We use the previous equality, (4.23), and the fact that $\widehat{\mathcal{G}}_m(\phi_m g) = 0$ for any function g to show that $\phi_m = \sqrt{2}\tilde{\phi}_m$. Indeed,

$$\tilde{h}_{m-1,m} = \mathcal{J}(y\tilde{\phi}_{m-1}\tilde{\phi}_m) = \frac{h_{m-1,m}}{\tilde{h}_{m-1,m}}\mathcal{J}(y\phi_{m-1}\phi_m) = \frac{h_{m-1,m}}{\tilde{h}_{m-1,m}}\,2\,\mathcal{I}(y\phi_{m-1}\phi_m) = \frac{2h_{m-1,m}^2}{\tilde{h}_{m-1,m}}.$$

The same relation holds for the other not necessarily vanishing entries of $\tilde{\boldsymbol{w}}_m$:

$$\tilde{h}_{m-j,m} = \mathcal{J}(y\tilde{\phi}_{m-j}\tilde{\phi}_m) = \frac{1}{\sqrt{2}}\,2\,\mathcal{I}(y\phi_{m-j}\phi_m) = \sqrt{2}h_{m-j,m}, \quad j = d,\ldots,2.$$

Finally,

$$\tilde{h}_{m,m} = \mathcal{J}(y\tilde{\phi}_m\tilde{\phi}_m) = 2\mathcal{I}(y\frac{\phi_m}{\sqrt{2}}\frac{\phi_m}{\sqrt{2}}) = h_{m,m}.$$

Therefore, the matrix $\tilde{H}_{m+1}$ associated with the rational anti-Gauss rule $\tilde{\mathcal{G}}_{m+1}$ is given by

$$\widetilde{H}_{m+1} = \begin{bmatrix} H_m & \sqrt{2}\boldsymbol{w}_m \\ \sqrt{2}\boldsymbol{w}_m^T & h_{m,m} \end{bmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}, \tag{4.24}$$

Analogously to formula (4.5), the rational anti-Gauss quadrature rule can be evaluated according to (4.20).

We are now in a position to provide some sufficient conditions for $\widehat{\mathcal{G}}_m(f)$ and $\widetilde{\mathcal{G}}_{m+1}(f)$ to bracket $\mathcal{I}(f)$. Assume that we can carry out N steps of Algorithm 1 without breakdown. This yields an orthonormal basis $\{\boldsymbol{v}_j\}_{j=0}^{N-1}$ of $\mathbb{R}^N$, and an associated sequence of orthonormal rational function $\{\phi_j\}_{j=0}^{N-1}$ determined by (3.11).

Theorem 4. *Let $\lambda(A)$ denote the spectrum of A, and consider the expansion of the integrand*

$$f(y) = \sum_{j=0}^{N-1} \omega_j\phi_j(y), \quad y \in \lambda(A), \tag{4.25}$$

in terms of the rational functions ϕ_j determined by (3.11)*, and assume that the coefficients ω_j in* (4.25) *are such that*

$$\left|\sum_{j=2m}^{2m+1} \omega_j\widehat{\mathcal{G}}_m(\phi_j)\right| \ge \max\left\{\left|\sum_{j=2m+2}^{N-1} \omega_j\widehat{\mathcal{G}}_m(\phi_j)\right|, \left|\sum_{j=2m+2}^{N-1} \omega_j\widetilde{\mathcal{G}}_{m+1}(\phi_j)\right|\right\}. \tag{4.26}$$

Then the quadrature rules $\widehat{\mathcal{G}}_m(f)$ and $\widetilde{\mathcal{G}}_{m+1}(f)$ bracket $\mathcal{I}(f)$.

Proof. Since

$$\mathcal{I}(f) = \omega_0 \mathcal{I}(\phi_0), \qquad \mathcal{I}(\phi_j) = 0, \quad \forall j > 0,$$

we have, in view of (4.3) and (4.19), that

$$\begin{aligned}\widehat{\mathcal{G}}_m(f) = \sum_{j=0}^{N-1} \omega_j \widehat{\mathcal{G}}_m(\phi_j) = \sum_{j=0}^{2m-1} \omega_j \widehat{\mathcal{G}}_m(\phi_j) + \sum_{j=2m}^{N-1} \omega_j \widehat{\mathcal{G}}_m(\phi_j)\\ = \mathcal{I}(f) + \omega_{2m} \widehat{\mathcal{G}}_m(\phi_{2m}) + \omega_{2m+1} \widehat{\mathcal{G}}_m(\phi_{2m+1}) + \sum_{j=2m+2}^{N-1} \omega_j \widehat{\mathcal{G}}_m(\phi_j) \qquad (4.27)\end{aligned}$$

and

$$\begin{aligned}\widetilde{\mathcal{G}}_{m+1}(f) = \sum_{j=0}^{N-1} \omega_j \widetilde{\mathcal{G}}_{m+1}(\phi_j) = \sum_{j=0}^{2m+1} \omega_j (2\mathcal{I} - \widehat{\mathcal{G}}_m)(\phi_j) + \sum_{j=2m+2}^{N-1} \omega_j \widetilde{\mathcal{G}}_{m+1}(\phi_j)\\ = \mathcal{I}(f) - \omega_{2m} \widehat{\mathcal{G}}_m(\phi_{2m}) - \omega_{2m+1} \widehat{\mathcal{G}}_m(\phi_{2m+1}) + \sum_{j=2m+2}^{N-1} \omega_j \widetilde{\mathcal{G}}_{m+1}(\phi_j). \qquad (4.28)\end{aligned}$$

Combining (4.27) and (4.28) shows (4.26). □

Theorem 4 shows that if the coefficients ω_j decay sufficiently rapidly with increasing index j, then rational Gauss and anti-Gauss rules provide quadrature errors that are of opposite sign and of roughly the same magnitude. Since it is difficult to verify for a given expression (1.1) whether the conditions of the theorem hold, we say that pairs of rational Gauss and rational anti-Gauss quadrature rules provide estimates of upper and lower bounds for (1.1).

