BIT manuscript No.
(will be inserted by the editor)

Multiple orthogonal polynomials applied to matrix function evaluation

Hessah Alqahtani · Lothar Reichel

Received: date / Accepted: date

Abstract Multiple orthogonal polynomials generalize standard orthogonal polynomials by requiring orthogonality with respect to several inner products. This paper discusses an application to the approximation of matrix functions and presents quadrature rules that generalize the anti-Gauss rules proposed by Laurie.

Keywords multiple orthogonal polynomials · multiple quadrature · matrix functions

Mathematics Subject Classification (2000) 65D32 · 65F60

1 Introduction

We are concerned with the evaluation of expressions of the form

$$F(A) := W^T f(A) V, \tag{1.1}$$

where $A \in \mathbb{R}^{n\times n}$ is a large symmetric matrix, f is a function that is analytic on the convex hull of the spectrum of A, $W \in \mathbb{R}^{n\times r}$ and $V \in \mathbb{R}^{n\times r'}$ with $1 \leq r, r' \ll n$ are "block vectors", and the superscript ^{T} denotes transposition. When A is small, the expression (1.1) can be evaluated by first explicitly evaluating $f(A)$ and multiplying this matrix by W^T and V. Many algorithms for evaluating matrix functions $f(A)$ for small to moderately sized matrices A are described and analyzed by Higham [17]. We are interested in the situation when the matrix A is so large that the explicit computation of $f(A)$

H. Alqahtani
Department of Mathematics, Faculty of Science and Arts, King Abdulaziz University, Rabigh, Saudi Arabia, and Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: hfq1981@gmail.com

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

requires too much computational effort or computer memory to be practical or attractive.

When $r = r' = 1$ and $V = W$, approximations of the expression (1.1) can be computed by first carrying out a few, say $m \ll n$, steps of the symmetric Lanczos algorithm applied to A with initial vector V. This yields a symmetric tridiagonal matrix $T_m \in \mathbb{R}^{m\times m}$. A commonly used approximation of (1.1) can be expressed as $c\, e_1^T f(T_m) e_1$, where $e_1 = [1, 0, \ldots, 0]^T$ is the first axis vector and c is a suitable constant depending on the norm of V; see, e.g., Golub and Meurant [15] or [4,7] for details and extensions. When $r = r' > 1$, analogous approximations can be computed with the aid of the symmetric block Lanczos algorithm; see [7,15]. Also the situation when $V, W \in \mathbb{R}^{n\times r}$ with $V \neq W$ is discussed in [4,7,15].

The present paper is concerned with the situation when $r \neq r'$. Let for the moment $1 = r' < r$. We will use the notation $v = V$ and $W = [w_1, w_2, \ldots, w_r]$. Introduce the spectral factorization

$$A = U\Lambda U^T, \tag{1.2}$$

where the nontrivial entries of $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n] \in \mathbb{R}^{n\times n}$ are the eigenvalues of A and the matrix $U \in \mathbb{R}^{n\times n}$ is orthogonal. Substituting the factorization (1.2) into (1.1) yields the sums

$$\begin{aligned} F(A) &= [w_1^T f(A)v, w_2^T f(A)v, \ldots, w_r^T f(A)v]^T \\ &= [\widetilde{w}_1^T f(\Lambda)\widetilde{v}, \widetilde{w}_2^T f(\Lambda)\widetilde{v}, \ldots, \widetilde{w}_r^T f(\Lambda)\widetilde{v}]^T \\ &= \left[\sum_{j=1}^n f(\lambda_j)\widetilde{w}_{1,j}\widetilde{v}_j, \sum_{j=1}^n f(\lambda_j)\widetilde{w}_{2,j}\widetilde{v}_j, \ldots, \sum_{j=1}^n f(\lambda_j)\widetilde{w}_{r,j}\widetilde{v}_j\right]^T . \end{aligned} \tag{1.3}$$

Here $\widetilde{v} = [\widetilde{v}_1, \widetilde{v}_2, \ldots, \widetilde{v}_n]^T = U^T v$ and $\widetilde{w}_j = [\widetilde{w}_{j,1}, \widetilde{w}_{j,2}, \ldots, \widetilde{w}_{j,n}]^T = U^T w_j$ for $1 \le j \le r$. We formally write these sums as Stieltjes integrals associated with measures $d\omega_j$ with piecewise constant distribution functions ω_j, $1 \le j \le r$. Thus,

$$\mathcal{I}_j f = \int f(t)\, d\omega_j(t), \qquad 1 \le j \le r,$$

and

$$F(A) = [\mathcal{I}_1 f, \mathcal{I}_2 f, \ldots, \mathcal{I}_r f]^T . \tag{1.4}$$

This shows that our computational task is to determine approximations of integrals of the function f with respect to r measures. We note that the measures $d\omega_j$ may be indefinite and this may cause numerical difficulties, such as breakdown of recursion relations. Further comments on this can be found at the end of Section 4.

One approach to approximate $F(A)$ is to compute approximations of each one of the r integrals in (1.4) independently. This can be achieved by approximating each one of the expressions $w_j^T f(A)v$ by applying the nonsymmetric Lanczos algorithm to A with initial vectors w_j and v independently for $1 \le j \le r$; see [4,7]. However, computed examples reported in Section 4 show

that it may be possible to determine an approximation of $F(A)$ of the same quality with fewer matrix-vector product evaluations by computing an approximation by Gauss-type quadrature rules associated with multiple orthogonal polynomials. These polynomials satisfy certain orthogonality conditions with respect to several inner products or bilinear forms determined by the different measures. When the matrix A is large, the main computational effort for determining approximations of $F(A)$ is the evaluation of matrix-vector products with A. Therefore, reducing the number of these products is beneficial. We remark that the cost of evaluating a matrix-vector product depends on the size, sparsity, and structure of A. Our method described for approximating $F(A)$ is attractive to use when the evaluation of $F(A)$ is much more expensive than the computation of a few matrix-vector products with A, and when the latter computations dominate other computational work. This is the case when the matrix A is large.

