Inexact Arnoldi residual estimates and decay properties for functions of non-Hermitian matrices

Published Version:
Inexact Arnoldi residual estimates and decay properties for functions of non-Hermitian matrices / Pozza S.; Simoncini V.. - In: BIT. - ISSN 0006-3835. - STAMPA. - 59:4(2019), pp. 969-986. [10.1007/s10543-019-00763-6]

Availability:
This version is available at: https://hdl.handle.net/11585/714895 since: 2020-01-18

Published:
DOI: http://doi.org/10.1007/s10543-019-00763-6

Inexact Arnoldi residual estimates and decay properties for functions of non-Hermitian matrices

Stefano Pozza[1,2] · **Valeria Simoncini**[3,4]

Abstract
This paper derives a priori residual-type bounds for the Arnoldi approximation of a matrix function together with a strategy for setting the iteration accuracies in the inexact Arnoldi approximation of matrix functions. Such results are based on the decay behavior of the entries of functions of banded matrices. Specifically, a priori decay bounds for the entries of functions of banded non-Hermitian matrices will be exploited, using Faber polynomial approximation. Numerical experiments illustrate the quality of the results.

Keywords Arnoldi algorithm · Inexact Arnoldi algorithm · Matrix functions · Faber polynomials · Decay bounds · Banded matrices

Mathematics Subject Classification 65F60 · 65F10

Communicated by Daniel Kressner.

This work has been supported by the FARB12SIMO grant, Università di Bologna, by INdAM-GNCS under the 2016 Project *Equazioni e funzioni di matrici con struttura: analisi e algoritmi*, by the INdAM-GNCS "Giovani ricercatori 2016" grant, and by Charles University Research program No. UNCE/SCI/023.

✉ Stefano Pozza
pozza@karlin.mff.cuni.cz

Valeria Simoncini
valeria.simoncini@unibo.it

[1] Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8, Prague, Czech Republic

[2] ISTI-CNR, Pisa, Italy

[3] Dipartimento di Matematica, Università di Bologna, Piazza di Porta San Donato 5, 40127 Bologna, Italy

[4] IMATI-CNR, Pavia, Italy

1 Introduction

Matrix functions have arisen as a reliable and a computationally attractive tool for solving a large variety of application problems; we refer the reader to [27] for a thorough discussion and references. Given a complex $n \times n$ matrix A and a sufficiently regular function f, we are interested in the approximation of the matrix function $f(A)$ times a vector $\mathbf{v}$, that is $f(A)\mathbf{v}$, where we assume that $\mathbf{v}$ has unit Euclidean norm. To this end, we consider the orthogonal projection onto a subspace $\mathscr{V}_m$ of dimension m much smaller than n, obtaining the approximation

$$f(A)\,\mathbf{v} \approx V_m f(H_m)\,\mathbf{w}, \tag{1.1}$$

with V_m an $n \times m$ matrix whose columns form an orthonormal basis of $\mathscr{V}_m$, $H_m = V_m^* A V_m$, and $\mathbf{w} = V_m^* \mathbf{v}$. In this paper, we will focus on the case in which $\mathscr{V}_m$ is the *Krylov subspace*

$$\mathscr{K}_m(A, \mathbf{v}) = \operatorname{span}\{\mathbf{v}, A\mathbf{v}, \ldots, A^{m-1}\mathbf{v}\}$$

and V_m is the orthogonal basis obtained by the Arnoldi algorithm; see, e.g., [27, chapter 13]. Arnoldi-type approximations for the matrix exponential have been deeply investigated, and estimates of the error norm $\|e^{-tA}\mathbf{v} - V_m e^{-tH_m}\mathbf{e}_1\|$ for A non-normal have been given for instance by Saad [38], by Lubich and Hochbruck in [28], and recently by Wang and Ye in [42,43]. Other methods related to the Arnoldi approximation can be found in [1,17,21,22] where *restarted* techniques are considered. Regarding rational Krylov approximations of matrix functions, we refer the reader to the review [25] and to the black-box rational Arnoldi variant given in [26].

