

## HARD CASES FOR CONJUGATE GRADIENT METHOD

REN-CANG LI

**Abstract.** The Conjugate Gradient (CG) method is often used to solve a positive definite linear system  $Ax = b$ . This paper analyzes two hard cases for CG or any Krylov subspace type methods by either analytically finding the residual formulas or tightly bound the residuals from above and below, in contrast to existing results which only bound residuals from above. The analysis is based on a general framework to estimate CG and GMRES residuals for certain linear systems with normal  $A$ , and the framework may potentially be useful elsewhere.

**Key Words.** linear system, conjugate gradient method, rate of convergence, MINRES, GMRES

### 1. Introduction

In 1960, Frank [7] tested Richardson’s method on an  $N \times N$  tridiagonal positive definite linear system  $Ax = b$  and experienced very slow convergence. Both Richardson’s method and the conjugate gradient method (CG) seek approximations within so-called Krylov subspaces, and satisfy certain optimal properties at each iteration step. What causes the slow convergence? In part, there is no secret about this because intuitively, it has become part of the folklore that CG would have a hard time to solve linear system  $Ax = b$  for which  $A$  has an eigenvalue distribution similar to that of the zeros of a Chebyshev polynomial (possible after a linear transformation). This is precisely the case for Frank’s example. In addition,  $A$  in Frank’s example has a condition number (the largest eigenvalue over smallest eigenvalue) growing proportional to  $N^2$ . The two combined contribute to extreme difficulty for CG to solve the system. Indeed Golub and O’Leary [9] called it “*the hardest case for conjugate gradients*”. But can we quantify this slow convergence phenomenon?

There is a well-known error bound for CG (see, e.g., [6, 12, 26, 28]): For the  $k$ th CG approximation  $x_k$  and the  $k$ th CG residual  $r_k = b - Ax_k$ ,

$$(1.1) \quad \frac{\|r_k\|_{A^{-1}}}{\|r_0\|_{A^{-1}}} \leq 2 [\Delta_\kappa^k + \Delta_\kappa^{-k}]^{-1}, \quad \Delta_\kappa \stackrel{\text{def}}{=} \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1},$$

where the  $M$ -vector norm  $\|z\|_M \stackrel{\text{def}}{=} \sqrt{z^* M z}$ ,  $\kappa \equiv \kappa(A) = \|A\|_2 \|A^{-1}\|_2$  is  $A$ ’s spectral condition number, generic norm  $\|\cdot\|_2$  is the usual  $\ell_2$  norm of a vector or the spectral norm of a matrix, and the superscript “ $*$ ” takes conjugate transpose. The widely cited Kaniel [14] (1966) gave a proof of (1.1) while saying “This result is known” with a pointer to Meinardus [22] (1963). Frank [7] also presented an error

---

Received by the editors April 28, 2006 and, in revised form, July 14, 2006.

2000 *Mathematics Subject Classification.* 65F10, 15A06.

This research was supported in part by the National Science Foundation CAREER award under Grant No. CCR-9875201 and by the National Science Foundation under Grant No. DMS-0510664.

analysis including error bounds for Richardson’s method, too. These bounds are in terms of  $A$ ’s condition number, and go to zero very slowly as iteration progresses for the example, as they should. Nevertheless they are upper bounds and, despite the folklore, it would be still of interest to be able to compute exactly or bound tightly the actual residuals from above and below.

In this paper, we first establish a general framework that can be used to compute exactly and/or bound tightly CG and GMRES (the generalized minimal residual method) residuals for certain linear systems with normal  $A$ . The key technical detail is very much along the lines of [15, 18], where examples are constructed to show CG error bounds by (1.1) and those of Kaniel-Saad [14, 24] for symmetric Lanczos method are sharp<sup>1</sup> in general, modulo a modest constant factor. It is worth noting that superlinear convergence, much faster than what the existing error bounds suggest, is often observed for both methods [3, 4, 27].

We then apply the framework to analyze the behaviors of CG on Frank’s example and a model problem (see, e.g., [6, p.267] and [5]). Exact residual expressions and tight lower and upper bounds are obtained. They show both problems are indeed difficult, much as expected. The model problem has been widely studied already, mostly as a teaching example to illustrate behaviors of various iterative methods for linear systems. Naiman, Babuška, and Elman [23] computed explicitly CG residuals for 3 special right-hand sides, while [4] Beckermann and Kuijlaars investigated the superlinear convergence behavior of CG when the right-hand sides are from discretizing certain very smooth functions. Our treatment to the model problem appears to be less complicated, with the help of the framework.

The rest of this paper is organized as follows. Section 2 establishes a general framework to analyze the convergence of CG and GMRES on certain linear systems. The framework is then applied to Frank’s example [7] in Section 3 and to the model problem [6] in Section 4. In both examples, exact residuals or tight bounds are obtained. Finally in Section 5 we present our concluding remarks.

**Notation.** Throughout this paper,  $\mathbb{K}^{n \times m}$  is the set of all  $n \times m$  matrices with entries in  $\mathbb{K}$ , where  $\mathbb{K}$  is  $\mathbb{C}$  (the set of complex numbers) or  $\mathbb{R}$  (the set of real numbers),  $\mathbb{K}^n = \mathbb{K}^{n \times 1}$ , and  $\mathbb{K} = \mathbb{K}^1$ .  $I_n$  (or simply  $I$  if its dimension is clear from the context) is the  $n \times n$  identity matrix, and  $e_j$  is its  $j$ th column.

The inequality  $X \leq Y$  is for two Hermitian matrices, meaning that  $Y - X$  is positive semidefinite. The superscript “ $\cdot^*$ ” takes conjugate transpose while “ $\cdot^T$ ” takes transpose only.

We shall also adopt MATLAB-like convention to access the entries of vectors and matrices. The set of integers from  $i$  to  $j$  inclusive is  $i : j$ . For vector  $u$  and matrix  $X$ ,  $u_{(j)}$  is  $u$ ’s  $j$ th entry,  $X_{(i,j)}$  is  $X$ ’s  $(i, j)$ th entry,  $\text{diag}(u)$  is the diagonal matrix with  $(\text{diag}(u))_{(j,j)} = u_{(j)}$ ;  $X$ ’s submatrices  $X_{(k:\ell,i:j)}$ ,  $X_{(k:\ell,:)}$ , and  $X_{(:,i:j)}$  consists of intersections of row  $k$  to row  $\ell$  and column  $i$  to column  $j$ , row  $k$  to row  $\ell$ , and column  $i$  to column  $j$ , respectively.

---

<sup>1</sup>It is well-known that CG error bounds by (1.1) are sharp (see, e.g., [2, 11, 14, 22]), but it is only claimed *locally* in the sense that for each iteration step  $k$  there is a linear system  $Ax = b$  (depending on  $k$  of course) on which the  $k$ th CG residual attains the bound. It turns out that for such  $Ax = b$ , CG computes the exact solution in the very next iteration [14]! The contribution in [15, 18] is that there is an  $Ax = b$  on which the CG residuals are comparable to the bounds by (1.1) for all iteration steps.

