Numerische Mathematik (2019) 143:605–648
https://doi.org/10.1007/s00211-019-01069-z

Numerische Mathematik

Generalized block anti-Gauss quadrature rules

Hessah Alqahtani[1,2] · **Lothar Reichel**[2]

Received: 7 June 2018 / Revised: 3 April 2019 / Published online: 30 July 2019

Abstract
Golub and Meurant describe how pairs of Gauss and Gauss–Radau quadrature rules can be applied to determine inexpensively computable upper and lower bounds for certain real-valued matrix functionals defined by a symmetric matrix. However, there are many matrix functionals for which their technique is not guaranteed to furnish upper and lower bounds. In this situation, it may be possible to determine upper and lower bounds by evaluating pairs of Gauss and anti-Gauss rules. Unfortunately, it is difficult to ascertain whether the values determined by Gauss and anti-Gauss rules bracket the value of the given real-valued matrix functional. Therefore, generalizations of anti-Gauss rules have recently been described, such that pairs of Gauss and generalized anti-Gauss rules may determine upper and lower bounds for real-valued matrix functionals also when pairs of Gauss and (standard) anti-Gauss rules do not. The available generalization requires the matrix that defines the functional to be real and symmetric. The present paper reviews available anti-Gauss and generalized anti-Gauss rules and extends them in several ways that allow applications in new situations. In particular, the genarlized anti-Gauss rules for a real-valued non-negative measure described in Pranić and Reichel (J Comput Appl Math 284:235–243, 2015) are extended to allow the estimation of the error in matrix functionals defined by a non-symmetric matrix, as well as to matrix-valued matrix functions. Modifications that give simpler formulas and thereby make the application of the rules both easier and applicable to a larger class of problems also are described.

Mathematics Subject Classification 65D32 · 65F60

✉ Lothar Reichel
reichel@math.kent.edu

Hessah Alqahtani
hfq1981@gmail.com

[1] Department of Mathematics, Faculty of Science and Arts, King Abdulaziz University, Rabigh, Saudi Arabia

[2] Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

1 Introduction

This paper reviews available methods and presents new ones for computing *bounds* or *estimates of bounds* for matrix functions of the form

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

where $A \in \mathbb{R}^{n\times n}$ is a large matrix and $V \in \mathbb{R}^{n\times k}$ is a tall skinny matrix that consists of $1 \leq k \ll n$ orthonormal columns. We will frequently refer to tall and skinny matrices as "block vectors" in order to distinguish them from large square matrices A. We derive expressions that yield upper and lower bounds for (1.1) when an expansion of $f(t)$ in terms of certain orthogonal or bi-orthogonal polynomials, defined by a measure that is determined by A and V, converges to $f(A)$ sufficiently quickly with increasing number of terms. However, generally, it is difficult to assess whether convergence is sufficiently fast to secure that the computed quantities furnish upper and lower bounds. We therefore refer to the expressions derived as estimates of upper and lower bounds.

The superscript ^{T} in (1.1) denotes transposition and the function f is assumed to be such that $f(A)$ is well defined; it suffices that f is analytic in an open simply connected set in the complex plane that contains the spectrum of A in its interior; see, e.g., Higham [27] for several ways to define $f(A)$. Matrix functions of the form (1.1) arise, e.g., in network analysis, when solving linear discrete ill-posed problems by Tikhonov regularization, and in certain solution methods for partial differential equations; see, e.g., [3,6,10,18,22,29] and references therein. Functions that appear in these applications include

$$f(t) := \exp(\alpha t), \quad f(t) := \cosh(t), \quad f(t) := \frac{1}{t+\beta},$$

where α and β are parameters. In some applications approximations of (1.1) are required for several values of the parameters α and β.

First consider the situation when $A \in \mathbb{R}^{n\times n}$ is a large symmetric matrix and $k > 1$. The evaluation of a function, such as the exponential, of a large matrix may be prohibitively expensive. Golub and Meurant [22] propose that an approximation of (1.1) be computed by application of a few steps of the symmetric block Lanczos algorithm to A with initial block vector V. This yields a fairly small block tridiagonal matrix T at which the function f has to be evaluated. Generally, T is small enough so that $f(T)$ can be computed fairly quickly, e.g., by using one of the methods described by Higham [27]. The computed approximation so obtained can be interpreted as a block Gauss quadrature rule; see Golub and Meurant [22] or Sect. 3 for details. This approach of approximating (1.1) is very attractive. However, it may be difficult to determine error estimates except in special situations and, therefore, it can be difficult to decide how many steps of the symmetric block Lanczos algorithm to carry out to determine an approximation of (1.1) of desired accuracy. Having a reliable estimate of the quadrature error makes it possible to assess how many steps of the block Lanczos algorithm should be carried out. Keeping the number of steps small, and thereby also keeping the number of matrix-block-vector products with the matrix A small,

is important when A is large, because then the evaluation of the matrix-block-vector products constitutes the dominating computational effort to evaluate an approximation of (1.1). We note that the simple and seemingly natural approach to estimate the error of a Gauss rule by evaluating the difference between two consecutive Gauss rules and using this difference as an error estimate may yield a very inaccurate error estimate. This is illustrated in [13] for Gauss quadrature rules defined by a real-valued positive measure. This suggests that the quadrature error should be estimated in a different manner.

Fenu et al. [18] describe block anti-Gauss quadrature rules that, when paired with a suitable block Gauss rule, give component-wise upper and lower bounds for (1.1) for certain functions f, matrices A, and block vectors V. These matrix-valued block anti-Gauss rules generalize the real-valued anti-Gauss rules introduced by Laurie [30] for the approximation of integrals with respect to a real-valued positive measure on the real axis. We remark that a reason for Laurie to define real-valued anti-Gauss rules is to be able to estimate the quadrature error in the associated Gauss rule. Anti-Gauss rules have the advantage over Gauss–Kronrod rules, which also are used for the same purpose, that they exist also when Gauss–Kronrod rules do not, and they are simpler to compute; see Notaris [33] for a nice recent survey of error estimation methods for real-valued Gauss quadrature rules associated with a positive measure on the real axis.

While it is possible to provide conditions that secure that pairs of suitable block Gauss and anti-Gauss rules yield component-wise upper and lower bounds for (1.1), it is difficult to verify whether these conditions hold for a given matrix function. Therefore, for block size $k = 1$, a generalization of real-valued anti-Gauss quadrature rules was presented in [36] that, together with a suitable real-valued Gauss rule, may bracket the functional (1.1) also when pairs of Gauss and (standard) anti-Gauss rules do not. It is one of the aims of this paper to describe block analogues of the generalized real-valued anti-Gauss rules discussed in [36].

Each step of the symmetric block Lanczos algorithm applied to a symmetric matrix $A \in \mathbb{R}^{n\times n}$ requires the evaluation of the product of A and a block vector of order $n \times k$. These matrix-block-vector product computations constitute the dominating computational effort for computing approximations of (1.1) when the matrix A is large. Every step of the symmetric block Lanczos algorithm determines a new block diagonal entry and a new block subdiagonal entry of a block tridiagonal matrix. The latter matrix defines a block Gauss quadrature rule and is also used to construct block tridiagonal matrices that define (standard) block anti-Gauss and generalized block anti-Gauss quadrature rules. The function f has to be evaluated at these block tridiagonal matrices; see below. After application of m steps of the symmetric block Lanczos algorithm, m diagonal blocks and m subdiagonal blocks of the block tridiagonal matrix are known. Commonly, the last subdiagonal block is discarded to obtain a square block tridiagonal matrix; see, e.g., [18]. We propose to include the last subdiagonal block in the quadrature rule and "guess" or estimate a suitable last diagonal block to obtain a square block tridiagonal matrix. This approach yields quadrature rules that are exact for polynomials of higher degree than the corresponding rules obtained when the last subdiagonal block is discarded. We refer to the block quadrature rules determined in this manner as "simplified," because they may be considered simplifications of Gauss quadrature rules that can be determined after execution of $m + 1$ steps of

the symmetric block Lanczos algorithm. Simplified anti-Gauss and block anti-Gauss quadrature rules have recently been described in [2]. The present paper introduces simplified generalized block anti-Gauss rules. The latter rules are simpler to derive and use than the corresponding "unsimplified" generalized block anti-Gauss rules, which are also described in this paper.

We remark that when A is symmetric, the block size k is one, and derivatives of the function f in (1.1) do not change sign in the convex hull of the spectrum of A, upper and lower bounds for the functional (1.1) can be determined by evaluating pairs of Gauss and suitable Gauss–Radau rules. This is described by Golub and Meurant [22]. The present paper is concerned with the computation of error estimates in situations when the approach by Golub and Meurant [22] is not guaranteed to yield upper and lower bounds for (1.1).

So far, we have discussed expressions (1.1) with a symmetric matrix A. This paper also considers functions of the form

$$F(A) := W^T f(A) V, \tag{1.2}$$

where the matrix $A \in \mathbb{R}^{n\times n}$ is allowed to be nonsymmetric, and the block vectors $V, W \in \mathbb{R}^{n\times k}$ are biorthonormal, i.e., $W^T V = I_k$, where I_k denotes the identity matrix of order k. An approximation of the function (1.2) can be computed by applying a few steps of the nonsymmetric block Lanczos algorithm to A with initial block vectors V and W. The approximation so obtained can be interpreted as a block Gauss quadrature rule. Fenu et al. [18] discuss this approach of approximating (1.2) and describe associated block anti-Gauss quadrature rules. This paper introduces generalized anti-Gauss rules that can be applied to bracket the expression (1.2) in a fairly inexpensive manner. These rules generalize those described in [36]. Also simplified versions of these quadrature rules are discussed.

This paper is organized as follows. Section 2 considers the approximation of the functions (1.1) and (1.2) when the block size k is one. A modification of the generalized anti-Gauss rules described in [36] for (1.1) with a symmetric matrix A is described. It allows one more moment to be matched for the same number of steps by the symmetric Lanczos algorithm than the method in [36] and, therefore, may yield higher accuracy for about the same computational effort. Analogous generalized anti-Gauss quadrature rules associated with the matrix function (1.2) with A nonsymmetric, and V and W vectors, are also described. Section 3 discusses block analogues of the quadrature rules considered in Sect. 2. An analysis of when pairs of block Gauss and generalized block anti-Gauss rules bracket the exact value is presented in Sect. 4, and a few computed examples are described in Sect. 5, including applications to network analysis. Concluding remarks can be found in Sect. 6.

2 Scalar-valued generalized anti-Gauss rules

This section first considers approximation of the functional (1.1) when $k = 1$ and the matrix A is symmetric, and subsequently discusses approximation of the functional

(1.2) with $k = 1$ and a nonsymmetric matrix A. In this section V in (1.1) is a unit vector, and V and W in (1.2) are biorthonormal vectors.

2.1 Scalar-valued functionals with a symmetric matrix

Let the symmetric matrix $A \in \mathbb{R}^{n\times n}$ have the spectral factorization

$$A = S\Lambda S^T, \quad \Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n], \tag{2.1}$$

where the eigenvector matrix $S \in \mathbb{R}^{n\times n}$ is orthogonal and the $\lambda_i \in \mathbb{R}$ are eigenvalues. We assume A to be so large that the computation of its spectral factorization is unfeasible or undesirable. The factorization (2.1) sheds light on properties of the approximation methods for (1.1), but does not have to be computed.

Substituting the spectral factorization (2.1) into (1.1) yields

$$F(A) = V^T Sf(\Lambda)S^T V = \sum_{j=1}^{n} f(\lambda_j)V^T Se_j e_j^T S^T V,$$

where $e_j = [0, \ldots, 0, 1, 0, \ldots, 0]^T \in \mathbb{R}^n$ denotes the jth axis vector. The sum on the right-hand side is a Stieltjes integral determined by a nondecreasing piece-wise constant distribution function ω defined on the real axis with a jump $(S^T V e_j)^2$ at λ_j for $1 \le j \le n$. We may write the sum as an integral,

$$\mathcal{I}f := \int f(x)d\omega(x), \tag{2.2}$$

with a measure $d\omega$ of total mass one.

Introduce the inner product

$$(f, g) := \mathcal{I}(fg)$$

for polynomials f and g of suitably low degrees, and let $\{p_j\}_{j=0,1,2,\ldots}$ denote the sequence of orthonormal polynomials with respect to this inner product. Thus, p_j is of degree j with a positive leading coefficient, and

$$(p_j, p_k) = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases}$$

The polynomials p_j satisfy a three-term recursion relation, which for the polynomials $p_0, p_1, \ldots, p_m$ can be expressed as

$$x \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{m-1}(x) \end{bmatrix} = T_m \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{m-1}(x) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ p_m(x) \end{bmatrix},$$

where T_m is a symmetric tridiagonal matrix defined by the recurrence coefficients for the orthonormal polynomials p_j,

$$T_m := \begin{bmatrix} \alpha_1 & \beta_1 & & & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & & & \\ & \beta_2 & \alpha_3 & \beta_3 & & & \\ & & \beta_3 & & \ddots & & \\ & & & \ddots & \ddots & \beta_{m-2} & \\ & & & & \beta_{m-2} & \alpha_{m-1} & \beta_{m-1} \\ 0 & & & & & \beta_{m-1} & \alpha_m \end{bmatrix} \in \mathbb{R}^{m\times m}. \tag{2.3}$$

This matrix can be computed by the symmetric Lanczos algorithm, which is described by Algorithm 1. Each step with this algorithm requires the evaluation of one matrix-vector product with the matrix A. Iteration j determines the entries α_j and β_j of the matrix (2.3). In applications, $m \ll n$. Moreover, we assume that m is small enough so that breakdown in Algorithm 1 does not occur within the first m steps and, therefore, the matrix (2.3) exists. Breakdown is rare, but fortuitous, because in case of breakdown the Gauss rule is exact for all polynomials. We will not dwell on this unusual situation further.

We note that the orthonormal polynomials p_j are implicitly defined by the relation $v_j = p_{j-1}(A)v_1$, where the v_j are determined by Algorithm 1. In particular, the vectors v_j and polynomials p_j satisfy the same recurrence relation.

Algorithm 1 The symmetric Lanczos process.

Input: symmetric matrix $A \in \mathbb{R}^{n\times n}$, initial unit vector $v \in \mathbb{R}^n$,
number of steps m.
$v_0 := 0 \in \mathbb{R}^n$, $\beta_0 := 0$, $v_1 := v$
for $j = 1$ **to** m
  $w := Av_j - v_{j-1}\beta_{j-1}$
  $\alpha_j := v_j^T w$
  $w := w - v_j\alpha_j$
  $\beta_j := \|w\|$; $v_{j+1} := w/\beta_j$
end for
Output: Entries $\alpha_1, \alpha_2, \ldots, \alpha_m$ and $\beta_1, \beta_2, \ldots, \beta_m$ of the matrix (2.3).

Let the function f be continuous on the convex hull of the spectrum of A. The m-point Gauss quadrature rule associated with the integral operator (2.2) can be written as

$$\mathcal{G}_m f = e_1^T f(T_m)e_1, \tag{2.4}$$

and is characterized by the property that

$$\mathcal{I} f = \mathcal{G}_m f \quad \forall f \in \mathbb{P}_{2m-1}, \tag{2.5}$$

where $\mathbb{P}_{2m-1}$ denotes the set of all polynomials of degree at most $2m-1$. The Gauss rule (2.4) can also be expressed as

$$\mathcal{G}_m f = \sum_{i=1}^{m} f(x_i^{(\mathcal{G})}) w_i^{(\mathcal{G})}, \tag{2.6}$$

where the nodes $x_i^{(\mathcal{G})}$ are the eigenvalues of T_m and the weights $w_i^{(\mathcal{G})}$ are the squares of the first components of normalized eigenvectors of T_m; it is well known that the eigenvalues of T_m are pairwise distinct, see, e.g., [21,22]. The equivalence of the expressions (2.4) and (2.6) follows by substituting the spectral factorization of T_m into (2.4). The nodes and weights in (2.6) can be computed efficiently with the Golub–Welsch algorithm [20–22] or by a method described by Laurie [31].

The following $(m+\ell)$-point quadrature rules for $\ell \geq 2$ were defined in [36]. They generalize the $(m+1)$-point anti-Gauss rule introduced by Laurie [30], which is obtained when $\ell = 1$.

Definition 2.1 The generalized anti-Gauss quadrature rule with nodes $\widetilde{x}_i^{(\ell)}$ and weights $\widetilde{w}_i^{(\ell)}$, $i = 1, 2, \ldots, m+\ell$,

$$\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f = \sum_{i=1}^{m+\ell} f(\widetilde{x}_i^{(\ell)}) \widetilde{w}_i^{(\ell)} \tag{2.7}$$

is an $(m+\ell)$-point quadrature rule such that

$$(\mathcal{I} - \widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}) f = -(\mathcal{I} - \mathcal{G}_m) f \quad \forall f \in \mathbb{P}_{2m+2\ell-1}. \tag{2.8}$$

It follows from (2.5) and (2.8) that

$$\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f = \mathcal{I} f \quad \forall f \in \mathbb{P}_{2m-1}. \tag{2.9}$$

We can express (2.8) as

$$\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f = (2\mathcal{I} - \mathcal{G}_m) f \quad \forall f \in \mathbb{P}_{2m+2\ell-1}, \tag{2.10}$$

which shows that $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}$ may be considered an $(m+\ell)$-point Gauss quadrature rule associated with the functional $2\mathcal{I} - \mathcal{G}_m$. This functional is said to be quasi-definite if the determinants of all Hankel matrices defined by the moments of the functional are either positive or negative; see, e.g., [12,34,35] for discussions on bilinear forms defined by a quasi-definite functional. When $2\mathcal{I} - \mathcal{G}_m$ is quasi-definite, there is a family of orthonormal polynomials $\{\widetilde{p}_i\}_{i=0,1,2,\ldots}$ associated with the bilinear form

$$\langle f, g \rangle := (2\mathcal{I} - \mathcal{G}_m)(fg);$$

see [12,34,35]. Thus,

$$\langle \widetilde{p}_j, \widetilde{p}_i \rangle = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases}$$

Assuming that the polynomials $\widetilde{p}_0, \widetilde{p}_1, \ldots, \widetilde{p}_{m+\ell}$ exist, they satisfy a three-term recursion relation, which can be written as

$$x \begin{bmatrix} \widetilde{p}_0(x) \\ \widetilde{p}_1(x) \\ \vdots \\ \widetilde{p}_{m+\ell-1}(x) \end{bmatrix} = \widetilde{T}_{m+\ell}^{(\ell)} \begin{bmatrix} \widetilde{p}_0(x) \\ \widetilde{p}_1(x) \\ \vdots \\ \widetilde{p}_{m+\ell-1}(x) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \widetilde{p}_{m+\ell}(x) \end{bmatrix},$$

where

$$\widetilde{T}_{m+\ell}^{(\ell)} = \begin{bmatrix} \widetilde{\alpha}_1 & \widetilde{\beta}_1 & & & & & & 0 \\ \widetilde{\beta}_1 & \widetilde{\alpha}_2 & \widetilde{\beta}_2 & & & & & \\ & \widetilde{\beta}_2 & \widetilde{\alpha}_3 & \widetilde{\beta}_3 & & & & \\ & & \widetilde{\beta}_3 & & \ddots & & & \\ & & & \ddots & \ddots & \widetilde{\beta}_{m-2} & & \\ & & & & \widetilde{\beta}_{m-2} & \widetilde{\alpha}_{m-1} & \widetilde{\beta}_{m-1} & \\ & & & & & \widetilde{\beta}_{m-1} & \ddots & \\ & & & & & \ddots & \ddots & \widetilde{\beta}_{m+\ell-1} \\ 0 & & & & & & \widetilde{\beta}_{m+\ell-1} & \widetilde{\alpha}_{m+\ell} \end{bmatrix} \tag{2.11}$$

is a tridiagonal matrix of order $m + \ell$. Its nontrivial entries are recursion coefficients for the polynomials $\widetilde{p}_j$. The matrix (2.11) can be chosen to be real and symmetric when the functional $2\mathcal{I} - \mathcal{G}_m$ is positive definite. In this case the nodes and weights of the generalized anti-Gauss rule (2.7) are the eigenvalues and the squares of the first components of normalized eigenvectors, respectively, of the matrix (2.11). When $2\mathcal{I} - \mathcal{G}_m$ is quasi-definite, the tridiagonal matrix (2.11) may be chosen to be real nonsymmetric. In any case, the $(m + \ell)$-point generalized anti-Gauss rule can be evaluated as

$$\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f = e_1^T f(\widetilde{T}_{m+\ell}^{(\ell)}) e_1, \tag{2.12}$$

which may not require explicit computation of the nodes and weights. For instance, when $f(t)$ is a rational function of t, it may be faster to compute $f(\widetilde{T}_{m+\ell}^{(\ell)})$ by evaluating matrix-vector products with the matrix $\widetilde{T}_{m+\ell}^{(\ell)}$ or by solving linear systems of equations with this matrix than to first compute the spectral factorization of $\widetilde{T}_{m+\ell}^{(\ell)}$ and then evaluate (2.7).