Multiple orthogonal polynomials arise in quadrature problems when the same function has to be integrated with respect to several measures and one would like to use the same nodes for the quadrature rules for each measure; see, e.g., Ismail [18, Chapter 23]. They also emerge in simultaneous Padé approximation when several functions are to be approximated by rational functions with the same denominator; see, e.g., [5,9–11,20,21,24]. An application to illumination models is described by Borges [3]. We are not aware of applications of multiple orthogonal polynomials and associated Gauss-type quadrature rules to the evaluation of matrix functions of the form (1.1). It is the purpose of the present paper to discuss this application.

The evaluation of our approximation of $F(A)$ requires the use of a generalized Lanczos algorithm. Such an algorithm can be derived from the recursion relations of multiple orthogonal polynomials; see, e.g., [5,18,20,21,24] and Section 2. One also can derive a generalized Lanczos algorithm by using linear algebra techniques, only, without reference to multiple orthogonal polynomials. Derivations of the latter kind are described by Aliaga et al. [1] and Freund [12]. These authors focus on the handling of breakdowns of the recursion relations. It is well known that the standard nonsymmetric Lanczos algorithm may suffer from breakdowns or near breakdowns; see, e.g., [2,23,25] for discussions. This also holds for the generalized Lanczos algorithm. This algorithm can be applied to model reduction; see Freund [13] for a nice survey.

This paper is organized as follows. Section 2 reviews results about multiple orthogonal polynomials and associated Gauss-type quadrature rules. In particular, we discuss the recurrence relation for multiple orthogonal polynomials; the recurrence formula can be expressed with a lower Hessenberg matrix, whose lower bandwidth depends on the number of measures. Section 3 presents quadrature rules of anti-Gauss type for estimating the error in quadrature rule values determined with the Gauss-type rules discussed in Section 2. Laurie [19] introduced anti-Gauss rules for the estimation of the error in (standard) Gauss rules associated with a nonnegative measure on the real axis; see Notaris [22] for a nice recent survey of techniques for estimating the error in (standard) Gauss rules. We generalize anti-Gauss rules to be ap-

plicable to estimate the error in Gauss-type quadrature rules associated with multiple orthogonal polynomials. A few computed examples are presented in Section 4. Concluding remarks can be found in Section 5.

2 Multiple orthogonal polynomials

This section provides an overview of multiple orthogonal polynomials associated with r distinct measures on the real line. For more details on multiple orthogonal polynomials, we refer to [5,10,11,18,20,21].

Let $r \geq 1$ be an integer and let $d\omega_1, d\omega_2, \ldots, d\omega_r$ denote measures with support on (part of) the real axis. Introduce the multi-index $\mathbf{m} = [m_1, m_2, ..., m_r]$, whose entries are nonnegative integers. The length of $\mathbf{m}$ is defined as $|\mathbf{m}| = m_1 + m_2 + \ldots + m_r$. We are concerned with type II multiple orthogonal polynomials. They are commonly labeled with a multi-index. Thus, the type II multiple orthogonal polynomial $p_{\mathbf{m}}$ is a monic orthogonal polynomial of degree $\leq |\mathbf{m}|$, not identically equal to zero, that satisfies the orthogonality conditions

$$\left.\begin{array}{ll} \int p_{\mathbf{m}}(t)t^k d\omega_1(t) = 0, & k = 0, 1, \ldots, m_1 - 1, \\ \int p_{\mathbf{m}}(t)t^k d\omega_2(t) = 0, & k = 0, 1, \ldots, m_2 - 1, \\ \vdots & \vdots \\ \int p_{\mathbf{m}}(t)t^k d\omega_r(t) = 0, & k = 0, 1, \ldots, m_r - 1. \end{array}\right\} \tag{2.1}$$

We assume the measures to be such that we can define r inner products or bilinear forms for functions f and g that are sufficiently smooth on the union of the supports of the measures,

$$(f, g)_k = \mathcal{I}_k(fg), \qquad 1 \leq k \leq r. \tag{2.2}$$

For $r = 1$, the monic multiple orthogonal polynomials simplify to monic polynomials that are orthogonal with respect to one inner product. The orthogonality conditions (2.1) yield a linear system of $|\mathbf{m}|$ equations. There are $|\mathbf{m}|$ unknown coefficients $\beta_{k,\mathbf{m}}$, $k = 0, 1, \ldots, |\mathbf{m}| - 1$, of the polynomial

$$p_{\mathbf{m}}(t) = t^{|\mathbf{m}|} + \sum_{k=0}^{|\mathbf{m}|-1} \beta_{k,\mathbf{m}} t^k. \tag{2.3}$$

Here we tacitly assume that the measures $d\omega_1, d\omega_2, \ldots, d\omega_r$ are such that the polynomial (2.3) of degree $|\mathbf{m}|$ exists and is uniquely determined. We will use index vectors of the form

$$\mathbf{m} = [\underbrace{\ell + 1, \ell + 1, \ldots, \ell + 1}_{j}, \underbrace{\ell, \ell, \ldots, \ell}_{r-j}],$$

where $|\mathbf{m}| = \ell r + j$ for some integers $\ell \geq 0$ and $0 \leq j < r$. Then the polynomial $p_{\mathbf{m}}$ can be identified with the polynomial p_m with the simple index $m = \ell r + j$. We will write $p_m = p_{\mathbf{m}}$.

The monic multiple orthogonal polynomials satisfy a recurrence relation of the form

$$tp_m(t) = p_{m+1}(t) + \sum_{i=0}^{\min\{r,m\}} \alpha_{m,r-i} p_{m-i}(t), \qquad m \geq 0. \tag{2.4}$$

