New error estimates for the conjugate gradient method

Hanan Almutairi

Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

Gérard Meurant

Paris, France

Lothar Reichel

Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

Miodrag M. Spalević

Department of Mathematics, University of Beograd, Faculty of Mechanical Engineering, Kraljice Marije 16, 11120 Belgrade 35, Serbia

Abstract

The conjugate gradient method is the default iterative method for the solution of linear systems of equations with a large symmetric positive definite matrix A. The development of techniques for estimating the norm of the error in iterates computed by this method has received considerable attention. Available methods for bracketing the A-norm of the error evaluate pairs of Gauss and Gauss-Radau quadrature rules to determine lower and upper bounds. The latter rules require a user to allocate a node (the Radau node) between the origin and the smallest eigenvalue of the system matrix. The determination of such a node generally demands further computations to estimate the location of the smallest eigenvalue; see, e.g., [14, 15, 16, 20, 21]. An approach that avoids the need for a lower bound of the smallest eigenvalue is to replace the Gauss-Radau quadrature rule by an anti-Gauss rule such as in [4]. However, this approach may sometimes yield inaccurate error norm estimates. This paper proposes the use of pairs of Gauss and associated optimal averaged Gauss quadrature rules to estimate the A-norm of the error in iterates determined by the conjugate gradient method.

Email addresses: halmutai@kent.edu (Hanan Almutairi), gerard.meurant@gmail.com (Gérard Meurant), reichel@math.kent.edu (Lothar Reichel), mspalevic@mas.bg.ac.rs (Miodrag M. Spalević)

Keywords: Linear system of equations, conjugate gradient method, error norm, Gauss quadrature rule, averaged Gauss quadrature rule, optimal averaged Gauss quadrature rule

AMS classification: 65D32, 65F10

1. Introduction

Linear systems of equations

$$Ax = b \tag{1}$$

with a large symmetric positive definite matrix $A \in \mathbb{R}^{n\times n}$ and right-hand side vector $b \in \mathbb{R}^n$ are often solved by the conjugate gradient (CG) iterative method. It is desirable to be able to estimate the norm of the error in iterates computed by the CG method so that the computations can be terminated as soon as an iterate that approximates the solution of (1) with required accuracy has been determined.

Let $x_* \in \mathbb{R}^n$ denote the solution of (1) and let $x_\ell \in \mathbb{R}^n$ stand for the ℓth iterate determined by the CG method when applied to the solution of this system. In many applications, e.g., in physics and structural engineering, since A is symmetric positive definite, it is meaningful to measure the error $\widetilde{e}_\ell = x_* - x_\ell$ with the A-norm,

$$\|\widetilde{e}_\ell\|_A = \sqrt{\widetilde{e}_\ell^T A \widetilde{e}_\ell}, \tag{2}$$

where the superscript ^{T} denotes transposition. The A-norm is often referred to as the energy norm.

The development of techniques for determining bounds or estimates of the A-norm of the error (2) has received considerable attention in the literature; see, e.g., [14, 15, 16, 20, 21, 23, 24, 34]. The CG method is closely related to Gauss-type quadrature rules, and one approach to calculate bounds for (2) evaluates pairs of Gauss and Gauss-Radau quadrature rules. To obtain an upper bound, the latter rule requires the user to specify a node (the Radau node) between the origin and the smallest eigenvalue of A. In many applications the allocation of this node is not straightforward, because further computations typically are required to determine a positive lower bound for the smallest eigenvalue of A, though there are exceptions; see, e.g., [13, 18, 27] for preconditioned linear systems of equations that arise from the discretization of certain boundary value problems for some partial

differential equations for which a lower bound for the smallest eigenvalue of A can be determined quite inexpensively.

To improve the accuracy of the Gauss lower bound, Golub and Strakoš [16] as well as Golub and Meurant [14, 15, 20] used a Gauss rule of higher degree of precision. Specifically, they introduce a "delay" of d conjugate gradient steps when computing an estimate of $\|\widetilde{e}_\ell\|_A$, i.e., they estimate $\|\widetilde{e}_\ell\|_A$ when $\ell + d$ steps of the CG method have been carried out. This adds an expense of d matrix-vector product evaluations that are not required to satisfy a specified error-tolerance for the A-norm. How expensive these d matrix-vector product evaluations are depends on the size of the matrix A, on how the matrix is stored, and on how the matrix-vector product is computed. It could be quite expensive; see [10, Example 1.1] for an illustration. Typical values of the delay d are $d = 5$, $d = 10$, or larger; see [14, 16, 20]. Large delays may sometimes be required. A heuristic algorithm for an adaptive choice of the delay during the CG iterations to obtain a required accuracy is proposed in [22]; see also [25].

The difficulties with the approaches mentioned above to estimate the error norm in the CG method lead Calvetti et al. [4] to propose to compute estimates of lower and upper bounds for the A-norm by evaluating pairs of Gauss and associated anti-Gauss quadrature rules. The latter rules, which were introduced by Laurie [19], do not require a user to specify a node between the origin and the smallest eigenvalue of A. A sufficient condition for pairs of Gauss and associated anti-Gauss rules to bracket the A-norm of the error is that the smallest eigenvalue of A is far away enough from the origin. This follows from the discussion at the end of [5, Section 2]. However, it is difficult to verify whether this condition holds for a given linear system (1), and it is quite easy to find linear systems for which the anti-Gauss rules yield error estimates of poor quality; see Example 4.3 in Section 4 for an illustration.

Laurie [19] also defined averaged Gauss quadrature rules, which are the average of a Gauss rule and an associated anti-Gauss rule. The difference between the Gauss rule and an associated averaged Gauss rule can be used to estimate the A-norm of the error. Subsequently, Spalević [33] introduced optimal averaged Gauss quadrature rules that are of higher degree of precision than the averaged Gauss rules by Laurie [19] with the same number of nodes. In this paper, we propose the use of pairs of Gauss and associated averaged or optimal averaged Gauss quadrature rules, or pairs of Gauss and shifted averaged or optimal averaged Gauss rules, to estimate the A-norm. The shift is analogous to the delay mentioned above. These quadrature rules do not require a user to explicitly allocate a node between the origin and

the smallest eigenvalue of A. Moreover, the shifted averaged and optimal averaged Gauss quadrature rules give error estimates of higher quality than using a delay of the same size as the shift for the Gauss rule. Our implementation of these rules uses an elegant approach described by Meurant and Tichý [23] for computing quadrature rules from the recurrence coefficients of the CG method.

This paper is organized as follows. Section 2 reviews properties of Gauss, averaged Gauss, and optimal averaged Gauss quadrature rules. The application of these rules to the estimation of the A-norm of the error in iterates computed by the CG method when applied to the solution of (1) is discussed in Section 3, where also implementations of these quadrature rules are described. Section 4 presents a few computed examples. This paper illustrates the performance of the computed error estimates for a few fixed values of d, i.e., d is the same for all iterations. The methods can be extended to allow d to depend on the iteration number ℓ, similarly as described in [22, 25] at the expense of a more complicated algorithm. Concluding remarks can be found in Section 5.

We conclude this section by noting that techniques based on extrapolation also can be applied to estimate the norm of the error in approximate solutions determined by iterative methods; see, e.g., Brezinski et al. [2, 3], Mitrouli and Roupa [26], and [28] for illustrations. A careful comparison of quadrature methods and extrapolation methods is a topic for forthcoming work.

2. Gauss, averaged Gauss, and optimal averaged Gauss quadrature rules

Let $\mathrm{d}\omega$ denote a non-negative real-valued measure with infinitely many points of support on the positive real axis such that all moments $\mu_k = \int t^k \mathrm{d}\omega(t)$, $k = 0, 1, 2, \ldots$, exist. We assume for notational simplicity in this section that $\mu_0 = 1$. Let the function f be continuous on the convex hull of the support of the measure and consider the approximation of Riemann-Stieltjes integrals of the form

$$\mathcal{I}(f) = \int f(t)\,\mathrm{d}\omega(t) \tag{3}$$

by an ℓ-node Gauss quadrature rule

$$\mathcal{G}_\ell(f) = \sum_{k=1}^{\ell} f(t_k^{(\ell)}) w_k^{(\ell)}. \tag{4}$$

It is well known that the nodes $t_k^{(\ell)}$ are real and distinct, and that the weights $w_k^{(\ell)}$ are positive. The Gauss rule has degree of precision $2\ell-1$, i.e., it satisfies

$$\mathcal{G}_\ell(f) = \mathcal{I}(f), \qquad \forall\ f \in \mathbb{P}_{2\ell-1},$$

where $\mathbb{P}_{2\ell-1}$ denotes the set of all polynomials of degree at most $2\ell-1$; see, e.g., Gautschi [12] for a proof.