Introduce the average quadrature rule

$$\widetilde{\mathcal{A}}_{m+\ell}^{(\ell)} := \frac{1}{2}(\mathcal{G}_m + \widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}), \tag{2.13}$$

which generalizes the average rule $\widetilde{\mathcal{A}}_{m+1}^{(1)}$ considered by Laurie [30]. We obtain from (2.10) that

$$\widetilde{\mathcal{A}}_{m+\ell}^{(\ell)} f = \mathcal{I} f \quad \forall f \in \mathbb{P}_{2m+2\ell-1}.$$

We turn to the determination of the entries of the matrix (2.11) and start with the case $\ell = 1$. The quadrature rule associated with the matrix $\widetilde{T}_{m+1}^{(1)}$ is the anti-Gauss rule described by Laurie [30]. It follows from (2.5) and (2.9) that

$$\widetilde{\alpha}_j = \alpha_j, \quad j = 1, \ldots, m,$$

and

$$\widetilde{\beta}_j = \beta_j, \quad j = 1, \ldots, m-1.$$

Therefore,

$$\widetilde{p}_j = p_j, \quad j = 1, \ldots, m-1.$$

Further,

$$\widetilde{\alpha}_{m+1} = \alpha_{m+1}, \quad \widetilde{\beta}_m = \sqrt{2}\beta_m, \quad \widetilde{p}_m = \frac{1}{\sqrt{2}} p_m,$$

and it follows that

$$\widetilde{p}_{m+1} = \frac{1}{\sqrt{2}\widetilde{\beta}_{m+1}}(\beta_{m+1} p_{m+1} - \beta_m p_{m-1});$$

see [30] for details.

The entries $\widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_m$ and $\widetilde{\beta}_1, \ldots, \widetilde{\beta}_m$ are available after execution of m steps of Algorithm 1. The determination of the entry $\widetilde{\alpha}_{m+1}$ of $\widetilde{T}_{m+1}^{(1)}$ requires the execution of one more step of the algorithm. Alternatively, we may provide an estimate for this entry without carrying out the extra step of Algorithm 1. For instance, we may set the last diagonal entry of the matrix $\widetilde{T}_{m+1}^{(1)}$ to $\widetilde{\alpha}_m$. Denote the "estimated" entry by $\widehat{\alpha}_{m+1}$ and let $\widehat{T}_{m+1}^{(1)}$ denote the matrix obtained by replacing the entry $\widetilde{\alpha}_{m+1}$ by $\widehat{\alpha}_{m+1}$ in $\widetilde{T}_{m+1}^{(1)}$. The quadrature rule $\widehat{\mathcal{G}}_{m+1}^{(1)}$ determined by the matrix $\widehat{T}_{m+1}^{(1)}$ satisfies

$$\widehat{\mathcal{G}}_{m+1}^{(1)} f = \widetilde{\mathcal{G}}_{m+1}^{(1)} f \quad \forall f \in \mathbb{P}_{2m}. \tag{2.14}$$

This is a consequence of the fact that the entry $\widetilde{\alpha}_{m+1}$ in $\widetilde{T}_{m+1}^{(1)}$ does not affect the integration of polynomials in $\mathbb{P}_{2m}$ by the quadrature rule $\widetilde{\mathcal{G}}_{m+1}^{(1)}$. Replacing the entry $\widetilde{\alpha}_{m+1}$ by $\widehat{\alpha}_{m+1}$ therefore does not affect the accuracy of the quadrature rule $\widehat{G}_{m+1}^{(1)}$ when applied to polynomials in $\mathbb{P}_{2m}$. A different proof is presented in Sect. 3 (Theorem 3.5). The rule $\widehat{\mathcal{G}}_{m+1}^{(1)}$ has already been considered in [2].

It follows from (2.14) that the average rule

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

satisfies

$$\widehat{\mathcal{A}}_{m+1}^{(1)} f = \mathcal{I} f \quad \forall f \in \mathbb{P}_{2m}. \tag{2.15}$$

We turn to the rare situation when the coefficient $\widetilde{\beta}_{m+1}$ vanishes, and the coefficients $\widetilde{\beta}_1, \ldots, \widetilde{\beta}_m$ are nonvanishing. Then the functional $\mathcal{I} - \mathcal{G}_m$ is quasi-definite on $\mathbb{P}_m$, but not on $\mathbb{P}_{m+1}$. It follows that the degree of exactness of $\widetilde{\mathcal{G}}_{m+1}^{(1)}$ is larger than $2m+1$.

For every positive integer ℓ, the leading $(m+\ell-1) \times (m+\ell-1)$ principal submatrix of $\widetilde{T}_{m+\ell}^{(\ell)}$ is $\widetilde{T}_{m+\ell-1}^{(\ell-1)}$. In particular, when $\ell = 2$ and $\widetilde{T}_{m+1}^{(1)}$ is known, we only have to derive expressions for the entries $\widetilde{\alpha}_{m+2}$ and $\widetilde{\beta}_{m+1}$ to determine $\widetilde{T}_{m+2}^{(2)}$. These expressions are shown in the following theorem.

Theorem 2.2 *Assume that $\widetilde{\beta}_{m+1} \neq 0$. Then the nontrivial entries in the last row of the matrix $\widetilde{T}_{m+2}^{(2)}$, defined by (2.11), satisfy*

$$\widetilde{\alpha}_{m+2} = \frac{\alpha_{m+2}\beta_{m+1}^2 - \alpha_m \beta_m^2}{\widetilde{\beta}_{m+1}^2}, \tag{2.16}$$

$$\widetilde{\beta}_{m+1}^2 = \beta_{m+1}^2 - \beta_m^2. \tag{2.17}$$

Proof An equivalent result is shown in [36], where the orthogonal polynomials are monic, i.e., have leading coefficient one. The analogues of the tridiagonal matrices (2.3) and (2.11) associated with monic orthogonal polynomials have all superdiagonal entries equal to one, and the formulas corresponding to (2.16) and (2.17) differ, but are equivalent. The formulas (2.16) and (2.17) can be shown by a modification of the proof of [36, Theorem 2.1]. □

When the right-hand side of (2.17) is positive, it is natural to define

$$\widetilde{\beta}_{m+1} = \sqrt{\beta_{m+1}^2 - \beta_m^2}.$$

This yields a real symmetric matrix $\widetilde{T}_{m+2}^{(2)}$. If, on the other hand, the right-hand side of (2.17) is negative, then we can let $\widetilde{T}_{m+2}^{(2)}$ be nonsymmetric and such that the product of the last off-diagonal entries equals the right-hand side of (2.17). For instance, we

may choose the last superdiagonal entry to be unity and the last subdiagonal entry to be equal to the right-hand side of (2.17). This is a consequence of the following result.

Proposition 2.3 *Let the matrix $C \in \mathbb{C}^{(m+2)\times(m+2)}$ be nonsingular and such that all entries but the diagonal one vanish in the first row and column. Assume that the spectral factorization*

$$\widetilde{T}_{m+2}^{(2)} = \widetilde{S}_{m+2}^{(2)} \widetilde{\Lambda}_{m+2}^{(2)} (\widetilde{S}_{m+2}^{(2)})^{-1} \tag{2.18}$$

exists. Then

$$\widetilde{\mathcal{G}}_{m+2}^{(2)} f = e_1^T f(C \widetilde{T}_{m+2}^{(2)} C^{-1}) e_1.$$

Proof We have to show that the value of the quadrature rule (2.12) with $\ell = 2$ is independent of the matrix C. This quadrature rule can be written as

$$\widetilde{\mathcal{G}}_{m+2}^{(2)} f = e_1^T \widetilde{S}_{m+2}^{(2)} f(\widetilde{\Lambda}_{m+2}^{(2)}) (\widetilde{S}_{m+2}^{(2)})^{-1} e_1,$$

see, e.g., Gragg [23] or Fenu et al. [18] for proofs. Due to the structure of the matrix C, we have

$$\widetilde{\mathcal{G}}_{m+2}^{(2)} f = e_1^T C \widetilde{S}_{m+2}^{(2)} f(\widetilde{\Lambda}_{m+2}^{(2)}) (\widetilde{S}_{m+2}^{(2)})^{-1} C^{-1} e_1 = e_1^T f(C \widetilde{T}_{m+2}^{(2)} C^{-1}) e_1.$$

The different choices of the last sub- and super-diagonal entries of $\widetilde{T}_{m+2}^{(2)}$ corresponds to different diagonal matrices C. □

Using results of Pozza et al. [34], the requirement that $\widetilde{T}_{m+2}^{(2)}$ has a spectral factorization (2.18) can be removed.

The computation of the entries of $\widetilde{T}_{m+2}^{(2)}$ requires that $m+2$ steps with Algorithm 1 be carried out. The matrix $\widetilde{T}_{m+2}^{(2)}$ defines the quadrature rule $\widetilde{\mathcal{G}}_{m+2}^{(2)}$. Analogously as above, we may carry out only $m+1$ steps of Algorithm 1 and use an estimate $\widehat{\alpha}_{m+2}$ of $\widetilde{\alpha}_{m+2}$. We denote the matrix obtained by replacing $\widetilde{\alpha}_{m+2}$ by $\widehat{\alpha}_{m+2}$ in $\widetilde{T}_{m+2}^{(2)}$ by $\widehat{T}_{m+2}^{(2)}$. For instance, we may let $\widetilde{\alpha}_{m+2} := \widetilde{\alpha}_{m+1}$. The matrix $\widehat{T}_{m+2}^{(2)}$ defines the quadrature rule $\widehat{\mathcal{G}}_{m+2}^{(2)}$. Similarly to (2.14), we have

$$\widehat{\mathcal{G}}_{m+2}^{(2)} f = \widetilde{\mathcal{G}}_{m+2}^{(2)} f \quad \forall f \in \mathbb{P}_{2m+2};$$

see Theorem 3.5 below for a proof.

We define the average rule

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

and obtain analogously to (2.15) that

$$\widehat{\mathcal{A}}_{m+2}^{(2)} f = \mathcal{I} f \quad \forall f \in \mathbb{P}_{2m+2}.$$

We turn to $\ell = 3$, and provide expressions for the entries in the last row of the matrix $\widetilde{T}_{m+3}^{(3)}$, assuming that a division by zero does not prevent these expressions from being evaluated. Analogous formulas for a nonsymmetric tridiagonal matrix that is analogous to $\widetilde{T}_{m+3}^{(3)}$ and is associated with monic orthogonal polynomials are derived in [36]. We therefore omit the proof.

Theorem 2.4 *Assume that $\widetilde{\beta}_{m+2}\widetilde{\beta}_{m+1} \neq 0$. Then the nontrivial entries in the last row of the matrix $\widetilde{T}_{m+3}^{(3)}$ defined by (2.11) satisfy*

$$\widetilde{\beta}_{m+2}^2 = \frac{\widetilde{\beta}_{m+1}^2(\beta_{m+2}^2\beta_{m+1}^2 - \beta_m^2\beta_{m-1}^2) - \beta_{m+1}^2\beta_m^2(\alpha_{m+2} - \alpha_m)^2}{\widetilde{\beta}_{m+1}^2} \tag{2.19}$$

and

$$\begin{aligned}\widetilde{\alpha}_{m+3} = {} & \frac{\beta_{m+2}^2\beta_{m+1}^2(\alpha_{m+3} + 2\alpha_{m+2} - 2\widetilde{\alpha}_{m+2}) + \beta_m^2\beta_{m-1}^2(2\widetilde{\alpha}_{m+2} - 2\alpha_m - \alpha_{m-1})}{\widetilde{\beta}_{m+2}^2\widetilde{\beta}_{m+1}^2} \\ & + \frac{\alpha_{m+2}\beta_{m+1}^2(\alpha_{m+2} - \widetilde{\alpha}_{m+2})^2 - \alpha_m\beta_m^2(\alpha_m - \widetilde{\alpha}_{m+2})^2}{\widetilde{\beta}_{m+2}^2\widetilde{\beta}_{m+1}^2}.\end{aligned}$$

We remark that if the right-hand side of (2.19) is positive, then we let $\widetilde{\beta}_{m+2}$ be the square root of this expression. If, instead, the right-hand side of (2.19) is negative, then we may choose $\widetilde{T}_{m+3}^{(3)}$ to be nonsymmetric with the last superdiagonal entry one and the last subdiagonal entry equal to the right-hand side of (2.19).

The matrix $\widetilde{T}_{m+3}^{(3)}$ determines the quadrature rule $\widetilde{\mathcal{G}}_{m+3}^{(3)}$. The computation of the entries of $\widetilde{T}_{m+3}^{(3)}$ requires $m+3$ steps with Algorithm 1, which demands the evaluation of $m+3$ matrix-vector products with the matrix A. Similarly as above, we may, in order to reduce the number of matrix-vector product evaluations with A by one, replace the entry $\widetilde{\alpha}_{m+3}$ by an estimate $\widehat{\alpha}_{m+3}$. For instance, we may use $\widetilde{\alpha}_{m+2}$ as an estimate of $\widetilde{\alpha}_{m+3}$. We denote the symmetric tridiagonal matrix that is obtained by replacing the entry $\widetilde{\alpha}_{m+3}$ of $\widetilde{T}_{m+3}^{(3)}$ by an estimate $\widehat{\alpha}_{m+3}$ by $\widehat{T}_{m+3}^{(3)}$, and denote the quadrature rule associated with the latter matrix by $\widehat{\mathcal{G}}_{m+3}^{(3)}$. We remark that $\widetilde{\alpha}_{m+3}$ may be replaced by some $\widehat{\alpha}_{m+3} \in \mathbb{R}$ also when the product $\widetilde{\beta}_{m+2}\widetilde{\beta}_{m+1}$ vanishes. Thus, the quadrature rule $\widehat{\mathcal{G}}_{m+3}^{(3)}$ can be defined also when the rule $\widetilde{\mathcal{G}}_{m+3}^{(3)}$ does not exist. The properties of the rule $\widehat{\mathcal{G}}_{m+3}^{(3)}$ are analogous to those of $\widehat{\mathcal{G}}_{m+2}^{(2)}$; see Theorem 3.5 below for details for the more general situation when the measure dw is matrix-valued. An average rule $\widehat{\mathcal{A}}_{m+3}^{(3)}$ can be defined similarly as $\widehat{\mathcal{A}}_{m+2}^{(2)}$; cf. (2.13).

2.2 Real-valued functionals with a nonsymmetric matrix

We extend the discussion of the previous subsection to matrix functionals with a large, possibly nonsymmetric, matrix $A \in \mathbb{R}^{n\times n}$, and vectors $V, W \in \mathbb{R}^n$ such that

$V^T W = 1$. Introduce the spectral factorization

$$A = S\Lambda S^{-1}, \quad \Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n], \tag{2.20}$$

and the vectors

$$[\nu_1, \nu_2, \ldots, \nu_n] := W^T S, \quad [\nu_1', \nu_2', \ldots, \nu_n'] := (S^{-1}V)^T.$$

We tacitly assume that the spectral factorization of A exists. This factorization is used in our derivation, but does not have to be computed. The rare situation when the matrix A does not have a spectral factorization is discussed by Pozza et al. [34,35].

The expression (1.2) can be written as

$$F(A) = W^T S f(\Lambda) S^{-1} V = \sum_{j=1}^{n} f(\lambda_j)\nu_j \nu_j'. \tag{2.21}$$

We may represent the right-hand side of (2.21) by an integral with a measure $d\omega$ with support in the complex plane, i.e.,

$$\mathcal{I} f = \int f(z) d\omega(z). \tag{2.22}$$

Introduce the the bilinear form

$$[f, g] := \mathcal{I}(fg)$$

for polynomials f and g of suitably low degree. Let $\{p_j\}_{j=0,1,2,\ldots}$ and $\{q_i\}_{i=0,1,2,\ldots}$ denote sequences of biorthonormal polynomials with respect to this bilinear form. Thus,

$$[p_j, q_i] = \begin{cases} 1, & j = i, \\ 0, & j \neq i, \end{cases}$$

where the polynomials p_j and q_j are of degree j. They are implicitly defined by the nonsymmetric Lanczos process described by Algorithm 2. In particular, this process determines how the p_j and q_j are scaled; see (2.24) below.

Gauss and anti-Gauss quadrature rules with respect to the measure $d\omega$ in (2.22) can be computed conveniently by the nonsymmetric Lanczos process applied to the matrix A with initial vectors $v_1 := V$ and $w_1 := W$; see, e.g., [11,18].

It is well known that the nonsymmetric Lanczos process may break down. In the present section we assume that the number of steps, $m \ll n$, of the nonsymmetric Lanczos process is small enough for breakdown not to take place. Insightful discussions on how to handle breakdowns are provided by Bai et al. [4], Brezinski et al. [8,9], Freund et al. [19], and Gutknecht [24,25].

Algorithm 2 The nonsymmetric Lanczos process.

Input: nonsymmetric matrix $A \in \mathbb{R}^{n\times n}$, initial vectors $V, W \in \mathbb{R}^n$ such that $V^T W = 1$, number of steps m.
$w_0 := 0 \in \mathbb{R}^n$, $v_0 := 0 \in \mathbb{R}^n$, $w_1 := W$, $v_1 := V$, $\beta_0 := 0$, $\beta_0' := 0$
for $j = 1$ **to** m
  $\alpha_j := v_j^T (A w_j - w_{j-1}\beta_j)$
  $r := A w_j - \alpha_{j-1} w_j - w_{j-1}\beta_{j-1}$
  $s := A^T v_j - \alpha_{j-1} v_j - v_{j-1}\beta'_{j-1}$
  $\beta_j := \|r\|$; $w_{j+1} := r/\beta_j$
  $\beta'_j := \|s\|$; $v_{j+1} := s/\beta'_j$
end for
Output: Entries $\alpha_1, \alpha_2, \ldots, \alpha_m$, $\beta_1, \beta_2, \ldots, \beta_m$, and $\beta'_1, \beta'_2, \ldots, \beta'_m$ of the matrix (2.23).

Application of m steps of the nonsymmetric Lanczos process to A with initial vectors $v_1 := V$ and $w_1 := W$ yields the decompositions

$$A V_m = V_m T_m + \beta_m v_{m+1} e_m^T,$$
$$A^T W_m = W_m T_m^T + \beta'_m w_{m+1} e_m^T,$$

where the matrices $V_m = [v_1, v_2, \ldots, v_m] \in \mathbb{R}^{n\times m}$, $W_m = [w_1, w_2, \ldots, w_m] \in \mathbb{R}^{n\times m}$, and the vectors $v_{m+1}, w_{n+1} \in \mathbb{R}^n$ satisfy $V_m^T W_m = D_m$, $v_{m+1}^T w_{m+1} = d_{m+1}$, $V_m^T w_{m+1} = 0$, $W_m^T v_{m+1} = 0$, and $\beta_m, \beta'_m \in \mathbb{R}$. The nontrivial entries of the diagonal matrix $D_m \in \mathbb{R}^{m\times m}$ are such that the columns of V_m and W_m are of unit norm; the scalar d_{m+1} is determined analogously. The matrix

$$T_m = \begin{bmatrix} \alpha_1 & \beta'_1 & & & & & 0 \\ \beta_1 & \alpha_2 & \beta'_2 & & & & \\ & \beta_2 & \alpha_3 & \beta'_3 & & & \\ & & \beta_3 & & \ddots & & \\ & & & \ddots & \ddots & \beta'_{m-2} & \\ & & & & \beta_{m-2} & \alpha_{m-1} & \beta'_{m-1} \\ 0 & & & & & \beta_{m-1} & \alpha_m \end{bmatrix} \in \mathbb{R}^{m\times m} \tag{2.23}$$

is tridiagonal and generally nonsymmetric. We remark that the nonsymmetric Lanczos process can be implemented by using coupled two-term recursion formulas, and such an implementation has been shown to have better numerical properties; see [26] for an analysis. Algorithm 2 describes a simpler implementation that uses three-term recursions for clarity of exposition.