We define $p_0(t) \equiv 1$ and $p_j(t) \equiv 0$ for $j = -1, -2, \ldots, -r$. The recurrence relations for the first $m+1$ monic multiple orthogonal polynomials can be expressed as

$$t \begin{bmatrix} p_0(t) \\ p_1(t) \\ \vdots \\ p_{m-1}(t) \end{bmatrix} = H_m \begin{bmatrix} p_0(t) \\ p_1(t) \\ \vdots \\ p_{m-1}(t) \end{bmatrix} + p_m(t) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{2.5}$$

where the matrix $H_m \in \mathbb{R}^{m \times m}$ is lower Hessenberg and banded,

$$H_m := \begin{bmatrix} \alpha_{0,r} & 1 & 0 & \cdots & & & & \cdots & 0 \\ \alpha_{1,r-1} & \alpha_{1,r} & 1 & \ddots & & & & & \vdots \\ \vdots & & \ddots & & & & & & \\ \alpha_{r,0} & \cdots & \alpha_{r,r-1} & \alpha_{r,r} & 1 & 0 & & & \\ 0 & \alpha_{r+1,0} & \cdots & \alpha_{r+1,r-1} & \alpha_{r+1,r} & 1 & & & \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & & \vdots \\ & & & & & & & 1 & 0 \\ \vdots & & & \alpha_{m-2,0} & \cdots & \alpha_{m-2,r-1} & & \alpha_{m-2,r} & 1 \\ 0 & \cdots & \cdots & 0 & \alpha_{m-1,0} & \cdots & & \alpha_{m-1,r-1} & \alpha_{m-1,r} \end{bmatrix}.$$

The entries of the above matrix can be determined by evaluating certain inner products or bilinear forms. We have

$$p_{j+1}(t) = (t - \alpha_{j,r}) p_j(t) - \sum_{i=1}^{\min\{r,j\}} \alpha_{j,r-i} p_{j-i}(t), \qquad j = 0, 1, \ldots .$$

Let $j = \ell r + v$, $\ell = \lfloor j/r \rfloor$, and $v \in \{0, 1, 2, \ldots, r-1\}$. Then

$$\alpha_{j,0} = \frac{(tp_j, p_{\lfloor (j-r)/r \rfloor})_{v+1}}{(p_{j-r}, p_{\lfloor (j-r)/r \rfloor})_{v+1}}, \qquad j = r, r+1, \ldots , \tag{2.6}$$

and

$$\alpha_{j,k} = \frac{(tp_j - \sum_{i=0}^{k-1} \alpha_{j,i} p_{j-r+i}, p_{\lfloor (j-r+k)/r \rfloor})_{1+((v+k) \bmod r)}}{(p_{j-r+k}, p_{\lfloor (j-r+k)/r \rfloor})_{1+((v+k) \bmod r)}} \tag{2.7}$$

for $k = \max\{r-j, 1\}, \max\{r-j, 1\} + 1, \ldots, r$ and $j = 0, 1, \ldots$. Here $\lfloor t \rfloor$ denotes the integer part of $t \geq 0$; when t is negative $\lfloor t \rfloor$ is the smallest integer

bounded below by t. Having computed the entry $\alpha_{0,r}$ of H_m, the first row of equation (2.5) yields the polynomial p_1, and we can evaluate the entries $\alpha_{1,r}$ and $\alpha_{1,r-1}$ in the second row of H_m. The latter entries determine the polynomial p_2. Using (2.5), we can proceed in the same manner to determine the entries of H_m row by row until all entries are known. This is the Stieltjes procedure for computing multiple orthogonal polynomials. It is described, e.g., by Gautschi [14, §2.2.3.1] for the situation when $r = 1$.

Algorithm 1 Generation of multiple orthogonal polynomials.

1: **Input:** Symmetric matrix $A \in \mathbb{R}^{n\times n}$, initial block vectors $V = v_0 \in \mathbb{R}^n$,
2: $W = [w_{0,1}, w_{0,2}, \ldots, w_{0,r}] \in \mathbb{R}^{n\times r}$,
3: $v_j := w_{0,j} := [0, 0, \ldots, 0]^T \in \mathbb{R}^n$ for $j = -1, -2, \ldots, -r$,
4: $\alpha_{r-1,0} := 0$, $\alpha_{0,r} := v_0^T A w_{0,1}/v_0^T w_{0,1}$,
5: number of steps m (a multiple of r).
6: $v_1 := Av_0 - \alpha_{0,r} v_0$
7: **for** $\ell = 1$ **to** r
8: $w_{1,\ell} := Aw_{0,\ell} - \alpha_{0,r} w_{0,\ell}$
9: **end for**
10: **for** $j = 1$ **to** m
11: $s_j := \lfloor (j - r)/r \rfloor$
12: $\beta_{j,0} := v_{j-r}^T w_{s_j, 1+(j \bmod r)}$
13: **if** $\beta_{j,0} \neq 0$ **then** $\alpha_{j,0} := (Av_j, w_{s_j,1+(j \bmod r)})/\beta_{j,0}$ **else** $\alpha_{j,0} := 0$ **end**
14: **for** $k = 1$ **to** r
15: $s_{j+k} := \lfloor (j - r + k)/r \rfloor$
16: $\beta_{j,k} := v_{j-r+k}^T w_{s_{j+k}, 1+((j+k) \bmod r)}$
17: **if** $\beta_{j,k} \neq 0$ **then**
18: $\alpha_{j,k} := (Av_j - \sum_{i=0}^{k-1} \alpha_{j,i} v_{j-r+i}, w_{s_{j+k},1+((j+k) \bmod r)})/\beta_{j,k}$
19: **else**
20: $\alpha_{j,k} := 0$
21: **end**
22: **end for**
23: **for** $\ell = 1$ **to** r
24: $w_{j+1,\ell} := Aw_{j,\ell} - \alpha_{j,r} w_{j,\ell} - \sum_{i=1}^{\min\{r,j\}} \alpha_{j,r-i} w_{j-i,\ell}$
25: **end for**
26: $v_{j+1} := Av_j - \alpha_{j,r} v_j - \sum_{i=1}^{\min\{r,j\}} \alpha_{j,r-i} v_{j-i}$
27: **end for**
28: **Output:** Entries $\alpha_{0,r}, \alpha_{1,r-1}, \alpha_{1,r}, \ldots, \alpha_{r,0}, \ldots, \alpha_{r,r-1}, \alpha_{r,r}, \ldots$, and
29: $\alpha_{m-1,0}, \ldots, \alpha_{m-1,r-1}, \alpha_{m-1,r}$ of the lower Hessenberg matrix H_m.