When V_m is the output of the Arnoldi algorithm, H_m is an upper Hessenberg matrix; that is a banded matrix with zero elements below the second lower diagonal. It can be shown that under certain assumptions the elements of $f(H_m)$ below the main diagonal are characterized by a decay behavior. Indeed, given a square banded matrix B, the entries of the matrix function $f(B)$ for a sufficiently regular function f are characterized by a—typically exponential—decay pattern as they move away from the main diagonal. This phenomenon has been known for a long time, and it is at the basis of approximations and estimation strategies in many fields, from signal processing to quantum dynamics and multivariate statistics; for a detailed description of relevant problems and a more comprehensive list of application fields where capturing the decay is particularly important we refer the reader to [3,4,7]. The interest in *a priori* estimates that can accurately predict the decay rate of matrix functions has significantly grown in the past decades, and it has mainly focused on Hermitian matrices [5,7,9,11,12,18,35,44]; the inverse and exponential functions have been given particular attention, due to their relevance in numerical analysis and other fields. Upper bounds usually take the form

$$|(f(B))_{k,\ell}| \le c\rho^{|k-\ell|}, \tag{1.2}$$

where $\rho \in (0, 1)$; both ρ and c depend on the spectral properties of B and on the domain of f, while ρ also strongly depends on the bandwidth of B.

In the case of a banded Hermitian matrix B, bounds of the Arnoldi approximation have been used to obtain upper estimates showing the decay phenomenon occurring in the entries of $f(B)$; see for instance [7] for the exponential function. Here we will exploit this connection but in the reverse direction. More precisely, we will first derive decay bounds for the entries of banded non-Hermitian matrices. Then we will apply such bounds to the matrix function $f(H_m)$, with H_m the upper Hessenberg matrix given by the Arnoldi algorithm, obtaining a priori bounds for the quality of the approximation (1.1), when a residual-based measure is used; these bounds complement available ones in the already mentioned literature for the Arnoldi approximation. Furthermore, we will use the described bounds in the inexact Krylov approximation of matrix functions; in particular, the bounds can be used to devise a priori relaxing thresholds for the inexact matrix-vector multiplications with A, whenever A is not available explicitly. These last results generalize the theory developed for $f(z) = z^{-1}$ and for the eigenvalue problem in [40] and [39], respectively; see also [14,31].

The analysis of the decay pattern for banded *non-Hermitian* matrices is significantly harder than in the Hermitian case, especially for non-normal matrices. In [6] Benzi and Razouk addressed this challenging case for diagonalizable matrices. They developed a bound of the type (1.2), where c also contains the eigenvector matrix condition number. In [33] the authors derive several qualitative bounds, mostly under the assumption that A is diagonally dominant. The exponential function provides a special setting, which has been explored in [29] and in [42,43]. In all these last articles, and also in our approach, bounds on the decay pattern of banded non-Hermitian matrices are derived that avoid the explicit reference to the possibly large condition number of the eigenvector matrix. Specialized off-diagonal decay results have been obtained for certain normal matrices; see, e.g., [11,20,23], and [3] for analytic functions of banded matrices over C^*-algebras.

Starting with the pioneering work [13], most estimates for the decay behavior of the entries have relied on Chebyshev and Faber polynomials as technical tool, for two main reasons. Firstly, polynomials of banded matrices are again banded matrices, although the bandwidth increases with the polynomial degree; see Fig. 1 below for a typical example. Secondly, sufficiently regular matrix functions can be written in terms of Chebyshev and Faber series, whose polynomial truncations enjoy nice approximation properties for a large class of matrices, from which an accurate description of the matrix function entries can be deduced. Using Faber polynomials, we will present an original derivation of a family of bounds for functions of banded non-Hermitian matrices. Such family can be adapted to several cases, depending on the function properties and on the matrix spectral properties. Very similar bounds can be obtained combining Theorem 10 in [3] with Theorem 3.7 in [6]. Another similar bound is given in [33, Theorem 2.6] for the case of multi-banded matrices and in [42, Theorem 3.8] for the exponential case. We also refer the reader to [36], where the bounds presented here have been extended to matrices with a more general sparsity pattern. Our bounds and the ones just cited make use of some approximation of the field of values (numerical range) of a matrix. An accurate approximation can be computationally quite expensive unless some structural properties can be exploited, as is the case for instance for

Toeplitz matrices ([16, Section 3]) or for network adjacency matrices ([36, Section 5.3]). Fortunately, for our purposes not-too-accurate field of value approximations can suffice, limiting the computational costs.