## 2. A framework to estimate CG and GMRES residuals

The  $k$ th Krylov subspace  $\mathcal{K}_k \equiv \mathcal{K}_k(A, b)$  of  $A$  on  $b$  is defined as

$$(2.1) \quad \mathcal{K}_k \equiv \mathcal{K}_k(A, b) \stackrel{\text{def}}{=} \text{span}\{b, Ab, \dots, A^{k-1}b\}.$$

We are interested in the “best” approximate solution in  $\mathcal{K}_k$  to  $Ax = b$ . For CG,  $A$  is positive definite, and the computed solution achieves [6, p.306]

$$(2.2) \quad \min_{y \in \mathcal{K}_k} \frac{\|A^{-1}b - y\|_A}{\|A^{-1}b\|_A} \equiv \min_{y \in \mathcal{K}_k} \frac{\|b - Ay\|_{A^{-1}}}{\|r_0\|_{A^{-1}}}.$$

where, without loss of generality, we set initially  $x_0 = 0$  and thus  $r_0 = b - Ax_0 = b$ . For GMRES,  $A$  may be non-Hermitian and the computed solution achieves [25]

$$(2.3) \quad \min_{y \in \mathcal{K}_k} \frac{\|b - Ay\|_2}{\|r_0\|_2}.$$

Historically, when  $A$  is Hermitian, GMRES is called MINRES (the minimal residual method). In what follows, we consider normal matrix  $A$  only. Let the eigen-decomposition of  $A$  be

$$(2.4) \quad A = Q\Lambda Q^*, \quad Q^*Q = I_N, \quad \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N),$$

where all  $\lambda_j > 0$  for CG and some  $\lambda_j$  may be complex for GMRES. It has been shown that both (2.2) and (2.3) can be restated as [13, 15, 20]

$$(2.5) \quad \min_{|u_{(1)}|=1} \frac{\|\text{diag}(g) V_{k+1,N}^T u\|_2}{\|g\|_2},$$

where  $g = \Lambda^{-1/2}Q^*b$  for CG, and  $g = Q^*b$  for GMRES, and

$$(2.6) \quad V_{k+1,N} \stackrel{\text{def}}{=} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^k & \lambda_2^k & \cdots & \lambda_n^k \end{pmatrix}$$

is the  $(k+1)$ -by- $N$  rectangular Vandermonde matrix with nodes  $\{\lambda_j\}_{j=1}^N$ . As a by-product, inequality (1.1) can now be restated in terms of  $V_{k+1,N}$ : Assume all  $\lambda_j > 0$  and set  $\kappa = \max_j \lambda_j / \min_j \lambda_j$ . Then for  $1 \leq k < N$

$$(2.7) \quad \min_{|u_{(1)}|=1} \frac{\|\text{diag}(g) V_{k+1,N}^T u\|_2}{\|g\|_2} \leq 2 [\Delta_\kappa^k + \Delta_\kappa^{-k}]^{-1}.$$

The quantity (2.5) points to a new direction to analyze the convergence behavior of CG or GMRES on certain linear systems. Frank’s example and the model problem are two of them.

Now we are ready to explain our general framework for solving (2.5). The key idea is similar to that in [15], i.e., seeking a decomposition, reminiscent of QR, for  $V_{k+1,N}^T$  but what presented here is in a more general term. The framework is built upon

- a family of polynomials  $p_m(t)$  of degree  $m$  in  $t$  for  $m = 0, 1, 2, \dots$ ,
- two parameters  $\omega \neq 0$  and  $\tau$  (complex or real), and
- $\Omega_N = \text{diag}(\omega_1, \omega_2, \dots, \omega_N)$  with all  $\omega_j > 0$ .

(2.8)

For the family of polynomials, we often pick Chebyshev polynomials of the 1st kind:

$$(2.9) \quad T_m(t) = \cos(m \arccos t) \quad \text{for } |t| \leq 1,$$

$$(2.10) \quad = \frac{1}{2} \left( t + \sqrt{t^2 - 1} \right)^m + \frac{1}{2} \left( t - \sqrt{t^2 - 1} \right)^m \quad \text{for } |t| \geq 1,$$

or, Chebyshev polynomials of the second kind:

$$(2.11) \quad U_m(t) = \frac{\sin((m+1) \arccos t)}{\sin(\arccos t)} \quad \text{for } |t| \leq 1,$$

$$(2.12) \quad = \frac{(t + \sqrt{t^2 - 1})^{m+1} - (t - \sqrt{t^2 - 1})^{m+1}}{2\sqrt{t^2 - 1}}, \quad \text{for } |t| \geq 1,$$

or, more generally any orthogonal polynomials; usually  $\Omega_N = I_N$  or take  $\{\omega_j\}_{j=1}^N$  to be weights in the associated Gaussian quadrature rule if  $p_m$  is an orthogonal polynomial; the selections of  $\omega \neq 0$  and  $\tau$  can be tricky and are obviously problem-dependent.

First we define the  $m$ th *Translated Polynomial* in  $z$  of degree  $m$  as

$$(2.13) \quad p_m(z; \omega, \tau) \stackrel{\text{def}}{=} p_m(z/\omega + \tau)$$

$$(2.14) \quad = a_{mm}z^m + a_{m-1m}z^{m-1} + \cdots + a_{1m}z + a_{0m},$$

where  $a_{jm} \equiv a_{jm}(\omega, \tau)$  are functions of  $\omega$  and  $\tau$ , and then define upper triangular  $R_m \in \mathbb{C}^{m \times m}$ , a matrix-valued function in  $\omega$  and  $\tau$ , too, as

$$(2.15) \quad R_m \equiv R_m(\omega, \tau) \stackrel{\text{def}}{=} \begin{pmatrix} a_{00} & a_{01} & a_{02} & \cdots & a_{0\ m-1} \\ & a_{11} & a_{12} & \cdots & a_{1\ m-1} \\ & & a_{22} & \cdots & a_{2\ m-1} \\ & & & \ddots & \vdots \\ & & & & a_{m-1\ m-1} \end{pmatrix},$$

i.e., the  $j$ th column consists of the coefficients of  $p_{j-1}(z; \omega, \tau)$ . Denote  $V_N = V_{N,N}$  and set

$$(2.16) \quad \mathbf{P}_N \stackrel{\text{def}}{=} \begin{pmatrix} p_0(t_1) & p_0(t_2) & \cdots & p_0(t_N) \\ p_1(t_1) & p_1(t_2) & \cdots & p_1(t_N) \\ \vdots & \vdots & & \vdots \\ p_{N-1}(t_1) & p_{N-1}(t_2) & \cdots & p_{N-1}(t_N) \end{pmatrix},$$

where

$$(2.17) \quad t_j = \lambda_j/\omega + \tau \quad \text{for } 1 \leq j \leq N.$$

Then

$$(2.18) \quad V_N^T R_N = \mathbf{P}_N^T,$$

Equation (2.18) yields  $V_N^T = \mathbf{P}_N^T R_N^{-1}$ . Extracting the first  $k+1$  columns from both sides of  $V_N^T = \mathbf{P}_N^T R_N^{-1}$  yields

$$(2.19) \quad V_{k+1,N}^T = \mathbf{P}_{k+1,N}^T R_{k+1}^{-1},$$

where  $\mathbf{P}_{k+1,N} = \mathbf{P}_{(1:k+1, \cdot)}$ . Next set, with matrix  $\Omega_N$  in (2.8),

$$(2.20) \quad \Upsilon \stackrel{\text{def}}{=} \mathbf{P}_N \Omega_N \mathbf{P}_N^*.$$

A key step that makes our framework work is the following assumption.

$$(2.21) \quad \boxed{\begin{array}{l} \text{there are } 0 < \gamma_{k+1} \leq \Gamma_{k+1} \text{ and } D_{k+1} = \text{diag}(d_1, d_2, \dots, d_{k+1}) \\ \text{with all } d_j > 0 \text{ such that} \\ \gamma_{k+1} D_{k+1}^{-1} \leq \Upsilon_{(1:k+1, 1:k+1)} \leq \Gamma_{k+1} D_{k+1}^{-1}. \end{array}}$$

Such  $\gamma_{k+1}$ ,  $\Gamma_{k+1}$ , and  $D_{k+1}$  always exist if  $\Upsilon_{(1:k+1, 1:k+1)}$  is nonsingular. For example,  $D_{k+1} = I_{k+1}$ ,  $\gamma_{k+1}$  and  $\Gamma_{k+1}$  are the smallest and largest eigenvalues of  $\Upsilon_{(1:k+1, 1:k+1)}$ .