It is natural to consider the average quadrature rule

$$\mathcal{A}_{2m+1}(f) := \frac{1}{2}(\widehat{\mathcal{G}}_m + \widetilde{\mathcal{G}}_{m+1})(f). \qquad (4.29)$$

It follows from (4.27) and (4.28) that

$$\mathcal{A}_{2m+1}(f) = \mathcal{I}(f) + \sum_{j=2m+2}^{N-1} \omega_j \mathcal{A}_{2m+1}(\phi_j).$$

This shows that

$$\mathcal{A}_{2m+1}(f) = \mathcal{I}(f), \quad f \in \mathbb{S}_{2m+2}.$$

This rule is an extension to the average rule defined by Laurie [21].

The computation of the matrix H_m that determines the rational rule $\widehat{\mathcal{G}}_m(f)$ requires that m steps of Algorithm 1 be carried out, while the calculation of the matrix $\widetilde{H}_{m+1}$ that defines the associated rational anti-Gauss rules $\widetilde{\mathcal{G}}_m(f)$ demands $m+1$ steps of Algorithm 1. The last step of the algorithm determines the last diagonal entry, $h_{m,m}$, of (4.24). We can

reduce the number of steps by replacing this entry by an arbitrary scalar, $\breve{h}$. This defines the matrix $\breve{H}_{m+1} \in \mathbb{R}^{(m+1)\times(m+1)}$. We refer to the quadrature rule so obtained,

$$\breve{\mathcal{G}}_{m+1}(f) = \int_0^\infty \boldsymbol{e}_1^T (tI + \breve{H}_{m+1})^{-1} \boldsymbol{e}_1 d\mu(t) = \boldsymbol{e}_1^T f(\breve{H}_{m+1}) \boldsymbol{e}_1, \tag{4.30}$$

as a simplified rational anti-Gauss rule. This rule is an extension to rational Gauss quadrature of the simplified anti-Gauss rules associated with (standard) Gauss rules discussed in [1]. In the computed examples of Section 5, we found the choice $\breve{h} = h_{m-1,m-1}$, where $h_{m-1,m-1}$ is the last diagonal element of the matrix H_m to yield good results.

The following result provides sufficient conditions for the quadrature rules $\widehat{\mathcal{G}}_m(f)$ and $\breve{\mathcal{G}}_{m+1}(f)$ to bracket (1.1), and holds for an arbitrary scalar $\breve{h}$.

Theorem 5. *The simplified rational anti-Gauss rule* (4.30) *satisfies*

$$\breve{\mathcal{G}}_{m+1}(f) = \mathcal{I}(f), \qquad \forall f \in \mathbb{S}_{2m}, \tag{4.31}$$

$$\breve{\mathcal{G}}_{m+1}(f) = (2\mathcal{I} - \widehat{\mathcal{G}}_m)(f), \qquad \forall f \in \mathbb{S}_{2m+1}. \tag{4.32}$$

Consider the expansion (4.25) *with the rational function* ϕ_j *determined by* (3.11)*, and assume that the coefficients* ω_j *in* (4.25) *are such that*

$$\left|\omega_{2m}\widehat{\mathcal{G}}_m(\phi_{2m})\right| \geq \max\left\{ \left| \sum_{j=2m+1}^{N-1} \omega_j \widehat{\mathcal{G}}_m(\phi_j) \right|, \left| \sum_{j=2m+1}^{N-1} \omega_j \breve{\mathcal{G}}_{m+1}(\phi_j) \right| \right\}. \tag{4.33}$$

Then the quadrature rules $\widehat{\mathcal{G}}_m(f)$ *and* $\breve{\mathcal{G}}_{m+1}(f)$ *bracket* (1.1).

Proof. The rational anti-Gauss rule $\widetilde{\mathcal{G}}_{m+1}(f)$ satisfies $\widetilde{\mathcal{G}}_{m+1}(f) = (2\mathcal{I} - \widehat{\mathcal{G}}_m)(f)$ for all rational functions in $\mathbb{S}_{2m+2}$. This rule is determined by the matrix $\widetilde{H}_{m+1}$, while the simplified rational anti-Gauss rule, $\breve{\mathcal{G}}_{m+1}(f)$, is defined by the matrix $\breve{H}_{m+1}$. These matrices have the same entries except for the last diagonal element. In the same way as in Subsection 4.1, we conclude that the simplified rational anti-Gauss rule satisfies

$$\breve{\mathcal{G}}_{m+1}(f) = \widetilde{\mathcal{G}}_{m+1}(f) = (2\mathcal{I} - \widehat{\mathcal{G}}_m)(f), \qquad \forall f \in \mathbb{S}_{2m+1}.$$