The computations are summarized by Algorithm 1. Each execution of the k-loop yields one new coefficient of the matrix H_m. The matrix-vector products $Aw_{s_j,1+(j \bmod r)}$ on line 13 and $Aw_{s_{j+k},1+((j+k) \bmod r)}$ on line 18 of the algorithm already have been computed previously and do not have to be recomputed if they have been stored. Here the matrix A is the same as in (1.1). Thus, execution of the algorithm requires $(r+1)m$ matrix-vector product evaluations. For notational simplicity, we assume that m is a multiple of r in the algorithm. This restriction easily can be removed.

We approximate the integrals in (1.4) by Gauss-type quadrature rules associated with the multiple orthogonal polynomials $p_0, p_1, \ldots$ and each one of

the measures $d\omega_1, d\omega_2, \ldots, d\omega_r$. For each $1 \leq j \leq r$, the m-point Gauss-type rule associated with the measure $d\omega_j$ can be expressed as

$$\mathcal{G}_{j,m} f = \sum_{\ell=1}^{m} w_{j,\ell,m} f(t_\ell) \tag{2.8}$$

and satisfies

$$\mathcal{G}_{j,m} f = \mathcal{I}_j f \qquad \forall f \in \mathbb{P}_{m+m_j-1}, \tag{2.9}$$

where $\mathbb{P}_j$ denotes the set of polynomials of degree at most j, and m_j is the jth component of the multi-index $\mathbf{m}$ associated with m, i.e., $m = |\mathbf{m}|$ with $\mathbf{m} = [m_1, m_2, \ldots, m_r]$. The quadrature nodes $t_1, t_2, \ldots, t_m$ are independent of j; they are the eigenvalues of the matrix H_m. We will assume that the nodes are pairwise distinct. A sufficient condition for this to hold is that the measures $d\omega_1, d\omega_2, \ldots, d\omega_r$ form an AT system; see [3,5,6,8].

Following Milovanović and Stanić [20], we compute the weights by solving r linear systems of equations with the same matrix and different right-hand sides. Let $t_1, t_2, \ldots, t_m$ be the eigenvalues of H_m. The columns of the Vandermonde-like matrix

$$P = \begin{bmatrix} p_0(t_1) & p_0(t_2) & \cdots & p_0(t_m) \\ p_1(t_1) & p_1(t_2) & \cdots & p_1(t_m) \\ \vdots & \vdots & & \vdots \\ p_{m-1}(t_1) & p_{m-1}(t_2) & \cdots & p_{m-1}(t_m) \end{bmatrix}$$

are eigenvectors of H_m. Since the polynomials p_j are monic, all entries of the first row equal one. The weights for the quadrature rule associated with the measure $d\omega_k$ satisfy

$$P \begin{bmatrix} w_{j,1,m} \\ w_{j,2,m} \\ \vdots \\ w_{j,m,m} \end{bmatrix} = \begin{bmatrix} (p_0, p_0)_j \\ (p_1, p_0)_j \\ \vdots \\ (p_{m-1}, p_0)_j \end{bmatrix}, \qquad 1 \leq j \leq r.$$

These systems can be solved by fast solution methods. For instance, the solution method by Higham [16] first computes an UL factorization of P in $\mathcal{O}(m^2)$ arithmetic floating point operations (flops) and then solves the system of equations by backward and forward substitution also in $\mathcal{O}(m^2)$ flops. A different approach to computing the weights is described in [5].

We have so far focused on the situation when $r' = 1$ and $r \geq 1$ in (1.1). It is straightforward to extend the discussion to the situation when $r' > 1$. Let $V = [v_1, v_2, \ldots, v_{r'}] \in \mathbb{R}^{n \times r'}$ and $W = [w_1, w_2, \ldots, w_r] \in \mathbb{R}^{n \times r}$. Then (1.1) can be expressed as

$$F(A) = \begin{bmatrix} w_1^T f(A) v_1 & w_1^T f(A) v_2 & \cdots & w_1^T f(A) v_{r'} \\ w_2^T f(A) v_1 & w_2^T f(A) v_2 & \cdots & w_2^T f(A) v_{r'} \\ \vdots & \vdots & & \vdots \\ w_r^T f(A) v_1 & w_r^T f(A) v_2 & \cdots & w_r^T f(A) v_{r'} \end{bmatrix}.$$

Substituting the spectral factorization (1.2) of A into $f(A)$ shows that each term can be expressed as a Stieltjes integral, analogously to (1.4). We can estimate these integrals using multiple orthogonal polynomials for rr' weight functions.

3 Anti-Gauss-type quadrature rules

Laurie [19] introduced so-called anti-Gauss quadrature rules for estimating the error of Gauss quadrature rules associated with a nonnegative measure with support on (part of) the real axis. This section defines analogous rules for estimating the error in Gauss-type quadrature rules (2.8) associated with the measures $d\omega_1, d\omega_2, \ldots, d\omega_r$. We define $(m+1)$-point anti-Gauss quadrature rules $\widehat{\mathcal{G}}_{j,m+1}$, for $1 \leq j \leq r$, by requiring them to satisfy

$$(\mathcal{I}_j - \widehat{\mathcal{G}}_{j,m+1})f = -(\mathcal{I}_j - \mathcal{G}_{j,m})f \qquad \forall f \in \mathbb{P}_{m+m_j+1}.$$

In particular, $\widehat{\mathcal{G}}_{j,m+1}$ is exact for all polynomials in $\mathbb{P}_{m+m_j-1}$. Moreover, for each j,

$$\widehat{\mathcal{G}}_{j,m+1}f = (2\mathcal{I}_j - \mathcal{G}_{j,m})f \qquad \forall f \in \mathbb{P}_{m+m_j+1}. \tag{3.1}$$

Hence, $\widehat{\mathcal{G}}_{j,m+1}$ is a quadrature rule associated with the bilinear form

$$\langle f, g\rangle_j = (2\mathcal{I}_j - \mathcal{G}_{j,m})(fg), \qquad 1 \leq j \leq r.$$

We may define monic multiple orthogonal polynomials $\widehat{p}_0, \widehat{p}_1, \ldots,$ with respect to these bilinear forms. These polynomials satisfy a recursion relation analogous to (2.4),

$$t\widehat{p}_j(t) = \widehat{p}_{j+1}(t) + \sum_{i=0}^{\min\{r,j\}} \widehat{\alpha}_{j,r-i}\widehat{p}_{j-i}(t), \qquad j \geq 0. \tag{3.2}$$