In its present general form, the next lemma was proved in [15, 18]. It was also implied by the proof of [13, Theorem 2.1]. See also [19].

**Lemma 2.1.** *If  $W$  has full column rank, then*

$$(2.22) \quad \min_{|u_{(1)}|=1} \|Wu\|_2 = [e_1^T (W^*W)^{-1} e_1]^{-1/2}.$$

*In particular if  $W$  is nonsingular,  $\min_{|u_{(1)}|=1} \|Wu\|_2 = \|W^{-*}e_1\|_2^{-1}$ .*

Assuming that  $\Omega_N^{1/2} V_{k+1, N}^T$  has full column rank, by Lemma 2.1 we have

$$(2.23) \quad \min_{|u_{(1)}|=1} \|\Omega_N^{1/2} V_{k+1, N}^T u\|_2 = [e_1^T (\bar{V}_{k+1, N} \Omega_N V_{k+1, N}^T)^{-1} e_1]^{-1/2},$$

where  $\bar{V}_{k+1, N}$  is the complex conjugate of  $V_{k+1, N}$ . By (2.19), we have

$$(2.24) \quad \begin{aligned} \bar{V}_{k+1, N} \Omega_N V_{k+1, N}^T &= R_{k+1}^{-*} \left[ \mathbf{P}_{k+1, N}^T \right]^* \Omega_N \mathbf{P}_{k+1, N}^T R_{k+1}^{-1} \\ &= R_{k+1}^{-*} (\bar{\mathbf{P}}_N \Omega_N \mathbf{P}_N^T)_{(1:k+1, 1:k+1)} R_{k+1}^{-1} \\ &= R_{k+1}^{-*} (\bar{\Upsilon})_{(1:k+1, 1:k+1)} R_{k+1}^{-1}, \end{aligned}$$

$$(2.25) \quad (\bar{V}_{k+1, N} \Omega_N V_{k+1, N}^T)^{-1} = R_{k+1} \left[ (\bar{\Upsilon})_{(1:k+1, 1:k+1)} \right]^{-1} R_{k+1}^*.$$

Consequently, by (2.21),

$$\frac{1}{\Gamma_{k+1}} e_1^T R_{k+1} D_{k+1} R_{k+1}^* e_1 \leq e_1^T (\bar{V}_{k+1, N} \Omega_N V_{k+1, N}^T)^{-1} e_1 \leq \frac{1}{\gamma_{k+1}} e_1^T R_{k+1} D_{k+1} R_{k+1}^* e_1$$

which implies

$$(2.26) \quad \frac{1}{\sqrt{\sum_{j=1}^N \omega_j}} \sqrt{\frac{\gamma_{k+1}}{\Phi_{k+1}}} \leq \min_{|u_{(1)}|=1} \frac{\|\Omega_N^{1/2} V_{k+1, N}^T u\|_2}{\sqrt{\sum_{j=1}^N \omega_j}} \leq \frac{1}{\sqrt{\sum_{j=1}^N \omega_j}} \sqrt{\frac{\Gamma_{k+1}}{\Phi_{k+1}}},$$

where

$$(2.27) \quad \Phi_{k+1} = e_1^T R_{k+1} D_{k+1} R_{k+1}^* e_1 = \sum_{j=0}^k d_j |a_{0j}|^2 = \sum_{j=0}^k d_j |p_j(\tau)|^2.$$

Occasionally it may also be possible to calculate  $\min_{|u_{(1)}|=1} \|\Omega_N^{1/2} V_{k+1, N}^T u\|_2$  exactly as it happens to the two examples we will be dealing with.

We summarize our findings into two mathematically equivalent theorems, in terms of  $V_{k+1, N}$  or the residual for  $Ax = b$ , respectively.

**Theorem 2.1.** *Given  $V_{k+1, N}$  defined as in (2.6), and given, as in (2.8), the polynomial family  $\{p_m\}_{m=0}^\infty$  and two prescribed numbers  $\omega \neq 0$  and  $\tau$  and  $\Omega_N$ . Let  $t_j$*

be as in (2.17),  $\mathbf{P}_N$  as in (2.16), and  $\Upsilon$  as in (2.20). Then for  $1 \leq k < N$

$$(2.28) \quad \min_{|u_{(1)}|=1} \frac{\|\Omega_N^{1/2} V_{k+1,N}^T u\|_2}{\sqrt{\sum_{j=1}^N \omega_j}} = \frac{1}{\sqrt{\sum_{j=1}^N \omega_j}} \frac{1}{\sqrt{y^* [\Upsilon_{(1:k+1,1:k+1)}]^{-1} y}},$$

where  $y = R_{k+1}^* e_1 = (p_0(\tau), p_1(\tau), \dots, p_k(\tau))^*$ . If (2.21) holds, then we have (2.26).

**REMARK 2.1.** Lower and upper bounds on (2.5) for any other  $g \in \mathbb{C}^N$  are immediately available once (2.5) with  $|g_{(j)}| = \omega_j^{1/2}$  ( $1 \leq j \leq N$ ) is known or bounded, since

$$\|\text{diag}(g) V_{k+1,N}^T u\|_2 = \|\text{diag}(g) \Omega_N^{-1/2} \Omega_N^{1/2} V_{k+1,N}^T u\|_2$$

which yields

$$(2.29) \quad \left( \min_j \frac{|g_{(j)}|}{\sqrt{\omega_j}} \right) \|\Omega_N^{1/2} V_{k+1,N}^T u\|_2 \leq \|\text{diag}(g) V_{k+1,N}^T u\|_2 \leq \left( \max_j \frac{|g_{(j)}|}{\sqrt{\omega_j}} \right) \|\Omega_N^{1/2} V_{k+1,N}^T u\|_2.$$

For this reason, in the rest of this paper, we shall simply consider (2.5) for  $g$  with  $|g_{(j)}| = \omega_j^{1/2}$ .

**Theorem 2.2.** Let  $A \in \mathbb{C}^{N \times N}$  be normal and have eigen-decomposition (2.4),  $b \in \mathbb{C}^N$ , and suppose (2.8). Define  $t_j$  as in (2.17),  $\mathbf{P}_N$  as in (2.16), and  $\Upsilon$  as in (2.20). For normal  $A$ ,  $q = 2$  and  $g = Q^* b$ , and for positive definite  $A$ , either  $q = 2$  and  $g = Q^* b$  or  $q = A^{-1}$  and  $g = \Lambda^{-1/2} Q^* b$ . If  $|g_{(j)}| = \omega_j^{1/2}$  for  $1 \leq j \leq N$ , then for  $1 \leq k < N$

$$(2.30) \quad \min_{y \in \mathcal{K}_k} \frac{\|b - Ay\|_q}{\|r_0\|_q} = \frac{1}{\sqrt{\sum_{j=1}^N \omega_j}} \frac{1}{\sqrt{y^* [\Upsilon_{(1:k+1,1:k+1)}]^{-1} y}},$$

where  $y = R_{k+1}^* e_1 = (p_0(\tau), p_1(\tau), \dots, p_k(\tau))^*$ . If (2.21) holds, then

$$(2.31) \quad \left( \min_j \frac{|g_{(j)}|}{\sqrt{\omega_j}} \right) \frac{1}{\|g\|_2} \sqrt{\frac{\gamma_{k+1}}{\Phi_{k+1}}} \leq \min_{y \in \mathcal{K}_k} \frac{\|b - Ay\|_q}{\|r_0\|_q} \leq \left( \max_j \frac{|g_{(j)}|}{\sqrt{\omega_j}} \right) \frac{1}{\|g\|_2} \sqrt{\frac{\Gamma_{k+1}}{\Phi_{k+1}}}.$$

**REMARK 2.2.** When  $t_j$  are the zeros of the  $N$ th polynomial from an orthogonal polynomial family, one should take  $\omega_j$  to be the weights of the associated Gaussian quadrature rule, and then  $\mathbf{P}_N \Omega_N \mathbf{P}_N^*$  is diagonal since the Gaussian quadrature rule is exact for polynomial of degree no higher than  $2N - 1$ . Then  $\min_{|u_{(1)}|=1} \|\Omega_N^{1/2} V_{k+1,N}^T u\|_2$  has a close formula solution.