This shows (4.32). Property (4.31) follows from (4.32) since $\widehat{\mathcal{G}}_m(f) = \mathcal{I}(f)$ for $f \in \mathbb{S}_{2m}$. Property (4.33) can be shown similarly as Theorem 4. □

Similarly to (4.29), we define the average quadrature rule

$$\breve{\mathcal{A}}_{2m+1}(f) := \frac{1}{2}(\widehat{\mathcal{G}}_m + \breve{\mathcal{G}}_{m+1})(f).$$

It follows from Theorem 5 that this rule satisfies

$$\breve{\mathcal{A}}_{2m+1}(f) = \mathcal{I}(f), \quad f \in \mathbb{S}_{2m+1}.$$

5. Computed examples. In this section, we illustrate the performance of the rational Gauss rules when applied to Stieltjes matrix functions of a symmetric matrix. All computations for this paper were carried out using MATLAB R2017b on a 64-bit MacBook Pro personal computer with about 15 significant decimal digits.

The examples of this section compare the performance of standard Gauss and rational Gauss rules. We also illustrate in Example 5.4 that rational Gauss rules (4.5) with several distinct poles may give higher accuracy than rational Gauss rules with a single pole at the origin with the same number of nodes.

In all examples, when $m \geq k + 2s$, where $k + 2s$ is the degree of $w(y)$ defined by (4.4), we observed that the derivative $(w^2 f)^{(2m)}$ in (4.9) is positive in an interval that contains the spectrum of A, and the derivative $(w^2 f)^{(2m+1)}$ in (4.14) is negative in an interval that contains the spectrum A and θ. It follows from (4.16) that pairs of rational Gauss and rational Gauss–Radau rules with a fixed node at $\theta \geq \lambda_N$ do not bracket $F(A)$, while pairs of rational Gauss and rational Gauss–Radau rules with a fixed node at $\theta \leq \lambda_1$ give lower and upper bounds for $F(A)$, respectively. We also will illustrate that error bounds for certain functionals (1.1) can be computed by pairs of rational Gauss–Radau rules, and estimates of upper and lower bounds can be determined by pairs of rational Gauss and simplified rational anti-Gauss quadrature rules. To determine the quadrature error we explicitly evaluate the functionals (1.1). This limits the size of the matrices A considered. To compute the quadrature rule (4.30), we require the elementary basis function ψ_{m-1} to be a monomial. In all examples, we let the elementary basis functions ψ_m and ψ_{m+1} be monomials.

Example 5.1. Consider the Stieltjes function (1.3) with $a = 1/2$. We would like to approximate the functional

$$F(A) := \boldsymbol{v}^T A^{-1/2} \boldsymbol{v}, \tag{5.1}$$

where $A \in \mathbb{R}^{1000\times 1000}$ is a symmetric Toeplitz matrix with first row $[1, 1/2, \ldots, 1/1000]$, and $\boldsymbol{v} = [1/\sqrt{1000}, \ldots, 1/\sqrt{1000}]^T \in \mathbb{R}^{1000}$. The smallest eigenvalue of A is $\lambda_1 = 0.3863$ and the largest one is $\lambda_{1000} = 12.1259$. The value of $F(A)$ is approximately 0.2897. Approximations of (5.1) determined by standard and rational Gauss rules, rational Gauss–Radau rules, rational anti-Gauss and simplified rational anti-Gauss quadrature rules are presented. The computations require the solution of linear system of equations with the symmetric positive definite Toeplitz matrices $A - \alpha_i I$, where the α_i are poles. We remark that fast algorithms for the solution of systems of equations with this kind of matrix are available; see, e.g., [3].

Consider the rational Krylov subspace

$$\begin{aligned}\mathbb{K}_m(A, \boldsymbol{v}) = \operatorname{span}\Big\{ & \boldsymbol{v}, A\boldsymbol{v}, (A-\alpha_1 I)^{-1}\boldsymbol{v}, A^2\boldsymbol{v}, \ldots, (A-\alpha_1 I)^{-k_1}\boldsymbol{v}, A^{k_1+1}\boldsymbol{v}, \\ & (A-\alpha_2 I)^{-1}\boldsymbol{v}, A^{k_1+2}\boldsymbol{v}, \ldots, (A-\alpha_2 I)^{-k_2}\boldsymbol{v}, A^{k_1+k_2+1}\boldsymbol{v}, \ldots, \\ & (A-\alpha_\ell I)^{-1}\boldsymbol{v}, \ldots\ , (A-\alpha_\ell I)^{-k_\ell}\boldsymbol{v}, A^{k+1}\boldsymbol{v}\Big\},\end{aligned} \tag{5.2}$$

where k is determined by (3.2). The Stieltjes function (1.3) is defined in the complex plane except for on the interval $(-\infty, 0]$. It therefore is natural to allocate poles on this interval. We consider rational Krylov subspaces $\mathbb{K}_m(A, \boldsymbol{v})$ with poles allocated on $(-\infty, 0]$ in three different ways:

(i) $\mathbb{K}_6(A, \boldsymbol{v})$ is determined by a simple pole, $\alpha_1 = -1/2$ of multiplicity two.

