Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 17/03

Andreas Frommer, Claudia Schimmel and Marcel Schweitzer

Bounds for the decay of the entries in inverses and Cauchy–Stieltjes functions of sparse, normal matrices

April 25, 2017

NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 0000; **00**:1–19
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nla

Bounds for the decay of the entries in inverses and Cauchy–Stieltjes functions of sparse, normal matrices

Andreas Frommer[1], Claudia Schimmel[1] and Marcel Schweitzer[2]

[1]*School of Mathematics and Natural Sciences, Bergische Universität Wuppertal, 42097 Wuppertal, Germany,* {frommer, schimmel}@math.uni-wuppertal.de

[2]*École Polytechnique Fédérale de Lausanne, Station 8, 1015 Lausanne, Switzerland,* marcel.schweitzer@epfl.ch

SUMMARY

It is known that in many functions of banded, and more generally, sparse Hermitian positive definite matrices, the entries exhibit a rapid decay away from the sparsity pattern. This is in particular true for the inverse, and based on results for the inverse, bounds for Cauchy–Stieltjes functions of Hermitian positive definite matrices have recently been obtained. We add to the known results by considering the more general case of normal matrices, for which fewer and typically less satisfactory results exist so far. Starting from a very general estimate based on approximation properties of Chebyshev polynomials on ellipses, we obtain as special cases insightful decay bounds for various classes of normal matrices, including (shifted) skew-Hermitian and Hermitian indefinite matrices. In addition, some of our results improve over known bounds when applied to the Hermitian positive definite case. Copyright © 0000 John Wiley & Sons, Ltd.

Received . . .

KEY WORDS: matrix function; banded matrices; normal matrices; off-diagonal decay; Chebyshev polynomials

1. INTRODUCTION

Off-diagonal decay behavior in functions of banded matrices has long been studied, beginning with special emphasis on the matrix inverse [10–13, 17, 19] and the matrix exponential [4, 6, 16, 18]. Further results on other classes of functions can be found in, e.g., [5,6], see also the recent survey [3].

There is much interest in finding bounds or estimates for the off-diagonal entries of matrix functions because these allow to efficiently find *sparse approximations* of quantities of interest in a variety of areas as, e.g., Markov chain queuing models [7, 8] and quantum dynamics [14]. Under certain conditions, e.g., when the decay behavior is independent of the matrix size n for a family of matrices $A_n \in \mathbb{C}^{n\times n}$, the knowledge of sharp decay bounds even allows the design of optimal, linearly scaling algorithms for matrix function computations [5, 9]. Thus, improving and extending results on off-diagonal decay in matrix functions is of great practical interest. All these results naturally generalize to general sparse matrices A to yield bounds on the decay of the matrix entries $[f(A)]_{ij}$ as a function of the distance of i and j in the graph describing the sparsity structure of A.

While most of the known results focus on Hermitian (positive definite) matrices, we investigate the more general case of normal matrices in this paper. There are also results for this class of matrices

*Correspondence to: Andreas Frommer, School of Mathematics and Natural Sciences, Bergische Universität Wuppertal, 42097 Wuppertal, Germany, frommer@math.uni-wuppertal.de

in the literature [3, 5, 18], but most of the results fail to be insightful or practically usable, as they typically depend on quantities that are very hard or impossible to compute in practice (e.g., because they require knowledge of the *complete* spectrum of the matrix A) or are very pessimistic and therefore do not capture the actual quantitative decay behavior well.

In this paper, we therefore try to improve over many of the bounds known from the literature by considering several special cases of normal matrices, in particular skew-Hermitian, shifted skew-Hermitian and Hermitian indefinite matrices, which allow to obtain better decay bounds by exploiting the fact that their spectrum is contained in a line segment in the complex plane and using approximation properties of complex Chebyshev polynomials on these line segments. In this manner, we obtain bounds that are in many cases (provably) sharper than those available in the literature so far, and in some cases our approach even allows to obtain improved results when applied to the Hermitian positive definite case.

The remainder of this paper is organized as follows. In section 2, we briefly introduce the most important concepts needed to develop our results in later sections, specifically the connection between the error of polynomial approximations for f and off-diagonal decay in $f(A)$, and the definition and basic properties of complex Chebyshev polynomials. Some known decay bounds from the literature, in particular those on the inverse from [11] and on Cauchy–Stieltjes functions from [6] are reviewed in section 3, as those bounds will (where applicable) be used as a comparison to judge the quality of our new bounds in later sections. Section 4 contains our main results on the matrix inverse, beginning with a very general result for banded, normal matrices before considering the special cases of Hermitian indefinite, skew-Hermitian and shifted skew-Hermitian matrices. The bounds derived in section 4 are then compared to known bounds from the literature in section 5, both theoretically and by numerical experiments. Building on the results obtained for the matrix inverse in section 4, we derive bounds for Cauchy–Stieltjes matrix functions in section 6. The quality of these bounds is then illustrated by further numerical experiments in section 7. We give concluding remarks in section 8.

2. BASICS

In this section, the basic concepts necessary for deriving our results are introduced.

2.1. *The relation between off-diagonal decay in banded matrices and polynomial approximation*

For ease of presentation, all of the results in this paper are formulated for banded matrices, with the natural extension to general sparse matrices being briefly sketched in Remark 2.1 below. We thus begin by clarifying our notation of bandwidth. We say that A has bandwidth $\beta \in \{1, \ldots, n-2\}$ (or, as a shorthand, that A is β-banded) if $A_{k\ell} = 0$ for $|k-\ell| > \beta$. Using this definition, a tridiagonal matrix has bandwidth $\beta = 1$, a pentadiagonal matrix has bandwidth $\beta = 2$ and so on.

For normal, banded matrices $A \in \mathbb{C}^{n\times n}$ and scalar functions f such that $f(A)$ is defined, it is possible to obtain decay bounds for the entries of $f(A)$ by exploiting knowledge of the error

$$\varepsilon(m) := \max_{z\in E} |f(z) - p_m(z)|$$

where $E \subset \mathbb{C}$ is a set containing $\sigma(A)$, the spectrum of A, and p_m is a polynomial approximation of f of degree at most m. To show this, consider the relation

$$\|f(A) - p_m(A)\|_2 = \max_{z\in\sigma(A)} |f(z) - p_m(z)|$$

which holds for any normal matrix A. Now assume that A is β-banded and write $|k-\ell| = m\beta + s$ for $m \geq 0$ and $s \in \{1, \ldots, \beta\}$. Now $|k-\ell| > m\beta$ and $p_m(A)$ has bandwidth $m\beta$, such that

DOI: 10.1002/nla

$[p_m(A)]_{k\ell} = 0$. Therefore we have

$$\begin{aligned} |[f(A)]_{k\ell}| &= |[f(A)]_{k\ell} - [p_m(A)]_{k\ell}| \\ &\le \|f(A) - p_m(A)\|_2 \\ &= \max_{z \in \sigma(A)} |f(z) - p_m(z)| \\ &\le \max_{z \in E} |f(z) - p_m(z)| = \varepsilon(m), \end{aligned}$$

i.e., the error $\varepsilon(m)$ gives an upper bound for all entries $[f(A)]_{k\ell}$ with $|k-\ell| > m\beta$.

Remark 2.1
The approach explained above is also applicable to general sparse matrices by using the geodesic distance $\mathrm{d}(k,\ell)$ of the nodes k and ℓ in the (directed) graph corresponding to the matrix A. For $m \ge 1$, it holds that $[A^m]_{k\ell} = 0$ if there is no path from node k to node ℓ with length at most m in the graph of A, so that in this case $[p_m(A)]_{k\ell} = 0$ as well for any polynomial p_m of degree at most m. Therefore, decay bounds for $[f(A)]_{k\ell}$ can be obtained in a similar manner as for banded matrices by considering the error $\varepsilon(\mathrm{d}(k,\ell)-1)$. Note that in case of a matrix A with bandwidth β and $A_{k\ell} \ne 0$ for $|k-\ell| \le \beta$, we have $\mathrm{d}(k,\ell) = \lceil \frac{|k-\ell|}{\beta} \rceil$, so that both approaches agree. For ease of presentation, we will formulate all results for banded matrices, keeping in mind that they apply to sparse matrices as well (in the sense that "off-diagonal decay" is replaced by "decay away from the sparsity pattern of A").

We begin our exposition by focusing on the function $f(z) = \frac{1}{z}$. By finding an appropriate polynomial p_m and corresponding error $\varepsilon(m)$ with help of Chebyshev polynomials, we can then obtain decay bounds for the inverse of banded normal matrices. To be able to do so, we first give a short introduction to Chebyshev polynomials in the next subsection.