**REMARK 2.3.** Besides yielding bounds on residuals for  $Ax = b$ , (2.25) may have other independent interests. For example, it relates  $\text{trace}(\bar{V}_{k+1,N} \Omega_N V_{k+1,N}^T)^{-1}$  to the coefficients of  $p_j(x; \omega, \tau)$  by

$$(2.32) \quad \Gamma_{k+1}^{-1} \|R_{k+1}\|_{\mathbb{F}}^2 \leq \text{trace}((\bar{V}_{k+1,N} \Omega_N V_{k+1,N}^T)^{-1}) \leq \gamma_{k+1}^{-1} \|R_{k+1}\|_{\mathbb{F}}^2.$$

It is natural to define the condition number in  $\Omega_N$ -Frobenius norm of  $V_{k+1,N}$  as

$$\kappa_{\mathbb{F}, \Omega_N}(V_{k+1,N}) \stackrel{\text{def}}{=} \|\Omega_N^{1/2} V_{k+1,N}\|_{\mathbb{F}} \sqrt{\text{trace}((\bar{V}_{k+1,N} \Omega_N V_{k+1,N}^T)^{-1})}.$$

$\|\Omega_N^{1/2} V_{k+1,N}\|_{\mathbb{F}}$  can be easily computed (estimated). If  $\|R_{k+1}\|_{\mathbb{F}}$  can be estimated, too, bounds on  $\kappa_{\mathbb{F}, \Omega_N}(V_{k+1,N})$  can be obtained. Indeed for  $p_j = T_j$ , if  $\tau$  is real and  $\tau = 0$  or  $|\tau| \geq 1$ , tight bounds on  $\|R_{k+1}\|_{\mathbb{F}}$  are indeed available [17].

### 3. Frank's example

In Frank's example [7],

$$(3.1) \quad A = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{N \times N}.$$

$A$  admits an analytical eigen-decomposition (2.4). Frank [7] credited this close expression

$$(3.2) \quad \theta_j = \frac{2j-1}{2N+1}\pi, \quad \lambda_j = 2 - 2\cos\theta_j \quad \text{for } 1 \leq j \leq N$$

to Young [29] who, presumably, found closed formulas for the associated eigenvectors, too. But we do not have access to Young [29]. In any case, it can be computed that the entries of the  $j$ th eigenvector are

$$Q_{(i,j)} = \frac{2}{\sqrt{2N+1}} \cos\left(\pi \frac{(2i-1)(2j-1)}{2(2N+1)}\right) \quad \text{for } 1 \leq i \leq N.$$

For this example, we shall take  $p_m = T_m$ ,  $\omega = -2$ ,  $\tau = 1$ , and  $\Omega_N = I_N$  in the framework in Section 2.

**Lemma 3.1.** For  $\theta_j = \frac{2j-1}{2N+1}\pi$  and integer  $\ell$ ,

$$(3.3) \quad \sum_{k=1}^N \cos \ell \theta_k = \begin{cases} (-1)^m N, & \text{if } \ell = m(2N+1) \text{ for some integer } m, \\ 1/2, & \text{if } \ell \text{ is odd, but } \ell \neq m(2N+1) \text{ for any integer } m, \\ -1/2, & \text{if } \ell \text{ is even, but } \ell \neq m(2N+1) \text{ for any integer } m. \end{cases}$$

*Proof.* If  $\ell = m(2N+1)$  for some integer  $m$ , then  $\ell \theta_k = m(2k-1)\pi$  and thus  $\cos \ell \theta_k = (-1)^m$ . Assume that  $\ell \neq m(2N+1)$  for any integer  $m$ . Set  $\phi = \ell\pi/(2N+1)$ . We have [10, p.30]

$$\sum_{k=1}^N \cos \ell \theta_k = \sum_{k=1}^N \cos(2k-1)\phi = \frac{1 \sin 2N\phi}{2 \sin \phi}.$$

Now notice  $\sin 2N\phi = \sin(\ell\pi - \phi) = -(-1)^\ell \sin \phi$  to conclude the proof.  $\square$

**Lemma 3.2.** Set  $p_m = T_m$ ,  $\omega = -2$ ,  $\tau = 1$ , and  $\Omega_N = I_N$  in (2.8). Let  $\Upsilon = \mathbf{P}_N \Omega_N \mathbf{P}_N^T$  be defined as in (2.20), and let  $w \in \mathbb{R}^n$  whose odd entries are 1 and even entries are  $-1$ . Then

$$(3.4) \quad \Upsilon = \Theta - \frac{1}{2}ww^T, \quad \Theta = \frac{2N+1}{4} \text{diag}(2, 1, 1, \dots).$$

Allow  $\Theta$  and  $w$  to have generic sizes but otherwise the same. We still have

$$(3.5) \quad \Upsilon_{(1:k+1, 1:k+1)} = \Theta - \frac{1}{2}ww^T = \Theta^{1/2} \left( I_{k+1} - \frac{1}{2}\Theta^{-1/2}ww^T\Theta^{-1/2} \right) \Theta^{1/2},$$

and thus

$$(3.6) \quad \frac{2(N-k)}{2N+1} \Theta \leq \Upsilon_{(1:k+1, 1:k+1)} \leq \Theta.$$

*Proof.* For  $0 \leq i, j \leq N-1$ ,

$$\begin{aligned}
(\mathbf{P}_N \mathbf{P}_N^T)_{(i+1, j+1)} &= \sum_{k=1}^N (\mathbf{P}_N)_{(i+1, k)} (\mathbf{P}_N^T)_{(k, j+1)} \\
&= \sum_{k=1}^N T_i(t_k) T_j(t_k) \\
&= \sum_{k=1}^N \cos i\theta_k \cos j\theta_k \\
(3.7) \qquad \qquad \qquad &= \frac{1}{2} \sum_{k=1}^N \cos(i+j)\theta_k + \frac{1}{2} \sum_{k=1}^N \cos(i-j)\theta_k.
\end{aligned}$$

Consider  $\ell = i \pm j$  in Lemma 3.1, and  $0 \leq i, j \leq N-1$ . Since  $0 \leq i+j \leq 2N-2$  and  $-N+1 \leq i-j \leq N-1$ , for some integer  $m$

$$(3.8) \qquad i+j = m(2N+1) \Leftrightarrow i=j=0;$$

$$(3.9) \qquad i-j = m(2N+1) \Leftrightarrow i=j.$$

It can then be verified that

$$(3.10) \qquad \Upsilon_{(i+1, j+1)} = \begin{cases} N, & \text{for } i=j=0, \\ \frac{2N-1}{4}, & \text{for } 0 \neq i=j, \\ -1/2, & \text{for } i \neq j, \text{ both odd or even,} \\ 1/2, & \text{for } i \neq j, \text{ one odd and one even.} \end{cases}$$

This proves (3.4) and (3.5). Next we notice

$$\|\Theta^{-1/2} w\|_2^2 = \frac{2}{2N+1} + k \times \frac{4}{2N+1} = \frac{2(2k+1)}{2N+1}$$

to get

$$I_{k+1} - \frac{1}{2} \Theta^{-1/2} w w^T \Theta^{-1/2} \geq \left(1 - \frac{1}{2} \|\Theta^{-1/2} w\|_2^2\right) I_{k+1} = \frac{2(N-k)}{2N+1} I_{k+1},$$

and thus  $\Upsilon_{(1:k+1, 1:k+1)} \geq \frac{2(N-k)}{2N+1} \Theta$ . On the other hand,  $\Upsilon_{(1:k+1, 1:k+1)} \leq \Theta$ .  $\square$

**Theorem 3.1.** For  $V_{k+1, N}$  with the eigenvalues (3.2) of Frank matrix (3.1) as its nodes,

$$(3.11) \qquad \frac{1}{\sqrt{N}} \sqrt{\frac{N-k}{2k+1}} \leq \min_{|u_{(1)}|=1} \frac{\|V_{k+1, N}^T u\|_2}{\sqrt{N}} \leq \frac{1}{\sqrt{N}} \sqrt{\frac{2N+1}{2(2k+1)}},$$

$$(3.12) \qquad \min_{|u_{(1)}|=1} \frac{\|V_{k+1, N}^T u\|_2}{\sqrt{N}} = \frac{1}{\sqrt{N}} \sqrt{\frac{(2N+1)(N-k)}{2(2k+1)(N-k)+1}}.$$

The right-hand side of (3.12) is very close to the upper bound in (3.11).