We also let $\widehat{p}_0(t) \equiv 1$ and $\widehat{p}_j(t) \equiv 0$ for $j = -1, -2, \ldots, -r$. The recurrence relations for the first $m+2$ monic multiple orthogonal polynomials can be written in the form

$$t\begin{bmatrix} \widehat{p}_0(t) \\ \widehat{p}_1(t) \\ \vdots \\ \widehat{p}_m(t) \end{bmatrix} = \widehat{H}_{m+1}\begin{bmatrix} \widehat{p}_0(t) \\ \widehat{p}_1(t) \\ \vdots \\ \widehat{p}_m(t) \end{bmatrix} + \widehat{p}_{m+1}(t)\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{3.3}$$

where the matrix $\widehat{H}_{m+1} \in \mathbb{R}^{(m+1)\times(m+1)}$ is lower Hessenberg and banded,

$$\widehat{H}_{m+1} := \begin{bmatrix} \widehat{\alpha}_{0,r} & 1 & 0 & \cdots & & & & \cdots & 0 \\ \widehat{\alpha}_{1,r-1} & \widehat{\alpha}_{1,r} & 1 & \ddots & & & & & \vdots \\ \vdots & & \ddots & & & & & & \\ \widehat{\alpha}_{r,0} & \cdots & \widehat{\alpha}_{r,r-1} & \widehat{\alpha}_{r,r} & 1 & & & & \\ 0 & \widehat{\alpha}_{r+1,0} & \cdots & \widehat{\alpha}_{r+1,r-1} & \widehat{\alpha}_{r+1,r} & 1 & & & \\ \vdots & & \ddots & \ddots & \ddots & \ddots & & & \vdots \\ & & & & & & & 1 & 0 \\ \vdots & & & \widehat{\alpha}_{m-1,0} & \cdots & \widehat{\alpha}_{m-1,r-1} & \widehat{\alpha}_{m-1,r} & & 1 \\ 0 & \cdots & \cdots & 0 & \widehat{\alpha}_{m,0} & \cdots & \widehat{\alpha}_{m,r-1} & & \widehat{\alpha}_{m,r} \end{bmatrix}.$$

Let $j = \ell r + v$, $\ell = \lfloor j/r \rfloor$, and $v \in \{0, 1, 2, \ldots, r-1\}$. Then we obtain analogously to (2.6) and (2.7) that

$$\widehat{\alpha}_{j,0} = \frac{\langle t\widehat{p}_j, \widehat{p}_{\lfloor (j-r)/r \rfloor}\rangle_{v+1}}{\langle \widehat{p}_{j-r}, \widehat{p}_{\lfloor (j-r)/r \rfloor}\rangle_{v+1}} \qquad j = r, r+1, \ldots \tag{3.4}$$

and for $k = \max\{r-j, 1\}, \max\{r-j, 1\}+1, \ldots, r$ and $j = 0, 1, \ldots$,

$$\widehat{\alpha}_{j,k} = \frac{\langle t\widehat{p}_j - \sum_{i=0}^{k-1} \widehat{\alpha}_{j,i}\widehat{p}_{j-r+i}, \widehat{p}_{\lfloor (j-r+k)/r \rfloor}\rangle_{1+((v+k) \bmod r)}}{\langle \widehat{p}_{j-r+k}, \widehat{p}_{\lfloor (j-r+k)/r \rfloor}\rangle_{1+((v+k) \bmod r)}}. \tag{3.5}$$

For notational simplicity, we assume in the following that $m = sr$ for some positive integer s. Then $\mathbf{m} = [m_1, m_2, \ldots, m_r] = [s, s, \ldots, s]$

Theorem 3.1 *The entries of the matrix $\widehat{H}_{m+1}$ satisfy*

$$\widehat{\alpha}_{j,k} = \alpha_{j,k}, \qquad 1 \le j < m, \quad k \ge 1 \tag{3.6}$$

and

$$\widehat{\alpha}_{m,0} = 2\alpha_{m,0}. \tag{3.7}$$

Moreover,

$$\widehat{p}_j(t) = p_j(t), \qquad 0 \le j \le m. \tag{3.8}$$

The entries $\widehat{\alpha}_{m,k}$ can be computed for increasing index $k = 1, 2, \ldots, r-1$ by

$$\widehat{\alpha}_{m,k} = \frac{2(tp_m, p_{\lfloor (m-r+k)/r \rfloor})_{1+k} - \sum_{i=0}^{k-1} \widehat{\alpha}_{m,i}(p_{m-r+i}, p_{\lfloor (m-r+k)/r \rfloor})_{1+k}}{(p_{m-r+k}, p_{\lfloor (m-r+k)/r \rfloor})_{1+(k \bmod r)}} \tag{3.9}$$

and by

$$\widehat{\alpha}_{m,r} = \frac{2(tp_m, p_s)_1 - \sum_{i=0}^{r-1} \widehat{\alpha}_{m,i}(p_{m-r+i}, p_s)_1}{2(p_m, p_s)_1}. \tag{3.10}$$

Proof The degree of the polynomial in the numerator of the expression (3.5) for $j \leq m-1$ is bounded by $1+j+\lfloor j/r \rfloor \leq m+\lfloor (m-1)/r \rfloor = m+s-1$. The quadrature rule $\mathcal{G}_{1+((v+k) \bmod r),m}$ is exact for all polynomials of degree up to $m+s-1$. Therefore the numerator equals

$$\int \left(t\widehat{p}_j(t) - \sum_{i=0}^{k-1} \widehat{\alpha}_{j,i} \widehat{p}_{j-r+i}(t) \widehat{p}_{\lfloor (j-r+k)/r \rfloor}(t) \right) d\omega_{1+((v+k) \bmod r)}(t),$$

which equals the numerator of (2.7) provided that the polynomials $\widehat{p}_j$ are equal to the polynomials p_j. The denominator of (3.5) is an inner product or bilinear form of polynomials of lower degree. It follows that the quadrature rule $\mathcal{G}_{1+((v+k) \bmod r),m}$ is exact also for the denominator. Hence, the expressions (3.5) and (2.7) are equal if the polynomials $\widehat{p}_j$ and p_j in these expressions are equal. These polynomials are defined by recursion relations (3.2) and (2.4) with the same initial values. By using these recursions, we see that the polynomials are, indeed, equal, and (3.6) and (3.8) follow.