The Gauss rule (4) can be associated with a symmetric tridiagonal matrix

$$T_\ell = \begin{bmatrix} \alpha_1 & \beta_1 & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{\ell-2} & \alpha_{\ell-1} & \beta_{\ell-1} \\ 0 & & & \beta_{\ell-1} & \alpha_\ell \end{bmatrix} \in \mathbb{R}^{\ell\times\ell} \tag{5}$$

with non-trivial entries $\alpha_k \in \mathbb{R}$ and $\beta_k > 0$. The α_k and β_k^2 are coefficients in the three-term recursion formula for the sequence of monic orthogonal polynomials $\{p_k\}_{k=0}^\infty$ (with $\deg(p_k) = k$) corresponding to the inner product

$$(g,h) = \int g(t)\,h(t)\,\mathrm{d}\omega(t).$$

Thus,

$$p_{k+1}(t) = (t-\alpha_{k+1})p_k(t) - \beta_k^2\,p_{k-1}(t), \quad k = 0,1,\ldots,$$

where $p_{-1}(t) \equiv 0$, $p_0(t) \equiv 1$, and

$$\alpha_{k+1} = \frac{(tp_k,p_k)}{(p_k,p_k)}, \qquad \beta_k^2 = \frac{(p_k,p_k)}{(p_{k-1},p_{k-1})};$$

see [12]. The eigenvalues of T_ℓ are the nodes $t_k^{(\ell)}$ and the squared first components of normalized eigenvectors are the weights $w_k^{(\ell)}$ of the Gauss rule (4); see [12, 15]. Using the spectral factorization of T_ℓ, it is easy to see that the Gauss rule (4) can be expressed as

$$\mathcal{G}_\ell(f) = e_1^T f(T_\ell) e_1; \tag{6}$$

throughout this paper $e_j = [0,\ldots,0,1,0,\ldots,0]^T$ denotes the jth column of an identity matrix of appropriate order.

Laurie [19] introduced the $(\ell+1)$-node anti-Gauss rule $\breve{\mathcal{G}}_{\ell+1}$ associated with the Gauss rule (4). It is characterized by the property

$$(\mathcal{I} - \breve{\mathcal{G}}_{\ell+1})(f) = -(\mathcal{I} - \mathcal{G}_\ell)(f), \qquad \forall\ f \in \mathbb{P}_{2\ell+1}. \tag{7}$$

The rule $\breve{\mathcal{G}}_{\ell+1}$ is related to the symmetric tridiagonal matrix

$$\breve{T}_{\ell+1} = \begin{bmatrix} \alpha_1 & \beta_1 & & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \beta_{\ell-2} & \alpha_{\ell-1} & \beta_{\ell-1} & \\ & & & \beta_{\ell-1} & \alpha_\ell & \sqrt{2}\beta_\ell \\ 0 & & & & \sqrt{2}\beta_\ell & \alpha_{\ell+1} \end{bmatrix} \in \mathbb{R}^{(\ell+1)\times(\ell+1)}. \tag{8}$$

Analogously to (6), we have

$$\breve{\mathcal{G}}_{\ell+1}(f) = e_1^T f(\breve{T}_{\ell+1}) e_1. \tag{9}$$

Moreover,

$$\breve{\mathcal{G}}_\ell(f) = \mathcal{I}(f), \qquad \forall\ f \in \mathbb{P}_{2\ell-1}.$$

Anti-Gauss rules have recently been analyzed by Díaz de Alba et al. [1].

If the magnitude of the coefficients of an expansion of the integrand f in terms of the orthogonal polynomials p_j decreases sufficiently quickly with increasing index, then the Gauss rule (4) and the anti-Gauss rule (9) bracket the integral (3); see [5] for a proof. We remark that in the application to estimating the error in the CG method, the integrand of interest is $f(t) = 1/t$ and the support of the measure is at the eigenvalues of the symmetric positive definite matrix A in (1); see Section 3. Then, the magnitude of the coefficients decays faster to zero the further away the support of the measure is from the origin.

However, given a measure $\mathrm{d}\omega$ and an integrand f, it is difficult to verify whether this condition on the coefficients holds. Therefore, it is typically not known to a user whether pairs of the rules (4) and (9) bracket the integral (3). Computed examples with application to the estimation of the A-norm of the error are presented in [4] as well as in Section 4, where we illustrate that anti-Gauss rules may determine estimates for the error norm of rather poor quality for some linear systems of equations.

Laurie [19] also defined the averaged Gauss rule

$$\mathcal{A}_{2\ell+1} = \frac{1}{2}(\mathcal{G}_\ell + \breve{\mathcal{G}}_{\ell+1}) \tag{10}$$

with $2\ell+1$ nodes. Property (7) suggests that the quadrature error for $\mathcal{A}_{2\ell+1}$ is smaller than the quadrature error for $\mathcal{G}_\ell$. Indeed, it follows from (7) that the degree of exactness of $\mathcal{A}_{2\ell+1}$ is at least $2\ell+1$. This suggests that the difference

$$\mathcal{A}_{2\ell+1}(f) - \mathcal{G}_\ell(f) \tag{11}$$

may be used to estimate the quadrature error

$$\mathcal{I}(f) - \mathcal{G}_\ell(f). \tag{12}$$

It is illustrated in [30] that for many integrands, the difference (11) yields a quite accurate estimate of the quadrature error (12) for various values of ℓ, where the accuracy of the estimates depends both on the integrand as well as on ℓ.

Spalević [33] showed that the averaged Gauss rule (10) can be represented by a single symmetric tridiagonal matrix of order $2\ell+1$. This observation lead him to modify this matrix to obtain the optimal averaged Gauss quadrature rule $\widehat{\mathcal{A}}_{2\ell+1}$ with higher degree of precision and the same number of nodes as the averaged Gauss rule (10). The symmetric tridiagonal matrix associated with the optimal averaged Gauss quadrature rule with $2\ell+1$ nodes is defined as follows. Introduce the reverse matrix

$$T'_\ell = \begin{bmatrix} \alpha_\ell & \beta_{\ell-1} & & & 0 \\ \beta_{\ell-1} & \alpha_{\ell-1} & \beta_{\ell-2} & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_2 & \alpha_2 & \beta_1 \\ 0 & & & \beta_1 & \alpha_1 \end{bmatrix} \in \mathbb{R}^{\ell\times\ell},$$

which is obtained by reversing the order of the rows and columns of the matrix (5). The nodes and weights of the optimal averaged Gauss rule are the eigenvalues and the squared first components of normalized eigenvectors, respectively, of the concatenated symmetric tridiagonal matrix

$$\widehat{T}_{2\ell+1} = \begin{bmatrix} T_\ell & \beta_\ell e_\ell & 0 \\ \beta_\ell e_\ell^T & \alpha_{\ell+1} & \beta_{\ell+1} e_1^T \\ 0 & \beta_{\ell+1} e_1 & T'_\ell \end{bmatrix} \in \mathbb{R}^{(2\ell+1)\times(2\ell+1)}.$$

We remark that ℓ eigenvalues of this matrix are eigenvalues of T_ℓ. Analogously to (6), we have

$$\widehat{\mathcal{A}}_{2\ell+1}(f) = e_1^T f(\widehat{T}_{2\ell+1}) e_1.$$

It is easy to see that this quadrature rule has degree of precision at least $2\ell+2$; see, e.g., [31] for details. This suggests that the quadrature error in the Gauss rule (4) be estimated as

$$\widehat{\mathcal{A}}_{2\ell+1}(f) - \mathcal{G}_\ell(f). \tag{13}$$

Computed examples reported in [30] show this estimate to be quite accurate for many integrands. Since this rule has higher degree of precision than Laurie's averaged Gauss rule (10), we expect the quadrature error estimate (13), generally, to be more accurate than the estimate (11) at least for integrands that are differentiable many times. Computed examples presented in [30] corroborate this expectation.