*Proof.* Set  $p_m = T_m$ ,  $\omega = -2$ ,  $\tau = 1$ , and  $\Omega_N = I_N$  in (2.8). Lemma 3.2 says that (2.21) holds with  $\gamma_{k+1} = \frac{2(N-k)}{2N+1}$ ,  $\Gamma_{k+1} = 1$ , and  $D_{k+1} = \Theta^{-1}$  which, substituted into (2.26) and (2.27), yield (3.11), since

$$\Phi_{k+1} = \sum_{j=0}^k d_j |T_j(1)|^2 = \frac{2(2k+1)}{2N+1}.$$

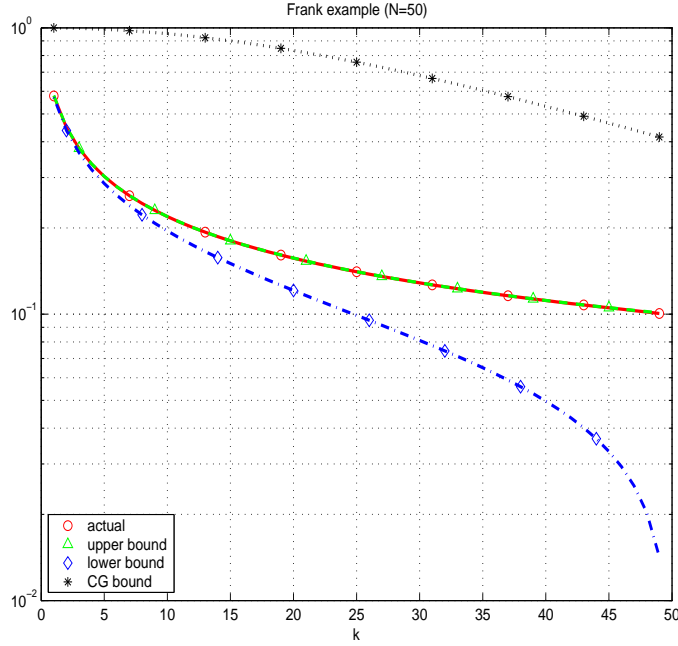


FIGURE 3.1.  $\min_{|u_{(1)}|=1} \frac{\|V_{k+1,N}^T u\|_2}{\sqrt{N}}$  vs. its lower and upper bounds for Frank's example: upper bounds nearly the same as actuals.

To prove (3.12), we have, by Sherman-Morrison formula [6, p.95],

$$\begin{aligned} [\Upsilon_{(1:k+1,1:k+1)}]^{-1} &= \Theta^{-1} + \frac{1}{2} \Theta^{-1} w w^T \Theta^{-1} \frac{1}{1 - w^T \Theta^{-1} w / 2} \\ &= \frac{4}{2N+1} \left( \text{diag}(2^{-1}, 1, 1, \dots) + \frac{1}{N-k} z z^T \right), \end{aligned}$$

where  $z = (2^{-1}, -1, 1, -1, 1, \dots)^T$ . Let  $y = R_{k+1}^T e_1 = (1, 1, \dots, 1)^T$ , and then  $y^T z = (-1)^k / 2$ . By (2.25), we have

$$\begin{aligned} e_1^T (V_{k+1,N} V_{k+1,N}^T)^{-1} e_1 &= y^T [\Upsilon_{(1:k+1,1:k+1)}]^{-1} y \\ &= \frac{4}{2N+1} \left( \frac{1}{2} + k + \frac{1}{N-k} \frac{1}{4} \right) \\ &= \frac{2(2k+1)(N-k) + 1}{(2N+1)(N-k)}, \end{aligned}$$

and hence (3.12) because of (2.23).  $\square$

(3.11) and (3.12) indicate slow convergence of any method that seeks approximations from  $\mathcal{K}_k(A, b)$  (if  $g = Q^*b$  or  $\Lambda^{-1/2}Q^*b$  does not have extremely tiny entries). Figure 3.1 plots the exact value in (3.12), its lower and upper bounds by (3.11), and the upper bound by (2.7). Notice that the relative residuals can only be reduced by as much as a factor of about  $10^{-1}$  at  $k = 49$  in Figure 3.1 (and of course the exact solution at  $k = 50$  since then the search space is the entire  $\mathbb{C}^N$  – a situation that is not very interesting practically).

We commented at the beginning of this paper that there are actually two reasons contributing to the slow convergence by any method on this  $A$  that seeking

approximations from  $\mathcal{K}_k(A, b)$  – the eigenvalue distribution and the condition number. Let us now examine the two. The zeros of translated Chebyshev polynomial  $T_N(x/\omega + \tau) = T_N(1 - x/2)$  are  $\{2 - 2 \cos \frac{2j-1}{2N} \pi\}_{j=1}^N$  which distribute on  $(0, 4)$  in much the same way as the eigenvalues  $\left\{2 - 2 \cos \frac{2j-1}{2N+1} \pi\right\}_{j=1}^N$  here in the example for large  $N$ . At the same time, the condition number  $\kappa(A) = \lambda_N/\lambda_1 \approx 16N^2/\pi^2$  rapidly increases proportionally to  $N^2$ .

#### 4. The model problem

Discretizing the following one-dimensional Poisson's equation

$$\begin{aligned} -\frac{d^2v(x)}{dx^2} &= f(x), \quad \text{for } 0 < x < 1, \\ v(0) &= v(1) = 0, \end{aligned}$$

leads to a linear system  $Ax = b$ , so-called *the model problem* [6, p.267], with

$$(4.1) \quad A = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix} \in \mathbb{R}^{N \times N}.$$

$A$  admits an analytical eigen-decomposition (2.4), too. In fact [6, 23],

$$(4.2) \quad \theta_j = \frac{j}{N+1} \pi, \quad \lambda_j = 2 - 2 \cos \theta_j \quad \text{for } 1 \leq j \leq N$$

and the corresponding  $j$ th eigenvector is

$$Q_{(i,j)} = \frac{2}{\sqrt{N+1}} \sin\left(\pi \frac{ij}{N+1}\right) \quad \text{for } 1 \leq i \leq N.$$

For this example, we shall take, in Section 2, either  $p_m = T_m$  or  $U_m$ ,  $\omega = -2$ ,  $\tau = 1$ , and either  $\Omega_N = I_N$  or  $\Omega_N = \text{diag}(\sin \theta_1, \dots, \sin \theta_N)$ .

In [23], three close expressions for CG residuals for three special right-hand sides were explicitly computed through clever, albeit complicated, trigonometric function manipulations. In our notation, equivalently, the computations there are for minimization problem (2.5) with

$$(4.3) \quad g_{(j)} = 2 \sin^{1-\beta} \frac{\theta_j}{2}, \quad 1 \leq j \leq N, \quad \text{for } \beta = 0, 1, 2.$$

In this section, we shall apply our general framework to solve (2.5) for<sup>2</sup>  $|g_{(j)}| = 2$  corresponding to  $\beta = 1$  [23]. This will provide a different approach that appears to be less complicated.