The biorthonormal polynomials p_j and q_j are implicitly defined by Algorithm 2 via

$$v_j = p_{j-1}(A) v_1, \quad w_j = q_{j-1}(A^T) w_1, \quad j = 1, 2, \ldots . \tag{2.24}$$

The m-point Gauss quadrature rule associated with the measure in (2.22) can be expressed as

$$\mathcal{G}_m f = e_1^T f(T_m) e_1, \tag{2.25}$$

where the matrix T_m is given by (2.23). This rule is characterized by the property that

$$\mathcal{I} f = \mathcal{G}_m f \quad \forall f \in \mathbb{P}_{2m-1}; \tag{2.26}$$

see [11,18] for details.

We turn to generalized anti-Gauss rules associated with the measure in (2.22). Consider the bilinear form

$$\langle f, g \rangle := (2\mathcal{I} - \mathcal{G}_m)(fg),$$

where $\mathcal{I}$ is defined by (2.22) and the Gauss rule by (2.25). There are two families of biorthonormal polynomials $\{\widetilde{p}_j\}_{j=0,1,2,\ldots}$ and $\{\widetilde{q}_i\}_{i=0,1,2,\ldots}$ associated with this bilinear form. Thus,

$$\langle \widetilde{p}_j, \widetilde{q}_i \rangle = \begin{cases} 1, & j = i, \\ 0, & j \neq i. \end{cases}$$

Assuming that the polynomials $\widetilde{p}_0, \widetilde{p}_1, \ldots, \widetilde{p}_{m+\ell}$ and $\widetilde{q}_0, \widetilde{q}_1, \ldots, \widetilde{q}_{m+\ell}$ exist, they satisfy three-term recursion relations, that can be written as

$$\begin{aligned} x\left[\widetilde{p}_0(x), \widetilde{p}_1(x), \ldots, \widetilde{p}_{m+\ell-1}(x)\right] &= \left[\widetilde{p}_0(x), \widetilde{p}_1(x), \ldots, \widetilde{p}_{m+\ell-1}(x)\right] \widetilde{T}_{m+\ell}^{(\ell)} \\ &\quad + \widetilde{\beta}_{m+\ell}\left[0, \ldots, 0, \widetilde{p}_{m+\ell}(x)\right], \\ x\left[\widetilde{q}_0(x), \widetilde{q}_1(x), \ldots, \widetilde{q}_{m+\ell-1}(x)\right] &= \left[\widetilde{q}_0(x), \widetilde{q}_1(x), \ldots, \widetilde{q}_{m+k-1}(x)\right] (\widetilde{T}_{m+\ell}^{(\ell)})^T \\ &\quad + \widetilde{\beta}'_{m+\ell}\left[0, \ldots, 0, \widetilde{q}_{m+\ell}(z)\right], \end{aligned} \tag{2.27}$$

where

$$\widetilde{T}_{m+\ell}^{(\ell)} = \begin{bmatrix} \widetilde{\alpha}_1 & \widetilde{\beta}'_1 & & & & & & 0 \\ \widetilde{\beta}_1 & \widetilde{\alpha}_2 & \widetilde{\beta}'_2 & & & & & \\ & \widetilde{\beta}_2 & \widetilde{\alpha}_3 & & & & & \\ & & & \ddots & & & & \\ & & & \ddots & \ddots & \widetilde{\beta}'_{m-2} & & \\ & & & & \widetilde{\beta}_{m-2} & \widetilde{\alpha}_{m-1} & \widetilde{\beta}'_{m-1} & \\ & & & & & \widetilde{\beta}_{m-1} & \ddots & \\ & & & & & \ddots & \ddots & \widetilde{\beta}'_{m+\ell-1} \\ 0 & & & & & & \widetilde{\beta}_{m+\ell-1} & \widetilde{\alpha}_{m+\ell} \end{bmatrix} \tag{2.28}$$

is a tridiagonal matrix of order $m + \ell$.

We turn to the computation of the generalized anti-Gauss rule

$$\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f = e_1^T f(\widetilde{T}_{m+\ell}^{(\ell)}) e_1 \tag{2.29}$$

determined by the matrix (2.28). Thus, we discuss the evaluation of the entries of this matrix. The rule (2.29) satisfies (2.8) with $\mathcal{I}$ given by (2.22) and the m-point Gauss rule $\mathcal{G}_m$ defined by (2.25). We first consider the entries of the matrix (2.28) for the situation when $\ell = 1$. It follow from (2.26) and (2.10) that

$$\widetilde{\alpha}_j = \alpha_j, \quad j = 1, \ldots, m,$$

and

$$\widetilde{\beta}_j = \beta_j, \quad \widetilde{\beta}'_j = \beta'_j, \quad j = 1, \ldots, m-1.$$

Therefore,

$$\widetilde{p}_j = p_j, \quad \widetilde{q}_j = q_j, \quad j = 1, \ldots, m-1,$$

and

$$\begin{aligned} \widetilde{\alpha}_{m+1} = \alpha_{m+1}, \quad \widetilde{\beta}_m = \sqrt{2}\beta_m, \quad \widetilde{\beta}'_m = \sqrt{2}\beta'_m, \\ \widetilde{p}_m = \frac{1}{\sqrt{2}} p_m, \quad \widetilde{q}_m = \frac{1}{\sqrt{2}} q_m. \end{aligned}$$

We obtain

$$\begin{aligned} \widetilde{p}_{m+1} &= \frac{1}{\sqrt{2}\widetilde{\beta}_{m+1}} (\beta_{m+1} p_{m+1} - \beta'_m p_{m-1}), \\ \widetilde{q}_{m+1} &= \frac{1}{\sqrt{2}\widetilde{\beta}'_{m+1}} (\beta'_{m+1} q_{m+1} - \beta_m q_{m-1}); \end{aligned} \tag{2.30}$$

see [11] for a proof. All entries of the matrix $\widetilde{T}_{m+1}^{(1)}$ can be determined by carrying out $m+1$ steps with Algorithm 2. Alternatively, we carry out only m steps and estimate the entry $\widetilde{\alpha}_{m+1}$; for instance, we may set this entry to $\widetilde{\alpha}_m$. The latter approach is discussed and illustrated in [2].

We turn to the situation when $\ell = 2$. For every $\ell \geq 1$, the leading $(m+\ell-1) \times (m+\ell-1)$ principal submatrix of $\widetilde{T}_{m+\ell}^{(\ell)}$ is $\widetilde{T}_{m+\ell-1}^{(\ell-1)}$. Assuming that the matrix $\widetilde{T}_{m+1}^{(1)}$ is known, we only have to determine expressions for the entries $\widetilde{\alpha}_{m+2}$, $\widetilde{\beta}_{m+1}$, and $\widetilde{\beta}'_{m+1}$ in the last row and column of $\widetilde{T}_{m+2}^{(2)}$, if they exist. It is convenient to sometimes write $\mathcal{I}(fg) = \mathcal{I}(f, g)$ and $\mathcal{G}_m(fg) = \mathcal{G}_m(f, g)$ for $\mathcal{I}$ defined by (2.22) and $\mathcal{G}_m$ by (2.25). Similar notation will be used in Sect. 3.

Theorem 2.5 *Assume that $\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1} \neq 0$. Then the nontrivial entries in the last row and column of the matrix $\widetilde{T}_{m+2}^{(2)}$ defined by (2.28) satisfy*

$$\widetilde{\alpha}_{m+2} = \frac{\alpha_{m+2}\beta_{m+1}\beta'_{m+1} - \alpha_m\beta_m\beta'_m}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}, \tag{2.31}$$

$$\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1} = \beta_{m+1}\beta'_{m+1} - \beta_m\beta'_m. \tag{2.32}$$

Proof We first show (2.32). The relation (2.30) and the fact that $\mathcal{I}(p_j, q) = 0$ for $q \in \mathbb{P}_{j-1}$ yield

$$\begin{aligned} 2\mathcal{I}(x\widetilde{p}_m, \widetilde{q}_{m+1}) &= 2\mathcal{I}\left(\frac{x}{\sqrt{2}}p_m, \frac{1}{\sqrt{2}\widetilde{\beta}'_{m+1}}(\beta'_{m+1}q_{m+1} - \beta_m q_{m-1})\right) \\ &= \frac{1}{\widetilde{\beta}'_{m+1}}\left(\mathcal{I}(xp_m, \beta'_{m+1}q_{m+1}) - \mathcal{I}(xp_m, \beta_m q_{m-1})\right) \\ &= \frac{1}{\widetilde{\beta}'_{m+1}}\left(\beta'_{m+1}\mathcal{I}(xp_m, q_{m+1}) - \beta_m\mathcal{I}(xp_m, q_{m-1})\right) \\ &= \frac{1}{\widetilde{\beta}'_{m+1}}(\beta_{m+1}\beta'_{m+1} - \beta_m\beta'_m). \end{aligned} \tag{2.33}$$

Similarly, the three-term recurrence relations (2.27), property (2.26), and the observation that $\mathcal{G}_m(p_m, q) = 0$ for all polynomials q, give

$$\begin{aligned} \mathcal{G}_m(x\widetilde{p}_m, \widetilde{q}_{m+1}) &= \mathcal{G}_m\left(\frac{x}{\sqrt{2}}p_m, \frac{1}{\sqrt{2}\widetilde{\beta}'_{m+1}}(\beta'_{m+1}q_{m+1} - \beta_m q_{m-1})\right) \\ &= \frac{1}{2}\mathcal{G}_m\left(xp_m, \frac{1}{\widetilde{\beta}'_{m+1}}(\beta'_{m+1}q_{m+1} - \beta_m q_{m-1})\right) = 0. \end{aligned} \tag{2.34}$$

The relations (2.33) and (2.34) now yield

$$\widetilde{\beta}_{m+1} = (2\mathcal{I} - \mathcal{G}_m)(x\widetilde{p}_m, \widetilde{q}_{m+1}) = \frac{1}{\widetilde{\beta}'_{m+1}}(\beta_{m+1}\beta'_{m+1} - \beta_m\beta'_m).$$

This shows (2.32).

We turn to (2.31). Analogously as above, we obtain

$$\begin{aligned} &2\mathcal{I}(x\widetilde{p}_{m+1}, \widetilde{q}_{m+1}) \\ &= 2\mathcal{I}\left(\frac{x}{\sqrt{2}\widetilde{\beta}_{m+1}}(\beta_{m+1}p_{m+1} - \beta'_m p_{m-1}), \frac{1}{\sqrt{2}\widetilde{\beta}'_{m+1}}(\beta'_{m+1}q_{m+1} - \beta_m q_{m-1})\right) \\ &= \frac{1}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}\mathcal{I}\left(x(\beta_{m+1}p_{m+1} - \beta'_m p_{m-1}), (\beta'_{m+1}q_{m+1} - \beta_m q_{m-1})\right) \\ &= \frac{1}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}\left(\beta'_{m+1}\beta_{m+1}\mathcal{I}\big(xp_{m+1}, q_{m+1}\big) + \beta'_m\beta_m\mathcal{I}\big(xp_{m-1}, q_{m-1}\big)\right) \\ &= \frac{1}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}\big(\beta'_{m+1}\beta_{m+1}\alpha_{m+2} + \beta'_m\beta_m\alpha_m\big) \end{aligned}$$

and

$$\begin{aligned}
&\mathcal{G}_m(x\widetilde{p}_{m+1},\widetilde{q}_{m+1}) = \mathcal{G}_m\bigg(\frac{x}{\widetilde{\beta}_{m+1}}((x-\widetilde{\alpha}_{m+1})\widetilde{p}_m - \widetilde{\beta}'_m\widetilde{p}_{m-1}),\\
&\quad\frac{1}{\widetilde{\beta}'_{m+1}}((x-\widetilde{\alpha}_{m+1})\widetilde{q}_m - \widetilde{\beta}_m\widetilde{q}_{m-1})\bigg)\\
&= \frac{\mathcal{G}_m\Big(x((x-\alpha_{m+1})\frac{1}{\sqrt{2}}p_m - \sqrt{2}\beta'_m p_{m-1}),\,((x-\alpha_{m+1})\frac{1}{\sqrt{2}}q_m - \sqrt{2}\beta_m q_{m-1})\Big)}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}\\
&= \frac{1}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}\mathcal{G}_m(x(-\sqrt{2}\beta'_m p_{m-1}),-\sqrt{2}\beta_m q_{m-1})\\
&= \frac{1}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}(2\beta'_m\beta_m\mathcal{I}_m(xp_{m-1},q_{m-1}))\\
&= \frac{1}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}(2\beta'_m\beta_m\alpha_m).
\end{aligned}$$

The above relations yield

$$\begin{aligned}
\widetilde{\alpha}_{m+2} &= (2\mathcal{I}-\mathcal{G}_m)(x\widetilde{p}_{m+1},\widetilde{q}_{m+1})\\
&= \frac{1}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}\big(\beta'_{m+1}\beta_{m+1}\alpha_{m+2} + \beta'_m\beta_m\alpha_m - 2\beta'_m\beta_m\alpha_m\big),
\end{aligned}$$

and the right-hand side simplifies to (2.31). □

The evaluation of the entries of the matrix $\widetilde{T}^{(2)}_{m+2}$ requires that $m+2$ steps with Algorithm 2 be carried out. If we estimate the value of the entry $\widetilde{\alpha}_{m+2}$, then it suffices to carry out $m+1$ steps. Moreover, the condition $\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1} \neq 0$ may be violated, in which case $\widetilde{\alpha}_{m+2}$ is not defined. It is then meaningful to replace $\widetilde{\alpha}_{m+2}$ in $\widetilde{T}^{(2)}_{m+2}$ by a well-defined entry $\widehat{\alpha}_{m+2}$. We refer to the matrix so defined as $\widehat{T}^{(2)}_{m+2}$ and to the quadrature rule associated with $\widehat{T}^{(2)}_{m+2}$ as $\widehat{\mathcal{G}}^{(2)}_{m+2}$. This rule is said to be a simplified generalized anti-Gauss rule. In computations, we choose the last subdiagonal and superdiagonal entries of $\widetilde{T}^{(2)}_{m+2}$ and $\widehat{T}^{(2)}_{m+2}$ to be $\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}$ and unity, respectively. This choice can be justified by a result analogous to Proposition 2.3.

Now let $\ell = 3$. We assume that the matrix $\widetilde{T}^{(2)}_{m+2}$ is known and seek to determine the entries $\widetilde{\beta}_{m+2}$ and $\widetilde{\beta}'_{m+2}$ in the last row and column of the matrix $\widehat{T}^{(3)}_{m+3}$. The last diagonal element of $\widehat{T}^{(3)}_{m+3}$ is set to $\widehat{\alpha}_{m+3} := \widetilde{\alpha}_{m+2}$. It is convenient not to use the matrix $\widetilde{T}^{(3)}_{m+3}$ because its last diagonal entry is tedious to evaluate. The quadrature rule associated with $\widehat{T}^{(3)}_{m+3}$ is denoted by $\widehat{\mathcal{G}}^{(3)}_{m+3}$. It satisfies

$$\widehat{\mathcal{G}}^{(3)}_{m+3}f = \widetilde{\mathcal{G}}^{(3)}_{m+3}f \quad \forall f \in \mathbb{P}_{2m+4},$$

where $\widetilde{\mathcal{G}}_{m+3}^{(3)}$ is the quadrature rule determined by the matrix $\widetilde{T}_{m+3}^{(3)}$; see Theorem 3.5 below. The following result yields properties of the last off-diagonal entries of the matrices $\widetilde{T}_{m+3}^{(3)}$ and $\widehat{T}_{m+3}^{(3)}$.

Theorem 2.6 *Assume that $\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}\widetilde{\beta}'_{m+2} \neq 0$. Then the last off-diagonal entries of the matrix $\widetilde{T}_{m+3}^{(3)}$ defined by (2.28) satisfy*

$$\widetilde{\beta}_{m+2}\widetilde{\beta}'_{m+2} = \frac{1}{\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}\Big(\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}\big(\beta_{m+2}\beta'_{m+2}\beta_{m+1}\beta'_{m+1} - \beta_m\beta'_m\beta_{m-1}\beta'_{m-1}\big) - \beta_{m+1}\beta'_{m+1}\beta_m\beta'_m(\alpha_{m+2}-\alpha_m)^2\Big).$$

Proof The last off-diagonal entries of $\widetilde{T}_{m+3}^{(3)}$ are defined by

$$\widetilde{\beta}_{m+2} = (2\mathcal{I} - \mathcal{G}_m)(x\widetilde{p}_{m+1}, \widetilde{q}_{m+2}), \quad \widetilde{\beta}'_{m+2} = (2\mathcal{I} - \mathcal{G}_m)(x\widetilde{p}_{m+2}, \widetilde{q}_{m+1}).$$

Express $\widetilde{q}_{m+2}$ in terms of polynomials q_j. We obtain from

$$\widetilde{q}_{m+2} = \frac{1}{\widetilde{\beta}'_{m+2}}((x-\widetilde{\alpha}_{m+1})\widetilde{q}_{m+1} - \widetilde{\beta}_{m+1}\widetilde{q}_m)$$

that

$$\widetilde{q}_{m+2} = \frac{1}{\sqrt{2}\widetilde{\beta}'_{m+2}\widetilde{\beta}'_{m+1}}\big(\beta'_{m+1}\beta'_{m+2}q_{m+2} + \beta'_{m+1}(\alpha_{m+2}-\widetilde{\alpha}_{m+2})q_{m+1} + \beta_m(\widetilde{\alpha}_{m+2}-\alpha_m)q_{m-1} - \beta_m\beta_{m-1}q_{m-2}\big).$$

Moreover,

$$2\mathcal{I}(x\widetilde{p}_{m+1}, \widetilde{q}_{m+2}) = \frac{1}{\widetilde{\beta}'_{m+2}\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}\Big(\beta_{m+1}\beta'_{m+1}\beta_{m+2}\beta'_{m+2} + \beta_{m+1}\beta'_{m+1}(\alpha_{m+2}^2 - \widetilde{\alpha}_{m+2}\alpha_{m+2}) - \beta_m\beta'_m(\widetilde{\alpha}_{m+2}\alpha_m - \alpha_m^2) + \beta_m\beta'_m\beta_{m-1}\beta'_{m-1}\Big)$$

and

$$\mathcal{G}_m(x\widetilde{p}_{m+1}, \widetilde{q}_{m+2}) = \frac{\big(2\beta_m\beta'_m\alpha_m^2 - 2\beta_m\beta'_m\widetilde{\alpha}_{m+2}\alpha_m + 2\beta_m\beta'_m\beta_{m-1}\beta'_{m-1}\big)}{\widetilde{\beta}'_{m+2}\widetilde{\beta}_{m+1}\widetilde{\beta}'_{m+1}}.$$

The theorem now follows after some extensive, but not difficult, computations. □

Assume that $\widetilde{\beta}_{m+2}\widetilde{\beta}'_{m+2} \neq 0$. Then, in computations, we choose the last subdiagonal and superdiagonal entries of $\widehat{T}_{m+3}^{(3)}$ to be $\widetilde{\beta}_{m+2}\widetilde{\beta}'_{m+2}$ and unity, respectively. This choice can be justified by a result analogous to Proposition 2.3.

We remark that it is possible, but tedious, to compute the entries of the last row and column of the nonsymmetric tridiagonal matrices $\widetilde{T}_{m+\ell}^{(\ell)}$ for $\ell \geq 4$.

3 Generalized block anti-Gauss rules

This section generalizes the discussion of Sect. 2 to the situation when V and W in (1.1) and (1.2) are block vectors with block size $1 < k \ll n$.