We turn to (3.7). It follows from the above reasoning that the denominators of the coefficients of $\widehat{\alpha}_{m,0}$ and $\alpha_{m,0}$ are equal. The numerator of $\widehat{\alpha}_{m,0}$ is given by

$$\begin{aligned}(2\mathcal{I}_{v+1} - \mathcal{G}_{v+1,m})(t\widehat{p}_m \widehat{p}_{\lfloor (m-r)/r \rfloor}) &= (2\mathcal{I}_{v+1} - \mathcal{G}_{v+1,m})(tp_m p_{\lfloor (m-r)/r \rfloor}) \\ &= 2\mathcal{I}_{v+1}(tp_m p_{\lfloor (m-r)/r \rfloor}),\end{aligned}$$

where the last equality follows from the fact that the nodes of $\mathcal{G}_{v+1,m}$ are the zeros of p_m. This shows (3.7).

Finally, by (3.5) and (3.8) we have, for $1 \leq k \leq r$,

$$\widehat{\alpha}_{m,k} = \frac{\langle tp_m - \sum_{i=0}^{k-1} \widehat{\alpha}_{m,i} p_{m-r+i}, p_{\lfloor (m-r+k)/r \rfloor} \rangle_{1+(k \bmod r)}}{\langle p_{m-r+k}, p_{\lfloor (m-r+k)/r \rfloor} \rangle_{1+(k \bmod r)}}. \tag{3.11}$$

Consider the denominator and assume that $k < r$. Then the degree of the polynomial $p_{m-r+k} p_{\lfloor (m-r+k)/r \rfloor}$ is bounded by $m+s-1$. Therefore, the quadrature rule $\mathcal{G}_{1+(k \bmod r),m}$ is exact for the polynomial $p_{m-r+k} p_{\lfloor (m-r+k)/r \rfloor}$. It follows that, for $1 \leq k < r$,

$$\begin{aligned}\langle p_{m-r+k}, p_{\lfloor (m-r+k)/r \rfloor} \rangle_{1+(k \bmod r)} &= \langle p_{m-r+k}, p_{\lfloor (m-r+k)/r \rfloor} \rangle_{1+k} \\ &= (p_{m-r+k}, p_{\lfloor (m-r+k)/r \rfloor})_{1+k}.\end{aligned} \tag{3.12}$$

For $k = r$ the denominator of (3.11) satisfies, with $m = sr$,

$$\langle p_m, p_s \rangle_1 = 2(p_m, p_s)_1. \tag{3.13}$$

We turn to the numerator of (3.11). We have

$$\begin{aligned}
&\langle tp_m - \sum_{i=0}^{k-1} \widehat{\alpha}_{m,i} p_{m-r+i}, p_{\lfloor (m-r+k)/r \rfloor} \rangle_{1+(k \bmod r)} \qquad (3.14)\\
&= \langle tp_m, p_{\lfloor (m-r+k)/r \rfloor} \rangle_{1+(k \bmod r)} - \langle \sum_{i=0}^{k-1} \widehat{\alpha}_{m,i} p_{m-r+i}, p_{\lfloor (m-r+k)/r \rfloor} \rangle_{1+(k \bmod r)}\\
&= 2(tp_m, p_{\lfloor (m-r+k)/r \rfloor})_{1+(k \bmod r)} - \sum_{i=0}^{k-1} \widehat{\alpha}_{m,i} (p_{m-r+i}, p_{\lfloor (m-r+k)/r \rfloor})_{1+(k \bmod r)},
\end{aligned}$$

where we have used the fact that, for $0 \leq i \leq r$,

$$\langle p_{m-r+i}, p_{\lfloor (m-r+k)/r \rfloor} \rangle_{1+(k \bmod r)} = (p_{m-r+i}, p_{\lfloor (m-r+k)/r \rfloor})_{1+(k \bmod r)},$$

which follows from the observation that the quadrature rules $\mathcal{G}_{1+(k \bmod r),m}$ are exact for the polynomials $p_{m-r+i} p_{\lfloor (m-r+k)/r \rfloor}$. Combining the relations (3.12)-(3.14) yields (3.9) and (3.10).

The anti-Gauss-type rules determined by the matrix $\widehat{H}_{m+1}$ can be evaluated similarly as the Gauss-type quadrature rules defined by the matrix H_m. We also define the average rules

$$\widehat{\mathcal{A}}_{j,m+1} f = \frac{1}{2} (\mathcal{G}_{j,m} + \widehat{\mathcal{G}}_{j,m+1}) f, \qquad 1 \leq j \leq r.$$

It follows from (3.1) that

$$\widehat{\mathcal{A}}_{j,m+1} f = \mathcal{I}_j f \qquad \forall f \in \mathbb{P}_{m+m_j+1}, \qquad 1 \leq j \leq r.$$

Analogous average rules have been introduced by Laurie [19] for the situation of only one nonnegative measure with support on the real axis (i.e., for $r = 1$).

4 Computed examples

We illustrate the application of Gauss-type quadrature rules associated with multiple orthogonal polynomials to the approximation of expressions of the form (1.1). We consider different matrices A and functions f in (1.1). The examples show that Gauss-type quadrature rules provide useful approximations and estimates of upper and lower bounds. We compare the computed approximations with the exact values determined by explicitly evaluating the matrix functions. All computations were carried out in MATLAB with about 15 significant decimal digits on a MacBook Pro laptop computer with a 2.6 GHz Intel Core i5 processor and 8 GB 1600 MHz DDR3 memory.