The paper is organized as follows. In Sect. 2 we use Faber polynomials to give a bound that can be adapted to approximate the entries of several functions of banded matrices; as an example we consider the functions e^A and $e^{-\sqrt{A}}$. In Sect. 3 and its subsections we first show that the derived estimates can be used for a residual-type bound in the approximation of $f(A)\mathbf{v}$, for certain functions f by means of the Arnoldi algorithm. Then we describe how to employ this bound to reliably estimate the quality of the approximation when in the Arnoldi iteration the accuracy in the matrix-vector product is relaxed. Numerical experiments illustrate the quality of the bounds. We conclude with some remarks in Sect. 4 and with technical proofs in the "Appendix".

All our numerical experiments were performed using Matlab (R2013b) [34]. In all our experiments, the computation of the field of values employed the code in [10].

2 Decay bounds for functions of banded matrices

We begin recalling the definition of matrix function and some of its properties. Matrix functions can be defined in several ways (see [27, section 1]). For our presentation, it is helpful to introduce the definition that employs the Cauchy integral formula.

Definition 2.1 Let $A \in \mathbb{C}^{n\times n}$ and f be an analytic function on some open $\Omega \subset \mathbb{C}$. Then

$$f(A) = \int_{\Gamma} f(z)\,(zI - A)^{-1}\,\mathrm{d}z,$$

where $\Gamma \subset \Omega$ is a Jordan curve (or a finite collection of Jordan curves) enclosing the eigenvalues of A exactly once, with mathematical positive orientation.

When f is analytic, Definition 2.1 is equivalent to other common definitions; see [37, section 2.3].

For $\mathbf{v} \in \mathbb{C}^n$, we denote with $||\mathbf{v}||$ the Euclidean vector norm, and for any matrix $A \in \mathbb{C}^{n\times n}$, with $||A||$ the induced matrix norm; that is, $||A|| = \sup_{||\mathbf{v}||=1} ||A\mathbf{v}||$. $\mathbb{C}^+$ denotes the open right-half complex plane. Moreover, we recall that the *field of values* (or *numerical range*) of A is defined as the set $W(A) = \{\mathbf{v}^* A\mathbf{v} \,|\, \mathbf{v} \in \mathbb{C}^n, ||\mathbf{v}|| = 1\}$, where $\mathbf{v}^*$ is the conjugate transpose of $\mathbf{v}$. We remark that the field of values of a matrix is a bounded convex subset of $\mathbb{C}$. Throughout the paper, $\sqrt{z}$ stands for the principal square root of $z \in \mathbb{C}$. Analogously $\sqrt{A}$ indicates the principal square root of the matrix A, which exists and is unique when A has no eigenvalues in $\mathbb{R}^-$; see, e.g., [27, Theorem 1.29].

The (k, ℓ) element of a matrix A is denoted by $(A)_{k,\ell}$. The set of banded matrices is defined as follows.

Definition 2.2 The notation $\mathscr{B}_n(\beta, \gamma)$ defines the set of banded matrices $A \in \mathbb{C}^{n\times n}$ with upper bandwidth $\beta \geq 0$ and lower bandwidth $\gamma \geq 0$, i.e., $(A)_{k,\ell} = 0$ for $\ell - k > \beta$ and $k - \ell > \gamma$.