3.1 Generalized block anti-Gauss rules for functions of a symmetric matrix

We consider the approximation of matrix functions of the form (1.1) with a large symmetric matrix $A \in \mathbb{R}^{n\times n}$ when $V \in \mathbb{R}^{n\times k}$ is a block vector with orthonormal columns. Substituting the spectral factorization (2.1) of A into (1.1) gives

$$V^T f(A)V = \widetilde{V} f(\Lambda)\widetilde{V}^T = \sum_{i=1}^{n} f(\lambda_i)\widetilde{v}_i\widetilde{v}_i^T = \int f(x)d\widetilde{w}(x) =: \mathcal{I}f, \tag{3.1}$$

where $\widetilde{V} = [\widetilde{v}_1, \ldots, \widetilde{v}_n] = V^T S \in \mathbb{R}^{k\times n}$ and $\widetilde{w} : \mathbb{R} \to \mathbb{R}^{k\times k}$ is a matrix-valued distribution with mass $\widetilde{v}_i\widetilde{v}_i^T$ at the eigenvalue λ_i of A. There is a sequence of polynomials p_j that are orthonormal with respect to the bilinear form

$$(f, g) := \mathcal{I}(f, g)$$

associated with the measure $d\widetilde{w}$ in (3.1) and have $k \times k$ matrix coefficients; see Golub and Meurant [22] or Fenu et al. [18] for details. This bilinear form can be defined for matrix-valued functions f and g; see (3.5) below.

The orthonormal polynomials p_j satisfy a three-term recursion relation of the form

$$\begin{aligned} xp_{j-1}(x) &= p_j(x)\Gamma_j + p_{j-1}(x)\Omega_j + p_{j-2}(x)\Gamma_{j-1}^T, \quad j = 1, 2, \ldots, \\ p_0(x) &:= I_k, \quad p_{-1}(x) := O_k, \end{aligned} \tag{3.2}$$

where the recursion coefficients Γ_j and Ω_j are $k \times k$ matrices with real entries. Moreover, Ω_j is symmetric and Γ_j can be chosen to be upper triangular. The matrix $O_k \in \mathbb{R}^{k\times k}$ denotes the zero matrix. We remark that the polynomials p_j are considered for theoretical purposes only; they are not explicitly computed.

Introduce the matrix

$$P_m(x) := \big[p_0(x), \ldots, p_{m-1}(x)\big] \in \mathbb{R}^{k\times km}.$$

Then the recursion relation (3.2) for the polynomials $p_0, \ldots, p_m$ can be expressed as

$$x P_m(x) = P_m(x)T_m + p_m(x)\Gamma_m E_m^T,$$

where

$$T_m := \begin{bmatrix} \Omega_1 & \Gamma_1^T & & & 0 \\ \Gamma_1 & \Omega_2 & \Gamma_2^T & & \\ & \ddots & \ddots & \ddots & \\ & & \Gamma_{m-2} & \Omega_{m-1} & \Gamma_{m-1}^T \\ 0 & & & \Gamma_{m-1} & \Omega_m \end{bmatrix} \in \mathbb{R}^{km\times km} \tag{3.3}$$

and $E_i := [e_{(i-1)k+1}, \ldots, e_{ik}]$ denotes a "block axis vector" of appropriate size with $k \times k$ blocks. Thus, the ith block of E_i is I_k and all other blocks vanish. The matrix T_m is symmetric, block tridiagonal, and has bandwidth $2k+1$. It is determined by m steps of the symmetric block Lanczos process with block size k. Algorithm 3 describes this process under the assumption that no breakdown takes place. In line 8 of the algorithm, one computes a QR factorization of the matrix R_j, i.e., $V_{j+1} \in \mathbb{R}^{n\times k}$ has orthonormal columns and $\Gamma_j \in \mathbb{R}^{k\times k}$ is upper triangular with positive diagonal entries.

Algorithm 3 The symmetric block Lanczos process.

1: **Input:** symmetric matrix $A \in \mathbb{R}^{n\times n}$, initial block vector $V \in \mathbb{R}^{n\times k}$ with
2: orthonormal columns, number of steps m.
3: $V_0 := O \in \mathbb{R}^{n\times k}$, $\Gamma_0 := O_k$, $V_1 := V$
4: **for** $j = 1$ **to** m
5: $B := AV_j - V_{j-1}\Gamma_{j-1}^T$
6: $\Omega_j := V_j^T B$
7: $R_j := B - V_j\Omega_j$
8: $V_{j+1}\Gamma_j := R_j$
9: **end for**
10: **Output:** Blocks $\Omega_1, \ldots, \Omega_m$ and $\Gamma_1, \ldots, \Gamma_m$ of the matrix (3.3).

The m-block Gauss rule associated with the matrix-valued measure $d\widetilde{w}$ in (3.1) is given by

$$\mathcal{G}_m f = E_1^T f(T_m) E_1, \tag{3.4}$$

where the matrix T_m is defined by (3.3). Thus,

$$\mathcal{G}_m f = \mathcal{I} f \quad \forall f \in \mathbb{P}_{2m-1},$$

where the function $\mathcal{I}$ is defined in (3.1); see, e.g., Fenu et al. [18] or Golub and Meurant [22] for proofs.

We turn to generalized block anti-Gauss rules. These rules are block generalizations of the generalized anti-Gauss rules described in Sect. 2. They also generalize the block anti-Gauss rules introduced in [18]. We say that $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}$ is a generalized $(m+\ell)$-block anti-Gauss rule associated with the matrix-valued measure $d\widetilde{w}$ in (3.1) if

$$\left(\mathcal{I} - \widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}\right) f = -\left(\mathcal{I} - \mathcal{G}_m\right) f \quad \forall f \in \mathbb{P}_{2m+2\ell-1},$$

where $\mathcal{I}$ is defined in (3.1).

It is convenient to extend $\mathcal{I}$ to allow matrix-valued function arguments $f, g \in \mathbb{R}^{k\times k}$. We define

$$\mathcal{I}(f, g) := \sum_{i=1}^{n} f^T(\lambda_i)\widetilde{v}_i\widetilde{v}_i^T g(\lambda_i). \tag{3.5}$$

The Gauss rule (3.4) can be extended similarly to allow arguments $f, g \in \mathbb{R}^{k\times k}$, i.e., we may define $\mathcal{G}_m(f, g)$; see [18] for details.

3.1.1 Generalized block anti-Gauss rules, $\ell = 1$

The block anti-Gauss rule $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}$ for $\ell = 1$ has previously been discussed in [18]. This rule satisfies

$$\left(\mathcal{I} - \widetilde{\mathcal{G}}_{m+1}^{(1)}\right) f = -\left(\mathcal{I} - \mathcal{G}_m\right) f \quad \forall f \in \mathbb{P}_{2m+1},$$

which can be written as

$$\widetilde{\mathcal{G}}_{m+1}^{(1)} f = \left(2\mathcal{I} - \mathcal{G}_m\right) f \quad \forall f \in \mathbb{P}_{2m+1}. \tag{3.6}$$

Hence, $\widetilde{\mathcal{G}}_{m+1}^{(1)}$ is the (standard) $(m+1)$-block Gauss quadrature rule with respect to the bilinear form defined by the matrix-valued function $2\mathcal{I} - \mathcal{G}_m$. It follows from (3.6) that the average rule

$$\mathcal{A}_{m+1} := \frac{1}{2}\left(\mathcal{G}_m + \widetilde{\mathcal{G}}_{m+1}^{(1)}\right) \tag{3.7}$$

is exact for all polynomials in $\mathbb{P}_{2m+1}$.

We provide a brief outline of the derivation of the block anti-Gauss rule $\widetilde{\mathcal{G}}_{m+1}^{(1)}$ for completeness; further details can be found in [18]. There is a sequence of orthonormal polynomials $\widetilde{p}_j$ associated with the bilinear from $2\mathcal{I} - \mathcal{G}_m$ defined analogously as (3.5). These polynomials satisfy a recurrence relation of the form

$$\begin{aligned} \lambda\widetilde{p}_{j-1}(\lambda) &= \widetilde{p}_j(\lambda)\widetilde{\Gamma}_j + \widetilde{p}_{j-1}(\lambda)\widetilde{\Omega}_j + \widetilde{p}_{j-2}(\lambda)\widetilde{\Gamma}_{j-1}^T, \quad j = 1, 2, \ldots, \\ \widetilde{p}_0(\lambda) &:= I_k, \quad \widetilde{p}_{-1}(\lambda) := O_k, \end{aligned} \tag{3.8}$$

with matrix-valued recurrence coefficients. The matrices $\widetilde{\Omega}_j \in \mathbb{R}^{k\times k}$ are symmetric and the matrices $\widetilde{\Gamma}_j \in \mathbb{R}^{k\times k}$ are upper triangular.

Introduce the symmetric block tridiagonal matrix

$$\widetilde{T}_{m+1}^{(1)} = \begin{bmatrix} \widetilde{\Omega}_1 & \widetilde{\Gamma}_1^T & & & 0 \\ \widetilde{\Gamma}_1 & \widetilde{\Omega}_2 & \widetilde{\Gamma}_2^T & & \\ & \ddots & \ddots & \ddots & \\ & & \widetilde{\Gamma}_{m-1} & \widetilde{\Omega}_m & \widetilde{\Gamma}_m^T \\ 0 & & & \widetilde{\Gamma}_m & \widetilde{\Omega}_{m+1} \end{bmatrix} \in \mathbb{R}^{k(m+1)\times k(m+1)} \tag{3.9}$$

defined by the matrix-valued recursion coefficients for the polynomials $\widetilde{p}_0, \ldots, \widetilde{p}_{m+1}$; cf. (3.8). The matrix (3.9) determines the block anti-Gauss rule $\widetilde{\mathcal{G}}_{m+1}^{(1)}$.

We can evaluate the block entries of the matrix (3.9) with almost no work from the matrix T_{m+1} that is associated with the $(m+1)$-block Gauss quadrature rule analogous to (3.4). It follows from (3.2) that the matrix recursion coefficients Ω_i and Γ_i are given by

$$\Omega_i = \mathcal{I}(p_{i-1}, \lambda p_{i-1}), \quad \Gamma_i = \mathcal{I}(p_i, \lambda p_{i-1}),$$

and, similarly, we obtain from (3.8) that

$$\widetilde{\Omega}_i = (2\mathcal{I} - \mathcal{G}_m)\,(\widetilde{p}_{i-1}, \lambda\widetilde{p}_{i-1})\,, \quad \widetilde{\Gamma}_i = (2\mathcal{I} - \mathcal{G}_m)\,(\widetilde{p}_i, \lambda\widetilde{p}_{i-1})\,.$$

These relations can be used to show that

$$\begin{aligned} \widetilde{\Omega}_i &= \Omega_i, \quad 1 \le i \le m, \\ \widetilde{\Gamma}_i &= \Gamma_i, \quad 1 \le i \le m-1, \\ \widetilde{p}_i &= p_i, \quad 0 \le i \le m-1; \end{aligned}$$

Moreover, one can choose

$$\widetilde{\Gamma}_m = \sqrt{2}\Gamma_m, \quad \widetilde{\Omega}_{m+1} = \Omega_{m+1}; \tag{3.10}$$

see [18] for details. In conclusion, the matrix $\widetilde{T}_{m+1}^{(1)}$ associated with the $(m+1)$-block anti-Gauss rule $\widetilde{\mathcal{G}}_{m+1}^{(1)}$ can be determined from the matrix T_{m+1} associated with the $(m+1)$-block Gauss rule $\mathcal{G}_{m+1}$ by multiplying the block entries Γ_m and Γ_m^T of the latter by $\sqrt{2}$. The quadrature rule $\widetilde{\mathcal{G}}_{m+1}^{(1)}$ allows the representation

$$\widetilde{\mathcal{G}}_{m+1}^{(1)} f = E_1^T f(\widetilde{T}_{m+1}^{(1)}) E_1,$$

which is analogous to (3.4).

3.1.2 Generalized block anti-Gauss rules, $\ell = 2$

For every positive integer ℓ, the leading $k(m+\ell-1)\times k(m+\ell-1)$ principal submatrix of $\widetilde{T}_{m+\ell}^{(\ell)}$ is $\widetilde{T}_{m+\ell-1}^{(\ell-1)}$. We derive explicit formulas for the coefficients $\widetilde{\Gamma}_{m+1}$ and $\widetilde{\Omega}_{m+2}$ in the last block row of $\widetilde{T}_{m+2}^{(2)}$, if they exist.

Theorem 3.1 *Assume that the matrix $\widetilde{\Gamma}_{m+1}$ is nonsingular. Then the nontrivial block entries in the last block row of the matrix $\widetilde{T}_{m+2}^{(2)}$ satisfy*

$$\widetilde{\Omega}_{m+2} = \widetilde{\Gamma}_{m+1}^{-T}(\Gamma_{m+1}^T\Omega_{m+2}\Gamma_{m+1} - \Gamma_m\Omega_m\Gamma_m^T)\widetilde{\Gamma}_{m+1}^{-1}, \tag{3.11}$$

$$\widetilde{\Gamma}_{m+1}^T\widetilde{\Gamma}_{m+1} = \Gamma_{m+1}^T\Gamma_{m+1} - \Gamma_m\Gamma_m^T. \tag{3.12}$$

Proof We first show (3.12). The relation (3.8) with $j = m+1$ reads

$$\widetilde{p}_{m+1}\widetilde{\Gamma}_{m+1} = \lambda\widetilde{p}_m - \widetilde{p}_m\widetilde{\Omega}_{m+1} - \widetilde{p}_{m-1}\widetilde{\Gamma}_m^T,$$

and from (3.10) and $\widetilde{p}_m = \left(1/\sqrt{2}\right) p_m$, we obtain

$$\widetilde{p}_{m+1}\widetilde{\Gamma}_{m+1} = \frac{\lambda}{\sqrt{2}}p_m - \frac{1}{\sqrt{2}}p_m\Omega_{m+1} - \sqrt{2}p_{m-1}\Gamma_m^T,$$

which yields

$$\begin{aligned}\widetilde{p}_{m+1} &= \frac{1}{\sqrt{2}}(\lambda p_m - p_m\Omega_{m+1} - 2p_{m-1}\Gamma_m^T)\widetilde{\Gamma}_{m+1}^{-1}\\ &= \frac{1}{\sqrt{2}}(p_{m+1}\Gamma_{m+1} - p_{m-1}\Gamma_m^T)\widetilde{\Gamma}_{m+1}^{-1}.\end{aligned} \tag{3.13}$$

Hence,

$$\widetilde{\Gamma}_{m+1} = (2\mathcal{I} - \mathcal{G}_m)\,(\widetilde{p}_{m+1}, \lambda\widetilde{p}_m) = 2\mathcal{I}\,(\widetilde{p}_{m+1}, \lambda\widetilde{p}_m) - \mathcal{G}_m\,(\widetilde{p}_{m+1}, \lambda\widetilde{p}_m)\,,$$

where $\mathcal{G}_m\,(\widetilde{p}_{m+1}, \lambda\widetilde{p}_m) = O_k$ due to [18, Theorem 2]. Moreover, it follows from (3.13) that

$$\begin{aligned}\widetilde{\Gamma}_{m+1} &= 2\mathcal{I}\left(\left(\frac{1}{\sqrt{2}}\right)(p_{m+1}\Gamma_{m+1} - p_{m-1}\Gamma_m^T)\widetilde{\Gamma}_{m+1}^{-1}, \left(\frac{1}{\sqrt{2}}\right)\lambda p_m\right)\\ &= \widetilde{\Gamma}_{m+1}^{-T}(\mathcal{I}\,(p_{m+1}\Gamma_{m+1}, \lambda p_m) - \mathcal{I}\left(p_{m-1}\Gamma_m^T, \lambda p_m\right)).\end{aligned}$$

Therefore,

$$\begin{aligned}\widetilde{\Gamma}_{m+1}^T\widetilde{\Gamma}_{m+1} &= \Gamma_{m+1}^T\mathcal{I}\,(p_{m+1}, \lambda p_m) - \Gamma_m\mathcal{I}\,(p_{m-1}, \lambda p_m)\\ &= \Gamma_{m+1}^T\Gamma_{m+1} - \Gamma_m\Gamma_m^T.\end{aligned}$$

We turn to (3.11). By using the relation (3.13) as above, we obtain

$$\begin{aligned}
2\mathcal{I}\,(\tilde{p}_{m+1}, \lambda\tilde{p}_{m+1}) &= 2\mathcal{I}\bigg(\frac{1}{\sqrt{2}}\big(p_{m+1}\Gamma_{m+1} - p_{m-1}\Gamma_m^T\big)\widetilde{\Gamma}_{m+1}^{-1}, \\
&\qquad \frac{\lambda}{\sqrt{2}}\big(p_{m+1}\Gamma_{m+1} - p_{m-1}\Gamma_m^T\big)\widetilde{\Gamma}_{m+1}^{-1}\bigg) \\
&= \widetilde{\Gamma}_{m+1}^{-T}(\mathcal{I}(p_{m+1}\Gamma_{m+1}, \lambda p_{m+1}\Gamma_{m+1}) \\
&\quad + \mathcal{I}(p_{m-1}\Gamma_m^T, \lambda p_{m-1}\Gamma_m^T))\widetilde{\Gamma}_{m+1}^{-1} \\
&= \widetilde{\Gamma}_{m+1}^{-T}(\Gamma_{m+1}^T\mathcal{I}(p_{m+1}, \lambda p_{m+1})\Gamma_{m+1} \\
&\quad + \Gamma_m\mathcal{I}(p_{m-1}, \lambda p_{m-1})\Gamma_m^T)\widetilde{\Gamma}_{m+1}^{-1} \\
&= \widetilde{\Gamma}_{m+1}^{-T}(\Gamma_{m+1}^T\Omega_{m+2}\Gamma_{m+1} + \Gamma_m\Omega_m\Gamma_m^T)\widetilde{\Gamma}_{m+1}^{-1}
\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}
\mathcal{G}_m\,(\tilde{p}_{m+1}, \lambda\tilde{p}_{m+1}) &= \mathcal{G}_m\bigg((\lambda\widetilde{p}_m - \widetilde{p}_m\widetilde{\Omega}_{m+1} - \widetilde{p}_{m-1}\widetilde{\Gamma}_m^T)\widetilde{\Gamma}_{m+1}^{-1}, \\
&\qquad \lambda(\lambda\widetilde{p}_m - \widetilde{p}_m\widetilde{\Omega}_{m+1} - \widetilde{p}_{m-1}\widetilde{\Gamma}_m^T)\widetilde{\Gamma}_{m+1}^{-1}\bigg) \\
&= \widetilde{\Gamma}_{m+1}^{-T}\bigg(\mathcal{G}_m\bigg(\frac{\lambda}{\sqrt{2}}p_m - \frac{1}{\sqrt{2}}p_m\Omega_{m+1} - \sqrt{2}p_{m-1}\Gamma_m^T), \\
&\qquad \lambda\bigg(\frac{\lambda}{\sqrt{2}}p_m - \frac{1}{\sqrt{2}}p_m\Omega_{m+1} - \sqrt{2}p_{m-1}\Gamma_m^T\bigg)\bigg)\bigg)\widetilde{\Gamma}_{m+1}^{-1} \\
&= \widetilde{\Gamma}_{m+1}^{-T}(\mathcal{G}_m\big(-\sqrt{2}p_{m-1}\Gamma_m^T, \lambda(-\sqrt{2}p_{m-1}\Gamma_m^T)\big))\widetilde{\Gamma}_{m+1}^{-1} \\
&= \widetilde{\Gamma}_{m+1}^{-T}\big(2\Gamma_m\mathcal{I}(p_{m-1}, \lambda p_{m-1})\Gamma_m^T\big)\widetilde{\Gamma}_{m+1}^{-1} \\
&= \widetilde{\Gamma}_{m+1}^{-T}\big(2\Gamma_m\Omega_m\Gamma_m^T\big)\widetilde{\Gamma}_{m+1}^{-1}.
\end{aligned} \tag{3.15}$$

Finally, combining (3.14) and (3.15) yields

$$\begin{aligned}
\widetilde{\Omega}_{m+2} &= (2\mathcal{I} - \mathcal{G}_m)(\widetilde{p}_{m+1}, \widetilde{p}_{m+1}) \\
&= \widetilde{\Gamma}_{m+1}^{-T}(\Gamma_{m+1}^T\Omega_{m+2}\Gamma_{m+1} + \Gamma_m\Omega_m\Gamma_m^T - 2\Gamma_m\Omega_m\Gamma_m^T)\widetilde{\Gamma}_{m+1}^{-1}.
\end{aligned}$$

The right-hand side simplifies to (3.11). □

When the right-hand side of (3.12) is symmetric positive definite, we let the matrix $\widetilde{\Gamma}_{m+1}$ be the Cholesky factor of the right-hand side; otherwise we choose the last sub- and super-diagonal blocks of $\widetilde{T}_{m+2}^{(2)}$ to be the right-hand side of (3.12) and the identity, respectively. Different choices of the last sub- and super-diagonal blocks correspond to different similarity transformations of the matrix $\widetilde{T}_{m+2}^{(2)}$. The quadrature rule $\widetilde{\mathcal{G}}_{m+2}^{(2)}$ can be shown to be invariant under these transformations. This follows similarly as Proposition 2.3.