2.2. *Chebyshev polynomials*

Chebyshev polynomials can be defined via the three-term recurrence

$$C_{m+1}(z) = 2zC_m(z) - C_{m-1}(z), \quad m \ge 1,$$

with $C_1(z) = z$ and $C_0(z) = 1$. For $\gamma \notin [-1,1]$ the normalized Chebyshev polynomial

$$P_m(z) = \frac{C_m(z)}{C_m(\gamma)} \tag{1}$$

solves the min-max problem

$$\min_{\substack{p_m \in \Pi_m \\ p_m(\gamma)=1}} \max_{z \in [-1,1]} |p_m(z)| \tag{2}$$

where Π_m denotes the space of all polynomials of degree at most m. The maximal value attained on $[-1,1]$ by this polynomial is

$$P_m(0) = \frac{1}{|C_m(\gamma)|}. \tag{3}$$

Using the transformation $t = 1 + 2\frac{z-b}{b-a}$ this result can be generalized to any interval $[a,b]$.

A similar result holds for ellipses in the complex plane. In the following we denote by $E(\rho, f_1, f_2)$ the ellipse with focal points f_1 and f_2 and semi-axes $\frac{\rho-\rho^{-1}}{2}$ and $\frac{\rho+\rho^{-1}}{2}$. As a short-hand, we further use the notation $E_\rho := E(\rho, -1, 1)$. Note in particular that $E(1, f_1, f_2)$ is a line segment connecting f_1 and f_2, as this is an important special case we will consider at various places in this paper.

An alternative representation of complex Chebyshev polynomials is by

$$C_m(z) = \frac{1}{2}(w^m + w^{-m}), \tag{4}$$

Prepared using nlaauth.cls
DOI: 10.1002/nla

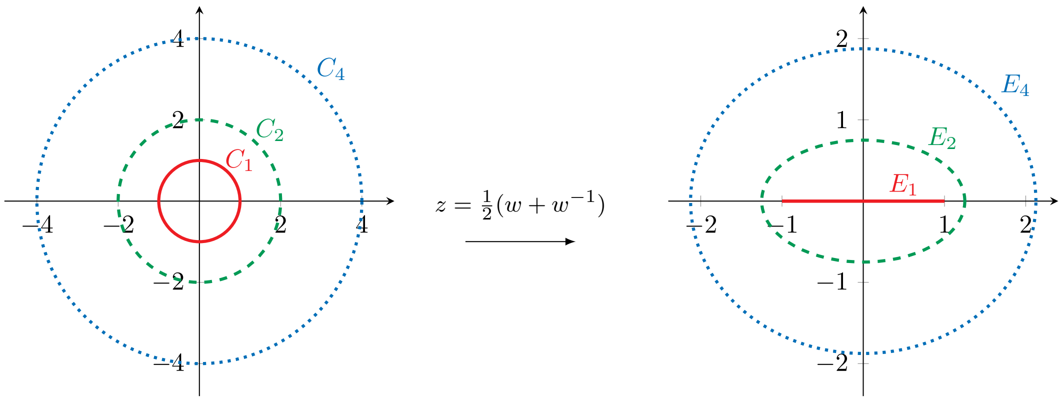


Figure 1. The Joukowski mapping $z = \frac{1}{2}(w + w^{-1})$.

where

$$z = \frac{1}{2}(w + w^{-1}), \tag{5}$$

see [20], e.g. . This gives, in particular,

$$C_m(z) = \cosh(m\xi), \quad \text{where} \quad \cosh(\xi) = z \text{ for } z \notin [-1, 1]. \tag{6}$$

The *Joukowski mapping* (5) maps the circle C_ρ of radius ρ centered at the origin to the ellipse E_ρ. This mapping is illustrated in Figure 1.

It is known that for the optimal polynomial of degree m with respect to the ellipse E_ρ the relation

$$\frac{\rho^m}{|w_\gamma|^m} \leq \min_{\substack{p_m \in \Pi_m \\ p_m(\gamma)=1}} \max_{z \in E_\rho} |p_m(z)| \leq \frac{\rho^m + \rho^{-m}}{|w_\gamma^m + w_\gamma^{-m}|} \tag{7}$$

holds, where w_γ is defined via $\gamma = \frac{1}{2}(w_\gamma + w_\gamma^{-1})$, and the upper bound in (7) is achieved by the normalized Chebyshev polynomial (1). This can be seen by the fact that

$$\max_{z \in E_\rho} |C_m(z)| = \max_{w \in C_\rho} \left| \frac{1}{2}(w^m + w^{-m}) \right| \leq \max_{w \in C_\rho} \frac{1}{2}(|w|^m + |w|^{-m}) = \frac{1}{2}(\rho^m + \rho^{-m}), \tag{8}$$

and this upper bound is reached for $w = \rho$. Since the difference between the upper and lower bound for the min-max problem in (7) tends to zero for increasing m, Chebyshev polynomials are *asymptotically* optimal for ellipses E_ρ.

By applying a variable transformation again, similarly to the real case, (7) can be generalized to ellipses $E(\rho, f_1, f_2)$.

The following lemma gives another useful property of Chebyshev polynomials, which we need for developing our results.

Lemma 2.2
Let C_m be the Chebyshev polynomial of degree m and $z \in \mathbb{R}$ of the form $z = 1 + 2x,\ x \in \mathbb{R}$. Then

$$C_m(z) \geq \frac{1}{2}\left(\sqrt{x} + \sqrt{x+1}\right)^{2m}.$$

Proof
See [20, Section 6.11.3]. □

3. PREVIOUS RESULTS

For the inverse of banded matrices, decay bounds were published in [11] for Hermitian positive definite matrices, which can be extended to also give bounds in the non-Hermitian case. The main results are summarized in the following theorem.

DOI: 10.1002/nla

Theorem 3.1
Let A be Hermitian positive definite and β-banded, with smallest eigenvalue $\lambda_{\min}$, largest eigenvalue $\lambda_{\max}$ and condition number $\kappa(A) = \lambda_{\max}/\lambda_{\min}$. Then

$$|[A^{-1}]_{k\ell}| \leq Cq^{\frac{|k-\ell|}{\beta}} \tag{9}$$

with

$$q = \frac{\sqrt{\kappa(A)}-1}{\sqrt{\kappa(A)}+1} \text{ and } C = \max\left\{\frac{1}{\lambda_{\min}}, \frac{\left(1+\sqrt{\kappa(A)}\right)^2}{2\lambda_{\max}}\right\}. \tag{10}$$

If A is just β-banded, then

$$|[A^{-1}]_{k\ell}| \leq C_1 q_1^{\frac{|k-\ell|}{\beta}}$$

with

$$q_1 = \sqrt{\frac{\kappa(A)-1}{\kappa(A)+1}}, \quad \kappa(A) = \|A\|_2\|A^{-1}\|_2$$

and

$$C_1 = (2\beta+1)q_1^{-2}\|A^{-1}\|_2\kappa(A)\max\left\{1, \left(\frac{1+\kappa(A)}{\kappa(A)}\right)^2/2\right\}. \tag{11}$$

The second assertion of Theorem 3.1 follows from the first one by using the fact that $A^{-1} = A^H(AA^H)^{-1}$. Note that the condition number of A appears as a multiplicative factor in the definition of the constant C_1 in the nonsymmetric case. Due to this factor, the entries of A^{-1} are highly overestimated by the bound (11) in many cases, especially for ill-conditioned problems.

Using bounds for the inverse, it is possible to obtain decay bounds for Cauchy–Stieltjes functions, see [6]. These are functions of the form

$$f(z) = \int_0^\infty \frac{\mathrm{d}\gamma(\tau)}{z+\tau} \tag{12}$$

where γ is a monotonically increasing, real-valued and non-negative function on $[0,\infty)$.

The bound

$$|[f(A)]_{k\ell}| \leq \int_0^\infty C(\tau)q(\tau)^{\frac{|k-\ell|}{\beta}}\ \mathrm{d}\gamma(\tau)$$

with $q(\tau) = (\sqrt{\kappa(\tau)}-1)/(\sqrt{\kappa(\tau)}+1)$, $\kappa(\tau) = (\lambda_{\max}+\tau)/(\lambda_{\min}+\tau)$ and

$$C(\tau) = \max\left\{\frac{1}{\lambda_{\min}+\tau}, \frac{(1+\sqrt{\kappa(\tau)})^2}{2(\lambda_{\max}+\tau)}\right\} \tag{13}$$

directly follows from (9) for Hermitian positive definite matrices A. Further, by exploiting the monotonicity of q, one finds

$$|[f(A)]_{k\ell}| \leq q(0)^{\frac{|k-\ell|}{\beta}} \int_0^\infty C(\tau)\ \mathrm{d}\gamma(\tau). \tag{14}$$

Equation (14) was given in [6], and for the special case $f(z) = z^{-1/2}$ the bound

$$|[A^{-\frac{1}{2}}]_{k\ell}| \leq \frac{2}{\pi}(C(0)+\widehat{C})\ q^{\frac{|k-\ell|}{\beta}} \tag{15}$$

with $\widehat{C} = (1+\sqrt{\kappa(0)})^2/2$ was derived there. The advantage of (15) over (14) is that no integral has to be evaluated. In section 6 we deal with the question whether it is possible to obtain similar bounds for general Cauchy–Stieltjes functions and other classes of matrices.