It is convenient to use the representation

$$\widehat{\mathcal{A}}_{2\ell+1}(f) = \frac{\beta_{\ell+1}^2}{\beta_\ell^2+\beta_{\ell+1}^2}\mathcal{G}_\ell(f) + \frac{\beta_\ell^2}{\beta_\ell^2+\beta_{\ell+1}^2}\mathcal{G}_{\ell+1}^*(f), \tag{14}$$

where

$$\mathcal{G}_{\ell+1}^*(f) = e_1^T f(T_{\ell+1}^*)e_1 \tag{15}$$

with

$$T_{\ell+1}^* = \begin{bmatrix} \alpha_1 & \beta_1 & & & & 0 \\ \beta_1 & \alpha_2 & \beta_2 & & & \\ & \ddots & \ddots & \ddots & & \\ & & \beta_{\ell-2} & \alpha_{\ell-1} & \beta_{\ell-1} & \\ & & & \beta_{\ell-1} & \alpha_\ell & \beta_\ell^* \\ 0 & & & & \beta_\ell^* & \alpha_{\ell+1} \end{bmatrix} \in \mathbb{R}^{(\ell+1)\times(\ell+1)}$$

and

$$\beta_\ell^* = \sqrt{\beta_\ell^2+\beta_{\ell+1}^2}.$$

The representation (14) is shown in [29]. It is analogous to the representation (10) of the averaged rule and simplifies to the latter when $\beta_{\ell+1} = \beta_\ell$. Some properties of the quadrature rule (15) have recently been shown in [11].

The A-norm of the error in approximate solutions determined by the CG method can be estimated by evaluating the quadrature error of the Gauss rule (4) when applied to the integrand $f(t) = 1/t$ for a certain measure, that is determined by the matrix A and right-hand side b in (1), with support on the positive real axis. The quality of the error estimates (11) and (13) may depend on how close to the origin a node of the averaged Gauss rules $\mathcal{A}_{2\ell+1}$ and optimal averaged Gauss rules $\widehat{\mathcal{A}}_{2\ell+1}$ is located, because the nodes are not required to be in the convex hull of the support of the measure that determines the Gauss rule. The location of the extreme nodes of these rules has been investigated for a variety of classical and other measures for which the recurrence coefficients for the orthogonal polynomials are explicitly known; see, e.g., [7, 8, 9] and references therein. The analyses show whether the nodes are in the convex hull of the support of the measure. However, these

results generally do not shed light on how the nodes for measures that arise in the application of this paper are allocated.

3. Error norm estimates for the conjugate gradient method

This section discusses how the averaged and optimal averaged Gauss rules described in the previous section can be used to estimate the A-norm of the error in iterates of the CG method when applied to the solution of the system (1). The system is assumed to be preconditioned if preconditioning is desired; preconditioning is discussed in, e.g., [23, 25]. We use the technique described by Meurant and Tichý [23] to evaluate the quadrature rules.

In this and the following sections, the initial iterate is set to $x_0 = 0$. We provide an algorithm for the CG method for completeness; see, e.g., [15, 17, 32] for discussions of this method.

Algorithm 1 The conjugate gradient method

1: **Input:** Symmetric positive definite matrix $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$, number of steps ℓ;
2: $x_0 := 0 \in \mathbb{R}^n$; $r_0 := b$; $p_0 := r_0$;
3: **for** $k = 1$ **to** ℓ
4: $\quad w := Ap_{k-1}$;
5: $\quad \gamma_{k-1} := r_{k-1}^T r_{k-1}/w^T p_{k-1}$;
6: $\quad x_k := x_{k-1} + \gamma_{k-1} p_{k-1}$;
7: $\quad r_k := r_{k-1} - \gamma_{k-1} w$;
8: $\quad \delta_k := r_k^T r_k / r_{k-1}^T r_{k-1}$;
9: $\quad p_k := r_k + \delta_k p_{k-1}$;
10: **end for**
11: **Output:** Approximate solution x_ℓ, recursion coefficients $\{\gamma_k\}_{k=0}^{\ell-1}$, $\{\delta_k\}_{k=1}^{\ell}$;

The conjugate gradient method applied to the solution of (1) generates a sequence of approximations $x_1, x_2, \ldots, x_\ell$ of the solution x_*. The kth iterate, x_k, minimizes the A-norm of the error,

$$\widetilde{e}_k = x_* - x_k, \tag{16}$$

over the Krylov subspace

$$\mathbb{K}_k(A, b) = \text{span}\{b, Ab, \ldots, A^{k-1}b\},$$

i.e., the error (16) satisfies

$$\|\widetilde{e}_k\|_A = \|x_* - x_k\|_A = \min_{x \in \mathbb{K}_k(A,b)} \|x_* - x\|_A; \tag{17}$$

see, e.g., [32, Chapter 6]. The residual vector $r_k = b - Ax_k$ satisfies

$$r_k^T A^{-1} r_k = \widetilde{e}_k A \widetilde{e}_k = \|\widetilde{e}_k\|_A^2.$$

Let f be a function such that $f(A)$ is well defined and introduce the spectral factorization

$$A = U \Lambda U^T,$$

where the matrix $U \in \mathbb{R}^{n \times n}$ is orthogonal and the diagonal entries of

$$\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_n]$$

are the eigenvalues of A. Let $b' = [b'_j]_{j=1}^n = U^T b$. Then

$$b^T f(A) b = \sum_{j=1}^n f(\lambda_j)(b'_j)^2 = \int f(t)\, \mathrm{d}\omega(t), \tag{18}$$

where the distribution function $\omega(t)$ associated with the measure $\mathrm{d}\omega(t)$ is non-decreasing, piece-wise constant, and has jumps $(b'_j)^2$ at $t = \lambda_j$ for $j = 1, 2, \ldots, n$. Define the lower bidiagonal matrix

$$L_\ell = \begin{bmatrix} 1 & & & & 0 \\ \sqrt{\delta_1} & 1 & & & \\ & \ddots & \ddots & & \\ & & \sqrt{\delta_{\ell-2}} & 1 & \\ 0 & & & \sqrt{\delta_{\ell-1}} & 1 \end{bmatrix} \in \mathbb{R}^{\ell \times \ell}, \tag{19}$$

the diagonal matrix

$$D_\ell = \mathrm{diag}[1/\gamma_0, 1/\gamma_1, \ldots, 1/\gamma_{\ell-1}] \in \mathbb{R}^{\ell \times \ell}, \tag{20}$$

and the symmetric tridiagonal matrix

$$T_\ell = L_\ell D_\ell L_\ell^T. \tag{21}$$

This is a Cholesky-type factorization. The entries of these matrices are determined by the CG method and can be evaluated when ℓ steps of Algorithm 1 have been carried out. We have that

$$\mathcal{G}_\ell(f) = \|b\|^2 e_1^T f(T_\ell) e_1$$

is the ℓ-node Gauss quadrature rule for the approximation of the integral on the right-hand side of (18).

In the context of approximating the A-norm, we let $f(t) = 1/t$ and define the matrix T_n that can be computed when n steps of the CG method have been executed. Then

$$\|\widetilde{e}_\ell\|_A^2 = \|b\|^2 e_1^T T_n^{-1} e_1 - \|b\|^2 e_1^T T_\ell^{-1} e_1; \tag{22}$$

see, e.g., [15, Theorem 12.1] or [21]. The matrix T_n is too expensive to compute. Instead we approximate the first term in the right-hand side of (22) by a quadrature rule that is more accurate than $\mathcal{G}_\ell(f)$.

We note for future reference that the entries of the matrices T_ℓ, L_ℓ, and D_ℓ are related as follows:

$$\beta_k = \frac{\sqrt{\delta_k}}{\gamma_{k-1}}, \quad \alpha_k = \frac{1}{\gamma_{k-1}} + \frac{\delta_{k-1}}{\gamma_{k-2}}, \quad k = 1, 2, \ldots, \ell, \tag{23}$$

with $\delta_0 = 0$ and $\gamma_{-1} = 1$; this follows from (21).