(ii) $\mathbb{K}_8(A, \boldsymbol{v})$ is determined by two distinct poles, $\alpha_1 = -0.4310$ of multiplicity two and $\alpha_2 = -0.9024$ of multiplicity one. These poles are the zeros of the Chebyshev polynomial of the first kind of degree two for the interval $[-1, -1/3]$.

(iii) $\mathbb{K}_{10}(A, \boldsymbol{v})$ is determined by four equidistant poles $\alpha_i \in \{0, -1/2, -1, -3/2\}$ of multiplicity one.

The poles in $\mathbb{K}_6(A, \boldsymbol{v})$ are the zeros of the polynomial $w(y)$ defined by (4.4). We have

$$w(y) = y^2 + y + \frac{1}{4}.$$

In the same manner, we can define $w(y)$ associated with $\mathbb{K}_8(A, \boldsymbol{v})$ and $\mathbb{K}_{10}(A, \boldsymbol{v})$. We evaluate (5.1) as $\boldsymbol{v}^T A^{-1/2} \boldsymbol{v}$, where the vector $A^{-1/2}\boldsymbol{v}$ is calculated by first computing the matrix square root and then solving a linear system of equations. The square root of the matrix is computed with the MATLAB function sqrtm. Evaluation of the standard Gauss rule

$$\boldsymbol{e}_1^T T_m^{-1/2} \boldsymbol{e}_1$$

requires the computation of m steps of the standard Lanczos process. The rational Gauss rule is evaluated as

$$\boldsymbol{e}_1^T H_m^{-1/2} \boldsymbol{e}_1,$$

where $H_m^{-1/2}\boldsymbol{e}_1$ is determined by first computing the matrix square root and then solving a linear system of equations. Analogously to the rational Gauss rule, the simplified rational anti-Gauss, and Gauss–Radau rules with a fixed node $\theta \leq \lambda_1$ or $\theta \geq \lambda_{1000}$ can be computed by carrying out m steps of the Algorithm 1, while the rational anti-Gauss rule is determined by $m+1$ steps of Algorithm 1.

Columns 2 and 3 of Table 5.1 display the errors in approximations determined by standard and rational Gauss rules. We observe that rational Gauss rules yield higher accuracy than the standard Gauss rules. Column 4 of Table 5.1 displays the errors achieved with the rational Gauss–Radau rules. We chose the Radau point $\theta = 0.3$. A comparison with the errors in column 3 shows that pairs of rational Gauss, $\widehat{\mathcal{G}}_m(f)$, and rational Gauss–Radau rules, $\widehat{\mathcal{R}}_m^{0.3}(f)$, provide lower and upper bounds for (5.1), respectively.

Columns 5 and 7 of Table 5.1 display the errors in approximations obtained by rational anti-Gauss and simplified rational anti-Gauss rules, respectively. It can be seen that the errors of these quadrature rules are of opposite sign and of about the same magnitude as the errors in the corresponding rational Gauss rules. In this example, we chose the last diagonal entry of the matrix $\breve{H}_{m+1}$ that determines the simplified rational anti-Gauss rules to be $\breve{h} = h_{m-1,m-1}$. We also observe that the choice of $\breve{h} = (h_{m-1,m-1} + h_{m-2,m-2})/2$ yields similar results. For instance, we found for $m = 8$ that

$$F(A) - \breve{\mathcal{G}}_{m+1}(f) = -9.21 \cdot 10^{-11}.$$

This illustrates that the results achieved with simplified rational anti-Gauss rules are fairly insensitive to the choice of $\breve{h}$. Therefore, simplified rational anti-Gauss rules can be used to reduce the computational cost. Table 5.1 also shows that the pairs of rules $\{\widehat{\mathcal{G}}_m, \widetilde{\mathcal{G}}_{m+1}\}$ and $\{\widehat{\mathcal{G}}_m, \breve{\mathcal{G}}_{m+1}\}$ yield tighter error bounds than the pairs of $\{\widehat{\mathcal{G}}_m, \widehat{\mathcal{R}}_{m+1}^{\theta}\}$.

Columns 6 and 8 of Table 5.1 show the errors in computed approximations determined by the average rules associated with rational Gauss and anti-Gauss rules, and rational Gauss and simplified rational anti-Gauss rules, respectively. These quadrature rules yield more accurate approximations of (5.1) than the corresponding rational Gauss rules.

Table 5.2 displays the errors in the computed rational Gauss–Radau quadrature rules with fixed nodes at $\theta = 0.3$ and $\theta = 13$. The table illustrates that pairs of rational Gauss–Radau rules provide upper and lower bounds for (5.1); we have

$$\widehat{\mathcal{R}}_{m+1}^{0.3}(f) \geq F(A) \geq \widehat{\mathcal{R}}_{m+1}^{13}(f), \quad \forall m.$$