$$A=\begin{bmatrix} * & * & * & & & & & \\ * & * & * & * & & & & \\ & * & * & * & * & & & \\ & & * & * & * & * & & \\ & & & * & * & * & * & \\ & & & & * & * & * & * \\ & & & & & * & * & * \\ & & & & & & * & * \end{bmatrix},\quad A^2=\begin{bmatrix} * & * & * & * & * & & & \\ * & * & * & * & * & * & & \\ * & * & * & * & * & * & * & \\ & * & * & * & * & * & * & * \\ & & * & * & * & * & * & * \\ & & & * & * & * & * & * \\ & & & & * & * & * & * \\ & & & & & * & * & * \end{bmatrix},\quad A^3=\begin{bmatrix} * & * & * & * & * & * & * & \\ * & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & * \\ & * & * & * & * & * & * & * \\ & & * & * & * & * & * & * \\ & & & * & * & * & * & * \\ & & & & * & * & * & * \end{bmatrix},$$

Fig. 1 Typical fill-in pattern of powers of a banded matrix $A \in \mathscr{B}_n(2, 1)$

We observe that if $A \in \mathscr{B}_n(\beta, \gamma)$ with $\beta, \gamma \neq 0$, for

$$\xi := \begin{cases} \lceil (\ell - k)/\beta \rceil, & \text{if } k < \ell \\ \lceil (k - \ell)/\gamma \rceil, & \text{if } k \geq \ell \end{cases} \tag{2.1}$$

it holds that

$$(A^m)_{k,\ell} = 0, \quad \text{for every } m < \xi; \tag{2.2}$$

see Fig. 1 for a typical fill-in pattern of A^m.

This characterization of banded matrices is a classical fundamental tool to prove the decay property of matrix functions, as sufficiently regular functions can be expanded in power series. Since we are interested in nontrivial banded matrices, in the following we shall assume that both β and γ are nonzero.

Faber polynomials extend the theory of power series to sets different from the disk, and can be effectively used to bound the entries of matrix functions. Let E be a continuum (i.e., a non-empty, compact and connected subset of $\mathbb{C}$) with connected complement. Then by Riemann's mapping theorem there exists a function ϕ that maps the exterior of E conformally onto $\{|z| > 1\}$ and such that

$$\phi(\infty) = \infty, \quad \lim_{z \to \infty} \frac{\phi(z)}{z} = d > 0.$$

Hence, ϕ can be expressed by a Laurent expansion $\phi(z) = dz + a_0 + \frac{a_1}{z} + \frac{a_2}{z^2} + \cdots$. Furthermore, for every $n > 0$ we have

$$(\phi(z))^n = dz^n + a_{n-1}^{(n)} z^{n-1} + \cdots + a_0^{(n)} + \frac{a_{-1}^{(n)}}{z} + \frac{a_{-2}^{(n)}}{z^2} + \cdots .$$

Then the Faber polynomial for the domain E is defined by (see, e.g., [41])

$$\Phi_n(z) = dz^n + a_{n-1}^{(n)} z^{n-1} + \cdots + a_0^{(n)}, \quad \text{for } n \geq 0.$$

If f is analytic on E, then it can be expanded in a series of Faber polynomials for E; that is,

$$f(z) = \sum_{j=0}^{\infty} f_j \Phi_j(z), \quad \text{for } z \in E;$$

see [41, Theorem 2, p. 52]. If the spectrum of A is contained in E and f is a function analytic in E, then the matrix function $f(A)$ can be expanded as follows (see, e.g., [41, p. 272])

$$f(A) = \sum_{j=0}^{\infty} f_j \Phi_j(A).$$

If, in addition, E contains the field of values $W(A)$, then for $n \geq 1$ we get

$$\|\Phi_n(A)\| \leq 2, \tag{2.3}$$

by Beckermann's Theorem 1.1 in [2].

By using the properties of Faber polynomials, in the following theorem we derive decay bounds for a large class of matrix functions. Notice that the estimate in [3, Theorem 10] combined with the results presented in [6, Theorem 3.7] results in similar bounds (see also [19]); moreover, in section 2 of [33], and in particular in Theorem 2.6, analogous results are discussed. Another similar bound can be found in [42, Theorem 3.8] for the case $f(z) = e^z$. The derivation we describe differs from the ones listed above by using inequality (2.3).