Following Meurant and Tichý [23], we evaluate the Gauss rule (4) with the integrand $f(t) = 1/t$ by using the Cholesky-type factorization (21),

$$\mathcal{G}_\ell(f) = \|b\|^2 \, e_1^T T_\ell^{-1} e_1 = \|b\|^2 \, e_1^T L_\ell^{-T} D_\ell^{-1} L_\ell^{-1} e_1 = \|b\|^2 \|D_\ell^{-1/2} L_\ell^{-1} e_1\|^2,$$

where the right-hand side is evaluated as

$$\mathcal{G}_\ell(f) = \|b\|^2 \left(\gamma_0 + \sum_{j=2}^{\ell} \gamma_{j-1} (\delta_{j-1} \delta_{j-2} \ldots \delta_1) \right). \tag{24}$$

From the definition of the δ_i's in Algorithm 1, it can be shown that $\delta_{j-1}\delta_{j-2}\ldots\delta_1 = \|r_{j-1}\|^2/\|b\|^2$; see [22].

After having carried out $\ell + d$ steps of the CG algorithm, we can evaluate the quadrature rule $\mathcal{G}_{\ell+d}(f)$ in a similar fashion to obtain

$$\mathcal{G}_{\ell+d}(f) - \mathcal{G}_\ell(f) = \|b\|^2 \sum_{j=\ell+1}^{\ell+d} \gamma_{j-1} (\delta_{j-1} \delta_{j-2} \ldots \delta_1).$$

The evaluation of the quadrature rules $\breve{\mathcal{G}}_{\ell+d+1}(f)$ and $\mathcal{G}^*_{\ell+d+1}(f)$ in (9) and (15), respectively, with ℓ replaced by $\ell + d$ is straightforward. We note that the Cholesky-type factorization of the matrix (8) (with ℓ replaced by $\ell + d$) is given by

$$\breve{T}_{\ell+d+1} = \breve{L}_{\ell+d+1} \breve{D}_{\ell+d+1} \breve{L}_{\ell+d+1}^T,$$

with lower bidiagonal matrix

$$\breve{L}_{\ell+d+1} = \begin{bmatrix} 1 & & & & 0 \\ \sqrt{\delta_1} & 1 & & & \\ & \ddots & \ddots & & \\ & & \sqrt{\delta_{\ell+d-1}} & 1 & \\ 0 & & & \sqrt{\breve{\delta}_{\ell+d}} & 1 \end{bmatrix} \in \mathbb{R}^{(\ell+d+1)\times(\ell+d+1)}$$

and diagonal matrix

$$\breve{D}_{\ell+d+1} = \mathrm{diag}[1/\gamma_0, \ldots, 1/\gamma_{\ell-1}, 1/\breve{\gamma}_{\ell+d}] \in \mathbb{R}^{(\ell+d+1)\times(\ell+d+1)},$$

where

$$\breve{\delta}_{\ell+d} = s^2\delta_{\ell+d}, \qquad \breve{\gamma}_{\ell+d} = \left(\frac{1}{\gamma_{\ell+d}} + \frac{(1-s^2)\delta_{\ell+d}}{\gamma_{\ell+d-1}}\right)^{-1}.$$

Here $\alpha_{\ell+d}$, $\alpha_{\ell+d+1}$, and $\beta_{\ell+d}$ are determined by (23) with ℓ replaced by $\ell+d$ and $s=\sqrt{2}$. In particular, the leading $(\ell+d)\times(\ell+d)$ principal submatrices of $\breve{L}_{\ell+d+1}$ and $\breve{D}_{\ell+d+1}$ agree with the matrices (19) and (20), respectively, with ℓ replaced by $\ell+d$. It follows that we obtain, analogously to (24),

$$\begin{aligned} \breve{\mathcal{G}}_{\ell+d+1}(f) &= \|b\|^2 \left(\gamma_0 + \sum_{j=2}^{\ell+d} \gamma_{j-1}(\delta_{j-1}\delta_{j-2}\ldots\delta_1)\right) \\ &+ \|b\|^2\breve{\gamma}_{\ell+d}(\breve{\delta}_{\ell+d}\delta_{\ell+d-1}\delta_{\ell+d-2}\ldots\delta_1). \end{aligned} \tag{25}$$

We are in a position to discuss the computation of the difference (11) with ℓ replaced by $\ell+d$ in the first term. It follows from (24) and (25) that

$$\begin{aligned} & \mathcal{A}_{2(\ell+d)+1}(f) - \mathcal{G}_\ell(f) \\ = \;& \frac{1}{2}(\mathcal{G}_{\ell+d}(f) + \breve{\mathcal{G}}_{\ell+d+1}(f)) - \mathcal{G}_\ell(f) \\ = \;& \|b\|^2 \left(\sum_{j=\ell+1}^{\ell+d} \gamma_{j-1}(\delta_{j-1}\delta_{j-2}\ldots\delta_1) + \frac{1}{2}\breve{\gamma}_{\ell+d}(\breve{\delta}_{\ell+d}\delta_{\ell+d-1}\ldots\delta_1)\right). \end{aligned}$$

Finally, we turn to the evaluation of the error estimate (13) with ℓ replaced by $\ell+d$ in the first term, and start with the computation of the quadrature rule (15). This rule is evaluated similarly as (9). Let

$$s = \left(1 + \frac{\delta_{\ell+d+1}}{\gamma_{\ell+d}^2} \cdot \frac{\gamma_{\ell+d-1}^2}{\delta_{\ell+d}}\right)^{1/2}$$

and

$$\delta^*_{\ell+d} = s^2\delta_{\ell+d}, \qquad \gamma^*_{\ell+d} = \left(\frac{1}{\gamma_{\ell+d}} + \frac{(1-s^2)\delta_{\ell+d}}{\gamma_{\ell+d-1}}\right)^{-1}$$

with $\alpha_{\ell+d}$, $\alpha_{\ell+d+1}$, and $\beta_{\ell+d}$ determined by (23). This gives analogously to (25),

$$\begin{aligned}\mathcal{G}^*_{\ell+d+1}(f) &= \|b\|^2 \left(\gamma_0 + \sum_{j=2}^{\ell+d} \gamma_{j-1}(\delta_{j-1}\delta_{j-2}\ldots\delta_1)\right) \\ &+ \|b\|^2\gamma^*_{\ell+d}(\delta^*_{\ell+d}\delta_{\ell+d-1}\ldots\delta_1).\end{aligned}$$

The error estimate (13), with ℓ replaced by $\ell+d$ in the first term, can now be computed by using the representation (14). Thus,

$$\begin{aligned}&\widehat{\mathcal{A}}_{2(\ell+d)+1}(f) - \mathcal{G}_\ell(f) \\ =\ & \frac{\beta^2_{\ell+d+1}}{\beta^2_{\ell+d}+\beta^2_{\ell+d+1}}\mathcal{G}_{\ell+d}(f) + \sigma_{(l,d)}\mathcal{G}^*_{\ell+d+1}(f) - \mathcal{G}_\ell(f) \\ =\ & \|b\|^2 \left(\sum_{j=\ell+1}^{\ell+d} \gamma_{j-1}(\delta_{j-1}\delta_{j-2}\ldots\delta_1) + \sigma_{(l,d)}\gamma^*_{\ell+d}(\delta^*_{\ell+d}\delta_{\ell+d-1}\ldots\delta_1)\right),\end{aligned}$$

where $\sigma_{(l,d)} = \frac{\beta^2_{\ell+d}}{\beta^2_{\ell+d}+\beta^2_{\ell+d+1}}$. It follows from (22) that an estimate of the A-norm of the error in the iterate x_ℓ is given by the square root of the above error estimate. The other estimates for the error in $\mathcal{G}_\ell(f)$ derived in this section can be used similarly.

We finally remark that the remainder formula for Gauss quadrature applied to the integrand $f(t) = 1/t$ on the positive real axis shows that the integration error (12) is positive, which means that the Gauss rule gives a lower bound; see [12, 15]. Moreover, Strakoš and Tichý [34] have shown that Gauss quadrature rules for increasing number of nodes ℓ can be evaluated accurately in finite precision arithmetic by a recursion formula. The difficulty is to compute an upper bound.

We have seen that a quite natural approach is to approximate the error (12) by

$$\mathcal{G}_{\ell+d}(f) - \mathcal{G}_\ell(f) \tag{26}$$

for some $d > 0$ as described in [14, 15, 16, 20]. It is desirable to choose the delay d as small as possible since, as already mentioned in Section 1, the computational effort required to evaluate $\mathcal{G}_{\ell+d}(f)$ increases with d; when the matrix A is large, the dominating computational work is the calculation

of $\ell + d$ matrix-vector products with A. This suggests that one should choose $d = 1$. However, Clenshaw and Curtis [6, p. 199] have illustrated that this choice may yield useless estimates of the quadrature error (12); in their example the actual quadrature error is more than a factor 100 larger than the computed estimate. The authors of [14, 15, 16, 20, 22, 25] use a larger value of d. This reduces the difficulty pointed out in [6]. The following section compares error estimates determined for several values of d to estimates obtained with the averaged and optimal averaged Gauss rules and shifted variants thereof.