Table 4.1 Example 4.1: $\mathcal{I}f = W^T(I+A)^{-1}V$, A a symmetric Toeplitz matrix, $\mathcal{I}f = [4.8263, 4.7883]^T \cdot 10^{-3}$.

m	$w_1^T(I+A)^{-1}v$	$w_2^T(I+A)^{-1}v$	$W^T(I+A)^{-1}v$
3	$5.78 \cdot 10^{-4}$	$2.28 \cdot 10^{-3}$	$[-1.20, -5.54]^T \cdot 10^{-5}$
5	$2.16 \cdot 10^{-5}$	$1.00 \cdot 10^{-4}$	$[-1.26, -5.66]^T \cdot 10^{-5}$

Example 4.1 This example shows that the use of multiple orthogonal polynomials and associated quadrature rules to approximate (1.1) with $V = v \in \mathbb{R}^{200}$ and $W = [w_1, w_2] \in \mathbb{R}^{200\times 2}$ may require fewer matrix-vector product evaluations than application of the nonsymmetric Lanczos process to A with initial vector pairs $\{w_1, v\}$ and $\{w_2, v\}$, independently. Table 4.1 shows the error of computed approximations of the functional

$$F(A) := W^T(I+A)^{-1}V,$$

where $A \in \mathbb{R}^{200\times 200}$ is a symmetric pentadiagonal Toeplitz matrix with first row $[1, 2, 3, 1, 1, \ldots, 1]$. The vector $V = v$ has normally distributed entries with zero mean and is normalized to be of unit norm, and the block vector W is given by $W = [e_1 + 3/2e_2 + 2e_3, e_1 + 2e_2 + 3e_3] + E$, where $e_j = [0, \ldots, 0, 1, 0, \ldots, 0]^T \in \mathbb{R}^{200}$ denotes the jth axis vector and $E = [e, e]$ with $e = [0, 0, 0, 1, 1, \ldots, 1]^T \in \mathbb{R}^{200}$. The exact value of $F(A)$ is provided in the header of Table 4.1. The first column of the table shows the error in the computed approximation of the first component of $F(A)$ using the nonsymmetric Lanczos process. Specifically, m steps of this process are applied to A with initial vector pair $\{w_1, v\}$ and the connection with Gauss quadrature is exploited to determine an approximation of the first component of $F(A)$. A detailed description of the nonsymmetric Lanczos process and its application to the evaluation of matrix functions can be found in [7]. The evaluation of the approximation so determined requires $2m$ matrix-vector product evaluations with A. An approximation of the second component of $F(A)$ is computed similarly, i.e., by applying m steps of the nonsymmetric Lanczos process to A with initial vector pair $\{w_2, v\}$. The error of the approximations so determined are found in the second column of Table 4.1. The evaluation of this approximation also requires the evaluation of $2m$ matrix-vector products with A. Thus, the computation of the first two columns of the table demands $4m$ matrix-vector product evaluations. The last column of Table 4.1 displays the error in approximations of the components of $F(A)$ determined by carrying out m steps with Algorithm 1. This requires $3m + 3$ matrix-vector product evaluations with A. Thus, Algorithm 1 gives higher accuracy and requires a smaller number of matrix-vector product evaluations than two applications of the nonsymmetric Lanczos process. □

Example 4.2 We would like to compute approximations of the functional

$$F(A) := W^T(I+A)^{-1}V,$$

Table 4.2 Example 4.2: $\mathcal{I}f = W^T(I+A)^{-1}V$, A a symmetric Toeplitz matrix, $\mathcal{I}f = [4.038, 4.038]^T \cdot 10^{-1}$.

m	$(\mathcal{I}-\mathcal{G}_m)f$	$(\mathcal{I}-\widehat{\mathcal{G}}_{m+1})f$	$(\mathcal{I}-\widetilde{\mathcal{A}}_{m+1})f$
3	$[\ 9.85,\ \ 9.87]^T \cdot 10^{-5}$	$[-3.33, -3.33]^T \cdot 10^{-4}$	$[-1.17, -1.17]^T \cdot 10^{-4}$
4	$[-2.14, -1.49]^T \cdot 10^{-6}$	$[\ 5.13,\ \ 5.79]^T \cdot 10^{-6}$	$[\ 1.49,\ \ 2.14]^T \cdot 10^{-6}$
5	$[-0.21, -9.71]^T \cdot 10^{-7}$	$[\ 5.13,\ \ 6.32]^T \cdot 10^{-6}$	$[\ 2.55,\ \ 2.67]^T \cdot 10^{-6}$

Table 4.3 Example 4.3: $\mathcal{I}f = W^T(\sqrt{A})V$, A a symmetric Toeplitz matrix. $\mathcal{I}f = [3.7482, 3.7482]^T$

m	$(\mathcal{I}-\mathcal{G}_m)f$	$(\mathcal{I}-\widehat{\mathcal{G}}_{m+1})f$	$(\mathcal{I}-\widetilde{\mathcal{A}}_{m+1})f$
3	$[\ 4.44,\ \ 8.88]^T \cdot 10^{-16}$	$[-2.66, -2.22]^T \cdot 10^{-15}$	$[-0.11, -6.66]^T \cdot 10^{-16}$
4	$[\ 2.22,\ \ 2.66]^T \cdot 10^{-15}$	$[-4.44, -4.44]^T \cdot 10^{-16}$	$[\ 8.88,\ \ 0.11]^T \cdot 10^{-16}$

with a symmetric Toeplitz matrix $A \in \mathbb{R}^{100\times 100}$, whose first row is given by $[1/2, 1/2^2, \ldots, 1/2^{100}]$. The columns of $W \in \mathbb{R}^{100\times 2}$ have normally distributed entries with zero mean and are normalized to be of unit norm, and this also holds for the vector $V \in \mathbb{R}^{100}$. Table 4.2 shows the desired value $\mathcal{I}f$ at the header of the table and the errors in computed approximations determined by the use of Gauss-type, anti-Gauss-type, and average quadrature rules with the integrand $f(t) = 1/(1+t)$. The table shows the errors in $\widehat{\mathcal{G}}_{m+1}f$ and $\mathcal{G}_m f$ to have opposite sign component-wise, and each component to be of about same magnitude. The average rules are seen to be most accurate. The table shows the Gauss-type and anti-Gauss-type rules to furnish bounds for $\mathcal{I}f$. □