3.1.3 Generalized block anti-Gauss rules, $\ell = 3$

It is quite tedious to compute the last block entries $\widetilde{\Gamma}_{m+2}$ and $\widetilde{\Omega}_{m+3}$ of the matrix $\widetilde{T}_{m+3}^{(3)}$. We therefore only outline how to determine

$$\widetilde{\Gamma}_{m+2} = (2\mathcal{I} - \mathcal{G}_m)\,(\widetilde{p}_{m+2}, \lambda\widetilde{p}_{m+1})$$

and propose to use a simplified generalized block anti-Gauss rule defined by a block tridiagonal matrix $\widehat{T}_{m+3}^{(3)}$, which is obtained by replacing the last diagonal block entry $\widetilde{\Omega}_{m+3}$ of $\widetilde{T}_{m+3}^{(3)}$ by $\widehat{\Omega}_{m+3} = \widetilde{\Omega}_{m+2}$. The quadrature rule $\widehat{\mathcal{G}}_{m+3}^{(3)}$ determined by the matrix $\widehat{T}_{m+3}^{(3)}$ satisfies

$$\widehat{\mathcal{G}}_{m+3}^{(3)} f = \widetilde{\mathcal{G}}_{m+3}^{(3)} f \quad \forall f \in \mathbb{P}_{2m+4};$$

see Theorem 3.5 below.

Theorem 3.2 *Assume that the matrix $\widetilde{\Gamma}_{m+1}\widetilde{\Gamma}_{m+2}$ is nonsingular. Then the last subdiagonal block entry $\widetilde{\Gamma}_{m+2}$ of the matrix $\widehat{T}_{m+3}^{(3)}$ satisfies*

$$\begin{aligned}\widetilde{\Gamma}_{m+2}^T\widetilde{\Gamma}_{m+2} = \widetilde{\Gamma}_{m+1}^{-T}(&\Gamma_{m+1}^T\Gamma_{m+2}^T\Gamma_{m+2}\Gamma_{m+1} - \Gamma_{m-1}^T\Gamma_m\Gamma_{m-1}^T\Gamma_m^T \\ &- (\widetilde{\Omega}_{m+2}-\Omega_{m+2})\Gamma_{m+1}^T\Omega_{m+2}\Gamma_{m+1} + (\widetilde{\Omega}_{m+2}-\Omega_m)\Gamma_m\Omega_m\Gamma_m^T)\widetilde{\Gamma}_{m+1}^{-1}.\end{aligned}$$

Proof We first express $\widetilde{p}_{m+2}$ in terms of polynomials p_j. The relation

$$\widetilde{p}_{m+2} = (\lambda\widetilde{p}_{m+1} - \widetilde{p}_{m+1}\widetilde{\Omega}_{m+2} - \widetilde{p}_m\widetilde{\Gamma}_{m+1})\widetilde{\Gamma}_{m+2}^{-1}$$

gives

$$\begin{aligned}\widetilde{p}_{m+2} = \frac{1}{\sqrt{2}}(&p_{m+2}\Gamma_{m+2}\Gamma_{m+1} + p_{m+1}\Gamma_{m+1}(\Omega_{m+2} - \widetilde{\Omega}_{m+2}) \\ &+ p_{m-1}\Gamma_m^T(\widetilde{\Omega}_{m+2} - \Omega_m) - p_{m-2}\Gamma_m^T\Gamma_{m-1}^T)\widetilde{\Gamma}_{m+1}^{-1}\widetilde{\Gamma}_{m+2}^{-1}.\end{aligned}$$

We obtain in the same manner as in Sect. 3.1.2 that

$$\begin{aligned}2\mathcal{I}(\widetilde{p}_{m+2}, \lambda\widetilde{p}_{m+1}) &= \widetilde{\Gamma}_{m+2}^{-T}\widetilde{\Gamma}_{m+1}^{-T}\Big(\Gamma_{m+1}^T\Gamma_{m+2}^T\Gamma_{m+2}\Gamma_{m+1} \\ &\quad - (\widetilde{\Omega}_{m+2} - \Omega_{m+2})\Gamma_{m+1}^T\Omega_{m+2}\Gamma_{m+1} - (\widetilde{\Omega}_{m+2} - \Omega_m)\Gamma_m\Omega_m\Gamma_m^T \\ &\quad + \Gamma_{m-1}\Gamma_m\Gamma_{m-1}^T\Gamma_m^T\Big)\widetilde{\Gamma}_{m+1}^{-1}, \\ \mathcal{G}_m(\widetilde{p}_{m+2}, \lambda\widetilde{p}_{m+2}) &= \widetilde{\Gamma}_{m+2}^{-T}\widetilde{\Gamma}_{m+1}^{-T}\big(2(\Omega_m - \widetilde{\Omega}_{m+2})\Gamma_m\Omega_m\Gamma_m^T + 2\Gamma_{m-1}\Gamma_m\Gamma_{m-1}^T\Gamma_m^T\big)\widetilde{\Gamma}_{m+1}^{-1}.\end{aligned}$$

The theorem now follows after some extensive computations. □

3.2 Generalized block anti-Gauss rules for functions of a nonsymmetric matrix

We extend the discussion of the previous subsection to matrix functions (1.2) with a large, possibly nonsymmetric, matrix $A \in \mathbb{R}^{n\times n}$ and block vectors $V, W \in \mathbb{R}^{n\times k}$ such that $V^T W = I_k$. Substituting the spectral factorization (2.20) into (1.2) yields

$$W^T f(A)V = \widetilde{W} f(\Lambda)\widetilde{V}^T = \sum_{i=1}^{n} f(\lambda_i)\widetilde{v}_i \widetilde{v}_i'^T = \int f(x)d\widetilde{w}(x) =: \mathcal{I}f, \quad (3.16)$$

where $\widetilde{W} = [\widetilde{v}_1, \ldots, \widetilde{v}_n] = W^T S \in \mathbb{R}^{n\times k}$, $\widetilde{V} = [\widetilde{v}_1', \ldots, \widetilde{v}_n'] = (S^{-1}V)^T \in \mathbb{R}^{n\times k}$, and $\widetilde{w} : \mathbb{R} \to \mathbb{R}^{b\times k}$ is a matrix-valued distribution with mass $\widetilde{v}_i \widetilde{v}_i'^T$ at the eigenvalue λ_i of A for $1 \le i \le n$.

There are two sequences of polynomials p_j and q_j, $j = 0, 1, \ldots$, that are biorthonormal with respect to a bilinear form associated with the function $\mathcal{I}$ in (3.16). This bilinear form is defined analogously as (3.5). The polynomials satisfy recursion relations

$$\begin{aligned} \lambda p_{j-1}(\lambda) &= p_j(\lambda)\Gamma_j + p_{j-1}(\lambda)\Omega_j + p_{j-2}(\lambda)\Delta_{j-1}^T, \\ \lambda q_{j-1}(\lambda) &= q_j(\lambda)\Delta_j + q_{j-1}(\lambda)\Omega_j^T + q_{j-2}(\lambda)\Gamma_{j-1}^T, \\ p_0(\lambda) &:= I_k, \quad q_0(\lambda) := I_k, \quad p_{-1}(\lambda) := O_k, \quad q_{-1}(\lambda) := O_k, \end{aligned} \quad (3.17)$$

for $j = 1, 2, \ldots$, whose the recursion coefficients Γ_j, Ω_j, and Δ_j are real $k \times k$ matrices; see [18] for details.

Letting

$$\begin{aligned} P_m(\lambda) &:= \big[p_0(\lambda), \ldots, p_{m-1}(\lambda)\big] \in \mathbb{R}^{k\times km}, \\ Q_m(\lambda) &:= \big[q_0(\lambda), \ldots, q_{m-1}(\lambda)\big] \in \mathbb{R}^{k\times km}, \end{aligned}$$

the recursion relations (3.17) for the polynomials $p_0, p_1, \ldots, p_m$ and $q_0, q_1, \ldots, q_m$ can be expressed as

$$\begin{aligned} \lambda P_m(\lambda) &= P_m(\lambda)T_m + p_m(\lambda)\Gamma_m E_m^T, \\ \lambda Q_m(\lambda) &= Q_m(\lambda)T_m^T + q_m(\lambda)\Delta_m E_m^T, \end{aligned}$$

where

$$T_m = \begin{bmatrix} \Omega_1 & \Delta_1^T & & & \\ \Gamma_1 & \Omega_2 & \Delta_2^T & & \\ & \ddots & \ddots & \ddots & \\ & & \Gamma_{m-2} & \Omega_{m-1} & \Delta_{m-1}^T \\ & & & \Gamma_{m-1} & \Omega_m \end{bmatrix} \in \mathbb{R}^{km\times km} \quad (3.18)$$

is a block tridiagonal matrix. It can be determined by m steps of the nonsymmetric block Lanczos process applied to $A \in \mathbb{R}^{n\times n}$ with initial block vectors $V, W \in \mathbb{R}^{n\times k}$.

The implementation described by Algorithm 4 is proposed by Bai et al. [4]. Line 9 of the algorithm computes QR factorizations of the matrices R_j and S_j, and in line 10 the singular value decomposition of $Q_S^T Q_R$ is evaluated. Thus, $\Sigma \in \mathbb{R}^{k\times k}$ in line 10 is a diagonal matrix, whose diagonal entries are assumed to be positive, and the matrices U and Z in the same line are orthogonal.

Algorithm 4 The nonsymmetric block Lanczos process.

1: **Input:** nonsymmetric matrix $A \in \mathbb{R}^{n\times n}$, initial block vectors $W, V \in \mathbb{R}^{n\times k}$
2: such that $V^T W = I_k$, number of steps m.
3: $W_0 := V_0 := O \in \mathbb{R}^{n\times k}$, $\Delta_0 := \Gamma_0 := O_k$, $V_1 := V$, $W_1 := W$
4: **for** $j = 1$ **to** m
5: $B := AV_j - V_{j-1}\Delta_{j-1}^T$
6: $\Omega_j := W_j^T B$
7: $R_j := B - V_j \Omega_j$
8: $S_j := A^T W_j - W_j \Omega_j^T - W_{j-1}\Gamma_{j-1}^T$
9: $Q_R R_R := R_j$, $Q_S R_S := S_j$
10: $U\Sigma Z^T := Q_S^T Q_R$
11: $V_{j+1} := Q_R Z \Sigma^{-\frac{1}{2}}$, $W_{j+1} := Q_S U \Sigma^{-\frac{1}{2}}$
12: $\Gamma_j := \Sigma^{\frac{1}{2}} Z^T R_R$, $\Delta_j := \Sigma^{\frac{1}{2}} U^T R_S$
13: **end for**
14: **Output:** Blocks $\Omega_1, \Omega_2, \ldots, \Omega_m, \Gamma_1, \Gamma_2, \ldots, \Gamma_{m-1}$, and $\Delta_1^T, \Delta_2^T, \ldots, \Delta_{m-1}^T$
15: of the matrix (3.18).

Consider the block quadrature rule

$$\mathcal{G}_m f = E_1^T f(T_m) E_1, \tag{3.19}$$

where the matrix T_m is given by (3.18). Fenu et al. [18, Section 5] show that

$$\mathcal{G}_m f = \mathcal{I} f \quad \forall f \in \mathbb{P}_{2m-1} \tag{3.20}$$

and, therefore, refer to $\mathcal{G}_m$ as an m-block Gauss quadrature rule associated with the operator $\mathcal{I}$ defined by (3.16).

3.2.1 Generalized block anti-Gauss rules, $\ell = 1$

We seek to determine a matrix-valued $(m+1)$-block anti-Gauss quadrature rule $\widetilde{\mathcal{G}}_{m+1}^{(1)}$ such that

$$\left(\mathcal{I} - \widetilde{\mathcal{G}}_{m+1}^{(1)}\right) f = -\left(\mathcal{I} - \mathcal{G}_m\right) f, \quad f \in \mathbb{P}_{2m+1}. \tag{3.21}$$

This relation implies, analogously to the discussion following (3.6), that $\widetilde{\mathcal{G}}_{m+1}^{(1)}$ is an $(m+1)$-block Gauss quadrature rule with respect to a bilinear form determined by the matrix-valued function $2\mathcal{I} - \mathcal{G}_m$. The average rule (3.7) with $\mathcal{G}_m$ and $\mathcal{G}_{m+1}^{(1)}$

determined by (3.19) and (3.21), respectively, is exact for all $p \in \mathbb{P}_{2m+1}$. A derivation of the, generally nonsymmetric, block tridiagonal matrix (3.23) below associated with the quadrature rule $\widetilde{\mathcal{G}}_{m+1}^{(1)}$ can be found in [18]. We therefore here only outline the derivation with the purpose of introducing notation that will be used subsequently.

There are sequences of polynomials $\widetilde{p}_j$ and $\widetilde{q}_j$, $j = 0, 1, \ldots$, that are biorthonormal with respect to a bilinear form determined by the matrix-valued function $2\mathcal{I} - \mathcal{G}_m$. These polynomials satisfy recursion relations for $j = 1, 2, \ldots$ of the form

$$\begin{aligned} \lambda \widetilde{p}_{j-1}(\lambda) &= \widetilde{p}_j(\lambda)\widetilde{\Gamma}_j + \widetilde{p}_{j-1}(\lambda)\widetilde{\Omega}_j + \widetilde{p}_{j-2}(\lambda)\widetilde{\Delta}_{j-1}^T, \\ \lambda \widetilde{q}_{j-1}(\lambda) &= \widetilde{q}_j(\lambda)\widetilde{\Delta}_j + \widetilde{q}_{j-1}(\lambda)\widetilde{\Omega}_j^T + \widetilde{q}_{j-2}(\lambda)\widetilde{\Gamma}_{j-1}^T, \\ \widetilde{p}_0(\lambda) &:= I_k, \quad \widetilde{q}_0(\lambda) := I_k, \quad \widetilde{p}_{-1}(\lambda) := O_k, \quad \widetilde{q}_{-1}(\lambda) := O_k. \end{aligned} \tag{3.22}$$

The recursion coefficients $\widetilde{\Omega}_j, \widetilde{\Gamma}_j, \widetilde{\Delta}_j^T$ are real $k \times k$ matrices, that define the block tridiagonal matrix

$$\widetilde{T}_{m+1}^{(1)} = \begin{bmatrix} \widetilde{\Omega}_1 & \widetilde{\Delta}_1^T & & & \\ \widetilde{\Gamma}_1 & \widetilde{\Omega}_2 & \widetilde{\Delta}_2^T & & \\ & \ddots & \ddots & \ddots & \\ & & \widetilde{\Gamma}_{m-1} & \widetilde{\Omega}_m & \widetilde{\Delta}_m^T \\ & & & \widetilde{\Gamma}_m & \widetilde{\Omega}_{m+1} \end{bmatrix} \in \mathbb{R}^{k(m+1)\times k(m+1)}, \tag{3.23}$$

which determines the $(m+1)$-block anti-Gauss rule

$$\widetilde{\mathcal{G}}_{m+1}^{(1)} f = E_1^T f(\widetilde{T}_{m+1}^{(1)}) E_1. \tag{3.24}$$

We turn to the computation of the block entries of the matrix (3.23). The recursions (3.17) and (3.22), together with (3.21) and [18, Corollary 8], yield

$$\begin{aligned} \widetilde{\Omega}_i &= \Omega_i, && && 1 \le i \le m+1, \\ \widetilde{\Gamma}_i &= \Gamma_i, && \widetilde{\Delta}_i = \Delta_i, && 1 \le i \le m-1, \\ \widetilde{p}_i &= p_i, && \widetilde{q}_i = q_i, && 0 \le i \le m-1, \end{aligned} \tag{3.25}$$

and $\widetilde{\Delta}_m^T \widetilde{\Gamma}_m = 2\Delta_m^T \Gamma_m$. Moreover, $\widetilde{\Omega}_{m+1} = \Omega_{m+1}$; see [18] for details. We choose

$$\widetilde{\Gamma}_m := \sqrt{2}\Gamma_m, \quad \widetilde{\Delta}_m := \sqrt{2}\Delta_m. \tag{3.26}$$

Thus, similarly as in Sect. 3.1.1, the matrix $\widetilde{T}_{m+1}^{(1)}$ given by (3.23) can be determined from the matrix T_{m+1} associated with the $(m+1)$-block Gauss rule (3.24) by multiplying the last off-diagonal blocks Γ_m and Δ_m by $\sqrt{2}$.

We note for future reference that the relations (3.25) yield

$$\begin{aligned} \widetilde{p}_m\widetilde{\Gamma}_m &= \lambda\widetilde{p}_{m-1} - \widetilde{p}_{m-1}\widetilde{\Omega}_m - \widetilde{p}_{m-2}\widetilde{\Delta}_{m-1}^T \\ &= \lambda p_{m-1} - p_{m-1}\Omega_m - p_{m-2}\Delta_{m-1}^T = p_m\Gamma_m, \\ \widetilde{q}_m\widetilde{\Delta}_m &= \lambda\widetilde{q}_{m-1} - \widetilde{q}_{m-1}\widetilde{\Omega}_m^T - \widetilde{q}_{m-2}\widetilde{\Gamma}_{m-1}^T \\ &= \lambda q_{m-1} - q_{m-1}\Omega_m^T - q_{m-2}\Gamma_{m-1}^T = q_m\Delta_m; \end{aligned} \tag{3.27}$$

see [18] for a detailed derivation.

3.2.2 Generalized block anti-Gauss rules, $\ell = 2$

We derive explicit formulas for the coefficients $\widetilde{\Gamma}_{m+1}$, $\widetilde{\Delta}_{m+1}$ and $\widetilde{\Omega}_{m+2}$ in the last block row and column of $\widetilde{T}_{m+2}^{(2)}$, if they exist. The matrix $\widetilde{T}_{m+2}^{(2)}$ is defined analogously as (3.23), and has the latter as $k(m+1)\times k(m+1)$ leading principal submatrix.

Theorem 3.3 *Assume that $\widetilde{\Gamma}_{m+1}\widetilde{\Delta}_{m+1}$ is nonsingular. Then the nontrivial block entries in the last block row and column of the matrix $\widetilde{T}_{m+2}^{(2)}$ satisfy*

$$\widetilde{\Omega}_{m+2} = \widetilde{\Delta}_{m+1}^{-T}(\Delta_{m+1}^T\Omega_{m+2}\Gamma_{m+1} - \Gamma_m\Omega_m\Delta_m^T)\widetilde{\Gamma}_{m+1}^{-1}, \tag{3.28}$$

$$\widetilde{\Delta}_{m+1}^T\widetilde{\Gamma}_{m+1} = \Delta_{m+1}^T\Gamma_{m+1} - \Gamma_m\Delta_m^T. \tag{3.29}$$

Proof The proof uses a formula analogous to (3.5) that allows the evaluation of the function $\mathcal{I}$ in (3.16) with two matrix-valued arguments. Similarly, a modification of the block Gauss rule (3.19) that allows two matrix-valued arguments is used; we refer to [18] for their exact definition.