4. Numerical experiments

This section presents a few computed examples that illustrate the quality of the error norm estimates discussed in the previous section. All computations were carried out with IEEE double precision arithmetic, i.e., with about 15 significant decimal digits, using MATLAB R2021a on a 64-bit personal computer. The initial approximate solution is $x_0 = 0$ in all examples. The iterations are terminated as soon as $\|\widetilde{e}_k\|_A < \varepsilon$, where $\varepsilon = 10^{-11}$. We observed that the error norm estimates perform well also for smaller values of ε, as long as the maximum attainable accuracy of the A-norm of the error has not been reached.

In the first three examples we compute estimates of the Riemann-Stieltjes integral (18) with $f(t) = 1/t$, that is, of the bilinear form $b^T A^{-1} b$. Notice that this bilinear form is the square of the initial A-norm of the error. The other examples are concerned with the estimates of the A-norm of the error $\widetilde{e}_k$ in the CG iterate x_k, which is given by (16).

EXAMPLE 4.1. Let the matrix $A \in \mathbb{R}^{n\times n}$ be symmetric tridiagonal with diagonal entries $2[1, 2, \ldots, n]$ and subdiagonal entries $[1, 2, \ldots, n-1]/2$. Then A is positive definite. We let $n = 500$, $x_* = [1, 1, \ldots, 1]^T/\sqrt{n} \in \mathbb{R}^n$, and $b = Ax_*$. The exact value of $\mathcal{F}(A) := b^T A^{-1} b$ is 750.50. The header MVP in the following tables stands for the number of matrix-vector products with the matrix A required to compute the iterates and the quadrature rules. This is the dominating computational cost when the matrix A is large.

Table 1 displays the magnitude of the relative quadrature errors incurred when approximating $b^T A^{-1} b$ by several quadrature rules. The $(\ell + 1)$-node anti-Gauss rule $\breve{\mathcal{G}}_{\ell+1}(f)$ is seen to yield a smaller error than the ℓ-node Gauss rule. The smallest errors are achieved by the averaged Gauss rules $\mathcal{A}_{2\ell+1}(f)$ and the optimal averaged Gauss rules $\widehat{\mathcal{A}}_{2\ell+1}(f)$ with $2\ell+1$ nodes; the latter rules give the smallest errors for $\ell = 30$ and $\ell = 40$, and close to the smallest error for $\ell = 20$. The fact that the quadrature errors achieved with the

Table 1: Example 4.1: Magnitude of the relative error of computed approximations of the integral (18) with $f(t) = 1/t$ for $A \in \mathbb{R}^{500\times 500}$, $b = Ax_*$, and $x_* = [1, 1, \ldots, 1]^T/\sqrt{500}$.

MVP		$\ell = 20$	$\ell = 30$	$\ell = 40$
ℓ	$\lvert\frac{\mathcal{G}_\ell(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$1.7668 \cdot 10^{-5}$	$3.5430 \cdot 10^{-6}$	$9.9117 \cdot 10^{-7}$
$\ell+1$	$\lvert\frac{\breve{\mathcal{G}}_{\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$1.1271 \cdot 10^{-5}$	$2.5176 \cdot 10^{-6}$	$8.6987 \cdot 10^{-7}$
$\ell+1$	$\lvert\frac{\mathcal{A}_{2\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$3.1985 \cdot 10^{-6}$	$5.1274 \cdot 10^{-7}$	$6.0650 \cdot 10^{-8}$
$\ell+1$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$3.4954 \cdot 10^{-6}$	$5.0020 \cdot 10^{-7}$	$5.1140 \cdot 10^{-8}$

averaged and optimal averaged Gauss rules are smaller than the errors for the Gauss rule makes the differences (11) and (13) useful estimates of the quadrature error (12) of the Gauss rule.

Table 2 reports the magnitude of the relative differences determined by (11), (13), and (26) for a few values of d. The evaluation of the quadrature rules $\widehat{\mathcal{A}}_{2(\ell+d)+1}(f)$ and $\mathcal{A}_{2(\ell+d)+1}(f)$ requires $\ell+d+1$ matrix-vector products (MVPs) with the matrix A. It is desirable to keep the number of MVPs small. We can observe that the error estimates determined by the averaged and optimal averaged Gauss rules are about the same for the values of d considered. Moreover, the error estimates determined by $\widehat{\mathcal{A}}_{2(\ell+d)+1}(f)$ and $\mathcal{A}_{2(\ell+d)+1}(f)$ are more accurate than those obtained by $\mathcal{G}_{\ell+d}(f)$ for the same values of d, i.e., they are closer to the actual errors shown in the first line of Table 1.

The error estimates in Table 2 are seen not to decrease monotonically as the delay d increases. This is due to the fact that the Gauss rule $\mathcal{G}_\ell(f)$ only furnishes an approximation of the exact value $\mathcal{F}(A)$. To illustrate that the quadrature errors of the rules $\mathcal{G}_{\ell+d}(f)$, $\widehat{\mathcal{A}}_{2(\ell+d)+1}(f)$, and $\mathcal{A}_{2(\ell+d)+1}(f)$ decrease monotonically as d increases, we show in Table 3 the quadrature errors for these rules for the same values of ℓ and d as in Table 2.

Example 4.2. Let $B \in \mathbb{R}^{100\times 100}$ be a symmetric matrix with randomly generated uniformly distributed real eigenvalues in the interval $[-5, 5]$ and a random orthogonal eigenvector matrix. Define the matrix $A = I_{100} + B^2$, where I_{100} denotes the identity matrix of order 100, and let the vector b in the linear system (1) be $b = Ax_*$, where the vector x_* is a random vector of unit norm.

Table 4 is similar to Table 1 and displays the magnitude of the relative quadrature errors in the computed approximations. We can observe that the

Table 2: Example 4.1: Magnitude of the relative error estimates determined by Gauss, averaged Gauss, and optimal averaged Gauss rules for the problem of Table 1. The entries should be compared with the entries of line 1 of Table 1.

		Error		
MVP		$\ell = 20$	$\ell = 30$	$\ell = 40$
$\ell + 1$	$\lvert\frac{\mathcal{G}_{\ell+1}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$3.0066 \cdot 10^{-6}$	$4.5295 \cdot 10^{-7}$	$1.1172 \cdot 10^{-7}$
$\ell + 4$	$\lvert\frac{\mathcal{G}_{\ell+4}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$8.9436 \cdot 10^{-6}$	$1.4605 \cdot 10^{-6}$	$3.7481 \cdot 10^{-7}$
$\ell + 8$	$\lvert\frac{\mathcal{G}_{\ell+8}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.2962 \cdot 10^{-5}$	$2.2803 \cdot 10^{-6}$	$6.0758 \cdot 10^{-7}$
$\ell + 1$	$\lvert\frac{\mathcal{A}_{2\ell+1}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.4469 \cdot 10^{-5}$	$3.0303 \cdot 10^{-6}$	$9.3052 \cdot 10^{-7}$
$\ell + 3$	$\lvert\frac{\mathcal{A}_{2\ell+5}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.5548 \cdot 10^{-5}$	$3.2084 \cdot 10^{-6}$	$9.6622 \cdot 10^{-7}$
$\ell + 5$	$\lvert\frac{\mathcal{A}_{2\ell+9}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.6208 \cdot 10^{-5}$	$3.3308 \cdot 10^{-6}$	$9.9189 \cdot 10^{-7}$
$\ell + 1$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+1}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.4172 \cdot 10^{-5}$	$3.0428 \cdot 10^{-6}$	$9.4003 \cdot 10^{-7}$
$\ell + 3$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+5}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.5226 \cdot 10^{-5}$	$3.1477 \cdot 10^{-6}$	$9.5042 \cdot 10^{-7}$
$\ell + 5$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+9}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.6009 \cdot 10^{-5}$	$3.2376 \cdot 10^{-6}$	$9.5490 \cdot 10^{-7}$

Table 3: Example 4.1: Magnitude of the relative error of computed approximations of the integral (18) with $f(t) = 1/t$ for $A \in \mathbb{R}^{500\times 500}$, $b = Ax_*$, and $x_* = [1, 1, \ldots, 1]^T/\sqrt{500}$.