DOI: 10.1002/nla

4. DECAY BOUNDS FOR THE INVERSE OF NORMAL MATRICES

In this section, we show how we can use Chebyshev polynomials for finding a polynomial approximation of $f(z) = 1/z$ with corresponding error $\varepsilon(m)$, which then yields decay bounds for the entries of the inverse A^{-1} of normal matrices A.

Lemma 4.1
There exists a polynomial q_m of degree m such that

$$\varepsilon(m) = \max_{z \in E(\rho, f_1, f_2)} \left| \frac{1}{z} - q_m(z) \right| \le \frac{\rho^{m+1} + \rho^{-(m+1)}}{2 \min_{z \in E(\rho, f_1, f_2)} |z| \, |C_{m+1}(\frac{f_1+f_2}{f_2-f_1})|}. \tag{16}$$

Proof
Let $P_{m+1}(z)$ be the normalized Chebyshev polynomial with $P_{m+1}(0) = 1$ with respect to the ellipse $E(\rho, f_1, f_2)$. Then $P_{m+1}(z)$ can be written as $P_{m+1}(z) = 1 - zq_m(z)$ for some $q_m \in \Pi_m$. The polynomial $q_m(z)$ can thus be interpreted as a polynomial approximation for $f(z) = 1/z$ on $E(\rho, f_1, f_2)$. The transformation $t = \frac{f_1+f_2-2z}{f_2-f_1}$ maps the ellipse $E(\rho, f_1, f_2)$ to E_ρ and therefore

$$P_{m+1}(z) = \frac{C_{m+1}(t)}{C_{m+1}(\frac{f_1+f_2}{f_2-f_1})}.$$

Now

$$\max_{z \in E(\rho, f_1, f_2)} \left| \frac{1}{z} - q_m(z) \right| = \max_{z \in E(\rho, f_1, f_2)} \left| \frac{P_{m+1}(z)}{z} \right| \le \frac{\max_{z \in E(\rho, f_1, f_2)} |P_{m+1}(z)|}{\min_{z \in E(\rho, f_1, f_2)} |z|}.$$

The assertion now follows by using (8). □

As discussed in section 2.1, the result of Lemma 4.1 already allows to obtain bounds for the entries of the inverses of normal matrices where the spectrum is contained in an ellipse excluding the origin. Specifically, $|[A^{-1}]_{ij}|$ is bounded by the right-hand side in (16) with $m = \lceil \frac{|k-\ell|}{\beta} \rceil - 1$.

In the following, we will show that for some important classes of normal matrices the bound (16) can be used as a basis for obtaining more insightful bounds in which the decay behavior is more apparent.

First, we focus on normal matrices whose spectrum is contained in a line segment in the complex plane which excludes the origin. Examples of matrices with this property are shifted skew-Hermitian matrices $A = M + sI$, $M = -M^H$, $s \in \mathbb{R}$ or complex shifted Hermitian matrices $A = M + i \cdot sI$, $M = M^H$, $s \in \mathbb{R}$.

In the following we denote by $[\lambda_1, \lambda_2]$ a complex line segment with end points λ_1 and λ_2.

Theorem 4.2
Let A be a nonsingular and β-banded normal matrix with $\sigma(A) \subset [\lambda_1, \lambda_2]$ for some complex line segment with $0 \notin [\lambda_1, \lambda_2]$. Then for $k \ne \ell$

$$|[A^{-1}]_{k\ell}| \le 2\, \|A^{-1}\|_2 \, \frac{1}{q^{|k-\ell|/\beta} - 1} \tag{17}$$

with

$$q = e^{\mathrm{Re}(\operatorname{arcosh}(x))} > 1 \text{ and } x = \frac{\lambda_1 + \lambda_2}{\lambda_2 - \lambda_1}.$$

For $k = \ell$

$$|[A^{-1}]_{k\ell}| \le \|A^{-1}\|_2. \tag{18}$$

Proof
Since $[\lambda_1, \lambda_2] = E(1, \lambda_1, \lambda_2)$ we know from Lemma 4.1 that

$$|[A^{-1}]_{k\ell}| \le \frac{1}{\min_{z \in [\lambda_1, \lambda_2]} |z|} \frac{1}{|C_{m+1}(x)|}.$$

DOI: 10.1002/nla

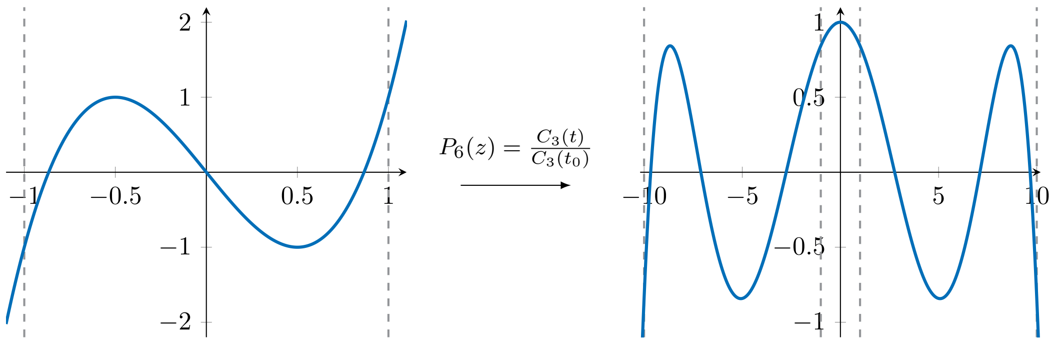


Figure 2. Transformation of the Chebyshev polynomial $C_3(t)$ into $P_6(z)$ according to (20) for the set $K = [-10, -1] \cup [1, 10]$.

Now

$$\frac{1}{\min_{z \in [\lambda_1, \lambda_2]} |z|} \leq \|A^{-1}\|_2,$$

and using the representation (6) of Chebyshev polynomials, we find

$$\begin{aligned} \frac{1}{|C_{m+1}(x)|} &= \frac{1}{|\cosh((m+1)\, \operatorname{arcosh}(x))|} \\ &\leq \frac{1}{|\sinh(\operatorname{Re}((m+1)\, \operatorname{arcosh}(x)))|} \\ &= \frac{2}{\left|\left(e^{\operatorname{Re}(\operatorname{arcosh}(x))}\right)^{m+1} - \left(e^{-\operatorname{Re}(\operatorname{arcosh}(x))}\right)^{m+1}\right|}. \end{aligned}$$

An elementary but somewhat tedious calculation shows that the assumption $0 \notin [\lambda_1, \lambda_2]$ implies $x \notin [-1, 1]$. Since $\operatorname{Re}(\operatorname{arcosh}(x)) \geq 0$ and $\operatorname{Re}(\operatorname{arcosh}(x)) = 0$ if and only if $x \in [-1, 1]$, this yields $q := e^{\operatorname{Re}(\operatorname{arcosh}(x))} > 1$. Thus

$$\frac{1}{|C_{m+1}(x)|} \leq \frac{2}{q^{m+1} - q^{-(m+1)}} \leq \frac{2}{q^{m+1} - 1}.$$

By writing $|k - \ell| = m\beta + s$ with $m \geq 0$ and $s \in \{1, \ldots, \beta\}$ for $k \neq \ell$ we know that $|k - \ell| \leq m\beta + \beta$ and therefore $\frac{|k-\ell|}{\beta} \leq m + 1$. Hence the bound (17) holds. Equation (18) holds for every entry and can thus in particular be used for the case $k = \ell$. □

The technique of proof used in Theorem 4.2 cannot be used for Hermitian indefinite matrices or skew-Hermitian matrices where the spectrum is distributed both above and below the origin on the imaginary axis, since then $0 \in [\lambda_1, \lambda_2]$ for any line segment $[\lambda_1, \lambda_2]$ containing $\sigma(A)$. Therefore we use a different approach for these cases.

We start with bounds for the Hermitian indefinite case first.