Example 4.3 We would like to compute approximations of the functional

$$F(A) := W^T(\sqrt{A})V,$$

with a symmetric Toeplitz matrix $A \in \mathbb{R}^{100\times 100}$ with first row $[1, 1/2, \ldots, 1/100]$. The vector $V \in \mathbb{R}^{100}$ has normally distributed entries with zero mean and is normalized to be of unit norm, and this also holds for the columns of $W \in \mathbb{R}^{100\times 2}$. Table 4.3 shows the errors in approximations determined by Gauss-type, anti-Gauss-type, and average quadrature rules for $m \in \{3, 4\}$. The desired value $\mathcal{I}f$ is shown in the header of Table 4.3 and lies between the values $\mathcal{G}_m f$ and $\widehat{\mathcal{G}}_{m+1}f$ for all components and all m. Moreover, the quadrature errors obtained with these rules are of about the same magnitude for all m. The most accurate quadrature rules are the average rule. □

We finally remark that for many expressions (1.1), the quadrature rules described in this paper provide useful and inexpensively computable approximations and error estimates. However, the sums (1.3) are bilinear forms that may have positive and negative coefficients $\widetilde{w}_{k,j}\widetilde{v}_j$. The recurrence relation (2.4) therefore is not guaranteed to exist for all degrees $m \geq 0$. Consequently, the banded lower Hessenberg matrix H_m in (2.5) may fail to exist, and then

the multiple Gauss-type rules cannot be evaluated as described. Recursion relations that are analogous to (2.4) also arise in the banded Lanczos algorithm and in this context remedies for the breakdown of the recursion formulas have been developed; see [1,12,13]. We are presently exploring how these remedies can be applied in the context of multiple Gauss-type and anti-Gauss quadrature rules.

5 Conclusion

The relation between a block Lanczos method for initial block vectors with different numbers of columns and multiple orthogonal polynomials is explored to compute approximations of certain matrix functions that may be interpreted as a vector or rectangular matrix of Stieltjes integrals. The connection between the block Lanczos method and multiple Gauss-type quadrature rules is exploited, and new anti-Gauss-type quadrature rules are developed. The latter help estimate the error in computed approximations determined by multiple Gauss-type rules. Computed examples illustrate the performance of the quadrature rules developed.

Acknowledgment

The authors would like to thank Ulises Fidalgo and Qiang Ye for valuable discussions, and the referees for comments. The research was supported in part by NSF grants DMS-1720259 and DMS-1729509.

References

1. J. I. Aliaga, D. L. Boley, R. W. Freund, and V. Hernández, A Lanczos-type method for multiple starting vectors, Math. Comp., 69 (1999), pp. 1577–1601.
2. Z. Bai, D. Day, and Q. Ye, ABLE: An adaptive block Lanczos method for non-Hermitian eigenvalue problems, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 1060–1082.
3. C. F. Borges, On a class of Gauss-like quadrature rules, Numer. Math., 67 (1994), pp. 271–288.
4. 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.
5. J. Coussement and W. Van Assche, Gaussian quadrature for multiple orthogonal polynomials, J. Comput. Appl. Math., 178 (2005), pp. 131–145.
6. E. Daems and A. B. J. Kuijlaars, Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions, J. Approx Theory, 146 (2007), pp. 91–114.
7. 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.
8. U. Fidalgo, S. Medina Peralta, and J. Mínguez Ceniceros, Mixed type multiple orthogonal polynomials: Perfectness and interlacing properties of zeros, Linear Algebra Appl., 438 (2013), pp. 1229–1239.
9. U. Fidalgo Prieto, J. Illán, and G. Lopéz Lagomasino, Hermite–Padé approximation and simultaneous quadrature formulas, J. Approx. Theory, 126 (2004), pp. 171–197.

10. U. Fidalgo Prieto and G. Lopéz Lagomasino, Rate of convergence of generalized Hermite–Padé approximants of Nikishin systems, Constr. Approx., 23 (2006), pp. 165–196.
11. U. Fidalgo Prieto and G. Lopéz Lagomasino, Nikishin systems are perfect, Constr. Approx., 34 (2011), pp. 297–356.
12. R. W. Freund, Band Lanczos methods, in Templets for the Solution of Algebraic Eigenvalue Problems, Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, eds., SIAM, Philadelphia, 2000, Sections 3.6 and 5.7.
13. R. W. Freund, Model reduction methods based on Krylov subspaces, Acta Numerica, 12 (2003), pp. 267–319.
14. W. Gautschi, Orthogonal Polynomials: Computation and Approximation, Oxford University Press, Oxford, 2004.
15. G. H. Golub and G. Meurant, Matrices, Moments and Quadrature with Applications, Princeton University Press, Princeton, 2010.
16. N. J. Higham, Stability analysis of algorithms for solving confluent Vandermonde-like systems, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 23–41.
17. N. J. Higham, Functions of Matrices: Theory and Computation, SIAM, Philadelphia, 2008
18. M. E. H. Ismail, Classical and Quantum Orthogonal Polynomials in One Variable, Cambridge University Press, Cambridge, 2005.
19. D. P. Laurie, Anti-Gauss quadrature formulas, Math. Comp., 65 (1996), pp. 739–747.
20. G. V. Milovanović and M. Stanić, Construction of multiple orthogonal polynomials by discretized Stieltjes–Gautschi procedure and corresponding Gaussian quadratures, Facta Univ. Ser. Math. Inform., 18 (2003), pp. 9–29.
21. G. V. Milovanović and M. Stanić, Multiple orthogonality and quadratures of Gaussian type, Rend. Circ. Mat. Palermo, Serie II, Suppl., 76 (2005), pp. 75–90.
22. S. E. Notaris, Gauss–Kronrod quadrature formulae – a survey of fifty years of research, Electron. Trans. Numer. Anal., 45 (2016), pp. 371–404.
23. Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd ed., SIAM, Philadelphia, 2003.
24. W. Van Assche, Padé and Hermite–Padé approximation and orthogonality, Surv. Approx. Theory, 2 (2006), pp. 61–91.
25. Q. Ye, A breakdown-free variation of the nonsymmetric Lanczos algorithm, Math. Comp., 62 (1994), pp. 179–207.