m	$F(A)-\mathcal{G}_m(f)$	$F(A)-\widehat{\mathcal{G}}_m(f)$	$F(A)-\widehat{\mathcal{R}}^{\theta}_{m+1}(f)$	$F(A)-\widetilde{\mathcal{G}}_{m+1}(f)$	$F(A)-\mathcal{A}_{2m+1}(f)$	$F(A)-\breve{\mathcal{G}}_{m+1}(f)$	$F(A)-\breve{\mathcal{A}}_{2m+1}(f)$
6	$5.79\cdot 10^{-7}$	$2.75\cdot 10^{-9}$	$-6.09\cdot 10^{-9}$	$-2.86\cdot 10^{-9}$	$-5.57\cdot 10^{-11}$	$-2.38\cdot 10^{-9}$	$1.85\cdot 10^{-10}$
8	$7.28\cdot 10^{-8}$	$3.95\cdot 10^{-11}$	$-1.16\cdot 10^{-10}$	$-4.10\cdot 10^{-11}$	$-7.65\cdot 10^{-13}$	$-3.45\cdot 10^{-11}$	$2.48\cdot 10^{-12}$
10	$9.20\cdot 10^{-9}$	$5.46\cdot 10^{-14}$	$-2.23\cdot 10^{-13}$	$-5.71\cdot 10^{-14}$	$-1.22\cdot 10^{-15}$	$-4.99\cdot 10^{-14}$	$2.38\cdot 10^{-15}$

TABLE 5.1

Example 5.1: Errors for computed approximations of $F(A)=\boldsymbol{v}^TA^{-1/2}\boldsymbol{v}$ with A a symmetric Toeplitz matrix. The Radau node is fixed at $\theta=0.3$

m	$F(A)-\widehat{\mathcal{R}}^{0.3}_{m+1}(f)$	$F(A)-\widehat{\mathcal{R}}^{13}_{m+1}(f)$
6	$-6.09\cdot 10^{-9}$	$2.21\cdot 10^{-9}$
8	$-1.16\cdot 10^{-10}$	$3.32\cdot 10^{-11}$
10	$-2.23\cdot 10^{-13}$	$4.61\cdot 10^{-14}$

TABLE 5.2

Example 5.1: Errors for computed approximations of $F(A)=\boldsymbol{v}^TA^{-1/2}\boldsymbol{v}$ by rational Gauss–Radau rules with A a symmetric Toeplitz matrix. The Radau nodes are $\theta=0.3$ and $\theta=13$.

□

Example 5.2. This example determines an approximation of the functional

$$F(A):=\boldsymbol{v}^T\log(A+I)A^{-1}\boldsymbol{v},$$

where $A\in\mathbb{R}^{1000\times 1000}$ is symmetric Toeplitz matrix with first row $[3,3/2,\ldots,3/1000]$. The vector $\boldsymbol{v}\in\mathbb{R}^{1000}$ and the rational Krylov subspaces $\mathbb{K}_6(A,\boldsymbol{v})$ and $\mathbb{K}_8(A,\boldsymbol{v})$ are defined to be the same as in Example 5.1. The subspace $\mathbb{K}_{10}(A,\boldsymbol{v})$ is determined by four equidistant poles $\alpha_i\in\{0,-1/4,-1/2,-1\}$ of multiplicity one. The smallest eigenvalue of A is $\lambda_1=1.1589$ and the largest one is $\lambda_{1000}=36.3776$. Thus, $F(A)$ is defined by the Stieltjes function

$$f(y)=\frac{\log(1+y)}{y}=\int_1^\infty\frac{t^{-1}}{t+y}\,dt.$$

The value of $F(A)$ is approximately 0.1009.

Columns 2 and 3 of Table 5.3 show the difference between the exact value and the approximations determined by the standard and rational Gauss rules. We note that the quadrature error for the rational Gauss rules is the smallest for all values of m. Column 4 of Table 5.3 displays the errors in approximations obtained by rational Gauss–Radau rules with a fixed node at $\theta=1.1$. The table illustrates that pairs of rational Gauss rules, $\widehat{\mathcal{G}}_m(f)$, and associated Gauss–Radau rules, $\widehat{\mathcal{R}}^{1.1}_{m+1}(f)$, bracket the exact value.

Columns 3 and 5 of Table 5.3 show the errors in the rational Gauss rules, $\widehat{\mathcal{G}}_m(f)$, and in rational anti-Gauss rules, $\widetilde{\mathcal{G}}_{m+1}(f)$, to have opposite sign and be of about the same magnitude. Similarly, Columns 3 and 7 of Table 5.3 show the errors in rational Gauss rules, $\widehat{\mathcal{G}}_m(f)$, and in simplified rational anti-Gauss rules, $\breve{\mathcal{G}}_{m+1}(f)$, with $\breve{h}=(h_{m-1,m-1}+h_{m-2,m-2})/2$ to be of opposite sign and of about the same magnitude. Columns 6 and 8 of Table 5.3 illustrate that the average rules yield the best approximations of $F(A)$.

Table 5.4 displays the errors in approximations obtained by rational Gauss–Radau quadrature rules. The table illustrates that

$$\widehat{\mathcal{R}}^{1.1}_{m+1}(f)\geq F(A)\geq\widehat{\mathcal{R}}^{37}_{m+1}(f),\quad\forall m.$$