Recently Liesen and Tichý [21] studied a slightly more general case:  $A$  is the same as in (4.1), except for its diagonal entries which are replaced by  $2(1+\delta)$ , where  $\delta \geq 0$ . They were able to establish bounds mostly for  $k = N - 1$ , the next-to-last iteration. Our technique below works equally well for this more general case for all iteration steps. Detail will be published elsewhere.

Recall Remark 2.2 for  $t_j$  being the zeros of an orthogonal polynomial. It turns out  $t_j$  here are the zeros of  $U_N(t)$  [1, p.889]. This enables us to produce close

<sup>2</sup>Since the constant factor 2 can and will be cancelled, we actually consider equivalently the case  $|g_{(j)}| = 1$ .

formulas for CG residuals for a particular  $g$ , other than those in (4.3) considered by [23], in a much simpler way. We shall do so in Theorem 4.2.

**Lemma 4.1.** For  $\theta_j = \frac{j}{N+1}\pi$  and integer  $\ell$ ,

$$(4.4) \quad \sum_{k=1}^N \cos \ell \theta_k = \begin{cases} N, & \text{if } \ell = 2m(N+1) \text{ for some integer } m, \\ 0, & \text{if } \ell \text{ is odd,} \\ -1, & \text{if } \ell \text{ is even, but } \ell \neq 2m(N+1) \text{ for any integer } m. \end{cases}$$

*Proof.* If  $\ell = 2m(N+1)$  for some integer  $m$ , then  $\ell \theta_k = 2mk\pi$  and thus  $\cos \ell \theta_k = 1$ . Assume that  $\ell \neq 2m(N+1)$  for any integer  $m$ . Set  $\phi = \ell\pi/(N+1)$ . We have [10, p.30]

$$\sum_{k=1}^N \cos \ell \theta_k = \sum_{k=1}^N \cos k\phi = \cos \frac{N+1}{2}\phi \times \frac{\sin \frac{N\phi}{2}}{\sin \frac{\phi}{2}}.$$

Now notice  $\cos \frac{N+1}{2}\phi = \cos \frac{\ell}{2}\pi = 0$  for odd  $\ell$  and  $(-1)^{\ell/2}$  for even  $\ell$ , and  $\sin \frac{N\phi}{2} = \sin(\frac{\ell}{2}\pi - \phi) = -(-1)^{\ell/2} \sin \phi$  for even  $\ell$  to conclude the proof.  $\square$

**Lemma 4.2.** Set  $p_m = T_m$ ,  $\omega = -2$ ,  $\tau = 1$ , and  $\Omega_N = I_N$  in (2.8). Let  $\Upsilon = \mathbf{P}_N \Omega_N \mathbf{P}_N^T$  be defined as in (2.20), and let  $e_{\text{odd}}$  and  $e_{\text{even}}$  be two column vectors defined as: all the odd entries of  $e_{\text{odd}}$  are ones and the even entries are zeros, and all the odd entries of  $e_{\text{even}}$  are zeros and the even entries are ones. Then

$$(4.5) \quad \Upsilon = \Theta - e_{\text{odd}} e_{\text{odd}}^T - e_{\text{even}} e_{\text{even}}^T, \quad \Theta = \frac{N+1}{2} \text{diag}(2, 1, 1, \dots).$$

Allow  $\Theta$ ,  $e_{\text{odd}}$ , and  $e_{\text{even}}$  to have generic sizes but otherwise the same. We still have

$$(4.6) \quad \Upsilon_{(1:k+1, 1:k+1)} = \Theta - e_{\text{odd}} e_{\text{odd}}^T - e_{\text{even}} e_{\text{even}}^T$$

$$(4.7) \quad = \Theta^{1/2} \left[ I_{k+1} - \frac{1}{2} \Theta^{-1/2} (e_{\text{odd}} e_{\text{odd}}^T + e_{\text{even}} e_{\text{even}}^T) \Theta^{-1/2} \right] \Theta^{1/2},$$

and thus

$$(4.8) \quad \frac{N-2\lfloor k/2 \rfloor}{N+1} \Theta \leq \Upsilon_{(1:k+1, 1:k+1)} \leq \Theta.$$

*Proof.* For  $0 \leq i, j \leq N-1$ , (3.7) with  $\theta_j$  defined here remains valid. Consider now  $\ell = i \pm j$ , and  $0 \leq i, j \leq N-1$ . Since  $0 \leq i+j \leq 2N-2$  and  $-N+1 \leq i-j \leq N-1$ , for some integer  $m$

$$(4.9) \quad i+j = 2m(N+1) \Leftrightarrow i=j=0;$$

$$(4.10) \quad i-j = 2m(N+1) \Leftrightarrow i=j.$$

It can then be verified that

$$(4.11) \quad \Upsilon_{(i+1, j+1)} = \begin{cases} N, & \text{for } i=j=0, \\ \frac{N-1}{2}, & \text{for } 0 \neq i=j, \\ -1, & \text{for } i \neq j, \text{ both odd or even,} \\ 0, & \text{for } i \neq j, \text{ one odd and one even.} \end{cases}$$

This gives (4.5) – (4.7). Let  $\lfloor k/2 \rfloor$  be the largest integer that is no bigger than  $k/2$ , and  $\lceil k/2 \rceil$  the smallest integer that is no less than  $k/2$ . Note that  $\Theta^{-1/2} e_{\text{odd}}$  and  $\Theta^{-1/2} e_{\text{even}}$  are orthogonal, and that

$$\begin{aligned} \|\Theta^{-1/2} e_{\text{odd}}\|_2^2 &= \frac{1}{N+1} + \frac{2}{N+1} \left\lfloor \frac{k}{2} \right\rfloor, \\ \|\Theta^{-1/2} e_{\text{even}}\|_2^2 &= \frac{2}{N+1} \left\lceil \frac{k}{2} \right\rceil, \end{aligned}$$

to get (notice that  $\|\Theta^{-1/2}e_{\text{odd}}\|_2^2 \geq \|\Theta^{-1/2}e_{\text{even}}\|_2^2$ )

$$\begin{aligned} I_{k+1} - \frac{1}{2}\Theta^{-1/2}(e_{\text{odd}}e_{\text{odd}}^T + e_{\text{even}}e_{\text{even}}^T)\Theta^{-1/2} &\geq \left(1 - \frac{2\lfloor k/2 \rfloor + 1}{N+1}\right) I_{k+1} \\ &= \frac{N - 2\lfloor k/2 \rfloor}{N+1} I_{k+1}, \end{aligned}$$

and thus  $\Upsilon_{(1:k+1, 1:k+1)} \geq \frac{N - 2\lfloor k/2 \rfloor}{N+1} \Theta$ . On the other hand,  $\Upsilon_{(1:k+1, 1:k+1)} \leq \Theta$ .  $\square$

**Theorem 4.1.** For  $V_{k+1, N}$  with the eigenvalues (4.2) of model matrix (4.1) as its nodes,

$$(4.12) \quad \frac{1}{\sqrt{N}} \sqrt{\frac{N - 2\lfloor k/2 \rfloor}{2k+1}} \leq \min_{|u_{(1)}|=1} \frac{\|V_{k+1, N}^T u\|_2}{\sqrt{N}} \leq \frac{1}{\sqrt{N}} \sqrt{\frac{N+1}{2k+1}},$$

(4.13)

$$\min_{|u_{(1)}|=1} \frac{\|V_{k+1, N}^T u\|_2}{\sqrt{N}} = \frac{1}{\sqrt{N}} \left[ \frac{2k+1}{N+1} + \frac{1}{N+1} \left( \frac{(k+1)^2}{N-k} + \frac{k^2}{N-k+1} \right) \right]^{-1/2}.$$

The right-hand side of (4.13) is very close to the lower bound in (4.12).