Theorem 2.3 *Let $A \in \mathscr{B}_n(\beta, \gamma)$ with field of values contained in a convex continuum E. Moreover, let ϕ be the conformal map sending the exterior of E onto the exterior of the unit disk, and let ψ be its inverse. For any $\tau > 1$ such that f is analytic on the level set G_τ defined as the complement of the set $\{\psi(z) \ : \ |z| > \tau\}$, it holds*

$$\left|(f(A))_{k,\ell}\right| \leq 2\frac{\tau}{\tau - 1} \max_{|z|=\tau} |f(\psi(z))| \left(\frac{1}{\tau}\right)^{\xi},$$

with ξ defined by (2.1).

Proof Properties (2.2) and (2.3) imply

$$|(f(A))_{k,\ell}| = \left|\sum_{j=0}^{\infty} f_j \left(\Phi_j(A)\right)_{k,\ell}\right| = \left|\sum_{j=\xi}^{\infty} f_j \left(\Phi_j(A)\right)_{k,\ell}\right| \leq 2\sum_{j=\xi}^{\infty} |f_j|,$$

where the Faber coefficients f_j are given by (see, e.g., [41, chapter III,Theorem 1])

$$f_j = \frac{1}{2\pi i} \int_{|z|=\tau} \frac{f(\psi(z))}{z^{j+1}} \, \mathrm{d}z.$$

Noticing that $|f_j| \le \frac{1}{(\tau)^j} \max_{|z|=\tau} |f(\psi(z))|$ gives

$$\left|(f(A))_{k,\ell}\right| \le 2 \max_{|z|=\tau} |f(\psi(z))| \sum_{j=\xi}^{\infty} \left(\frac{1}{\tau}\right)^j = 2\, \frac{\tau}{\tau-1} \max_{|z|=\tau} |f(\psi(z))| \left(\frac{1}{\tau}\right)^{\xi}.$$

□

The choice of τ in Theorem 2.3, and thus the sharpness of the derived estimate, depends on the trade-off between the possible large size of f on the given region, and the exponential decay of $(1/\tau)^{\xi}$, and thus it produces an infinite family of bounds depending on the problem considered. In our examples, we apply Theorem 2.3 to the approximation of the functions $f(z) = e^z$ and $f(z) = e^{-\sqrt{z}}$, with z in a properly chosen domain.

Corollary 2.4 *Let $A \in \mathscr{B}_n(\beta, \gamma)$ with field of values contained in a closed set E whose boundary is a horizontal ellipse with semi-axes $a \ge b > 0$ and center $c = c_1 + ic_2 \in \mathbb{C}$, $c_1, c_2 \in \mathbb{R}$. Then*

$$\left|\left(e^A\right)_{k,\ell}\right| \le 2e^{c_1} \frac{\xi + \sqrt{\xi^2 + a^2 - b^2}}{\xi + \sqrt{\xi^2 + a^2 - b^2} - (a+b)} \left(\frac{a+b}{\xi} \frac{e^{q(\xi)}}{1 + \sqrt{1 + (a^2 - b^2)/\xi^2}}\right)^{\xi},$$

for $\xi > b$, with $q(\xi) = 1 + \frac{a^2-b^2}{\xi^2+\xi\sqrt{\xi^2+a^2-b^2}}$ and ξ as in (2.1).

The proof is postponed to the "Appendix". Notice that for ξ large enough, the decay rate is of the form $((a+b)/(2\xi))^{\xi}$; that is, the decay is super-exponential. Moreover, in the Hermitian case, we can let $b \to 0$ in Corollary 2.4, thus obtaining a bound that is asymptotically equivalent—up to a multiplicative constant—to the one derived in [7, Theorem 4.2(ii)].

The function $f(z) = e^{-\sqrt{z}}$ is not analytic in the whole complex plane. This property has crucial effects on the approximation.