		Error		
MVP		$\ell = 20$	$\ell = 30$	$\ell = 40$
$\ell+1$	$\lvert\frac{\mathcal{G}_{\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$1.4661 \cdot 10^{-5}$	$3.0901 \cdot 10^{-6}$	$8.7945 \cdot 10^{-7}$
$\ell+4$	$\lvert\frac{\mathcal{G}_{\ell+4}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$8.7240 \cdot 10^{-6}$	$2.0825 \cdot 10^{-6}$	$6.1636 \cdot 10^{-7}$
$\ell+8$	$\lvert\frac{\mathcal{G}_{\ell+8}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$4.7051 \cdot 10^{-6}$	$1.2628 \cdot 10^{-6}$	$3.8359 \cdot 10^{-7}$
$\ell+1$	$\lvert\frac{\mathcal{A}_{2\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$3.1985 \cdot 10^{-6}$	$5.1274 \cdot 10^{-7}$	$6.0650 \cdot 10^{-8}$
$\ell+3$	$\lvert\frac{\mathcal{A}_{2\ell+5}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$2.1197 \cdot 10^{-6}$	$3.3468 \cdot 10^{-7}$	$2.4946 \cdot 10^{-8}$
$\ell+5$	$\lvert\frac{\mathcal{A}_{2\ell+9}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$1.4598 \cdot 10^{-6}$	$2.1227 \cdot 10^{-7}$	$7.2648 \cdot 10^{-10}$
$\ell+1$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$3.4954 \cdot 10^{-6}$	$5.0020 \cdot 10^{-7}$	$5.1140 \cdot 10^{-8}$
$\ell+3$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+5}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$2.4416 \cdot 10^{-6}$	$3.9533 \cdot 10^{-7}$	$4.0747 \cdot 10^{-8}$
$\ell+5$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+9}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$1.6585 \cdot 10^{-6}$	$3.0548 \cdot 10^{-7}$	$3.6264 \cdot 10^{-8}$

quadrature errors achieved with the optimal averaged rules are of somewhat smaller magnitude than the errors achieved with the averaged rules that require the same number of matrix-vector product evaluations.

Table 5 is analogous to Table 2 and shows the relative differences

$$|(\mathcal{A}_{2(\ell+d)+1}(f) - \mathcal{G}_\ell(f))/\mathcal{F}(A)| \quad \text{and} \quad |(\widehat{\mathcal{A}}_{2(\ell+d)+1}(f) - \mathcal{G}_\ell(f))/\mathcal{F}(A)|$$

to provide more accurate estimates of the relative errors in $\mathcal{G}_\ell(f)$ than $|(\mathcal{G}_{\ell+d}(f) - \mathcal{G}_\ell(f))/\mathcal{F}(A)|$ with fewer matrix-vector product evaluations.

Example 4.3. Let the matrix $A \in \mathbb{R}^{48\times 48}$ in the linear system (1) be of the form $A = WDW^T$, where W is a random orthogonal matrix and D is a diagonal matrix with diagonal entries

$$d_i = \gamma + \frac{i-1}{47}(\rho - \gamma)\sigma^{48-i}, \quad i \in \{1, 2, \ldots, 48\}$$

with $\rho = 100$, $\sigma = 0.875$, and $\gamma = 0.1$. Hence, A is a symmetric positive definite dense matrix with eigenvalues in $[0.1, 100]$. The vector b is chosen so that $x_* = [1, 1, \ldots, 1]^T$ solves the linear system (1). A similar example can be found in [34].

Table 4: Example 4.2: Magnitude of the relative error of computed approximations of the integral (18) with $f(t) = 1/t$ for $A \in \mathbb{R}^{100\times 100}$, $b = Ax_*$, and x_* a random vector.

MVP		$\ell = 10$	$\ell = 20$
ℓ	$\lvert\frac{\mathcal{G}_\ell(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$1.0883 \cdot 10^{-4}$	$3.2009 \cdot 10^{-8}$
$\ell+1$	$\lvert\frac{\breve{\mathcal{G}}_{\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$1.5541 \cdot 10^{-4}$	$8.8694 \cdot 10^{-8}$
$\ell+1$	$\lvert\frac{\mathcal{A}_{2\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$2.3288 \cdot 10^{-5}$	$2.8343 \cdot 10^{-8}$
$\ell+1$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+1}(f)-\mathcal{F}(A)}{\mathcal{F}(A)}\rvert$	$7.1428 \cdot 10^{-6}$	$1.6473 \cdot 10^{-9}$

Table 5: Example 4.2: Magnitude of the relative error estimates determined by Gauss, averaged Gauss, and optimal averaged Gauss rules for the problem of Table 4.

		Error	
MVP		$\ell = 10$	$\ell = 20$
$\ell+1$	$\lvert\frac{\mathcal{G}_{\ell+1}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$6.3832 \cdot 10^{-5}$	$2.0127 \cdot 10^{-8}$
$\ell+4$	$\lvert\frac{\mathcal{G}_{\ell+4}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.0490 \cdot 10^{-4}$	$3.1407 \cdot 10^{-8}$
$\ell+8$	$\lvert\frac{\mathcal{G}_{\ell+8}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.0871 \cdot 10^{-4}$	$3.1983 \cdot 10^{-8}$
$\ell+1$	$\lvert\frac{\mathcal{A}_{2\ell+1}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.3212 \cdot 10^{-4}$	$6.0352 \cdot 10^{-8}$
$\ell+3$	$\lvert\frac{\mathcal{A}_{2\ell+5}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.1072 \cdot 10^{-4}$	$3.2372 \cdot 10^{-8}$
$\ell+5$	$\lvert\frac{\mathcal{A}_{2\ell+9}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.0792 \cdot 10^{-4}$	$3.1930 \cdot 10^{-8}$
$\ell+1$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+1}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.0169 \cdot 10^{-4}$	$3.3656 \cdot 10^{-8}$
$\ell+3$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+5}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.1049 \cdot 10^{-4}$	$3.2124 \cdot 10^{-8}$
$\ell+5$	$\lvert\frac{\widehat{\mathcal{A}}_{2\ell+9}(f)-\mathcal{G}_\ell(f)}{\mathcal{F}(A)}\rvert$	$1.0872 \cdot 10^{-4}$	$3.1933 \cdot 10^{-8}$

Figure 1 displays the estimates given by the Gauss rule $\mathcal{G}_\ell(f)$, the anti-Gauss rule $\breve{\mathcal{G}}_{\ell+1}(f)$, and the exact value $\mathcal{F}(A) = b^T A^{-1} b$ as functions of the number of iterations ℓ. The Gauss quadrature rules achieve values smaller than the value of $b^T A^{-1} b$, but pairs of Gauss and anti-Gauss rules can be seen not to bracket the value of the integral.

EXAMPLE 4.4. Let the matrix $A \in \mathbb{R}^{n\times n}$ be defined by

$$A = WDW^T, \qquad D = \mathrm{diag}[d_1, d_2, \ldots, d_n],$$

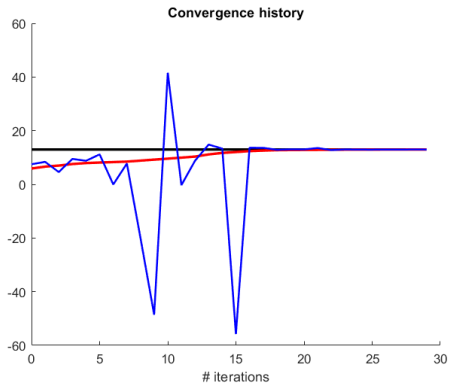
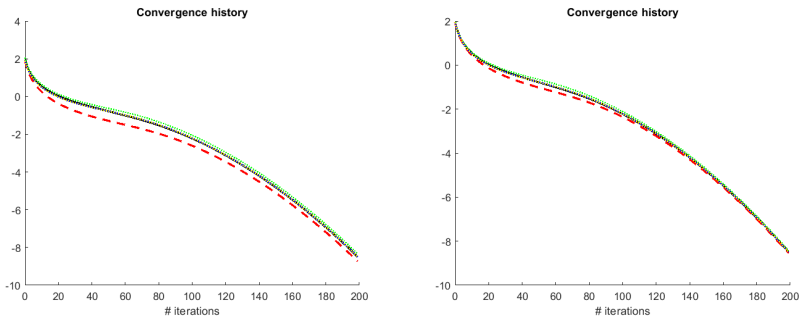


Figure 1: Example 4.3: Graphs for $\mathcal{G}_\ell(f)$ (solid red curve), $\breve{\mathcal{G}}_{\ell+1}(f)$ (dash-dotted blue curve), and the value of the integral $\mathcal{F}(A)$ (straight black line) for $f(t) = 1/t$.

where the eigenvector matrix $W \in \mathbb{R}^{n\times n}$ either is the identity or a random orthogonal matrix, and $n = 1000$. Thus, the matrix A is symmetric and either diagonal or dense. Let the diagonal elements of D be $d_j = j/7$. Then A is positive definite. The right-hand side vector b is chosen so that the vector $x_* = [1, 1, \ldots, 1]^T$ solves the linear system (1).