*Proof.* Set  $p_m = T_m$ ,  $\omega = -2$ ,  $\tau = 1$ , and  $\Omega_N = I_N$  in (2.8). Lemma 4.2 says that (2.21) holds with  $\gamma_{k+1} = \frac{N - 2\lfloor k/2 \rfloor}{N+1}$ ,  $\Gamma_{k+1} = 1$ , and  $D_{k+1} = \Theta^{-1}$  which, substituted into (2.26) and (2.27), yield (4.12) since  $\Phi_{k+1} = \sum_{j=0}^k d_j |T_j(1)|^2 = \frac{2k+1}{N+1}$ .

For  $k \geq 1$ , by Sherman-Morrison-Woodbury formula [6, p.95],

$$[\Upsilon_{(1:k+1, 1:k+1)}]^{-1} = \Theta^{-1} + \Theta^{-1}W(I_2 - W^T\Theta^{-1}W)^{-1}W^T\Theta^{-1},$$

where  $W = (e_{\text{odd}}, e_{\text{even}})$ . Let  $y = R_{k+1}^T e_1 = (1, 1, \dots, 1)^T$ . We need to evaluate  $y^T [\Upsilon_{(1:k+1, 1:k+1)}]^{-1} y$ . To this end, we notice

$$\begin{aligned} y^T \Theta^{-1} &= \frac{2}{N+1} (2^{-1}, 1, 1, \dots), \\ y^T \Theta^{-1} y^T &= \frac{2}{N+1} \left( \frac{1}{2} + k \right) = \frac{2k+1}{N+1}, \\ y^T \Theta^{-1} W &= \frac{2}{N+1} \left( \frac{1}{2} + \left\lfloor \frac{k}{2} \right\rfloor, \left\lceil \frac{k}{2} \right\rceil \right) \\ &= \frac{1}{N+1} \left( 2 \left\lfloor \frac{k}{2} \right\rfloor + 1, 2 \left\lceil \frac{k}{2} \right\rceil \right), \\ W^T \Theta^{-1} W &= \frac{2}{N+1} \begin{pmatrix} \frac{1}{2} + \lfloor \frac{k}{2} \rfloor & 0 \\ 0 & \lceil \frac{k}{2} \rceil \end{pmatrix}, \\ I_2 - W^T \Theta^{-1} W &= \frac{1}{N+1} \begin{pmatrix} n - 2 \lfloor \frac{k}{2} \rfloor & 0 \\ 0 & n - 2 \lceil \frac{k}{2} \rceil + 1 \end{pmatrix}. \end{aligned}$$

Finally

$$\begin{aligned} y^T [\Upsilon_{(1:k+1, 1:k+1)}]^{-1} y &= \frac{2k+1}{N+1} + \frac{1}{N+1} \left( \frac{(2\lfloor k/2 \rfloor + 1)^2}{N - 2\lfloor k/2 \rfloor} + \frac{4\lceil k/2 \rceil^2}{N - 2\lceil k/2 \rceil + 1} \right) \\ &= \frac{2k+1}{N+1} + \frac{1}{N+1} \left( \frac{(k+1)^2}{N-k} + \frac{k^2}{N-k+1} \right), \end{aligned}$$

and hence (4.13).  $\square$

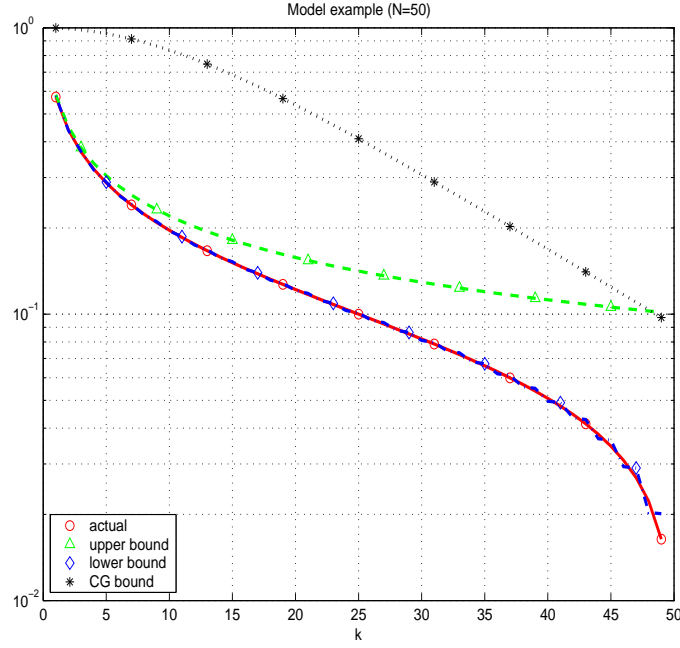


FIGURE 4.1.  $\min_{|u_{(1)}|=1} \frac{\|V_{k+1,N}^T u\|_2}{\sqrt{N}}$  vs. its lower and upper bounds for the model example: lower bounds nearly the same as actuals.

Equation (4.13) has already been obtained via an direct optimization approach. (4.12) and (4.13) also show slow convergence of any method that seeks approximations from  $\mathcal{K}_k(A, b)$ . Figure 4.1 plots the exact value in (4.13), its lower and upper bounds by (4.12), and the upper bound by (2.7).

**Theorem 4.2.** Let  $\Omega_N = \text{diag}(\sin^2 \theta_1, \dots, \sin^2 \theta_N)$ . For  $V_{k+1,N}$  with the eigenvalues (4.2) of model matrix (4.1) as its nodes,

$$(4.14) \quad \min_{|u_{(1)}|=1} \frac{\|\Omega_N^{1/2} V_{k+1,N}^T u\|_2}{\sqrt{(N+1)/2}} = \sqrt{\frac{6}{(2k+3)(k+2)(k+1)}}.$$

*Proof.* Set  $p_m = U_m$ ,  $\omega = -2$ ,  $\tau = 1$ , and  $\Omega_N = \text{diag}(\sin^2 \theta_1, \dots, \sin^2 \theta_N)$  in (2.8).  $\Upsilon = \mathbf{P}_N \Omega_N \mathbf{P}_N^T = \frac{N+1}{2} I$  because  $\omega_j = \sin^2 \theta_j$  are the scaled weights in the associated Gaussian quadrature rule [1, p.889]. Similarly to (2.24) and (2.25), we have

$$\begin{aligned} e_1^T (\bar{V}_{k+1,N} \Omega_N V_{k+1,N}^T)^{-1} e_1 &= \frac{2}{N+1} e_1^T R_{k+1} R_{k+1}^T e_1 \\ &= \frac{2}{N+1} \sum_{j=0}^k |U_j(1)|^2 \\ &= \frac{2}{N+1} \sum_{j=0}^k (j+1)^2 \\ &= \frac{2}{N+1} \frac{(2k+3)(k+2)(k+1)}{6}. \end{aligned}$$

Use the summation formula [10, p.30] to get  $\sum_{j=1}^N \sin^2 \theta_j = \frac{N+1}{2}$ . (4.14) is now a consequence of (2.28).  $\square$

## 5. Concluding Remarks

This paper analyzes two examples that are hard for CG, or any method that seeks (best possible) approximate solutions within the Krylov subspace  $\mathcal{K}_k(A, b)$ . Frank's example has a reputation of being "the hardest case for" CG [9], and the model example is in Demmel's textbook<sup>3</sup> [6] to demonstrate effectiveness of various numerical methods for positive definite linear systems. Analytical formulas or extremely tight lower and upper bounds can be found for their residuals by CG or MINRES. That both problems are hard for CG is well-known, owing to their eigenvalue distribution and growing condition numbers in  $N$ . One of our contributions here is the quantification of the fact.