□

m	$F(A)-\mathcal{G}_m(f)$	$F(A)-\widehat{\mathcal{G}}_m(f)$	$F(A)-\widehat{\mathcal{R}}^{\theta}_{m+1}(f)$	$F(A)-\widetilde{\mathcal{G}}_{m+1}(f)$	$F(A)-\mathcal{A}_{2m+1}(f)$	$F(A)-\breve{\mathcal{G}}_{m+1}(f)$	$F(A)-\breve{\mathcal{A}}_{2m+1}(f)$
6	$9.65\cdot 10^{-8}$	$1.88\cdot 10^{-9}$	$-7.92\cdot 10^{-9}$	$-1.91\cdot 10^{-9}$	$-1.57\cdot 10^{-11}$	$-3.13\cdot 10^{-9}$	$-6.25\cdot 10^{-10}$
8	$5.93\cdot 10^{-9}$	$1.32\cdot 10^{-11}$	$-3.98\cdot 10^{-11}$	$-1.33\cdot 10^{-11}$	$-8.45\cdot 10^{-14}$	$-2.01\cdot 10^{-11}$	$-3.44\cdot 10^{-12}$
10	$3.56\cdot 10^{-10}$	$1.99\cdot 10^{-13}$	$-5.21\cdot 10^{-13}$	$-2.01\cdot 10^{-13}$	$-1.05\cdot 10^{-15}$	$-2.97\cdot 10^{-13}$	$-4.87\cdot 10^{-14}$

TABLE 5.3

Example 5.2: Errors for computed approximations of $F(A) := \boldsymbol{v}^T \log(A+I)A^{-1}\boldsymbol{v}$ with A a symmetric Toeplitz matrix. The Radau node is $\theta = 1.1$.

m	$F(A)-\widehat{\mathcal{R}}^{1.1}_{m+1}(f)$	$F(A)-\widehat{\mathcal{R}}^{37}_{m+1}(f)$
6	$-7.92\cdot 10^{-9}$	$1.23\cdot 10^{-9}$
8	$-3.98\cdot 10^{-11}$	$8.60\cdot 10^{-12}$
10	$-5.21\cdot 10^{-13}$	$1.31\cdot 10^{-13}$

TABLE 5.4

Example 5.2: Errors for computed approximations of $F(A) := \boldsymbol{v}^T \log(A+I)A^{-1}\boldsymbol{v}$ by rational Gauss–Radau rules with A a symmetric Toeplitz matrix. The Radau nodes are fixed at $\theta = 1.1$ and $\theta = 37$.

Example 5.3. We would like to approximate the functional

$$F(A) := \boldsymbol{v}^T(\log(I+A))^{-1}\boldsymbol{v},$$

where the matrix $A \in \mathbb{R}^{1000\times 1000}$ is the same as in Example 5.2. The vector $\boldsymbol{v} \in \mathbb{R}^{1000}$ has normally distributed random entries with zero mean and is normalized to be of unit norm. The value of $F(A)$ is approximately 0.9472. We consider the rational Krylov subspace (5.2) with poles α_i in $(-\infty, 0]$ allocated in three different ways:

(i) $\mathbb{K}_8(A,\boldsymbol{v})$ is determined by two equidistant poles, $\alpha_1 = 0$ of multiplicity two and $\alpha_2 = -1/2$ of multiplicity one.

(ii) $\mathbb{K}_{10}(A,\boldsymbol{v})$ is determined by four equidistant poles $\alpha_i \in \{0,-1,-2,-3\}$ of multiplicity one.

(iii) $\mathbb{K}_{14}(A,\boldsymbol{v})$ is determined by three equidistant poles $\alpha_i \in \{0,-1/4,-1/2\}$ of multiplicity two.

Columns 2 and 3 of Table 5.5 report the errors in approximations determined by the standard and rational Gauss rules. Columns 4 and 5 of Table 5.5 show the approximations determined by rational Gauss–Radau and simplified rational anti-Gauss rules. The table illustrates that pairs of rational Gauss and Gauss–Radau rules with fixed node at $\theta = 1.1$, or pairs of rational Gauss and simplified rational anti-Gauss rules with $\breve{h} = h_{m-1,m-1}$, bracket the exact value. Column 6 of Table 5.5 illustrates that the average rules can yield much higher accuracy than rational Gauss and anti-Gauss rules. Table 5.6 shows that the values determined by rational Gauss–Radau quadrature rules bracket $F(A)$.

□

m	$F(A)-\mathcal{G}_m(f)$	$F(A)-\widehat{\mathcal{G}}_m(f)$	$F(A)-\widehat{\mathcal{R}}^{\theta}_{m+1}(f)$	$F(A)-\breve{\mathcal{G}}_{m+1}(f)$	$F(A)-\breve{\mathcal{A}}_{2m+1}(f)$
8	$1.10\cdot 10^{-3}$	$1.81\cdot 10^{-8}$	$-1.11\cdot 10^{-7}$	$-1.36\cdot 10^{-8}$	$2.24\cdot 10^{-9}$
10	$1.81\cdot 10^{-4}$	$2.18\cdot 10^{-12}$	$-4.13\cdot 10^{-12}$	$-1.80\cdot 10^{-12}$	$1.93\cdot 10^{-13}$
14	$5.84\cdot 10^{-6}$	$3.60\cdot 10^{-14}$	$-1.70\cdot 10^{-13}$	$-3.91\cdot 10^{-14}$	$-1.55\cdot 10^{-15}$

TABLE 5.5

Example 5.3: Errors for computed approximations of $F(A) := \boldsymbol{v}^T(\log(I+A))^{-1}\boldsymbol{v}$ with A a symmetric Toeplitz matrix. The Radau node is $\theta = 1.1$.

m	$F(A)-\widehat{\mathcal{R}}_{m+1}^{1.1}(f)$	$F(A)-\widehat{\mathcal{R}}_{m+1}^{37}(f)$
8	$-1.11\cdot 10^{-7}$	$1.14\cdot 10^{-8}$
10	$-4.13\cdot 10^{-12}$	$1.24\cdot 10^{-12}$
14	$-1.70\cdot 10^{-13}$	$2.29\cdot 10^{-14}$

TABLE 5.6

Example 5.3: Errors for computed approximations of $F(A) := \boldsymbol{v}^T(\log(I+A))^{-1}\boldsymbol{v}$ with A a symmetric Toeplitz matrix. The Radau nodes are $\theta = 1.1$ and $\theta = 37$.