Figure 2 displays the 10-logarithm of the error norm $\|\widetilde{e}_\ell\|_A$ as a function of the iteration number ℓ, as well as error estimates determined by several quadrature rules when the matrix A is diagonal. Panel (a) shows the optimal averaged Gauss rule to determine accurate error estimates. Also, all quadrature rules used for panel (b) give accurate estimates. The curves obtained from the optimal averaged rule $\widehat{\mathcal{A}}_{2\ell+1}$, the averaged rule $\mathcal{A}_{2\ell+1}$, and the anti-Gauss rule are very close in all panels. Figure 3 shows results for the situation when A is a dense matrix. The lack of smoothness of the curves of panels (a) and (b) depends on oscillations of the CG residual vectors. The optimal averaged rule $\widehat{\mathcal{A}}_{2\ell+1}$ gives the best error estimates in panel (a), while all quadrature rules determine accurate estimates in panel (b).

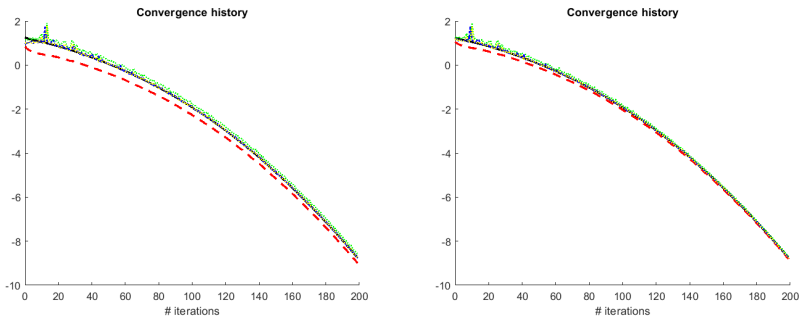
Example 4.5. Let $n = 200$ and consider the symmetric Toeplitz matrix $T \in \mathbb{R}^{n\times n}$ with first row $[1, 1/2, 1/2^2, \ldots 1/2^{n-1}]$. Define $A = T^2$. Then the matrix A is symmetric and positive definite. Let $x_* \in \mathbb{R}^{200}$ be a vector with normally distributed random entries (with zero mean and variance one), and let $b = Ax_*$. Figure 4 is analogous to Figure 3 and shows that the optimal averaged rule $\widehat{\mathcal{A}}_{2\ell+1}$ gives the best error estimates in panel (a), while all quadrature rules determine accurate estimates in panel (b). Similarly, as in Example 4.4, we notice that the three curves obtained from the averaged, the optimal averaged, and the anti-Gauss rules are almost the same.



(a) (b)

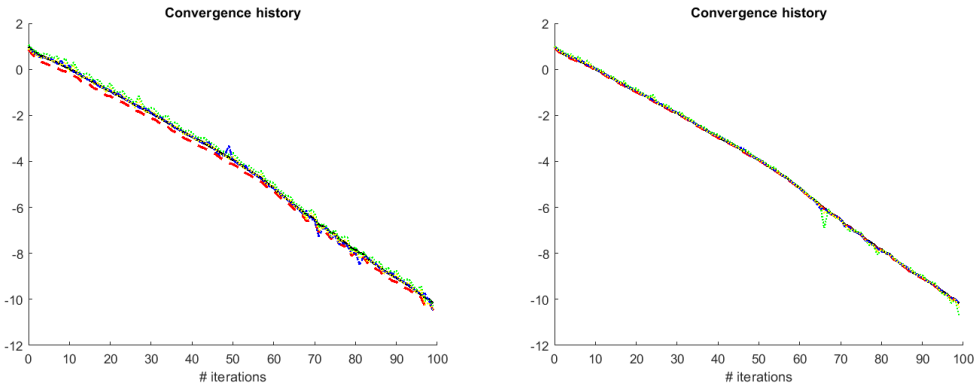
Figure 2: Example 4.4: Diagonal matrix A. (a) 10-logarithm of $\|\widetilde{e}_\ell\|_A$ (solid black graph) and 10-logarithm of estimates of $\|\widetilde{e}_\ell\|_A$ given by (i) $|\mathcal{G}_{\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dashed red curve), (ii) $|\breve{\mathcal{G}}_{\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dotted green curve), (iii) $|\widehat{\mathcal{A}}_{2\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dash-dotted blue curve), and (iv) $|\mathcal{A}_{2\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dotted yellow curve) for $f(t) = 1/t$. (b) 10-logarithm of $\|\widetilde{e}_\ell\|_A$ (solid black curve) and 10-logarithm of estimates of $\|\widetilde{e}_\ell\|_A$ given by (i) $|\mathcal{G}_{\ell+4}(f) - \mathcal{G}_\ell(f)|$ (dashed red curve), (ii) $|\breve{\mathcal{G}}_{\ell+4}(f) - \mathcal{G}_\ell(f)|$ (dotted green curve), (iii) $|\widehat{\mathcal{A}}_{2\ell+7}(f) - \mathcal{G}_\ell(f)|$ (dash-dotted blue curve), and (iv) $|\mathcal{A}_{2\ell+7}(f) - \mathcal{G}_\ell(f)|$ (dotted yellow curve) for $f(t) = 1/t$.

Finally, Figure 5 is analogous to Figure 1 and displays the convergence of Gauss and anti-Gauss rules to the value of the integral $b^T A^{-1} b$ as the number of nodes increases. Here the Gauss rule and anti-Gauss rule can be seen to bracket the exact value.



(a) (b)

Figure 3: Example 4.4: Dense matrix A. (a) 10-logarithm of $\|\widetilde{e}_\ell\|_A$ (solid black curve) and 10-logarithm of estimates of $\|\widetilde{e}_\ell\|_A$ given by (i) $|\mathcal{G}_{\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dashed red curve), (ii) $|\breve{\mathcal{G}}_{\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dotted green curve), (iii) $|\widehat{\mathcal{A}}_{2\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dash-dotted blue curve), and (iv) $|\mathcal{A}_{2\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dotted yellow curve) for $f(t) = 1/t$. (b) 10-logarithm of $\|\widetilde{e}_\ell\|_A$ (solid black curve) and 10-logarithm of estimates of $\|\widetilde{e}_\ell\|_A$ given by (i) $|\mathcal{G}_{\ell+4}(f) - \mathcal{G}_\ell(f)|$ (dashed red curve), (ii) $|\breve{\mathcal{G}}_{\ell+4}(f) - \mathcal{G}_\ell(f)|$ (dotted green curve), (iii) $|\widehat{\mathcal{A}}_{2\ell+7}(f) - \mathcal{G}_\ell(f)|$ (dash-dotted blue curve), and (iv) $|\mathcal{A}_{2\ell+7}(f) - \mathcal{G}_\ell(f)|$ (dotted yellow curve) for $f(t) = 1/t$.