We established a general framework that could be potential useful for analyzing certain linear systems. These include the case when  $A$ 's eigenvalues are  $\omega(t_j - \tau)$  ( $1 \leq j \leq N$ ), where  $\omega \neq 0$  and  $\tau$  are two (complex) numbers, and  $t_j$  given as one of the following.

- $t_j = \cos \frac{2j-1}{2N} \pi$  ( $1 \leq j \leq N$ ), zeros of  $T_N(t)$ . Take  $p_m = T_m$  and  $\Omega_N = I_N$ , and then  $\mathbf{P}_N \Omega_N \mathbf{P}_N^T = \frac{N}{2} \text{diag}(2, 1, 1, \dots)$  [15, 16, 18].
- $t_j = \cos \frac{j-1}{N-1} \pi$  ( $1 \leq j \leq N$ ), the extreme points of  $T_{N-1}(t)$  in  $[-1, 1]$ . Take  $p_m = T_m$  and  $(\Omega_N)_{(1,1)} = (\Omega_N)_{(N,N)} = 1/2$ ,  $(\Omega_N)_{(j,j)} = 1$  for  $1 < j < N$ , and then  $\mathbf{P}_N \Omega_N \mathbf{P}_N^T = \frac{N-1}{2} \Omega_N^{-1}$  [15, 16, 18].
- $t_j$  ( $1 \leq j \leq N$ ) are zeros of the  $N$ th orthogonal polynomial in an orthogonal polynomial family. Take  $p_m$  to be the  $m$ th orthogonal polynomial and  $\omega_j$  the weights in the associate Gaussian quadrature rule, and then  $\mathbf{P}_N \Omega_N \mathbf{P}_N^T$  is diagonal [8, 16].
- $t_j = \cos \frac{2j-1}{2N+1} \pi$  ( $1 \leq j \leq N$ ) as in Section 3.
- $t_j = \exp(\iota \frac{j-1}{N} \pi)$  ( $1 \leq j \leq N$ ), where  $\iota = \sqrt{-1}$ . Take  $p_m(t) = t^m$  and  $\Omega_N = I_N$ , and then  $\mathbf{P}_N \Omega_N \mathbf{P}_N^T$  is diagonal.

It is possible to construct many more problems that are hard for CG or any Krylov subspace type method and justifiable by our framework. For example, if  $A$  has eigenvalues  $1 + t_j$  ( $1 \leq j \leq N$ ) or  $1 - t_j$  ( $1 \leq j \leq N$ ), where  $t_j$ 's are zeros of  $T_N(t)$ , then

$$\min_{|u_{(1)}|=1} \frac{\|V_{k+1,N}^T u\|_2}{\sqrt{N}} = \sqrt{\frac{1}{2k+1}}.$$

## References

- [1] Abramowitz, M. and Stegun, I. A. (editors), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, Inc., New York, 9th printing edition, 1970.
- [2] Axelsson, O., *Iterative Solution Methods*. Cambridge University Press, New York, 1996.
- [3] Beckermann, B. and Kuijlaars, A. B. J., Superlinear convergence of conjugate gradients. *SIAM J. Numer. Anal.*, 39(1):300–329, 2001.
- [4] Beckermann, B. and Kuijlaars, A. B. J., Superlinear CG convergence for special right-hand sides. *Electron. Trans. Numer. Anal.*, 14:1–19, 2002.
- [5] Briggs, W. L., Henson, V. E., and McCormick, S. F., *A Multigrid Tutorial*. SIAM, Philadelphia, Pennsylvania, 2nd edition, 2000.
- [6] Demmel, J., *Applied Numerical Linear Algebra*. SIAM, Philadelphia, 1997.

<sup>3</sup>Demmel focused more on the 2-dimensional Poisson equation.

- [7] Frank, W. L., Solution of linear systems by Richardson's method. *J. ACM*, 7(3):274–286, 1960.
- [8] Gautschi, W., The condition of Vandermonde-like matrices involving orthogonal polynomials. *Linear Algebra Appl.*, 52/53:293–300, 1983.
- [9] Golub, G. H. and O'Leary, D. P., Some history of the conjugate gradient and Lanczos algorithms: 1948-1976. *SIAM Rev.*, 31(1):50–102, March 1989.
- [10] Gradshteyn, I. S. and Ryzhik, I. M., *Table Of Integrals, Series, and Products*. Academic Press, New York, 1980. Corrected and Enlarged Edition prepared by A. Jeffrey, incorporated the fourth edition prepared by Yu. V. Geronimus and M. Yu. Tseytlin, translated from the Russian by Scripta Technica, Inc.
- [11] Greenbaum, A., Comparison of splittings used with the conjugate gradient algorithm. *Numer. Math.*, 33:181–194, 1979.
- [12] Greenbaum, A., *Iterative Methods for Solving Linear Systems*. SIAM, Philadelphia, 1997.
- [13] Ipsen, I. C. F., Expressions and bounds for the GMRES residual. *BIT*, 40(3):524–535, September 2000.
- [14] Kaniel, S., Estimates for some computational techniques in linear algebra. *Math. Comp.*, 20(95):369–378, July 1966.
- [15] Li, Ren-Cang, Sharpness in rates of convergence for CG and symmetric Lanczos methods. Technical Report 2005-01, Department of Mathematics, University of Kentucky, 2005. Available at <http://www.ms.uky.edu/~math/MAreport/>.
- [16] Li, Ren-Cang, Vandermonde matrices with Chebyshev nodes. *Linear Algebra Appl.*, 2006. to appear.
- [17] Li, Ren-Cang, Asymptotically optimal lower bounds for the condition number of a real Vandermonde matrix. *SIAM J. Matrix Anal. Appl.*, 28(3):829–844, 2006.
- [18] Li, Ren-Cang, On Meinardus' examples for the conjugate gradient method. *Math. Comp.*, 77(261):335–352, 2008.
- [19] Liesen, J., Rozložník, M., and Strakos, Z., Least squares residuals and minimal residual methods. *SIAM J. Sci. Comput.*, 23(5):1503–1525, 2002.
- [20] Liesen, J. and Tichý, P., The worst-case GMRES for normal matrices. *BIT*, 44(1):79–98, January 2004.
- [21] Liesen, J. and Tichý, P., On the worst-case convergence of MR and CG for symmetric positive definite tridiagonal Toeplitz matrices. *Electron. Trans. Numer. Anal.*, 20:180–197, 2005.
- [22] Meinardus, G., Über eine Verallgemeinerung einer Ungleichung von L. V. Kantorowitsch. *Numer. Math.*, 5:14–23, 1963.
- [23] Naiman, A. E., Babuška, I. M., and Elman, H. C., A note on conjugate gradient convergence. *Numer. Math.*, 76(2):209–230, April 1997.
- [24] Saad, Y., On the rates of convergence of the Lanczos and the block-Lanczos methods. *SIAM J. Numer. Anal.*, 15(5):687–706, October 1980.
- [25] Saad, Y. and Schultz, M., GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7:856–869, 1986.
- [26] Saad, Y., *Iterative Methods for Sparse Linear Systems*. SIAM, Philadelphia, 2nd edition, 2003.
- [27] Sleijpen, G. L. G. and van der Sluis, A., Further results on the convergence behavior of conjugate-gradients and Ritz values. *Linear Algebra Appl.*, 246:233–378, October 1996.
- [28] Trefethen, L. N. and Bau III, D., *Numerical Linear Algebra*. SIAM, Philadelphia, 1997.
- [29] Young, D., Computation of eigenvalues and eigenvectors of a certain matrix. Numerical Analysis Notes NN-18, Space Technology Laboratories, August 1956.

Department of Mathematics, University of Texas at Arlington, P.O. Box 19408, Arlington, TX 76019-0408, USA

*E-mail:* rcli@uta.edu