We first show (3.29). From (3.27), we have

$$\widetilde{q}_m\widetilde{\Delta}_m = \lambda\widetilde{q}_{m-1} - \widetilde{q}_{m-1}\widetilde{\Omega}_m - \widetilde{q}_{m-2}\widetilde{\Gamma}_{m-1}^T,$$

and by (3.26) and $\widetilde{q}_m = \left(1/\sqrt{2}\right) q_m$ it follows that

$$\widetilde{q}_{m+1}\widetilde{\Delta}_{m+1} = \frac{\lambda}{\sqrt{2}}q_m - \frac{1}{\sqrt{2}}q_m\Omega_{m+1} - \sqrt{2}q_{m-1}\Gamma_m^T$$

and

$$\begin{aligned} \widetilde{q}_{m+1} &= \left(\frac{1}{\sqrt{2}}(\lambda q_m - q_m\Omega_{m+1} - 2q_{m-1}\Gamma_m^T)\widetilde{\Delta}_{m+1}^{-1}\right) \\ &= \frac{1}{\sqrt{2}}(q_{m+1}\Delta_{m+1} - q_{m-1}\Gamma_m^T)\widetilde{\Delta}_{m+1}^{-1}. \end{aligned} \tag{3.30}$$

Hence,

$$\widetilde{\Gamma}_{m+1} = (2\mathcal{I} - \mathcal{G}_m)\,(\widetilde{q}_{m+1}, \lambda\widetilde{p}_m) = 2\mathcal{I}\,(\widetilde{q}_{m+1}, \lambda\widetilde{p}_m) - \mathcal{G}_m\,(\widetilde{q}_{m+1}, \lambda\widetilde{p}_m)\,,$$

where $\mathcal{G}_m\,(\widetilde{q}_{m+1}, \lambda\widetilde{p}_m) = O_k$ due to [18, Theorem 4]. In view of (3.30), we also have

$$\widetilde{\Gamma}_{m+1} = 2\mathcal{I}\left(\frac{1}{\sqrt{2}}(q_{m+1}\Delta_{m+1} - q_{m-1}\Gamma_m^T)\widetilde{\Delta}_{m+1}^{-1}, \frac{1}{\sqrt{2}}\lambda p_m\right)$$

and

$$\begin{aligned}
\widetilde{\Gamma}_{m+1} &= \widetilde{\Delta}_{m+1}^{-T}\left(\mathcal{I}\,(q_{m+1}\Delta_{m+1}, \lambda p_m) - \mathcal{I}\left(q_{m-1}\Gamma_m^T, \lambda p_m\right)\right),\\
\widetilde{\Delta}_{m+1}^T\widetilde{\Gamma}_{m+1} &= \mathcal{I}\,(q_{m+1}\Delta_{m+1}, \lambda p_m) - \mathcal{I}\left(q_{m-1}\Gamma_m^T, \lambda p_m\right)\\
&= \Delta_{m+1}^T\mathcal{I}\,(q_{m+1}, \lambda p_m) - \Gamma_m\mathcal{I}\,(q_{m-1}, \lambda p_m)\\
&= \Delta_{m+1}^T\Gamma_{m+1} - \Gamma_m\Delta_m^T.
\end{aligned}$$

We turn to (3.28). Using the relation (3.30) as above, we obtain

$$\begin{aligned}
2\mathcal{I}\,(\widetilde{q}_{m+1}, \lambda\widetilde{p}_{m+1}) &= 2\mathcal{I}\Big(\frac{1}{\sqrt{2}}(q_{m+1}\Delta_{m+1} - q_{m-1}\Gamma_m^T)\widetilde{\Delta}_{m+1}^{-1},\\
&\qquad \frac{\lambda}{\sqrt{2}}(p_{m+1}\Gamma_{m+1} - p_{m-1}\Delta_m^T)\widetilde{\Gamma}_{m+1}^{-1}\Big)\\
&= \widetilde{\Delta}_{m+1}^{-T}\Big(\mathcal{I}(q_{m+1}\Delta_{m+1}, \lambda p_{m+1}\Gamma_{m+1}) + \mathcal{I}(q_{m-1}\Gamma_m^T, \lambda p_{m-1}\Delta_m^T)\Big)\widetilde{\Gamma}_{m+1}^{-1}\\
&= \widetilde{\Delta}_{m+1}^{-T}\Big(\Delta_{m+1}^T\mathcal{I}(q_{m+1}, \lambda p_{m+1})\Gamma_{m+1} + \Gamma_m\mathcal{I}(q_{m-1}, \lambda p_{m-1})\Delta_m^T\Big)\widetilde{\Gamma}_{m+1}^{-1}\\
&= \widetilde{\Delta}_{m+1}^{-T}\Big(\Delta_{m+1}^T\Omega_{m+2}\Gamma_{m+1} + \Gamma_m\Omega_m\Delta_m^T\Big)\widetilde{\Gamma}_{m+1}^{-1} \qquad (3.31)
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{G}_m\,(\widetilde{q}_{m+1}, \lambda\widetilde{p}_{m+1}) &= \mathcal{G}_m\Big(\big(\lambda\widetilde{q}_m - \widetilde{q}_m\widetilde{\Omega}_{m+1} - \widetilde{q}_{m-1}\widetilde{\Gamma}_m^T\big)\widetilde{\Delta}_{m+1}^{-1},\\
&\qquad \lambda\big(\lambda\widetilde{p}_m - \widetilde{p}_m\widetilde{\Omega}_{m+1} - \widetilde{p}_{m-1}\widetilde{\Delta}_m^T\big)\widetilde{\Gamma}_{m+1}^{-1}\Big)\\
&= \widetilde{\Delta}_{m+1}^{-T}\mathcal{G}_m\Big(\frac{\lambda}{\sqrt{2}}q_m - \frac{1}{\sqrt{2}}q_m\Omega_{m+1} - \sqrt{2}q_{m-1}\Gamma_m^T),\\
&\qquad \Big(\lambda\Big(\frac{\lambda}{\sqrt{2}}p_m - \frac{1}{\sqrt{2}}p_m\Omega_{m+1} - \sqrt{2}p_{m-1}\Delta_m^T\Big)\Big)\widetilde{\Gamma}_{m+1}^{-1}\\
&= \widetilde{\Delta}_{m+1}^{-T}\mathcal{G}_m\big(-\sqrt{2}q_{m-1}\Gamma_m^T, \lambda(-\sqrt{2}p_{m-1}\Delta_m^T)\big)\widetilde{\Gamma}_{m+1}^{-1}\\
&= \widetilde{\Delta}_{m+1}^{-T}\big(2\Gamma_m\mathcal{I}(q_{m-1}, \lambda p_{m-1})\Delta_m^T\big)\widetilde{\Gamma}_{m+1}^{-1}\\
&= \widetilde{\Delta}_{m+1}^{-T}(2\Gamma_m\Omega_m\Delta_m^T)\widetilde{\Gamma}_{m+1}^{-1}. \qquad (3.32)
\end{aligned}$$

It now follows from (3.31) and (3.32) that

$$\begin{aligned}
\widetilde{\Omega}_{m+2} &= (2\mathcal{I} - \mathcal{G}_m)(\widetilde{q}_{m+1}, \lambda\widetilde{p}_{m+1})\\
&= \widetilde{\Delta}_{m+1}^{-T}(\Delta_{m+1}^T\Omega_{m+2}\Gamma_{m+1} + \Gamma_m\Omega_m\Delta_m^T - 2\Gamma_m\Omega_m\Delta_m^T)\widetilde{\Gamma}_{m+1}^{-1},
\end{aligned}$$

where the right-hand side simplifies to (3.28). □

The last off-diagonal block entries of $\widetilde{T}_{m+2}^{(2)}$ can be chosen arbitrarily so that (3.29) holds. The different choices correspond to similarity transformations of $\widetilde{T}_{m+2}^{(2)}$ under which the quadrature rule $\widetilde{\mathcal{G}}_{m+2}^{(2)}$ is invariant. The invariance can be shown analogously as Proposition 2.3.

3.2.3 Generalized block anti-Gauss rules, $\ell = 3$

We derive a relation for the last subdiagonal and superdiagonal block entries $\widetilde{\Gamma}_{m+2}$ and $\widetilde{\Delta}_{m+2}^T$ of $\widetilde{T}_{m+3}^{(3)}$.

Theorem 3.4 *Assume that the matrices $\widetilde{\Gamma}_{m+1}$ and $\widetilde{\Delta}_{m+1}\widetilde{\Delta}_{m+2}$ are nonsingular. Then the last subdiagonal block entries $\widetilde{\Gamma}_{m+2}$ and $\widetilde{\Delta}_{m+2}$ of $\widetilde{T}_{m+3}^{(3)}$ satisfy*

$$\begin{aligned}\widetilde{\Delta}_{m+2}^T\widetilde{\Gamma}_{m+2} = \widetilde{\Delta}_{m+1}^{-T}\Big(&\Delta_{m+1}^T\Delta_{m+2}^T\Gamma_{m+2}\Gamma_{m+1} - \Gamma_{m-1}\Gamma_m\Delta_{m-1}^T\Delta_m^T\\ &+ (\widetilde{\Omega}_{m+2} - \Omega_{m+2})^T\Delta_{m+1}^T\Omega_{m+2}\Gamma_{m+1}\\ &- (\widetilde{\Omega}_{m+2} - \Omega_m)^T\Gamma_m\Omega_m\Delta_m^T\Big)\widetilde{\Gamma}_{m+1}^{-1}.\end{aligned}$$

Proof We first express $\widetilde{q}_{m+2}$ in terms of polynomials q_j. It follows from

$$\widetilde{q}_{m+2} = \big(\lambda\widetilde{q}_{m+1} - \widetilde{q}_{m+1}\widetilde{\Omega}_{m+2} - \widetilde{p}_m\widetilde{\Gamma}_{m+1}\big)\widetilde{\Delta}_{m+2}^{-1}$$

that

$$\begin{aligned}\widetilde{q}_{m+2} = \frac{1}{\sqrt{2}}\Big(&q_{m+2}\Delta_{m+2}\Delta_{m+1} - q_{m+1}\Delta_{m+1}(\widetilde{\Omega}_{m+2} - \Omega_{m+2})\\ &+ q_{m-1}\Gamma_m^T(\widetilde{\Omega}_{m+2} - \Omega_m) - q_{m-2}\Gamma_m^T\Gamma_{m-1}^T\Big)\widetilde{\Delta}_{m+1}^{-1}\widetilde{\Delta}_{m+2}^{-1},\end{aligned}$$

and we obtain

$$\begin{aligned}2\mathcal{I}(\widetilde{q}_{m+2}, \lambda\widetilde{p}_{m+1}) = \widetilde{\Delta}_{m+2}^{-T}\widetilde{\Delta}_{m+1}^{-T}\Big(&\Delta_{m+1}^T\Delta_{m+2}^T\Gamma_{m+2}\Gamma_{m+1}\\ &+ (\widetilde{\Omega}_{m+2} - \Omega_{m+2})^T\Delta_{m+1}^T\Omega_{m+2}\Gamma_{m+1}\\ &- (\widetilde{\Omega}_{m+2} - \Omega_m)^T\Gamma_m\Omega_m\Delta_m^T\\ &+ \Gamma_{m-1}\Gamma_m\Delta_{m-1}^T\Delta_m^T\Big)\widetilde{\Gamma}_{m+1}^{-1},\\ \mathcal{G}_m(\widetilde{q}_{m+2}, \lambda\widetilde{p}_{m+1}) = \widetilde{\Delta}_{m+2}^{-T}\widetilde{\Delta}_{m+1}^{-T}\big(&2(\Omega_m - \widetilde{\Omega}_{m+2})^T\Gamma_m\Omega_m\Delta_m^T\\ &+ 2\Gamma_{m-1}\Gamma_m\Delta_{m-1}^T\Delta_m^T\big)\widetilde{\Gamma}_{m+1}^{-1}.\end{aligned}$$

The theorem now follows after some computations. They are tedious but fairly straightforward. We therefore omit the details. □

The formula for the last diagonal block, $\widetilde{\Omega}_{m+3}$, of $\widetilde{T}_{m+3}^{(3)}$ is quite complicated. We therefore set this block to $\widetilde{\Omega}_{m+2}$. This defines the block tridiagonal matrix $\widehat{T}_{m+3}^{(3)}$ and the associated quadrature rule $\widehat{\mathcal{G}}_{m+3}^{(3)}$. This rule satisfies

$$\widehat{\mathcal{G}}_{m+3}^{(3)} f = \widetilde{\mathcal{G}}_{m+3}^{(3)} f \quad \forall f \in \mathbb{P}_{2m+4}; \tag{3.33}$$

see below.

We conclude this section with a proof of (3.33). Let $\mathcal{G}_m f$ denote the m-block Gauss rule (3.19) associated with the operator $\mathcal{I}$ defined by (3.16). This quadrature rule can be expressed with the, generally nonsymmetric, block tridiagonal matrix T_m, which is defined by (3.18); cf. (3.19). Consider the associated $(m+\ell)$-block generalized anti-Gauss rule $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f$. It is defined by a block tridiagonal matrix

$$\widetilde{T}_{m+\ell}^{(\ell)} = \begin{bmatrix} \widetilde{\Omega}_1 & \widetilde{\Delta}_1^T & & & & \\ \widetilde{\Gamma}_1 & \widetilde{\Omega}_2 & \widetilde{\Delta}_2^T & & & \\ & \ddots & \ddots & \ddots & & \\ & & \widetilde{\Gamma}_{m-1} & \widetilde{\Omega}_m & \widetilde{\Delta}_{m+\ell-1}^T & \\ & & & \widetilde{\Gamma}_{m+\ell-1} & \widetilde{\Omega}_{m+\ell} \end{bmatrix} \in \mathbb{R}^{k(m+\ell)\times k(m+\ell)}, \tag{3.34}$$

i.e.,

$$\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f = E_1^T f(\widetilde{T}_{m+\ell}^{(\ell)}) E_1.$$

The matrix (3.34) is a block analogue of (2.11). We assume that m and ℓ are small enough so that the matrix (3.34) exists. Its leading $km \times km$ principal submatrix is the block tridiagonal matrix T_m defined by (3.18). The other blocks of $\widetilde{T}_{m+\ell}^{(\ell)}$ are determined so that the quadrature rule $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f$ satisfies

$$(\mathcal{I} - \widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}) f = -(\mathcal{I} - \mathcal{G}_m) f \quad \forall f \in \mathbb{P}_{2m+2\ell-1}. \tag{3.35}$$

This is the block analogue of equation (2.8).

When the last diagonal block entry $\widetilde{\Omega}_{m+\ell} \in \mathbb{R}^{k\times k}$ of (3.34) is replaced by another matrix $\widehat{\Omega}_{m+\ell} \in \mathbb{R}^{k\times k}$, we obtain the block tridiagonal matrix $\widehat{T}_{m+\ell}^{(\ell)}$ and the associated quadrature rule

$$\widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f = E_1^T f(\widehat{T}_{m+\ell}^{(\ell)}) E_1. \tag{3.36}$$

The following theorem collects some results for this quadrature rule.

Theorem 3.5 *The quadrature rule* (3.36) *satisfies*

$$\widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f = \mathcal{I} f \quad \forall f \in \mathbb{P}_{2m-1}, \tag{3.37}$$

$$\widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f = \widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f \quad \forall f \in \mathbb{P}_{2m+2\ell-2}. \tag{3.38}$$

Proof The nonsymmetric Lanczos process (Algorithm 4) generates the block entries of the matrix (3.18) in the order $\Omega_1, \Gamma_1, \Delta_1, \Omega_2, \Gamma_2, \Delta_2, \Omega_3, \ldots$. Each diagonal block Ω_j and each pair of off-diagonal blocks $\{\Gamma_j, \Delta_j\}$ in the matrix (3.18) increase the degree of the polynomials that are integrated exactly by the quadrature rule (3.19) by one. The matrix (3.18) has m diagonal blocks and $m-1$ pairs of off-diagonal blocks and, therefore, the rule (3.19) is exact for all polynomials in $\mathbb{P}_{2m-1}$. If we would replace the diagonal block Ω_m by an arbitrary $k \times k$ matrix, then the quadrature rule defined by the matrix so obtained would be exact for all polynomials in $\mathbb{P}_{2m-2}$.

We turn to (3.37). Since the matrix $\widehat{T}_{m+\ell}^{(\ell)}$ associated with the rule $\widehat{\mathcal{G}}_{m+\ell}^{(\ell)}$ is block tridiagonal and has the matrix (3.18) as its $km \times km$ leading principal submatrix, it integrates all polynomials exactly that the rule (3.19) integrates exactly. This shows (3.37).

Consider the rule $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}$ exact. It is defined by the block tridiagonal matrix $\widetilde{T}_{m+\ell}^{(\ell)}$. The rule $\widehat{\mathcal{G}}_{m+\ell}^{(\ell)}$ is associated with a block tridiagonal matrix $\widehat{T}_{m+\ell}^{(\ell)}$, which differs from $\widetilde{T}_{m+\ell}^{(\ell)}$ only in the last diagonal block. Thus, the matrices $\widetilde{T}_{m+\ell}^{(\ell)}$ and $\widehat{T}_{m+\ell}^{(\ell)}$ share the first $m+\ell-1$ diagonal block entries and all $m+\ell-1$ pairs of off-diagonal block entries. They therefore integrate all polynomials in $\mathbb{P}_{2m+2\ell-2}$ in the same way. This shows (3.38). □

Corollary 3.6 *Let the block quadrature rules $\mathcal{G}_m$ and $\widehat{\mathcal{G}}_{m+\ell}^{(\ell)}$ be defined by* (3.19) *and* (3.36)*, respectively. Then the average rule*

$$\widehat{\mathcal{A}}_{m+\ell}^{(\ell)} := \frac{1}{2}(\mathcal{G}_m + \widehat{\mathcal{G}}_{m+\ell}^{(\ell)}) \tag{3.39}$$

satisfies

$$\widehat{\mathcal{A}}_{m+\ell}^{(\ell)} f = \mathcal{I} f \quad \forall f \in \mathbb{P}_{2m+2\ell-2}. \tag{3.40}$$

Hence, the average rule is exact for polynomials of higher degree than the Gauss and simplified generalized anti-Gauss rules that determine it.

Proof The rule $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}$ is constructed to satisfy

$$\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f = (2\mathcal{I} - \mathcal{G}_m) f \quad \forall f \in \mathbb{P}_{2m+2\ell-1}.$$

Since the block tridiagonal matrices that are associated with the block quadrature rules $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)}$ and $\widehat{\mathcal{G}}_{m+\ell}^{(\ell)}$ only differ in the last block diagonal entry, we obtain similarly as in the proof of Theorem 3.5 that

$$\widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f = (2\mathcal{I} - \mathcal{G}_m) f \quad \forall f \in \mathbb{P}_{2m+2\ell-2}.$$

Substitution into (3.39) shows (3.40). □

We remark that it is straightforward to show an analogue of Corollary 3.6 for the average rule $\widetilde{\mathcal{A}}_{m+\ell}^{(\ell)} := \frac{1}{2}(\mathcal{G}_m + \widetilde{\mathcal{G}}_{m+\ell}^{(\ell)})$. We therefore omit the details.