(a) (b)

Figure 4: Example 4.5: (a) 10-logarithm of $\|\widetilde{e}_\ell\|_A$ (solid black curve) and 10-logarithm of the estimate of $\|\widetilde{e}_\ell\|_A$ given by (i) $|\mathcal{G}_{\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dashed red curve), (ii) $|\breve{\mathcal{G}}_{\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dotted green curve), (iii) $|\widehat{\mathcal{A}}_{2\ell+1}(f) - \mathcal{G}_\ell(f)|$ (dash-dotted blue curve), and (iv) $|\mathcal{A}_{2\ell+1}(f) - \mathcal{G}_\ell(f)|$ (point yellow curve) for $f(t) = 1/t$. (b) 10-logarithm of $\|\widetilde{e}_\ell\|_A$ (solid black curve) and 10-logarithm of the estimate of $\|\widetilde{e}_\ell\|_A$ given by (i) $|\mathcal{G}_{\ell+4}(f) - \mathcal{G}_\ell(f)|$ (dashed red curve), (ii) $|\breve{\mathcal{G}}_{\ell+4}(f) - \mathcal{G}_\ell(f)|$ (dotted green curve), (iii) $|\widehat{\mathcal{A}}_{2\ell+7}(f) - \mathcal{G}_\ell(f)|$ (dash-dotted blue curve), and (iv) $|\mathcal{A}_{2\ell+7}(f) - \mathcal{G}_\ell(f)|$ (point yellow curve) for $f(t) = 1/t$.

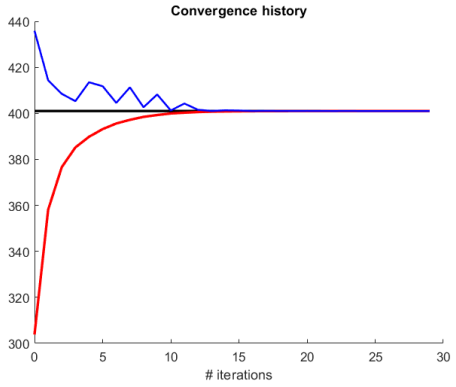


Figure 5: Example 4.5: Graphs for $\mathcal{G}_\ell(f)$ (solid red curve), $\breve{\mathcal{G}}_{\ell+1}(f)$ (dash-dotted blue curve), and the value of the integral $\mathcal{F}(A)$ (straight black line) for $f(t) = 1/t$.

5. Conclusion

This paper describes new approaches to estimate the A-norm of the error in iterates determined by the conjugate gradient method when applied to the solution of linear systems of equations with a symmetric positive definite matrix. The averaged and optimal averaged Gauss rules and shifted variants thereof are found to determine more accurate error estimates than the Gauss and anti-Gauss rules. Since the optimal averaged Gauss rules, or shifted variants thereof, are of higher degree of precision, than the corresponding averaged Gauss rules and their evaluation requires about the same computational effort, we propose the application of the former rules.

Acknowledgments

The authors would like to thank the referees for comments. The research of M. M. Spalević is supported in part by the Serbian Ministry of Science, Technological Development, and Innovations, according to Contract 451-03-65/2024-03/200105 dated on February 5, 2024.

References

[1] P. Díaz de Alba, L. Fermo, and G. Rodriguez, Solution of second kind Fredholm integral equations by means of Gauss and anti-Gauss quadrature rules, Numer. Math., 146 (2020), pp. 699–728.

[2] C. Brezinski, Error estimates for the solution of linear systems, SIAM J. Sci. Comput., 21 (1999), pp. 764–781.

[3] C. Brezinski, G. Rodriguez, and S. Seatzu, Error estimates for linear systems with applications to regularization, Numer. Algorithms, 49 (2008), pp. 85–104.

[4] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari, Computable error bounds and estimates for the conjugate gradient method, Numer. Algorithms, 25 (2000), pp. 75–88.

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

[6] C. W. Clenshaw and A. R. Curtis, A method for numerical integration on an automatic computer, Numer. Math., 2 (1960), pp. 197–205.

[7] D. Lj. Djukić, R. M. Mutavdzić Djukić, L. Reichel, and M. M. Spalević, Weighted averaged Gaussian quadrature rules for modified Chebyshev measures, Appl. Numer. Math., 200 (2024), pp. 195–208.

[8] D. Lj. Djukić, L. Reichel, M. M. Spalević, and J. D. Tomanović, Internality of generalized averaged Gaussian quadratures and their truncations for Bernstein-Szegő weights, Electron. Trans. Numer. Anal., 45 (2016), pp. 405–419.

[9] D. Lj. Djukić, L. Reichel, M. M. Spalević, and J. D. Tomanović, Internality of generalized averaged Gaussian quadratures and truncated variants for modified Chebyshev measures of the second kind, J. Comput. Appl. Math., 345 (2019), pp. 70–85.

[10] N. Eshghi, T. Mach, and L. Reichel, New matrix function approximations and quadrature rules based on the Arnoldi process, J. Comput. Appl. Math., 391 (2021), Art. 113442.

[11] L. Fermo, L. Reichel, G. Rodriguez, and M. M. Spalević, Averaged Nyström interpolants for the solution of Fredholm integral equations of the second kind, Appl. Math. Comput., 467 (2024), Art. 128482.

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

[13] T. Gergelits, B. F. Nielsen, and Z. Strakoš, Numerical approximation of the spectrum of self-adjoint operators in operator preconditioning, Numer. Algorithms, 91 (2022), pp. 301–325.

[14] G. H. Golub and G. Meurant, Matrices, moments and quadrature II: How to compute the norm of the error in iterative methods, BIT Numer. Math., 37 (1997), pp. 687–705.

[15] G. H. Golub and G. Meurant, Matrices, Moments and Quadrature, Princeton University Press, Princeton, 2010.

[16] G. H. Golub and Z. Strakoš, Estimates in quadratic formulas, Numer. Algorithms, 8 (1994), pp. 241–268.

[17] G. H. Golub and C. F. Van Loan, Matrix Computations, 4th ed., Johns Hopkins University Press, Baltimore, 2013.

[18] M. Ladecky, I. Pultarova, and J. Zeman, Guaranteed two-sided bounds on all eigenvalues of preconditioned diffusion and elasticity problems solved by the finite element method, Appl. Math., 66 (2021), pp. 21–42.

[19] D. P. Laurie, Anti-Gaussian quadrature formulas, Math. Comp., 65 (1996), pp. 739–747.

[20] G. Meurant, The computation of bounds for the norm of the error in the conjugate gradient algorithm, Numer. Algorithms, 16 (1997), pp. 77–87.

[21] G. Meurant, Numerical experiments in computing bounds for the norm of the error in the preconditioned conjugate gradient algorithm, Numer. Algorithms, 22 (1999), pp. 353–365.

[22] G, Meurant, J. Papež, and P. Tichý, Accurate error estimation in CG, Numer. Algorithms, 88 (2021), pp. 1337–1359.

[23] G. Meurant and P. Tichý, On computing quadrature-based bounds for the A-norm of the error in conjugate gradients, Numer. Algorithms, 62 (2013), pp. 163–191.

[24] G. Meurant and P. Tichý, Approximating the extreme Ritz values and upper bounds for the A-norm of the error in CG, Numer. Algorithms, 82 (2019), pp. 937–968.

[25] G. Meurant and P. Tichý, Error Norm Estimation in the Conjugate Gradient Algorithm, SIAM, Philadelphia, 2024.

[26] M. Mitrouli and P. Roupa, Vector estimates for $f(A)b$ via extrapolation, Electron. Trans. Numer. Anal., 47 (2017), pp. 179–196.

[27] I. Pultarova and M. Ladecky, Two-sided guaranteed bounds to individual eigenvalues of preconditioned finite element and finite difference problems, Numer. Linear Algebra Appl., 28 (2021), Art. e2382.

[28] L. Reichel, G. Rodriguez, and S. Seatzu, Error estimates for large-scale ill-posed problems, Numer. Algorithms, 51 (2009), pp. 341–361.

[29] L. Reichel and M. M. Spalević, A new representation of generalized averaged Gauss quadrature rules, Appl. Numer. Math., 165 (2021), pp. 614–619.

[30] L. Reichel and M. M. Spalević, Averaged Gauss quadrature formulas: Properties and applications, J. Comput. Appl. Math., 410 (2022), Art. 114232.

[31] L. Reichel, M. M. Spalević, and T. Tang, Generalized averaged Gauss quadrature rules for the approximation of matrix functionals, BIT Numer. Math., 56 (2016), pp. 1045–1067.

[32] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd ed., SIAM, Philadelphia, 2003.

[33] M. M. Spalević, On generalized averaged Gaussian formulas, Math. Comp., 76 (2007), pp. 1483–1492.

[34] Z. Strakoš and P. Tichý, On error estimation in the conjugate gradient method and why it works in finite precision computations, Electron. Trans. Numer. Anal., 13 (2003), pp. 53–80.