Theorem 4.3
Let A be Hermitian indefinite and nonsingular with bandwidth β. Define $m = \lfloor \frac{|k-\ell|-1}{\beta} \rfloor$. Then the entries of A^{-1} can be bounded by

$$|[A^{-1}]_{k\ell}| \leq C \cdot \begin{cases} q^{\frac{|k-\ell|}{\beta}} & \text{for } k \neq \ell \text{ and } m \text{ odd} \\ q^{\frac{|k-\ell|}{\beta} - 1} & \text{for } k \neq \ell \text{ and } m \text{ even} \\ \frac{1}{2} & \text{for } k = \ell \end{cases} \tag{19}$$

Prepared using nlaauth.cls
DOI: 10.1002/nla

where

$$q = \sqrt{\frac{\kappa(A)-1}{\kappa(A)+1}} \text{ and } C = 2\,\|A^{-1}\|_2.$$

Proof
Let $K = [-b,-a] \cup [a,b]$ with $a := \min_{\lambda \in \sigma(A)} |\lambda|$ and $b := \max_{\lambda\in\sigma(A)} |\lambda|$. The set K is mapped to the interval $[-1,1]$ by the transformation $t = 1 + 2\frac{a^2-z^2}{b^2-a^2}$. For $p > 0$ define the polynomial of degree $2p$ as

$$P_{2p}(z) = P_p(t) = \frac{C_p(t)}{C_p(t_0)} \text{ with } t_0 = 1 + 2\frac{a^2}{b^2-a^2}. \tag{20}$$

P_{2p} is even in z. We illustrate the construction of P_{2p} in Figure 2. By construction, $P_{2p}(z)$ minimizes the value

$$\max_{z\in K} |p_{2p}(z)|$$

among all even polynomials $p_{2p} \in \Pi_{2p}$ with $p_{2p}(0) = 1$. The polynomial q_{2p-1} defined by $P_{2p}(z) = 1 - zq_{2p-1}(z)$ can be interpreted as a polynomial approximation for $f(z) = \frac{1}{z}$ on K. For m odd, i.e., $m = 2p-1$, we bound the corresponding error via

$$\begin{aligned}\varepsilon(m) &= \max_{z\in K}\left|\frac{1}{z} - q_m(z)\right| = \max_{z\in K}\left|\frac{P_{m+1}(z)}{z}\right| \\ &\le \frac{\max_{z\in K}|P_{m+1}(z)|}{\min_{z\in K}|z|} = \|A^{-1}\|_2 \cdot \frac{1}{|C_{\frac{m+1}{2}}(t_0)|}.\end{aligned}$$

With Lemma 2.2

$$C_{\frac{m+1}{2}}(t_0) = C_{\frac{m+1}{2}}\left(1 + 2\frac{a^2}{b^2-a^2}\right) \ge \frac{1}{2}\left(\sqrt{\frac{a^2}{b^2-a^2}} + \sqrt{\frac{a^2}{b^2-a^2}+1}\right)^{m+1}.$$

An easy calculation shows that

$$\left(\sqrt{\frac{a^2}{b^2-a^2}} + \sqrt{\frac{a^2}{b^2-a^2}+1}\right)^2 = \frac{\kappa(A)+1}{\kappa(A)-1},$$

since $\kappa(A) = \|A\|_2\,\|A^{-1}\|_2 = b/a$. Summarizing, we have

$$\varepsilon(m) \le Cq^{m+1}$$

for $m = 2p-1$ and

$$\varepsilon(m) \le \varepsilon(m-1) \le Cq^m$$

for $m = 2p$. As in the proof of Theorem 4.2, for $k \ne \ell$ we write $|k-\ell| = m\beta + s$ with $m \ge 0$ and $s \in \{1,\dots,\beta\}$. Then $|k-\ell| \le m\beta + \beta$ and therefore $\frac{|k-\ell|}{\beta} - 1 \le m$ which gives (19) for $k \ne \ell$. Since A is normal, the bound

$$|[A^{-1}]_{k\ell}| \le \|A^{-1}\|_2$$

holds for all k, ℓ and can be used to give (19) for the case $k = \ell$. □

With this result we easily obtain decay bounds for the skew-Hermitian case as well.

Corollary 4.4
Let A be skew-Hermitian and nonsingular with bandwidth β. Then the bound (19) of Theorem 4.3 holds for the entries $|[A^{-1}]_{k\ell}|$.

Proof
We can write A as $A = iB$ with $B = B^H$. Then

$$|[A^{-1}]_{k\ell}| = |[(iB)^{-1}]_{k\ell}| = |-i[B^{-1}]_{k\ell}| = |-i|\,|[B^{-1}]_{k\ell}| = |[B^{-1}]_{k\ell}|. \tag{21}$$

Since B fulfills the assumptions of Theorem 4.4 and $\kappa(A) = \kappa(B)$, we can apply the bound (19). □

We point out that the result of Corollary 4.4 can also be obtained by applying the transformation $t = 1 + 2\frac{a^2+z^2}{b^2-a^2}$ and then using the same techniques as in the proof of Theorem 4.3.

5. COMPARISON TO PREVIOUS BOUNDS FOR THE INVERSE AND NUMERICAL EXPERIMENTS

In this section, we compare the new decay bounds for the matrix inverse to the ones from [11]—see Theorem 3.1—for different classes of normal matrices and also illustrate the quality of the new bounds for an application problem from in quantum electrodynamics.

The first case we consider is that of a shifted skew-Hermitian matrix, i.e., $A = M + sI$, where $M = -M^H$ and $s \in \mathbb{R}$. In this case, the decay factor obtained in Theorem 3.1 is

$$q_1^{\frac{|k-\ell|}{\beta}}, \text{ where } q_1 = \sqrt{\frac{\kappa(A)-1}{\kappa(A)+1}}$$

and the constant factor is

$$C_1 = (2\beta+1)q_1^{-2}\|A^{-1}\|_2\kappa(A)\max\left\{1, \left(\frac{1+\kappa(A)}{\kappa(A)}\right)^2/2\right\}.$$

In contrast, the new bound from Theorem 4.2 gives the decay factor

$$\frac{1}{q^{|k-\ell|/\beta}-1}, \text{ where } q = e^{\mathrm{Re}(\operatorname{arcosh}(x))} > 1, \quad x = \frac{\lambda_1+\lambda_2}{\lambda_2-\lambda_1}$$

with constant

$$C = 2\|A^{-1}\|_2.$$

We begin by comparing the constants C and C_1 and observe that

$$\frac{C}{C_1} = \frac{q_1^2}{(\beta+\frac{1}{2})\kappa(A)\max\{1, (\frac{1+\kappa(A)}{\kappa(A)})^2/2\}}.$$

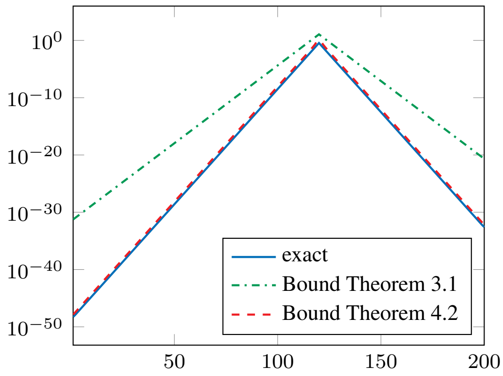
For large condition number $\kappa(A)$, the terms q_1^2 and $\max\{1, (\frac{1+\kappa(A)}{\kappa(A)})^2/2\}$ both tend to one, so that in this case the ratio between the constants becomes

$$\frac{C}{C_1} \approx \frac{1}{(\beta+\frac{1}{2})\kappa(A)},$$

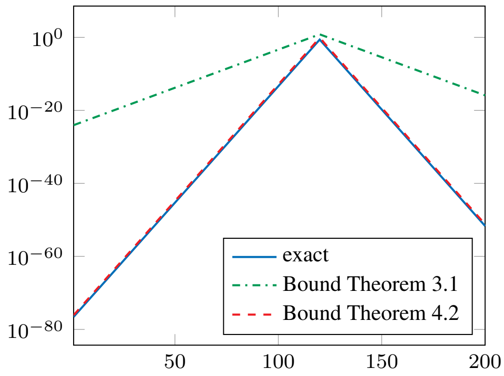
showing that the constant C of Theorem 4.2 is smaller by a factor which depends both on the bandwidth *and* the condition number of A.

Concerning the decay factors, we note that the decay factor obtained by our approach in the shifted skew-Hermitian case is in general independent of the condition number of $A = M + sI$ if the matrix M has eigenvalues both on the positive and negative imaginary axis. The line segment defining the decay factor in Theorem 4.2 only depends on the eigenvalue of largest modulus with negative imaginary part and the eigenvalue of largest modulus with positive imaginary part. The bound thus does not depend on the eigenvalue of A closest to the real axis (or, equivalently, the

DOI: 10.1002/nla



(a) A =tridiag$(1, i, -1)$, $s = 2$



(b) A =tridiag$(-1, 4, -1)$, $s = 2i$

Figure 3. Bounds for $(A + sI)^{-1}_{k\ell}$, $n = 200$, $\ell = 120$

eigenvalue of M closest to the origin). This is a clear advantage over the decay factor of Theorem 9, which depends on $\kappa(A)$. Therefore, we can expect the bound of Theorem 4.2 to be much sharper than that of Theorem 9, especially for matrices with large condition number.