Corollary 2.5 *Let $A \in \mathscr{B}_n(\beta, \gamma)$ with field of values contained in a closed set $E \subset \mathbb{C}^+$, whose boundary is a horizontal ellipse with semi-axes $a \ge b > 0$ and center $c \in \mathbb{C}$. Then,*

$$\left|\left(e^{-\sqrt{A}}\right)_{k,\ell}\right| \le 2q_2(a,b,c) \left(\frac{a+b}{|c|} \frac{1}{|1 + \sqrt{1 - (a^2-b^2)/c^2}|}\right)^{\xi},$$

with ξ defined by (2.1) and

$$q_2(a,b,c) = \frac{\left|c + \sqrt{c^2 - (a^2 - b^2)}\right|}{\left|c + \sqrt{c^2 - (a^2 - b^2)}\right| - (a+b)}.$$

The proof is given in the "Appendix". If c is not real, then the bound in Corollary 2.5 can be further improved since the ellipses considered in the proof are not the maximal one.

Remark 2.6 For the sake of simplicity, in the previous corollaries horizontal ellipses were employed. However, more general convex sets E may be considered. The previous bounds will change accordingly, since the optimal value for τ in Theorem 2.3 does depend on the parameters associated with E. For instance, for the exponential function and a *vertical* ellipse, we can derive the same bound as in Corollary 2.4 by letting $b > a$. Notice that this is different from exchanging the role of a and b in the bound. The proof of this fact is non-trivial but technical, and it is not reported.

3 Residual bounds for Arnoldi and inexact Arnoldi methods

3.1 The Arnoldi method

Given a matrix $A \in \mathbb{C}^{n\times n}$ and a vector $\mathbf{v} \in \mathbb{C}^n$, for $m \geq 1$ the mth step of the Arnoldi algorithm determines an orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ for the Krylov subspace $\mathscr{K}_m(A, \mathbf{v})$, the subsequent orthonormal basis vector $\mathbf{v}_{m+1}$, an $m \times m$ upper Hessenberg matrix H_m, and a nonnegative scalar $h_{m+1,m}$ such that

$$AV_m = V_m H_m + h_{m+1,m}\mathbf{v}_{m+1}\mathbf{e}_m^T,$$

where $V_m = [\mathbf{v}_1, \dots, \mathbf{v}_m]$; note that $h_{m+1,m} = 0$ if and only if the algorithm stops, i.e., $\mathscr{K}_m(A, \mathbf{v})$ is an invariant subspace of A. Due to the orthogonality of the columns of $[V_m, v_{m+1}]$, the matrix H_m is the projection and restriction of A onto $\mathscr{K}_m(A, \mathbf{v})$; that is, $H_m = V_m^* A V_m$. Throughout the paper we assume exact arithmetic. As commonly performed, in our numerical computations we generated the matrix V_m by means of the *modified* Gram-Schmidt method with reorthogonalization, which ensures good orthogonality properties of the constructed basis in finite precision arithmetic; see, e.g., [24]. Without loss of generality assume that $\|\mathbf{v}\| = 1$. Then the Arnoldi approximation to $f(A)\mathbf{v}$ is given as $V_m f(H_m)\mathbf{e}_1$; see, e.g., [27, chapter 13]. The quantity

$$|\mathbf{e}_m^T f(H_m)\mathbf{e}_1| = |(f(H_m))_{m,1}|$$

– the last entry of the first column of $|f(H_m)|$ – is commonly employed to monitor the accuracy of the approximation $\|f(A)\mathbf{v} - V_m f(H_m)\mathbf{e}_1\|$; see, e.g., [38] and a related discussion in [30]. In the case of the exponential, $e^{-tA}\mathbf{v}$, the quantity

$$r_m(t) = |h_{m+1,m}\mathbf{e}_m^T e^{-tH_m}\mathbf{e}_1|$$

can be interpreted as the "residual" norm of an associated differential equation; see [8] and references therein. This interpretation can be shown to be true also for other functions; see, e.g., [15, section 6]). Indeed, assume that $\mathbf{y}(t) = f(tA)\mathbf{v}$ is the solution to the differential equation $y^{(d)} = Ay$ for some dth derivative, $d \in \mathbb{N}$ and specified initial conditions for $t = 0$. Let $\mathbf{y}_m(t) = V_m f(tH_m)\mathbf{e}_1 =: V_m \widehat{\mathbf{y}}_m(t)$. The vector $\widehat{\mathbf{y}}_m(t)$ is the solution to the projected differential equation $\widehat{\mathbf{y}}_m^{(d)} = H_m \widehat{\mathbf{y}}_m$ with initial condition $\widehat{\mathbf{y}}_m(0) = \mathbf{e}_1$. The differential equation residual $\mathbf{r}_m = A\mathbf{y}_m - \mathbf{y}_m^{(d)}$ can be used to monitor the accuracy of the approximate solution as follows: using the definition of $\mathbf{y}_m$ and the Arnoldi relation, we obtain