Example 5.4. In our last example, we compute an approximation of

$$F(A) := \boldsymbol{v}^T(\pi(I+\sqrt{A})^{-1})\boldsymbol{v}, \tag{5.3}$$

where the matrix A is obtained from the discretization of the self-adjoint differential operator $L(u) = \frac{1}{10}u_{xx}+u_{yy}$ in the unit square. Each derivative is approximated by the standard three-point stencil with 40 equally spaced interior nodes in each space-dimension. Homogeneous boundary conditions are used. This gives a symmetric positive definite matrix $A \in \mathbb{R}^{1600\times 1600}$. The vector $\boldsymbol{v}$ is given by $\boldsymbol{v} = \boldsymbol{e}_1 \in \mathbb{R}^{1600}$. The extreme eigenvalues of A are $\lambda_1 = 0.0646$ and $\lambda_{1600} = 43.9354$. Define the Stieltjes function

$$f(y) = \frac{\pi}{1+\sqrt{y}} = \int_0^\infty \frac{1}{(y+t)}\left(\frac{\sqrt{t}}{1+t}\right)dt,$$

and consider the subspace (5.2) with a single pole, $\alpha_1 = -1/2$, of high multiplicity.

We compare the performance of the methods of this paper with rational Gauss rules that are presented in [19]. The latter rule is exact for Laurent polynomials, which are rational functions, whose only finite pole is at the origin, and it is known as a Gauss–Laurent quadrature rule. We will denote these rules by $\widehat{\mathcal{G}}_m^{\mathcal{L}}(f)$; they are described in [19]. Algorithm 1 requires the solution of linear systems of equations with the matrix $(A-\alpha_1 I)$, where $\alpha_1 = -1/2$. An algorithm for computing an approximation of (5.3) by Gauss–Laurent rules is presented in [18]. The computation of this rule requires the solution of linear systems of equations with the matrix A.

Columns 2 and 3 of Table 5.7 display the difference between the exact value, $F(A) \approx 0.5983$, and the approximations obtained by rational Gauss rules, $\widehat{\mathcal{G}}_m(f)$, and Gauss–Laurent rules, $\widehat{\mathcal{G}}_m^{\mathcal{L}}(f)$. We find that rational Gauss rules associated with the rational Krylov subspace (5.2) give higher accuracy than Gauss–Laurent rules for all values of m. Columns 3, 4, and 5 of Table 5.7 show the pairs $\{\widehat{\mathcal{G}}_m(f), \widehat{\mathcal{R}}_{m+1}^{0.05}(f)\}$ and $\{\widehat{\mathcal{G}}_m(f), \breve{\mathcal{G}}_{m+1}(f)\}$ to bracket $F(A)$. The average rules, which are displayed in column 6 of Table 5.7, are seen to be quite accurate.

Table 5.8 display the errors in approximations obtained by the rational Gauss–Radau rules. The table illustrates that

$$\widehat{\mathcal{R}}_{m+1}^{0.05}(f) \geq F(A) \geq \widehat{\mathcal{R}}_{m+1}^{45}(f), \quad \forall m.$$

□

6. Conclusion. This paper discusses the approximation of the expression $\boldsymbol{v}^T f(A)\boldsymbol{v}$, where A is a symmetric positive definite matrix and f is a Stieltjes function, by rational Gauss rules with preselected poles. Associated rational Gauss–Radau and anti-Gauss rules are introduced. Computed examples show that when the integrand f has singularities close to the spectrum of A, rational Gauss rules with poles at or close to these singularities give

m	$F(A)-\widehat{\mathcal{G}}_m^{\mathcal{L}}(f)$	$F(A)-\widehat{\mathcal{G}}_m(f)$	$F(A)-\widehat{\mathcal{R}}_{m+1}^{\theta}(f)$	$F(A)-\breve{\mathcal{G}}_{m+1}(f)$	$F(A)-\breve{\mathcal{A}}_{2m+1}(f)$
8	$1.70\cdot 10^{-5}$	$3.85\cdot 10^{-7}$	$-1.99\cdot 10^{-6}$	$-3.90\cdot 10^{-7}$	$-2.82\cdot 10^{-9}$
10	$3.17\cdot 10^{-6}$	$2.28\cdot 10^{-8}$	$-1.24\cdot 10^{-7}$	$-2.33\cdot 10^{-8}$	$-2.75\cdot 10^{-10}$
14	$9.77\cdot 10^{-8}$	$1.09\cdot 10^{-10}$	$-4.67\cdot 10^{-10}$	$-1.13\cdot 10^{-10}$	$-1.83\cdot 10^{-12}$