In Figure 3(a), we compare the bounds for the tridiagonal matrix

$$A = \text{tridiag}(1, i, -1) + 2 \cdot I \in \mathbb{C}^{200\times 200},$$

by showing the exact decay behavior of the 120th column of A compared to the bounds obtained by Theorem 9 and 4.2, respectively. We observe that the decay predicted by Theorem 9 is too slow, while the bound of Theorem 4.2 almost completely agrees with the exact decay behavior, both in terms of decay rate and in terms of magnitude of the entries. As the above discussion applies in the same way for Hermitian matrices with complex shift, we also illustrate the bounds for this case in Figure 3(b) by repeating the experiment with the matrix

$$A = \text{tridiag}(1, 4, -1) + 2i \cdot I \in \mathbb{C}^{200\times 200}.$$

This time, the prediction of the decay rate from Theorem 9 is even worse, while the bound from Theorem 4.2 again agrees almost completely with the exact values.

In case of a Hermitian indefinite or a skew-Hermitian matrix (without shift) both Theorem 9 and Theorem 4.3 (or Corollary 4.4) give the decay factor

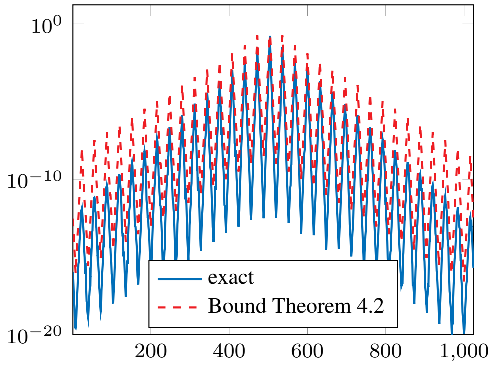
$$\sqrt{\frac{\kappa(A) - 1}{\kappa(A) + 1}},$$

so that our approach does not yield a better decay rate than Theorem 9. However, we still have the advantage (exactly as discussed before for the shifted case) that the constant in the decay bound of Theorem 4.3 is smaller by a factor that depends both on the bandwidth and the condition number of A. So the new approach using complex Chebyshev polynomials is still preferable, although in this case it does not improve over the classical results as much as before.

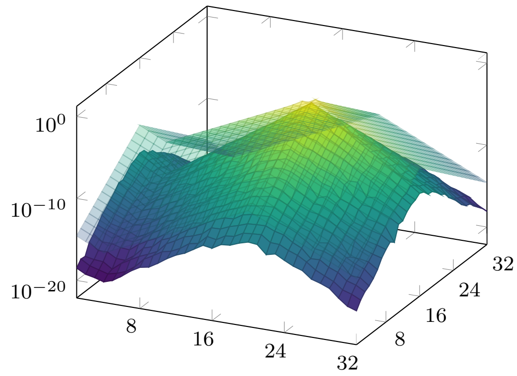
In order to also show a numerical example for a non-banded sparse matrix, we consider the staggered Schwinger discretization arising in quantum electrodynamics, the basic quantum field theory for the interaction of electrons and photons according to the standard model of Theoretical Physics. The discretization we are considering here uses a periodic two-dimensional lattice, where at each lattice site $x = (i, j)$ the unknown $\psi_{i,j}$ couples with its direct neighbors as

$$\begin{aligned} &\mu\psi_{i,j} + u^1_{i,j}\psi_{i+1,j} + \eta_{i,j}u^2_{i,j}\psi_{i,j+1} - \overline{u}^1_{i-1,j}\psi_{i-1,j} - \eta_{i,j}\overline{u}^2_{i,j-1}\psi_{i,j-1} = \phi_{i,j}, \\ &\quad i, j = 1, \dots, N, \;\; \eta_{i,j} = (-1)^i. \end{aligned} \tag{22}$$

DOI: 10.1002/nla



(a) Lexicographic ordering



(b) Underlying $N\ timesN$ grid

Figure 4. Bounds for $(\mu I + D)^{-1}_{k\ell}$, $n = 1024$, $\ell = 504$ for the staggered Schwinger discretization.

Herein, the indices $i-1, i+1, j-1, j+1$ are to be understood modulo N to account for the periodicity. The numbers $u^1_{i,j}$ and $u^2_{i,j}$ represent the $SU(1)$ background field, i.e. they are randomly distributed complex numbers of modulus 1. Clearly, (22) results in a system

$$(\mu I + D)\psi = \phi,$$

with D skew-Hermitian. The spectrum of $D = -D^H$ is not only purely imaginary, but also symmetric with respect to the origin due to the odd-even structure of the coupling. The graph underlying D is the periodic $N \times N$ mesh, so that we are this time in the presence of the general case where the graph distance enters into the decay bounds. As the spectrum of $\mu I + D$ does not contain the origin for any $\mu > 0$, we can apply Theorem 4.2. For $N = 32$, Figure 4 shows the the exact decay for the 504th column of the inverse corresponding to the point $(16, 24)$ on the mesh and the bounds given by (17). The left part arranges the values according to a one-dimensional, lexicographic ordering of the mesh points, whereas the right part gives the same information arranged on the underlying two-dimensional mesh.

6. DECAY BOUNDS FOR CAUCHY–STIELTJES FUNCTIONS

Recalling (12), the definition of a Stieltjes function, we can obtain bounds for Cauchy–Stieltjes matrix functions of banded (and, more generally, sparse) normal matrices by exploiting decay bounds for $(A + \tau I)^{-1}$, the inverses of shifted versions of A. The following theorem shows that it is possible to obtain explicit bounds this way, i.e., bounds in which no integrals appear anymore, for *any* Stieltjes function, similar to what was done in [6] for the special case of the inverse square root.

Theorem 6.1
Let f be a Cauchy–Stieltjes function of the form (12) and A Hermitian positive definite with bandwidth β. Then

$$|f(A)_{k\ell}| \leq 2f(\lambda_{\min})\, q^{\frac{|k-\ell|}{\beta}} \text{ with } q = \frac{\sqrt{\kappa(A)}-1}{\sqrt{\kappa(A)}+1}. \tag{23}$$

Proof
From (14) we know that

$$|f(A)_{k\ell}| \leq q^{\frac{|k-\ell|}{\beta}} \int_0^\infty C(\tau)\, \mathrm{d}\gamma(\tau),$$

where $C(\tau)$ is defined in (13). The second argument of the maximum in (13) is bounded by

$$\frac{(1+\sqrt{\kappa(\tau)})^2}{2(\lambda_{\max}+\tau)} = \frac{1}{2(\lambda_{\max}+\tau)} + \frac{1}{\sqrt{(\lambda_{\min}+\tau)(\lambda_{\max}+\tau)}} + \frac{1}{2(\lambda_{\min}+\tau)} \leq \frac{2}{\lambda_{\min}+\tau},$$

which obviously is also an upper bound for the first argument in (13), such that

$$\int_0^\infty C(\tau)\,\mathrm{d}\gamma(\tau) \leq \int_0^\infty \frac{2}{\lambda_{\min}+\tau}\,\mathrm{d}\gamma(\tau) = 2f(\lambda_{\min}).$$

□

The simple bound (23) holds for general Cauchy–Stieltjes functions. In addition a straightforward but lengthy calculation, which we refrain from presenting here, shows that (23) is sharper than (15) for the case of the inverse square root, $f(z) = z^{-1/2}$.

With the help of the following lemma we can give similar bounds for other classes of normal matrices with $\sigma(A) \subset \mathbb{C} \setminus \mathbb{R}_0^-$, like (shifted) skew-Hermitian matrices.