4 Bracketing with Gauss and generalized anti-Gauss rules

It is the purpose of this section to provide sufficient conditions for pairs of Gauss and associated simplified generalized anti-Gauss rules to bracket $\mathcal{I}f$ component-wise. Let $\{s_j\}_{j=0}^{\infty}$ be a family of polynomials with each polynomial s_j of exact degree j, and such that the expansion

$$f(x) = \sum_{j=0}^{\infty} \eta_j s_j(x) \tag{4.1}$$

converges on the convex hull of the support of the measure $d\widetilde{w}$ defined in (3.16). We also would like the s_j be such that the terms $\eta_j s_j$ converge to zero. For instance, when the convex hull of the support of the measure is an interval, then the s_j can be chosen to be Chebyshev polynomials for this interval. Alternatively, we may choose the s_j to be Chebyshev or Faber polynomials for the field of values of A or for some ε-pseudospectrum of A; see Trefethen and Embree [38] for discussions on the latter. We obtain

$$\begin{aligned}(\mathcal{I} - \mathcal{G}_m)f &= \sum_{j=0}^{\infty} \eta_j (\mathcal{I} - \mathcal{G}_m)s_j = \sum_{j=2m}^{\infty} \eta_j (\mathcal{I} - \mathcal{G}_m)s_j = \eta_{2m}(\mathcal{I} - \mathcal{G}_m)s_{2m} + \cdots \\ &\quad + \eta_{2m+2\ell-2}(\mathcal{I} - \mathcal{G}_m)s_{2m+2\ell-2} + \sum_{j=2m+2\ell-1}^{\infty} \eta_j (\mathcal{I} - \mathcal{G}_m)s_j, \end{aligned} \tag{4.2}$$

where we have used (3.20). Similarly,

$$\begin{aligned}(\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})f &= \sum_{j=0}^{\infty} \eta_j (\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})s_j = \sum_{j=2m}^{\infty} \eta_j (\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})s_j \\ &= \eta_{2m}(\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})s_{2m} + \cdots + \eta_{2m+2\ell-2}(\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})s_{2m+2\ell-2} \\ &\quad + \sum_{j=2m+2\ell-1}^{\infty} \eta_j (\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})s_j \\ &= -\eta_{2m}(\mathcal{I} - \mathcal{G}_m)s_{2m} - \cdots - \eta_{2m+2\ell-2}(\mathcal{I} - \mathcal{G}_m)s_{2m+2\ell-2} \\ &\quad + \sum_{j=2m+2\ell-1}^{\infty} \eta_j (\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})s_j. \end{aligned} \tag{4.3}$$

Here we have applied (3.37), (3.35), and (3.38) in order.

If the terms $\eta_j(\mathcal{I}-\mathcal{G}_m)s_j$ decay in norm sufficiently rapidly with increasing j, then the right-hand sides of

$$(\mathcal{I}-\mathcal{G}_m)f \approx \eta_{2m}(\mathcal{I}-\mathcal{G}_m)s_{2m} + \cdots + \eta_{2m+2\ell-2}(\mathcal{I}-\mathcal{G}_m)s_{2m+2\ell-2},$$
$$(\mathcal{I}-\widehat{\mathcal{G}}^{(\ell)}_{m+\ell})f \approx -\eta_{2m}(\mathcal{I}-\mathcal{G}_m)s_{2m} - \cdots - \eta_{2m+2\ell-2}(\mathcal{I}-\mathcal{G}_m)s_{2m+2\ell-2}$$

are accurate approximations of the left-hand sides. Hence,

$$\mathcal{G}_m f \approx \mathcal{I}f - \eta_{2m}(\mathcal{I}-\mathcal{G}_m)s_{2m} - \cdots - \eta_{2m+2\ell-2}(\mathcal{I}-\mathcal{G}_m)s_{2m+2\ell-2},$$
$$\widehat{\mathcal{G}}^{(\ell)}_{m+\ell} f \approx \mathcal{I}f + \eta_{2m}(\mathcal{I}-\mathcal{G}_m)s_{2m} + \cdots + \eta_{2m+2\ell-2}(\mathcal{I}-\mathcal{G}_m)s_{2m+2\ell-2}.$$

This suggests that the component-wise errors of the quadrature rules $\mathcal{G}_m f$ and $\widehat{\mathcal{G}}^{(\ell)}_{m+\ell} f$ are roughly equal in magnitude and of opposite sign, and then the components of $\mathcal{G}_m f$ and $\widehat{\mathcal{G}}^{(\ell)}_{m+\ell} f$ bracket the components of $\mathcal{I}f$. The magnitudes of the entries of $\eta_j s_j$ decay quickly to zero when j increases if f is analytic in a large simply connected region in the complex plane that contains the support of the measure $d\widetilde{w}$ and has its boundary far away from the support. The following theorem provides sufficient conditions for $\mathcal{G}_m f$ and $\widehat{\mathcal{G}}^{(\ell)}_{m+\ell} f$ to bracket $\mathcal{I}f$.

Theorem 4.1 *Consider the expansion* (4.1) *in terms of the polynomials* s_j *and assume that for some* $1 \le q, r \le k$,

$$\left|\left[\sum_{j=2m}^{2m+2\ell-2} \eta_j(\mathcal{I}-\mathcal{G}_m)s_j\right]_{q,r}\right| \ge \max\left\{\left|\left[\sum_{j=2m+2\ell-1}^{\infty} \eta_j(\mathcal{I}-\mathcal{G}_m)s_j\right]_{q,r}\right|, \left|\left[\sum_{j=2m+2\ell-1}^{\infty} \eta_j(\mathcal{I}-\widehat{\mathcal{G}}^{(\ell)}_{m+\ell})s_j\right]_{q,r}\right|\right\}, \tag{4.4}$$

where $[M]_{q,r}$ *denotes the* (q,r)*-entry of the matrix* $M \in \mathbb{R}^{k\times k}$. *Then* $[\mathcal{G}_m f]_{q,r}$ *and* $[\widehat{\mathcal{G}}_{m+\ell} f]_{q,r}$ *bracket* $[\mathcal{I}f]_{q,r}$.

Proof It follows from the expansions (4.2) and (4.3) that

$$\mathcal{G}_m f = \mathcal{I}f - \sum_{j=2m}^{2m+2\ell-2} \eta_j(\mathcal{I}-\mathcal{G}_m)s_j - \sum_{j=2m+2\ell-1}^{\infty} \eta_j(\mathcal{I}-\mathcal{G}_m)s_j, \tag{4.5}$$

$$\widehat{\mathcal{G}}^{(\ell)}_{m+\ell} f = \mathcal{I}f + \sum_{j=2m}^{2m+2\ell-2} \eta_j(\mathcal{I}-\mathcal{G}_m)s_j - \sum_{j=2m+2\ell-1}^{\infty} \eta_j(\mathcal{I}-\widehat{\mathcal{G}}^{(\ell)}_{m+\ell})s_j. \tag{4.6}$$

This shows (4.4). □

It is difficult to assess whether the bound (4.4) holds for a given function f, matrix A, and pair of quadrature rules $\{\mathcal{G}_m, \widehat{\mathcal{G}}^{(\ell)}_{m+\ell}\}$. Nevertheless, Theorem 4.1 indicates that

if the terms in the expansions (4.2) and (4.3) converge to zero quickly, then the pair of quadrature rule values $\{\mathcal{G}_m f, \widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f\}$ is likely to bracket $\mathcal{I} f$ component-wise. This is in agreement with computational experience. For instance, when $f(t) = \exp(t)$, the quadrature rules $\mathcal{G}_m f$ and $\widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f$ bracket $\mathcal{I} f$ component-wise for many matrices A. This is illustrated in Sect. 5. Our interest in generalized and simplified generalized anti-Gauss rules (with $\ell > 1$) stems from the fact that the pair of rules $\{\mathcal{G}_m f, \widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f\}$ may bracket $\mathcal{I} f$ when $\ell > 1$ also when this is not the case for $\ell = 1$; see Sect. 5. When several of the generalized or simplified generalized anti-Gauss rules $\widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f$ for, say, $\ell = 1, 2, 3$, all are larger or all are smaller than $\mathcal{G}_m f$, this suggests that the pairs $\{\mathcal{G}_m f, \widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f\}$, $\ell = 1, 2, 3$, indeed may bracket $\mathcal{I} f$.

Corollary 4.2 *Let the integrand f have the expansion (4.1). Then the average rule (3.39) satisfies*

$$\widehat{\mathcal{A}}_{m+\ell}^{(\ell)} f = \mathcal{I} f - \sum_{j=2m+2\ell-1}^{\infty} \eta_j (\mathcal{I} - \widehat{\mathcal{A}}_{m+\ell}^{(\ell)}) s_j.$$

In particular, if $f \in \mathbb{P}_{2m+2\ell-2}$, then $\eta_j = 0$ for $j \geq 2m + 2\ell - 1$, and therefore $\widehat{A}_{m+\ell}^{(\ell)} f = \mathcal{I} f$, in agreement with (3.40).

Proof The result follows by substituting the expansions (4.5) and (4.6) into (3.39). □

5 Computed examples

We illustrate the performance of real-valued generalized anti-Gauss rules, simplified real-valued generalized anti-Gauss rules, and their block versions. 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. We first describe a few applications of these quadrature rules to the approximation of real-valued functionals, and subsequently discuss applications to matrix-valued functions. Examples from network analysis also are presented.

Example 5.1 We would like to compute approximations of the functional

$$\mathcal{I} f = F(A) := u^T (I + A^2)^{-1} u,$$

with a symmetric Toeplitz matrix $A \in \mathbb{R}^{200\times 200}$ with first row $[1, 1/2, 1/3, \ldots, 1/200]$. Thus, $f(t) = (1 + t^2)^{-1}$ in (1.1). The vector u has normally distributed entries with zero mean and is normalized to be of unit norm. The desired value is $F(A) \approx 1.36 \cdot 10^{-2}$. Table 1 shows the pairs $\{\mathcal{G}_m f, \widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f\}$ as well as the pairs $\{\mathcal{G}_m f, \widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f\}$ to bracket $\mathcal{I} f$ for $\ell \in \{1, 2, 3\}$ and $m \in \{3, 6\}$. The simplified generalized anti-Gauss rules $\widehat{\mathcal{G}}_{m+\ell}^{(\ell)} f$ are seen to give slightly more accurate approximations of $\mathcal{I} f$ than the (standard) generalized anti-Gauss rules $\widetilde{\mathcal{G}}_{m+\ell}^{(\ell)} f$. We remark that pairs of Gauss and Gauss–Radau quadrature rules are not guaranteed to furnish upper and lower bounds

Table 1 Example 5.1: $F(A) = u^T (I + A^2)^{-1} u$, A a symmetric Toeplitz matrix

m	ℓ	$(\mathcal{I} - \mathcal{G}_m) f$	$(\mathcal{I} - \widetilde{\mathcal{G}}^{(\ell)}_{m+\ell}) f$	$(\mathcal{I} - \widetilde{\mathcal{A}}^{(\ell)}_{m+\ell}) f$	$(\mathcal{I} - \widehat{\mathcal{G}}^{(\ell)}_{m+\ell}) f$	$(\mathcal{I} - \widehat{\mathcal{A}}^{(\ell)}_{m+\ell}) f$
3	1	$4.48 \cdot 10^{-5}$	$-4.82 \cdot 10^{-5}$	$-1.69 \cdot 10^{-6}$	$-4.16 \cdot 10^{-5}$	$1.56 \cdot 10^{-6}$
3	2		$-4.76 \cdot 10^{-5}$	$-1.39 \cdot 10^{-6}$	$-4.07 \cdot 10^{-5}$	$2.04 \cdot 10^{-6}$
3	3		$-4.73 \cdot 10^{-5}$	$-1.29 \cdot 10^{-6}$	$-4.06 \cdot 10^{-5}$	$2.07 \cdot 10^{-6}$
6	1	$-3.64 \cdot 10^{-7}$	$3.66 \cdot 10^{-7}$	$1.09 \cdot 10^{-9}$	$3.42 \cdot 10^{-7}$	$-1.10 \cdot 10^{-8}$
6	2		$3.64 \cdot 10^{-7}$	$-6.47 \cdot 10^{-11}$	$3.40 \cdot 10^{-7}$	$-1.22 \cdot 10^{-8}$
6	3		$3.64 \cdot 10^{-7}$	$-6.18 \cdot 10^{-11}$	$3.40 \cdot 10^{-7}$	$-1.22 \cdot 10^{-8}$

for $\mathcal{I} f$, because higher derivatives of the integrand f change sign on the convex hull of the spectrum of A. □

Example 5.2 We determine approximations of the functional

$$\mathcal{I} f = F(A) := u^T (I + A^2)^{-1} v,$$

where A is a 200×200 real nonsymmetric Toeplitz matrix with first row and column $[1, 1/2, 1/3, \ldots, 1/200]$ and $[1, 1, \ldots, 1]^T$, respectively, and $f(t) = (1 + t^2)^{-1}$ in (1.2). The vectors u and v have normally distributed random entries with zero mean; they are scaled so that $u^T v = 1$. The exact value is $F(A) \approx 5.36 \cdot 10^{-3}$. Since the matrix A is nonsymmetric, the technique described by Golub and Meurant [22] of evaluating pairs of Gauss and Gauss–Radau quadrature rules is not guaranteed to furnish upper and lower bounds for $\mathcal{I} f$. Table 2 shows the errors in approximations determined by Gauss, generalized anti-Gauss, simplified generalized anti-Gauss, and average quadrature rules. Pairs of Gauss rules $\mathcal{G}_m f$ and generalized anti-Gauss rules $\widetilde{\mathcal{G}}^{(\ell)}_{m+\ell} f$ and pairs of Gauss rules and simplified generalized anti-Gauss rules $\widehat{\mathcal{G}}^{(\ell)}_{m+\ell} f$ bracket $F(A)$ for $\ell \in \{1, 2, 3\}$ and $m = 4$, as well as for $\ell \in \{1, 3\}$ and $m = 5$, but not for $\ell = 2$ and $m = 5$. This illustrates that it may be beneficial to compute quadrature rules $\widetilde{\mathcal{G}}^{(\ell)}_{m+\ell} f$ or $\widehat{\mathcal{G}}^{(\ell)}_{m+\ell} f$ for several values of m and ℓ. □

Table 2 Example 5.2: $F(A) = u^T (I + A^2)^{-1} v$, A a nonsymmetric Toeplitz matrix

m	ℓ	$(\mathcal{I} - \mathcal{G}_m) f$	$(\mathcal{I} - \widetilde{\mathcal{G}}^{(\ell)}_{m+\ell}) f$	$(\mathcal{I} - \widetilde{\mathcal{A}}^{(\ell)}_{m+\ell}) f$	$(\mathcal{I} - \widehat{\mathcal{G}}^{(\ell)}_{m+\ell}) f$	$(\mathcal{I} - \widehat{\mathcal{A}}^{(\ell)}_{m+\ell}) f$
4	1	$-3.71 \cdot 10^{-3}$	$2.64 \cdot 10^{-3}$	$-5.41 \cdot 10^{-4}$	$4.41 \cdot 10^{-3}$	$3.47 \cdot 10^{-4}$
4	2		$2.43 \cdot 10^{-3}$	$-6.42 \cdot 10^{-4}$	$8.46 \cdot 10^{-3}$	$2.37 \cdot 10^{-3}$
4	3		$2.64 \cdot 10^{-3}$	$-5.41 \cdot 10^{-4}$	$4.41 \cdot 10^{-3}$	$3.47 \cdot 10^{-4}$
5	1	$1.46 \cdot 10^{-3}$	$-4.07 \cdot 10^{-3}$	$-1.31 \cdot 10^{-3}$	$-2.68 \cdot 10^{-3}$	$-6.09 \cdot 10^{-4}$
5	2		$1.54 \cdot 10^{-3}$	$1.50 \cdot 10^{-3}$	$1.44 \cdot 10^{-3}$	$1.45 \cdot 10^{-3}$
5	3		$-4.07 \cdot 10^{-3}$	$-1.31 \cdot 10^{-3}$	$-2.68 \cdot 10^{-3}$	$-6.09 \cdot 10^{-4}$

Example 5.3 We would like to approximate

$$\mathcal{I}f = F(A) := u^T \exp(A)v,$$

where A is the nonsymmetric Toeplitz matrix of Example 5.2. Thus, $f(t) = \exp(t)$. The vectors u and v also are determined as in Example 5.2. Table 3 shows the pairs $\{\mathcal{G}_{14}f, \widetilde{\mathcal{G}}^{(\ell)}_{14+\ell}f\}$ to bracket $F(A)$ for $\ell = 3$, but not for $\ell \in \{1, 2\}$. While the absolute quadrature errors are large, the relative errors are quite small since $\mathcal{I}f \approx 4.04 \cdot 10^{11}$. □

Example 5.4 This example discusses an application of generalized block anti-Gauss rules and simplified generalized block anti-Gauss rules to the evaluation of matrix functions of the form

$$\mathcal{I}f = F(A) := U^T \exp(A)U,$$

where $A \in \mathbb{R}^{200\times 200}$ is the symmetric Toeplitz matrix of Example 5.1 and $U \in \mathbb{R}^{200\times 2}$ is a block vector with random orthonormal columns. Thus, $f(t) = \exp(t)$. Table 4 displays the block Gauss rule $\mathcal{G}_5 f$ and the generalized block anti-Gauss rules $\widetilde{\mathcal{G}}^{(\ell)}_{5+\ell}f$ for $\ell = 1, 2, 3$ to give quadrature errors of component-wise opposite sign and of about the same magnitude. The desired value is

$$F(A) \approx \begin{bmatrix} 6.2667 & -0.2506 \\ -0.2506 & 0.0208 \end{bmatrix} \cdot 10^3.$$

This value is computed with the MATLAB function `expm`. Thus, the relative errors achieved with the quadrature rules of Tables 4 and 5 are quite small.

Table 3 Example 5.3: $F(A) = u^T \exp(A)v$, A a nonsymmetric Toeplitz matrix

m	ℓ	$(\mathcal{I} - \mathcal{G}_m)f$	$(\mathcal{I} - \widetilde{\mathcal{G}}^{(\ell)}_{m+\ell})f$	$(\mathcal{I} - \widetilde{\mathcal{A}}^{(\ell)}_{m+\ell})f$
14	1	$-6.51 \cdot 10^1$	$-8.42 \cdot 10^1$	$-7.47 \cdot 10^1$
14	2		$-9.13 \cdot 10^1$	$-7.82 \cdot 10^1$
14	3		$1.57 \cdot 10^3$	$7.53 \cdot 10^2$

Table 4 Example 5.4: $F(A) = U^T \exp(A)U$, A symmetric Toeplitz matrix

m	ℓ	$(\mathcal{I} - \mathcal{G}_m)f$	$(\mathcal{I} - \widetilde{\mathcal{G}}^{(\ell)}_{m+\ell})f$	$(\mathcal{I} - \widetilde{\mathcal{A}}^{(\ell)}_{m+\ell})f$
5	1	$\begin{bmatrix} -8.98 & -1.00 \\ -1.00 & -1.49 \end{bmatrix} \cdot 10^{-4}$	$\begin{bmatrix} 8.96 & 1.00 \\ 1.00 & 1.49 \end{bmatrix} \cdot 10^{-4}$	$\begin{bmatrix} -9.84 & 1.95 \\ 1.95 & -0.26 \end{bmatrix} \cdot 10^{-7}$
5	2		$\begin{bmatrix} 9.19 & 1.03 \\ 1.03 & 1.54 \end{bmatrix} \cdot 10^{-4}$	$\begin{bmatrix} 1.04 & 0.16 \\ 0.16 & 0.25 \end{bmatrix} \cdot 10^{-5}$
5	3		$\begin{bmatrix} 9.19 & 1.03 \\ 1.03 & 1.54 \end{bmatrix} \cdot 10^{-4}$	$\begin{bmatrix} 1.06 & 1.60 \\ 0.16 & 0.24 \end{bmatrix} \cdot 10^{-5}$

Table 5 Example 5.4: $F(A) = U^T \exp(A)U$, A symmetric Toeplitz matrix

m	ℓ	$(\mathcal{I} - \mathcal{G}_m)f$	$(\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})f$	$(\mathcal{I} - \widehat{\mathcal{A}}_{m+\ell}^{(\ell)})f$
5	1	$\begin{bmatrix} -8.98 & -1.00 \\ -1.00 & -1.49 \end{bmatrix} \cdot 10^{-4}$	$\begin{bmatrix} 8.96 & 1.00 \\ 1.00 & 1.49 \end{bmatrix} \cdot 10^{-4}$	$\begin{bmatrix} -9.84 & 1.95 \\ 1.95 & -0.26 \end{bmatrix} \cdot 10^{-7}$
5	2		$\begin{bmatrix} 9.23 & 1.04 \\ 1.04 & 1.55 \end{bmatrix} \cdot 10^{-4}$	$\begin{bmatrix} 1.23 & 0.20 \\ 0.20 & 0.30 \end{bmatrix} \cdot 10^{-5}$
5	3		$\begin{bmatrix} 9.25 & 1.04 \\ 1.04 & 1.56 \end{bmatrix} \cdot 10^{-4}$	$\begin{bmatrix} 1.34 & 0.20 \\ 0.20 & 0.32 \end{bmatrix} \cdot 10^{-5}$

Table 5 shows the component-wise difference between the exact value and approximations determined by simplified generalized block anti-Gauss and associated average rules. They are defined by setting $\widehat{\Omega}_{m+\ell} = \Omega_{m+\ell-1}$. The simplified generalized block anti-Gauss and the block Gauss rule also bracket $F(A)$ component-wise. The average rules can be seen to give the best approximations of $F(A)$ in both Tables 4 and 5 . □

Our last examples illustrate the application of the quadrature rules of this paper to quantities of interest in network analysis. A network is identified by a graph $G = \{\mathcal{V}, \mathcal{E}\}$, which is defined by a set of vertices (also referred to as nodes) $\mathcal{V}$ and a set of edges $\mathcal{E}$. Let G be an unweighted graph with m nodes, and assume that G has no self-loops or multiple edges. We consider both undirected graphs, in which travel can occur in both directions along each edge, and directed graphs, in which some or all edges are "one way streets". Networks arise in many scientific and industrial applications, including genetics, epidemiology, energy distribution, and telecommunication; see, e.g., [15,32]. The adjacency matrix $A = [A_{ij}] \in \mathbb{R}^{n\times n}$ associated with a graph G with n nodes has the entry $A_{ij} = 1$ if there is an edge from node i to node j, and $A_{ij} = 0$ otherwise. Thus, A is symmetric if and only if G is undirected. The importance of node i can be assessed by evaluating the element (i, i) of a matrix function f, such as the exponential function or a resolvent, of the adjacency matrix. Similarly, the ease of communication between node i and node j can be determined by computing the entry (i, j) of the exponential function or a resolvent of the adjacency matrix; see, e.g., [16] for a justification in terms of walks in a graph. The computational task is to evaluate expressions of the form (1.1) or (1.2); see [16,18]. When the graph has many nodes and, therefore, the adjacency matrix A is large, it is expensive to evaluate functions of A. The use of quadrature rules to estimate desired entries of functions of A is cheaper and therefore attractive.