TABLE 5.7

Example 5.4: Errors for computed approximations of $F(A) := \boldsymbol{v}^T(\pi(I+\sqrt{A})^{-1})\boldsymbol{v}$ when A is a discretization of a differential operator. The Radau node is fixed at $\theta = 0.05$

m	$F(A)-\widehat{\mathcal{R}}_{m+1}^{0.05}(f)$	$F(A)-\widehat{\mathcal{R}}_{m+1}^{45}(f)$
8	$-1.99\cdot 10^{-6}$	$2.51\cdot 10^{-7}$
10	$-1.24\cdot 10^{-7}$	$1.51\cdot 10^{-8}$
14	$-4.67\cdot 10^{-10}$	$8.33\cdot 10^{-11}$

TABLE 5.8

Example 5.4: Errors for computed approximations of $F(A) := \boldsymbol{v}^T(\pi(I+\sqrt{A})^{-1})\boldsymbol{v}$ when A is a discretization of a differential operator. The Radau nodes are $\theta = 0.05$ and $\theta = 45$

higher accuracy than standard Gauss rules and Gauss–Laurent rules with a pole at the origin with the same number of nodes. The examples also illustrate that pairs of rational Gauss and Gauss–Radau rules, or pairs of rational Gauss and rational anti-Gauss rules, or simplified rational anti-Gauss rules, provide error bounds, or estimates of bounds.

Acknowledgment. The authors would like to thank the referees for comments. 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), pp. 577–602.

[2] H. Alqahtani and L. Reichel, *Generalized block anti-Gauss quadrature rules*, Numer. Math., 143 (2019), pp. 605–648.

[3] G. S. Ammar and W. B. Gragg, *Superfast solution of real positive definite Toeplitz systems*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 61–76.

[4] B. Beckermann and L. Reichel, *Error estimation and evaluation of matrix functions via the Faber transform*, SIAM J. Numer. Anal., 47 (2009), pp. 3849–3883.

[5] 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), pp. 144–163.

[6] C. Berg, *Stieltjes-Pick-Bernstein-Schoenberg and their connection to complete monotonicity*, in Positive Definite Functions. From Schoenberg to Space-Time Challenges, S. Mateu and E. Porcu, eds., Dept. of Mathematics, University Jaume I, Castellón de la Plana, Spain, 2008.

[7] D. Calvetti, L. Reichel, and F. Sgallari, *Application of anti-Gauss quadrature rules in linear algebra*, in Applications and Computation of Orthogonal Polynomials, W. Gautschi, G. H. Golub, and G. Opfer, eds., Birkhäuser, Basel, 1999, pp. 41–56.

[8] K. Deckers and A. Bultheel, *Recurrence and asymptotics for orthonormal rational functions on an interval*, IMA J. Numer. Anal., 29 (2009), pp. 1–23.

[9] K. Deckers and A. Bultheel, *The existence and construction of rational Gauss-type rules*, Appl. Math. Comput., 218 (2012), pp. 10299–10320.

[10] 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), pp. 1655–1684.

[11] A. Frommer, *Monotone convergence of the Lanczos approximations to matrix functions of Hermitian matrices*, Electron. Trans. Numer. Anal., 35 (2009), pp. 118–128.

[12] A. Frommer and M. Schweitzer, *Error bounds and estimates for Krylov subspace approximations of Stieltjes matrix functions*, BIT Numer. Math., 56 (2016), pp. 865–892.

[13] W. Gautschi, *Orthogonal Polynomials: Computation and Approximation*, Oxford University Press, Oxford, 2004.

[14] G. H. Golub and G. Meurant, *Matrices, Moments and Quadrature with Applications*, Princeton University Press, Princeton, 2010.
[15] A. A. Gonchar and G. López Lagomasino, *On Markov's theorem for multipoint Padé approximants*, Math. USSR Sb., 34 (1978), pp. 449–459.
[16] P. Henrici, *Applied and Computational Complex Analysis*, Vol. 2, Wiley, New York, 1977.
[17] 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.
[18] C. Jagels and L. Reichel, *Recursion relations for the extended Krylov subspace method*, Linear Algebra Appl., 434 (2011), pp. 1716–1732.
[19] C. Jagels and L. Reichel, *The structure of matrices in rational Gauss quadrature*, Math. Comp., 82 (2013), pp. 2035–2060.
[20] G. Keady, N. Khajohnsaksumeth, and B. Wiwatanapataphee, *On functions and inverses, both positive, decreasing and convex: and Stieltjes functions*, Cog. Math. Stat., 5 (2018), Art. ID 1477543.
[21] D. P. Laurie, *Anti-Gaussian quadrature formulas*, Math. Comp., 65 (1996), pp. 735–747.
[22] S. Massei and L. Robol, *Rational Krylov for Stieltjes matrix functions: Convergence and pole selection*, BIT Numer. Math., 61 (2021), pp. 237–273.
[23] M. S. Pranić and L. Reichel, *Generalized anti-Gauss quadrature rules*, J. Comput. Appl. Math., 284 (2015), pp. 235–243.
[24] M. S. Pranić and L. Reichel, *Recurrence relations for orthogonal rational functions*, Numer. Math., 123 (2013), pp. 629–642.
[25] M. S. Pranić and L. Reichel, *Rational Gauss quadrature*, SIAM J. Numer. Anal., 52, (2014), pp. 832–851.
[26] A. D. Sokol, *Real-variables characterization of generalized Stieltjes functions*, Expo. Math., 28 (2010), pp. 179–185.