Lemma 6.2
Let $A = M + sI$ with $M^H = -M$ and $s \geq 0$. Define

$$\widehat{\lambda} := \operatorname{argmin}_{\lambda\in\sigma(A)}|\lambda|.$$

Then

$$\|(A+\tau I)^{-1}\|_2 \leq \frac{\sqrt{2}}{|\widehat{\lambda}|+\tau} \text{ for } \tau \in \mathbb{R}_0^+. \tag{24}$$

Proof
Since A and thus $A + \tau I$ is (shifted) skew-Hermitian, we have

$$\|(A+\tau I)^{-1}\|_2 = \frac{1}{\min_{\lambda\in\sigma(A)}|\lambda+\tau|} = \frac{1}{|\widehat{\lambda}+\tau|}.$$

The function $g(\tau) = \frac{|\widehat{\lambda}|+\tau}{|\widehat{\lambda}+\tau|}$ attains its maximum at $\tau = |\widehat{\lambda}|$, hence

$$g(\tau) \leq \frac{2|\widehat{\lambda}|}{|\widehat{\lambda}+|\widehat{\lambda}|\,|} = \frac{2}{\left|\frac{\widehat{\lambda}}{|\widehat{\lambda}|}+1\right|} \leq \frac{2}{\sqrt{2}} = \sqrt{2},$$

where the second inequality holds since $\operatorname{Re}(\widehat{\lambda}) \geq 0$. The assertion now follows because $g(\tau) \leq \sqrt{2}$ implies $\frac{1}{|\widehat{\lambda}+\tau|} \leq \frac{\sqrt{2}}{|\widehat{\lambda}|+\tau}$. □

Theorem 6.3
Let A be skew-Hermitian and nonsingular with bandwidth β and let f be a Cauchy–Stieltjes function of the form (12). Define $m = \lfloor\frac{|k-\ell|-1}{\beta}\rfloor$. Then

$$|[f(A)]_{k\ell}| \leq 2\sqrt{2}\,f(\|A^{-1}\|_2^{-1}) \cdot \begin{cases} q^{\frac{|k-\ell|}{\beta}} & \text{for } k \neq \ell \text{ and } m \text{ odd} \\ q^{\frac{|k-\ell|}{\beta}-1} & \text{for } k \neq \ell \text{ and } m \text{ even} \\ \frac{1}{2} & \text{for } k = \ell \end{cases}$$

with

$$q = \sqrt{\frac{\kappa(A)-1}{\kappa(A)+1}}.$$

Prepared using nlaauth.cls
DOI: 10.1002/nla

Proof
We know from Corollary 4.4 that

$$|[f(A)]_{k\ell}| = \left|\int_0^\infty [(A+\tau I)^{-1}]_{k\ell}\,\mathrm{d}\gamma(\tau)\right| \le \int_0^\infty 2\,\|(A+\tau I)^{-1}\|_2\, q(\tau)^{\frac{|k-\ell|}{\beta}}\,\mathrm{d}\gamma(\tau) \tag{25}$$

with

$$q(\tau) = \sqrt{\frac{\kappa(A+\tau I)-1}{\kappa(A+\tau I)+1}}$$

for $k \neq \ell$ and m odd. Let $a := \min_{\lambda\in\sigma(A)}|\lambda|$ and $b := \max_{\lambda\in\sigma(A)}|\lambda|$. Then for $\tau \ge 0$

$$\frac{\kappa(A+\tau I)-1}{\kappa(A+\tau I)+1} = \frac{\sqrt{\frac{\tau^2+b^2}{\tau^2+a^2}}-1}{\sqrt{\frac{\tau^2+b^2}{\tau^2+a^2}}+1} = 1-\frac{2}{\sqrt{\frac{\tau^2+b^2}{\tau^2+a^2}}+1} \le 1-\frac{2}{\frac{b}{a}+1} = \frac{\kappa(A)-1}{\kappa(A)+1},$$

the inequality holding because $\frac{\tau^2+b^2}{\tau^2+a^2} = 1+\frac{b^2-a^2}{\tau^2+a^2}$ is monotonically decreasing in τ. This gives $q(0) \ge q(\tau)$ for all $\tau \ge 0$, and from (25) we therefore obtain

$$|[f(A)]_{k\ell}| \le q^{\frac{|k-\ell|}{\beta}} \int_0^\infty 2\,\|(A+\tau I)^{-1}\|_2\,\mathrm{d}\gamma(\tau),$$

with $q = q(0)$, and with Lemma 6.2 it follows

$$\int_0^\infty 2\,\|(A+\tau I)^{-1}\|_2\,\mathrm{d}\gamma(\tau) \le \int_0^\infty \frac{2\sqrt{2}}{a+\tau}\,\mathrm{d}\gamma(\tau) = 2\sqrt{2}f(a) = 2\sqrt{2}f(\|A^{-1}\|_2^{-1}).$$

The other cases can be shown in an analogous manner. □

A similar result can be obtained for normal matrices where the spectrum is contained in a line segment $[\lambda_1, \lambda_2]$.

Theorem 6.4
Let A be a normal matrix with bandwidth β and $\sigma(A) \subset [\lambda_1, \lambda_2]$, where $[\lambda_1, \lambda_2] \cap \mathbb{R}_0^- = \emptyset$. Then there exist $\gamma \in [1, \infty)$ and $\tau^* \in \mathbb{R}_0^+$ such that for all $k \neq \ell$

$$|[f(A)]_{k\ell}| \le \int_0^\infty \|(A+\tau I)^{-1}\|_2\,\frac{2}{q(\tau)^{|k-\ell|/\beta}-1}\,\mathrm{d}\gamma(\tau) \tag{26}$$

$$\le \gamma\, f(\|A^{-1}\|_2^{-1})\frac{2}{q(\tau^*)^{|k-\ell|/\beta}-1} \tag{27}$$

where

$$q(\tau) = e^{\mathrm{Re}(\operatorname{arcosh}(x(\tau)))} > 1 \qquad \text{and} \qquad x(\tau) = \frac{\lambda_1+\lambda_2+2\tau}{\lambda_2-\lambda_1}.$$

For $k = \ell$

$$|[f(A)]_{k\ell}| \le \gamma\, f(\|A^{-1}\|_2^{-1}). \tag{28}$$

Proof
The inequality (26) immediately follows by applying Theorem 4.2 to $A+\tau I$. In particular, $[\lambda_1, \lambda_2] \cap \mathbb{R}_0^- = \emptyset$ implies that $0 \notin [\lambda_1+\tau, \lambda_2+\tau]$ for all $\tau \in \mathbb{R}_0^+$, so $q(\tau) > 1$ for all such τ.
Postponing the proof to the end, let us assume that we already know that $q(\tau)$ has a minimum on $\mathbb{R}_0^+$ and denote $\tau^* = \operatorname{argmin}_{\tau\in\mathbb{R}_0^+} q(\tau)$. Then

$$\int_0^\infty \|(A+\tau I)^{-1}\|_2\,\frac{2}{q(\tau)^{|k-\ell|/\beta}-1}\,\mathrm{d}\gamma(\tau) \le \frac{2}{q(\tau^*)^{|k-\ell|/\beta}-1}\int_0^\infty \|(A+\tau I)^{-1}\|_2\,\mathrm{d}\gamma(\tau).$$

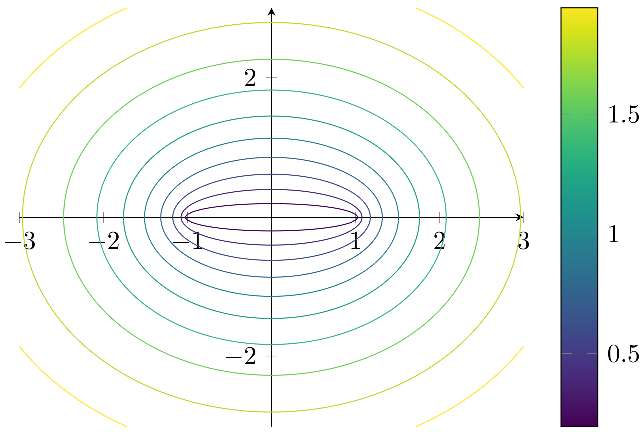


Figure 5. Contour lines of the function $h(z) = \mathrm{Re}(\operatorname{arcosh}(z))$.

Define $|\widehat{\lambda}| := \min_{\lambda\in\sigma(A)} |\lambda| = \|A^{-1}\|_2^{-1}$. Then

$$\|(A+\tau I)^{-1}\|_2 = \frac{1}{\min_{\lambda\in\sigma(A)} |\lambda+\tau|} \leq \frac{\gamma}{|\widehat{\lambda}|+\tau}$$

for some $\gamma \in [1,\infty)$ as can be seen from the equivalent formulation

$$g(\tau) := \frac{|\widehat{\lambda}|+\tau}{\min_{\lambda\in\sigma(A)} |\lambda+\tau|} \leq \gamma,$$

where this upper bound γ exists since g is continuous, $g(0) = 1$ and $\lim_{\tau\to\infty} g(\tau) = 1$.