The total communicability $F(A) = e^T f(A)e$, where $e = [1, 1, \ldots, 1]^T$, measures the ease of communication in a network; see [7,32]. Here $f(A)$ is the exponential function or a resolvent of the adjacency matrix A. The total communicability can conveniently be approximated by quadrature rules. The remaining tables of this section show results for anti-Gauss, generalized anti-Gauss, and simplified anti-Gauss rules with $\ell \in \{1, 2\}$ applied to evaluating quantities of interest when determining properties of a network or of nodes in a network.

Example 5.5 We consider the network Yeast, which is represented by an undirected graph with 2114 vertices and 4480 edges. It describes the protein interaction of yeast.

Table 6 Example 5.5: $\mathcal{I}f = F(A) = e^T \exp(A)e$, A the symmetric adjacency matrix for the Yeast network, $e = [1, 1, \ldots, 1]^T$

m	ℓ	$(\mathcal{I} - \mathcal{G}_m)f$	$(\mathcal{I} - \widetilde{\mathcal{G}}^{(\ell)}_{m+\ell})f$	$(\mathcal{I} - \widetilde{\mathcal{A}}^{(\ell)}_{m+\ell})f$	$(\mathcal{I} - \widehat{\mathcal{G}}^{(\ell)}_{m+\ell})f$	$(\mathcal{I} - \widehat{\mathcal{A}}^{(\ell)}_{m+\ell})f$
8	1	$2.03 \cdot 10^{-4}$	$-2.03 \cdot 10^{-4}$	$-2.79 \cdot 10^{-7}$	$-2.03 \cdot 10^{-4}$	$-2.79 \cdot 10^{-7}$
8	2		$-1.95 \cdot 10^{-4}$	$4.02 \cdot 10^{-6}$	$-2.03 \cdot 10^{-4}$	$-1.95 \cdot 10^{-7}$
10	1	$1.97 \cdot 10^{-7}$	$-1.97 \cdot 10^{-7}$	$8.77 \cdot 10^{-10}$	$-1.95 \cdot 10^{-7}$	$8.77 \cdot 10^{-10}$
10	2		$-1.97 \cdot 10^{-7}$	$-2.72 \cdot 10^{-11}$	$-1.97 \cdot 10^{-7}$	$-1.80 \cdot 10^{-10}$

Table 7 Example 5.5: $\mathcal{I}f = F(A) = e^T \exp(A)e$, A a nonsymmetric adjacency matrix for the modified Yeast network, $e = [1, 1, \ldots, 1]^T$

m	ℓ	$(\mathcal{I} - \mathcal{G}_m)f$	$(\mathcal{I} - \widetilde{\mathcal{G}}^{(\ell)}_{m+\ell})f$	$(\mathcal{I} - \widetilde{\mathcal{A}}^{(\ell)}_{m+\ell})f$	$(\mathcal{I} - \widehat{\mathcal{G}}^{(\ell)}_{m+\ell})f$	$(\mathcal{I} - \widehat{\mathcal{A}}^{(\ell)}_{m+\ell})f$
7	1	$-4.12 \cdot 10^{-3}$	$4.18 \cdot 10^{-3}$	$2.58 \cdot 10^{-5}$	$3.75 \cdot 10^{-3}$	$-1.84 \cdot 10^{-4}$
7	2		$4.28 \cdot 10^{-3}$	$7.67 \cdot 10^{-5}$	$3.87 \cdot 10^{-3}$	$-1.26 \cdot 10^{-4}$
8	1	$-2.02 \cdot 10^{-4}$	$2.03 \cdot 10^{-4}$	$2.98 \cdot 10^{-7}$	$2.07 \cdot 10^{-4}$	$2.25 \cdot 10^{-6}$
8	2		$2.03 \cdot 10^{-4}$	$4.08 \cdot 10^{-7}$	$2.07 \cdot 10^{-4}$	$2.27 \cdot 10^{-6}$
10	1	$-1.96 \cdot 10^{-7}$	$1.95 \cdot 10^{-7}$	$-8.43 \cdot 10^{-10}$	$1.98 \cdot 10^{-7}$	$1.02 \cdot 10^{-9}$
10	2		$1.99 \cdot 10^{-7}$	$1.37 \cdot 10^{-9}$	$2.04 \cdot 10^{-7}$	$3.69 \cdot 10^{-9}$

Each edge represents an interaction between two proteins [28,37]. The data set is available at [5]. We would like to determine the total communicability $F(A) = e^T \exp(A)e$, where $e = [1, 1, \ldots, 1]^T$, of the network; see [7,32]. We use quadrature rules of Sect. 2 to provide estimates. The exact value is approximately 221. Table 6 shows the errors in computed approximations of the total communicability determined by Gauss, generalized anti-Gauss, simplified generalized anti-Gauss, and average quadrature rules. Quadrature rules with only a few nodes are seen to yield approximations with higher accuracy than what typically is required in applications. Pairs of Gauss and generalized anti-Gauss, as well as pairs of Gauss and simplified generalized anti-Gauss rules are seen to bracket the exact value.

The adjacency matrix for the Yeast network is symmetric and all derivatives of the integrand are positive. Therefore, the technique of Golub and Meurant [22], based on evaluating pairs of a Gauss rule and a suitable Gauss–Radau rule, can be applied to compute upper and lower bounds for $F(A)$. This technique requires that an upper bound for the largest eigenvalue of A be available in order to determine a suitable Gauss–Radau rule. The application of pairs of Gauss and generalized anti-Gauss rules or simplified generalized anti-Gauss rules does not require eigenvalue bounds. We therefore believe the application of these rules may be of interest also when the technique by Golub and Meurant [22] can be used.

It is easy to modify the Yeast network to obtain a directed network with a nonsymmetric adjacency matrix. Then the technique by Golub and Meurant is not guaranteed to determine bounds for the total communicability. We replace the superdiagonal entries in rows 2 to 4 by zero. This implies that we replace "two-way streets" to the nodes 2 to 4, by "one-way streets". Table 7 shows that pairs of Gauss and generalized

Table 8 Example 5.6: subgraph centrality for the vertices 100, 224, and 1000 of the modified Yeast network

i	$[f(A)]_{ii}$	$[f(A)]_{ii} - \mathcal{G}_m f$
100	$3.919 \cdot 10^{0}$	$-1.54 \cdot 10^{-6}$
224	$9.124 \cdot 10^{2}$	$-6.71 \cdot 10^{-6}$
1000	$1.031 \cdot 10^{1}$	$-1.58 \cdot 10^{-5}$

Table 9 Example 5.6: subgraph centrality for the vertices 100, 224, and 1000 of the modified Yeast network

i	ℓ	$[f(A)]_{ii} - \widetilde{\mathcal{G}}^{(\ell)}_{m+\ell} f$	$[f(A)]_{ii} - \widetilde{\mathcal{A}}^{(\ell)}_{m+\ell} f$	$[f(A)]_{ii} - \widehat{\mathcal{G}}^{(\ell)}_{m+\ell} f$	$[f(A)]_{ii} - \widehat{\mathcal{A}}^{(\ell)}_{m+\ell} f$
100	1	$1.61 \cdot 10^{-6}$	$3.26 \cdot 10^{-8}$	$1.14 \cdot 10^{-6}$	$-2.02 \cdot 10^{-7}$
100	2	$2.03 \cdot 10^{-6}$	$2.40 \cdot 10^{-7}$	$1.48 \cdot 10^{-6}$	$-3.22 \cdot 10^{-8}$
224	1	$6.72 \cdot 10^{-6}$	$4.27 \cdot 10^{-9}$	$5.87 \cdot 10^{-6}$	$-4.20 \cdot 10^{-7}$
224	2	$6.72 \cdot 10^{-6}$	$4.62 \cdot 10^{-9}$	$5.87 \cdot 10^{-6}$	$-4.19 \cdot 10^{-7}$
1000	1	$1.60 \cdot 10^{-5}$	$1.03 \cdot 10^{-7}$	$1.82 \cdot 10^{-5}$	$1.23 \cdot 10^{-6}$
1000	2	$1.63 \cdot 10^{-5}$	$2.96 \cdot 10^{-7}$	$1.87 \cdot 10^{-5}$	$1.45 \cdot 10^{-6}$

anti-Gauss or simplified generalized anti-Gauss rules provide upper and lower bounds for the total communicability for this directed network. □

Example 5.6 Let $f(t) = \exp(t)$. The subgraph centrality of a the node i is defined as the ith diagonal entry, $[f(A)]_{ii}$, of $f(A)$. A (relatively) large value indicates that node i is important; see [16] for a discussion. We compute approximations of the subgraph centralities $[f(A)]_{ii}$ for $i \in \{100, 224, 1000\}$ for the modified Yeast network described in Example 5.5. Thus, the adjacency matrix is nonsymmetric. Table 8 displays the exact values and the errors obtained by using the Gauss rule $\mathcal{G}_8 f$. The application of this rule to approximate $[f(A)]_{ii}$ requires 8 steps of the nonsymmetric Lanczos process applied to A with the axis vector $e_i = [0, \ldots, 0, 1, 0, \ldots, 0]^T$ as initial vectors. Table 9 shows the errors in approximations determined by corresponding generalized anti-Gauss and simplified generalized anti-Gauss rules. The table also shows the errors for the associated average rules. Table 8 shows node 224 to be more important than the nodes 100 and 1000. A preprocessing technique for determining which ones of all the vertices in a network may have the largest subgraph centrality is described in [17]. □

Example 5.7 Consider the metabolic network Celegans with 306 nodes and 2345 edges of the nematode (roundworm) caenorhabditis elegans [14]. The data set is available at [1]. The network is directed, i.e., the adjacency matrix is nonsymmetric. Table 10 shows the errors of computed approximations of the total communicability determined by Gauss, generalized anti-Gauss, simplified generalized anti-Gauss, and average quadrature rules. The exact value is about $\mathcal{I} f \approx 1.09 \cdot 10^4$. This example illustrates that accuracy that typically is sufficient in applications can be achieved already

Table 10 Example 5.7: $\mathcal{I}f = F(A) = e^T \exp(A)e$, A a nonsymmetric adjacency matrix for the Celegans network, $e = [1, 1, \ldots, 1]^T$

m	ℓ	$(\mathcal{I} - \mathcal{G}_m)f$	$(\mathcal{I} - \widetilde{\mathcal{G}}_{m+\ell}^{(\ell)})f$	$(\mathcal{I} - \widetilde{\mathcal{A}}_{m+\ell}^{(\ell)})f$	$(\mathcal{I} - \widehat{\mathcal{G}}_{m+\ell}^{(\ell)})f$	$(\mathcal{I} - \widehat{\mathcal{A}}_{m+\ell}^{(\ell)})f$
3	1	$8.49 \cdot 10^2$	$-9.83 \cdot 10^2$	$-6.67 \cdot 10^1$	$-2.39 \cdot 10^3$	$-7.74 \cdot 10^2$
3	2		$-4.35 \cdot 10^3$	$-1.75 \cdot 10^3$	$-7.65 \cdot 10^3$	$-3.40 \cdot 10^3$
4	1	$3.59 \cdot 10^0$	$-4.89 \cdot 10^0$	$-6.49 \cdot 10^{-1}$	$-4.24 \cdot 10^0$	$-3.27 \cdot 10^{-1}$
4	2		$-1.15 \cdot 10^2$	$-5.60 \cdot 10^1$	$-8.85 \cdot 10^1$	$-4.24 \cdot 10^1$

with quadrature rules with very few nodes. The quadrature rules are seen to bracket the exact value. □

6 Conclusion

Laurie [30] introduced a class of quadrature rules, that he referred to as anti-Gauss rules, for the estimation of the error in associated Gaussian quadrature rules. We derive new generalized anti-Gauss quadrature rules, and discuss their properties and computation. Also a family of simplified generalized anti-Gauss quadrature rules is introduced. The latter rules are easier to evaluate and may exist also when generalized anti-Gauss quadrature rules do not. Computed examples, which include applications to network analysis, illustrate the performance of the new quadrature rules.

Acknowledgements We would like to thank Miroslav Pranić and the referees for comments that lead to improvements of the presentation. This research is supported in part by NSF Grants DMS-1720259 and DMS-1729509.

References

1. Alex Arenas's data sets. http://deim.urv.cat/~aarenas/data/welcome.htm. Accessed Jan 2018
2. Alqahtani, H., Reichel, L.: Simplified anti-Gauss quadrature rules with applications in linear algebra. Numer. Algorithms **77**, 577–602 (2018)
3. Baglama, J., Fenu, C., Reichel, L., Rodriguez, G.: Analysis of directed networks via partial singular value decomposition and Gauss quadrature. Linear Algebra Appl. **456**, 93–121 (2014)
4. Bai, Z., Day, D., Ye, Q.: ABLE: an adaptive block Lanczos method for non-Hermitian eigenvalue problems. SIAM J. Matrix Anal. Appl. **20**, 1060–1082 (1999)
5. Batagelj, V., Mrvar, A.: Pajek data sets (2006). http://vlado.fmf.uni-lj.si/pub/networks/data/. Accessed Jan 2018
6. Benzi, M., Boito, P.: Quadrature rule-based bounds for functions of adjacency matrices. Linear Algebra Appl. **433**, 637–652 (2010)
7. Benzi, M., Klymko, C.: Total communicability as a centrality measure. J. Complex Netw. **1**, 1–26 (2013)
8. Brezinski, C., Redivo Zaglia, M., Sadok, H.: Avoiding breakdown and near-breakdown in Lanczos type algorithms. Numer. Algorithms **1**, 261–284 (1991)
9. Brezinski, C., Redivo Zaglia, M., Sadok, H.: Addendum to "Avoiding breakdown and near-breakdown in Lanczos type algorithms". Numer. Algorithms **2**, 133–136 (1992)
10. Calvetti, D., Reichel, L.: Tikhonov regularization of large linear problems. BIT **43**, 263–283 (2003)

11. Calvetti, D., Reichel, L., Sgallari, F.: Application of anti-Gauss quadrature rules in linear algebra. In: Gautschi, W., Golub, G.H., Opfer, G. (eds.) Applications and Computation of Orthogonal Polynomials, pp. 41–56. Basel, Birkhäuser (1999)
12. Chihara, T.S.: An Introdution to Orthogonal Polynomials. Gordon and Breach, New York (1978)
13. Clenshaw, C.W., Curtis, A.R.: A method for numerical integration on an automatic computer. Numer. Math. **2**, 197–205 (1960)
14. Duch, J., Arenas, A.: Community identification using extremal optimization. Phys. Rev. E **72**, 027104 (2005)
15. Estrada, E.: The Structure of Complex Networks. Oxford University Press, Oxford (2012)
16. Estrada, E., Higham, D.J.: Network properties revealed through matrix functions. SIAM Rev. **52**, 696–714 (2010)
17. Fenu, C., Martin, D., Reichel, L., Rodriguez, G.: Network analysis via partial spectral factorization and Gauss quadrature. SIAM J. Sci. Comput. **35**, A2046–A2068 (2013)
18. Fenu, C., Martin, D., Reichel, L., Rodriguez, G.: Block Gauss and anti-Gauss quadrature with application to networks. SIAM J. Matrix Anal. Appl. **34**, 1655–1684 (2013)
19. Freund, R.W., Gutknecht, M.H., Nachtigal, N.M.: An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices. SIAM J. Sci. Comput. **14**, 137–158 (1993)
20. Gautschi, W.: The interplay between classical analysis and (numerical) linear algebra—a tribute to Gene H. Golub. Electron. Trans. Numer. Anal. **13**, 119–147 (2002)
21. Gautschi, W.: Orthogonal Polynomials: Approximation and Computation. Oxford University Press, Oxford (2004)
22. Golub, G.H., Meurant, G.: Matrices, Moments and Quadrature with Applications. Princeton University Press, Princeton (2010)
23. Gragg, W.B.: Matrix interpretations and applications of the continued fraction algorithm. Rocky Mt. J. Math. **4**, 213–225 (1974)
24. Gutknecht, M.H.: A completed theory of the unsymmetric Lanczos process and related algorithms, part I. SIAM J. Matrix Anal. Appl. **13**, 594–639 (1992)
25. Gutknecht, M.H.: A completed theory of the unsymmetric Lanczos process and related algorithms, part II. SIAM J. Matrix Anal. Appl. **15**, 15–58 (1994)
26. Gutknecht, M.H., Strakoš, Z.: Accuracy of two three-term and three two-term recurrences for Krylov space solvers. SIAM J. Matrix Anal. **22**, 213–229 (2000)
27. Higham, N.J.: Functions of Matrices: Theory and Computation. SIAM, Philadelphia (2008)
28. Jeong, H., Mason, S., Barabási, A.-L., Oltvai, Z.N.: Lethality and centrality of protein networks. Nature **411**, 41–42 (2001)
29. Lambers, J.V.: Enhancement of Krylov subspace spectral methods by block Lanczos iteration. Electron. Trans. Numer. Anal. **31**, 86–109 (2008)
30. Laurie, D.P.: Anti-Gauss quadrature formulas. Math. Comput. **65**, 739–747 (1996)
31. Laurie, D.P.: Computation of Gauss-type quadrature formulas. J. Comput. Appl. Math. **127**, 201–217 (2001)
32. Newman, M.E.J.: Networks: An Introduction. Oxford University Press, Oxford (2010)
33. Notaris, S.: Gauss–Kronrod quadrature formulae—a survey of fifty years of research. Electron. Trans. Numer. Anal. **45**, 371–404 (2016)
34. Pozza, S., Pranić, M.S., Strakoš, Z.: Gauss quadrature for quasi-definite linear functionals. IMA J. Numer. Anal. **37**, 1468–1495 (2017)
35. Pozza, S., Pranić, M.S., Strakoš, Z.: The Lanczos algorithm and complex Gauss quadrature. Electron. Trans. Numer. Anal. **50**, 1–19 (2018)
36. Pranić, M., Reichel, L.: Generalized anti-Gauss quadrature rules. J. Comput. Appl. Math. **284**, 235–243 (2015)
37. Sun, S., Ling, L., Zhang, N., Li, G., Chen, R.: Topological structure analysis of the protein–protein interaction network in budding yeast. Nucleic Acids Res. **31**, 2443–2450 (2003)
38. Trefethen, L.N., Embree, M.: Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators. Princeton University Press, Princeton (2005)