Overall, we have the estimate

$$\begin{aligned}\int_0^\infty \|(A+\tau I)^{-1}\|_2 \,\mathrm{d}\gamma(\tau) &\leq \int_0^\infty \frac{\gamma}{\|A^{-1}\|_2^{-1}+\tau} \,\mathrm{d}\gamma(\tau)\\ &= \gamma\, f(\|A^{-1}\|_2^{-1}).\end{aligned}$$

It remains to show that the minimizer τ^* of $q(\tau)$ on $\mathbb{R}_0^+$ exists. Since the exponential is monotonically increasing on $\mathbb{R}$, we have

$$\tau^* = \operatorname*{argmin}_{\tau\in\mathbb{R}_0^+} q(\tau) = \operatorname*{argmin}_{\tau\in\mathbb{R}_0^+} \mathrm{Re}(\operatorname{arcosh}(x(\tau))).$$

We investigate the function

$$h : \mathbb{C} \to \mathbb{R}, \quad h(z) = \mathrm{Re}(\operatorname{arcosh}(z)). \tag{29}$$

Figure 5 shows the contour lines of h, which are confocal ellipses with focal points -1 and 1, i.e., ellipses E_ρ with $\rho \geq 1$. The values of h on these ellipses are monotonically increasing with increasing ρ. With

$$t(z,\tau) = \frac{\lambda_1+\lambda_2-2(z-\tau)}{\lambda_2-\lambda_1}$$

the shifted line segments $[\lambda_1+\tau, \lambda_2+\tau]$ are transformed to the interval $[-1,1]$, and $x(\tau) = t(0,\tau)$ maps the negative real axis to a half-line in the complex plane. Figure 6 shows the line segment $[\lambda_1, \lambda_2]$ and the corresponding transformation $x(\tau)$ of $\mathbb{R}_0^-$. This half-line does not intersect $[-1,1]$ and thus also can not be arbitrarily close to $[-1,1]$. Further,

$$\lim_{\tau\to\infty} h(x(\tau)) = \infty.$$

DOI: 10.1002/nla

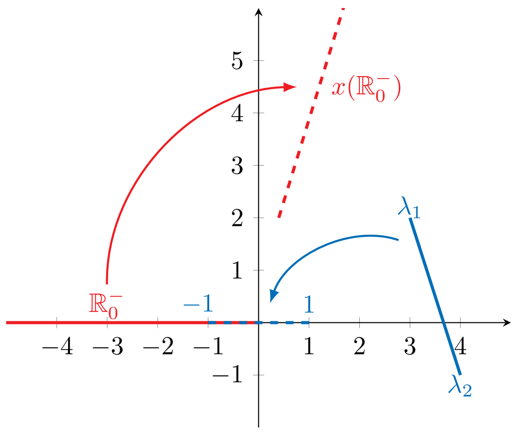


Figure 6. Transformation of the line segment $[\lambda_1, \lambda_2] = [3 + 2i, 4 - i]$ to the interval $[-1, 1]$ and corresponding transformation of the negative real axis.

Thus $q(\tau)$ can not be arbitrarily close to 1, and

$$\lim_{\tau\to\infty} q(\tau) = \infty,$$

which shows that the function $q : [0, \infty) \to (1, \infty)$ must have a minimum $q(\tau^*)$. $\square$

Calculating γ and τ^* explicitly might be too expensive in general but Theorem 6.4 shows that the integral (26) exists and can thus, e.g., be approximated numerically. In the next theorem we show that τ^* can be calculated easily for shifted skew-Hermitian matrices and Hermitian matrices with a complex shift.

Theorem 6.5
Let A be β-banded and of the form $A = M + i \cdot sI$, where $M = M^H$, or $A = M + sI$, where $M = -M^H$, with $s \in \mathbb{R}$. Further let $\sigma(A) \subset [\lambda_1, \lambda_2]$ with $[\lambda_1, \lambda_2] \cap \mathbb{R}_0^- = \emptyset$. Then there exists $\gamma \in [1, \infty)$ such that for all $k \neq \ell$

$$|[f(A)]_{k\ell}| \leq \gamma\, f(\|A^{-1}\|_2^{-1}) \frac{2}{q(\tau^*)^{|k-\ell|/\beta} - 1} \tag{30}$$

with

$$\tau^* = \max\left\{0, -\tfrac{\mathrm{Re}(\lambda_1)+\mathrm{Re}(\lambda_2)}{2}\right\} \tag{31}$$

where $q(\tau)$ is defined in Theorem 6.4. For $k = \ell$

$$|[f(A)]_{k\ell}| \leq \gamma\, f(\|A^{-1}\|_2^{-1}). \tag{32}$$

Proof
The assertion is proven by showing

$$\operatorname*{argmin}_{\tau\in\mathbb{R}_0^+} \mathrm{Re}(\mathrm{arcosh}(x(\tau))) = \max\left\{0, -\tfrac{\mathrm{Re}(\lambda_1)+\mathrm{Re}(\lambda_2)}{2}\right\}.$$

First let A be Hermitian with complex shift $i \cdot s$, i.e., $\mathrm{Im}(\lambda_1) = \mathrm{Im}(\lambda_2) = s$. Then the imaginary part of $x(\tau)$ is constant and

$$\mathrm{Re}(x(\tau)) = \frac{\mathrm{Re}(\lambda_1) + \mathrm{Re}(\lambda_2) + 2\tau}{\mathrm{Re}(\lambda_2) - \mathrm{Re}(\lambda_1)},$$

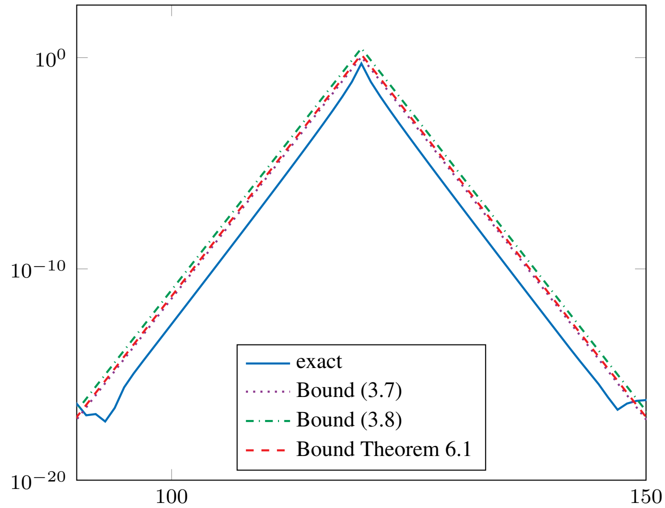


Figure 7. Bounds for $[A^{-1/2}]_{k\ell}$, $\ell = 120$ for the matrix $A = \text{tridiag}(-1, 4, -1)$.

hence $x(\mathbb{R}_0^+)$ is a half-line which is parallel to the real axis. Therefore, if $x(\mathbb{R}_0^+)$ does not cross the imaginary axis, then $\mathrm{Re}(\operatorname{arcosh}(x(\tau)))$ attains it minimal value for $\tau = 0$. Otherwise $q(\tau)$ is minimal for the intersection with the imaginary axis, i.e., for $\mathrm{Re}(x(\tau)) = 0$, which is the case when $\tau = -(\mathrm{Re}(\lambda_1) + \mathrm{Re}(\lambda_2))/2$. Thus, $x(\mathbb{R}_0^+)$ intersects the imaginary axis if and only if $-(\mathrm{Re}(\lambda_1) + \mathrm{Re}(\lambda_2))/2 \geq 0$.

Now consider the shifted skew-Hermitian case. Then $\mathrm{Re}(\lambda_1) = \mathrm{Re}(\lambda_2) = s$ and the real part of $x(\tau)$ is constant. For the imaginary part we have

$$\mathrm{Im}(x(\tau)) = \frac{2s + 2\tau}{\mathrm{Im}(\lambda_1) - \mathrm{Im}(\lambda_2)}.$$

With similar arguments as above, $q(\tau)$ is either minimal if $\tau = 0$ or at the intersection of $x(\mathbb{R}_0^+)$ with the real axis, i.e., if $\tau = -s = -(\mathrm{Re}(\lambda_1) + \mathrm{Re}(\lambda_2))/2 \geq 0$. □

When imposing further conditions on A, it is possible to specify the constant γ. For instance, if $A = M + sI$ with $M^H = -M$ and $s \in \mathbb{R}^+$, we know from Lemma 6.2 that $\gamma = \sqrt{2}$, hence the entries of A are bounded by

$$|[f(A)]_{k\ell}| \leq f(\|A^{-1}\|_2^{-1}) \frac{2\sqrt{2}}{q(0)^{|k-\ell|/\beta} - 1} \tag{33}$$

for $k \neq \ell$. For Cauchy–Stieltjes functions of the form $f(z) = z^{-\alpha}$ with $\alpha \in (0,1)$ the bound (33) provides bounds for shifted Hermitian matrices $A = M + i \cdot sI$, $M^H = M$ as well, since there exists a shifted skew-Hermitian matrix B with $A = iB$ and

$$|[A^{-\alpha}]_{k\ell}| = |[(iB)^{-\alpha}]_{k\ell}| = |i^{-\alpha}[B^{-\alpha}]_{k\ell}| = |[B^{-\alpha}]_{k\ell}|$$

.

7. NUMERICAL EXPERIMENTS FOR CAUCHY–STIELTJES FUNCTIONS

In this section we illustrate some of the bounds obtained in section 6 for Cauchy–Stieltjes functions. We begin by investigating the Hermitian positive definite case, where we can compare our bounds to those from [6]. In particular, we begin by comparing the decay bound from Theorem 6.1 for $A^{-1/2}$

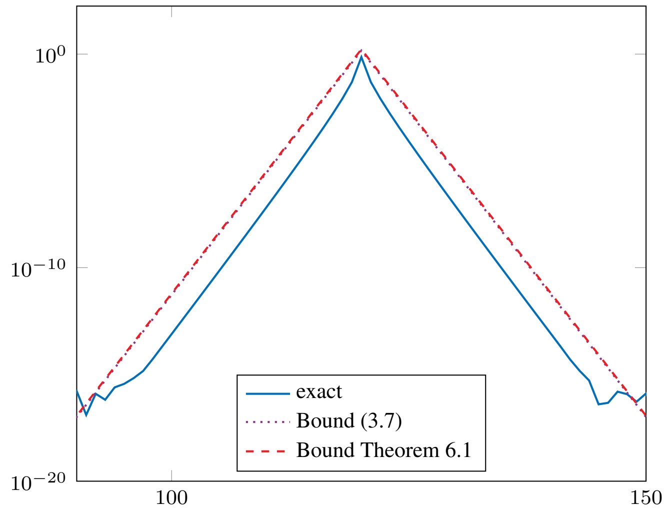
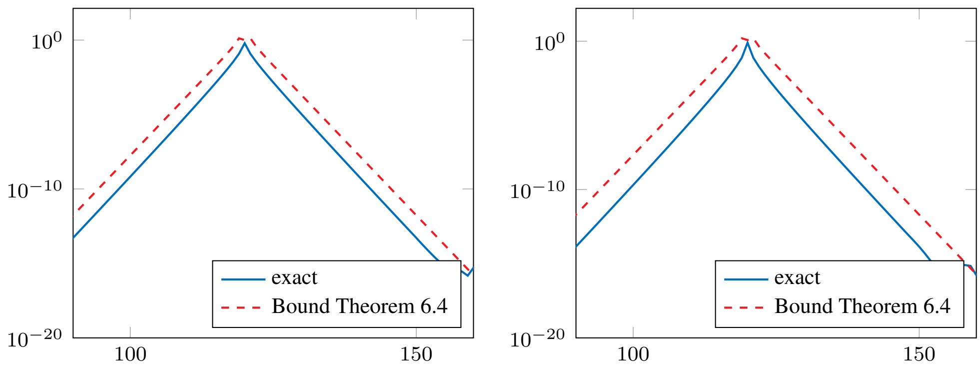


Figure 8. Bounds for $[A^{-1/4}]_{k\ell}$, $\ell = 120$ for the matrix $A = \text{tridiag}(-1, 4, -1)$.



(a) $f(z) = z^{-1/2}$ (b) $f(z) = z^{-1/4}$

Figure 9. Bounds for $[A^{-1/2}]_{k\ell}$ (left) and $[A^{-1/4}]_{k\ell}$ (right), $\ell = 120$ for the matrix $A = \text{tridiag}(1, i, -1) + 2 \cdot I$.

to the bound (15), using the matrix

$$A = \text{tridiag}(-1, 4, -1) \in \mathbb{C}^{200 \times 200}.$$

The exact values of the 120th column of $A^{-1/2}$ as well as the bounds obtained by (14) (evaluated by numerical quadrature), (15) and Theorem 6.1 are depicted in Figure 7. Of course, all approaches give the same decay rate, but the constant obtained in Theorem 6.1 is slightly smaller (and thus better) than the one in (15), giving a sharper bound, which almost agrees with the quadrature based bound (14) (which is the sharpest of the bounds, as the explicit bounds are obtained by bounding and estimating the terms appearing in the integral in (14)). We repeat the experiment, replacing the function $f(z) = z^{-1/2}$ by $f(z) = z^{-1/4}$, see Figure 8. In this case, no explicit bound was obtained in [6], so that we can just compare the bound of Theorem 6.1 to the bound (14) which has to be evaluated by numerical quadrature. Of course, the quadrature based bound is again sharper than the bound from Theorem 6.1, but we see that the estimates we applied for obtaining an explicit bound do not significantly increase the value of the bound, so that we still obtain good results without needing to use numerical quadrature.

In a second series of experiments, we compute bounds for the entries of the matrix functions $A^{-1/2}$ and $A^{-1/4}$, where A is now the shifted skew-Hermitian matrix

$$A = \text{tridiag}(1, i, -1) + 2 \cdot I \in \mathbb{C}^{200\times 200},$$

that we already considered in section 5 for illustrating bounds for the inverse. For the shifted skew-Hermitian case, no bounds were provided in [6], so that we only compare our bound from Theorem 6.5 to the exact value. The results of these experiments are shown in Figure 9. We again observe a very good approximation of the actual decay rate and the magnitude of the entries is only slightly overestimated, giving sharp bounds overall.

8. CONCLUSION

We have presented new decay bounds for inverses and Cauchy–Stieltjes functions of several classes of banded (and, more generally, sparse), normal matrices. Our bounds are based on approximation properties of complex Chebyshev polynomials on ellipses, and we have shown both by theoretical considerations and numerical experiments that these new bounds often give a significant improvement over known bounds from the literature. Particularly good results were obtained for the classes of shifted skew-Hermitian matrices and Hermitian matrices with complex shift. Inverting shifted skew-Hermitian matrices is, e.g., needed in Hermitian/skew-Hermitian splitting methods [1, 2, 15], so that an interesting topic for future research is to use our bounds in this setting for constructing sparse approximations for the inverse.

ACKNOWLEDGEMENT

The authors would like to thank Karsten Kahl and Francesco Knechtli for explaining the staggered Schwinger discretization and providing the data used in the corresponding numerical example in section 5.

REFERENCES

1. Z.-Z. Bai, G. H. Golub, and M. K. Ng, *Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 603–626.
2. M. Benzi, *A generalization of the Hermitian and skew-Hermitian splitting iteration*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 360–374.
3. ———, *Localization in Matrix Computations: Theory and Applications*, tech. rep., 2016.
4. M. Benzi and G. H. Golub, *Bounds for the entries of matrix functions with applications to preconditioning*, BIT, 39 (1999), pp. 417–438.
5. M. Benzi and N. Razouk, *Decay bounds and $O(n)$ algorithms for approximating functions of sparse matrices*, Electron. Trans. Numer. Anal., 28 (2007), pp. 16–39.
6. M. Benzi and V. Simoncini, *Decay bounds for functions of Hermitian matrices with banded or Kronecker structure*, SIAM J. Matrix Anal. Appl., 36 (2015), pp. 1263–1282.
7. D. Bini, S. Dendievel, G. Latouche, and B. Meini, *Computing the exponential of large block-triangular block-toeplitz matrices encountered in fluid queues*, Linear Algebra Appl., 502 (2016), pp. 387–419.
8. D. A. Bini, G. Latouche, and B. Meini, *Numerical Methods for Structured Markov Chains*, Oxford University Press, 2005.
9. D. R. Bowler and T. Miyazaki, *$O(N)$ methods in electronic structure calculations*, Rep. Prog. Phys., 75 (2012), p. 036503.
10. S. Demko, *Inverses of band matrices and local convergence of spline projections*, SIAM J. Numer. Anal., 14 (1977), pp. 616–619.
11. S. Demko, W. F. Moss, and W. Smith, *Decay rates for inverses of banded matrices*, Math. Comput., 43 (1984), pp. 491–499.
12. V. Eijkhout and B. Polman, *Decay rates of inverses of banded M-matrices that are near to Toeplitz matrices*, Linear Algebra Appl., 109 (1988), pp. 247–277.
13. N. J. Ford, D. V. Savostyanov, and N. L. Zamarashkin, *On the decay of the elements of inverse triangular Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 35 (2014), pp. 1288–1302.
14. P.-L. Giscard, K. Lui, S. Thwaite, and D. Jaksch, *An exact formulation of the time-ordered exponential using path-sums*, J. Math. Phys., 56 (2015), p. 053503.
15. G. H. Golub and D. Vanderstraeten, *On the preconditioning of matrices with skew-symmetric splittings*, Numer. Algorithms, 25 (2000), pp. 223–239.
16. A. Iserles, *How large is the exponential of a banded matrix*, N. Z. J. Math., 29 (2000), pp. 177–192.

17. D. KERSHAW, *Inequalities on the elements of the inverse of a certain tridiagonal matrix*, Math. Comput., (1970), pp. 155–158.
18. L. LOPEZ AND A. PUGLIESE, *Decay behaviour of functions of skew-symmetric matrices*, Proceedings of HERCMA, (2005), pp. 22–24.
19. G. MEURANT, *A review on the inverse of symmetric tridiagonal and block tridiagonal matrices*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 707–728.
20. Y. SAAD, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, 2nd edition ed., 2003.

Prepared using nlaauth.cls
DOI: 10.1002/nla