$$
\begin{aligned}
\mathbf{r}_m(t) &= A\mathbf{y}_m - \mathbf{y}_m^{(d)} = AV_m f(tH_m)\mathbf{e}_1 - \mathbf{y}_m^{(d)} \\
&= V_m H_m f(tH_m)\mathbf{e}_1 - V_m (f(tH_m))^{(d)}\mathbf{e}_1 + \mathbf{v}_{m+1} h_{m+1,m} \mathbf{e}_m^T f(tH_m)\mathbf{e}_1 \\
&= V_m (H_m \widehat{\mathbf{y}}_m - \widehat{\mathbf{y}}_m^{(d)}) + \mathbf{v}_{m+1} h_{m+1,m} \mathbf{e}_m^T f(tH_m)\mathbf{e}_1 \\
&= \mathbf{v}_{m+1} h_{m+1,m} \mathbf{e}_m^T f(tH_m)\mathbf{e}_1 .
\end{aligned}
$$

Therefore $r_m(t) = \|\mathbf{r}_m(t)\|$.

Without loss of generality, in the following we consider $t = 1$. Hence, for simplicity, we denote $r_m = r_m(1)$, and $\mathbf{r}_m = \mathbf{r}_m(1)$. We remark that the property $H_m = V_m^* A V_m$ ensures that the field of values of H_m is contained in that of A, so that our theory can be applied using A as reference matrix to individuate the spectral region of interest. We also remark that the inclusion of $h_{m+1,m}$ in $r_m(t)$ does not influence the actual behavior of the quantity. On the one hand, it holds that $h_{m+1,m} \le \|A\|$, so that $h_{m+1,m}$ could in principle be eliminated from the bound. On the other hand, $h_{m+1,m}$ is not going to be small, unless the Krylov subspace is close to an invariant subspace of A, so that $AV_m \approx V_m H_m$. The strength of Krylov subspaces precisely relies on being able to obtain good approximations to the sought after quantities far before an invariant subspace is determined. Hence our analysis is of interest for m such that the Krylov subspace is still far from being an invariant subspace of A, for which $h_{m+1,m}$ is not small. This implies that the behavior of $h_{m+1,m}\mathbf{e}_m^T f(tH_m)\mathbf{e}_1$ is fully determined by the quantity under examination; that is, $|\mathbf{e}_m^T f(tH_m)\mathbf{e}_1|$.

Let a, b be the semi-axes and $c = c_1 + ic_2$ the center of an elliptical region E containing the field of values of A. For the entry $(k, \ell) \equiv (m, 1)$ of $f(tH_m)$ and lower bandwidth $\beta = 1$ of H_m, the definition in (2.1) yields $\xi = m - 1$. Hence, from Corollary 2.4 and $m > b + 1$ we deduce the inequality

$$
|r_m| \le h_{m+1,m} 2e^{-c_1} p(m) \left(\frac{e^{q(m-1)}(a+b)}{m - 1 + \sqrt{(m-1)^2 + (a^2 - b^2)}} \right)^{m-1}, \qquad (3.1)
$$

with

$$
q(m-1) = 1 + \frac{(a^2 - b^2)}{(m-1)^2 + (m-1)\sqrt{(m-1)^2 + (a^2 - b^2)}}
$$

and

$$
p(m) = \frac{m - 1 + \sqrt{(m-1)^2 + (a^2 - b^2)}}{m - 1 + \sqrt{(m-1)^2 + (a^2 - b^2)} - (a+b)}.
$$

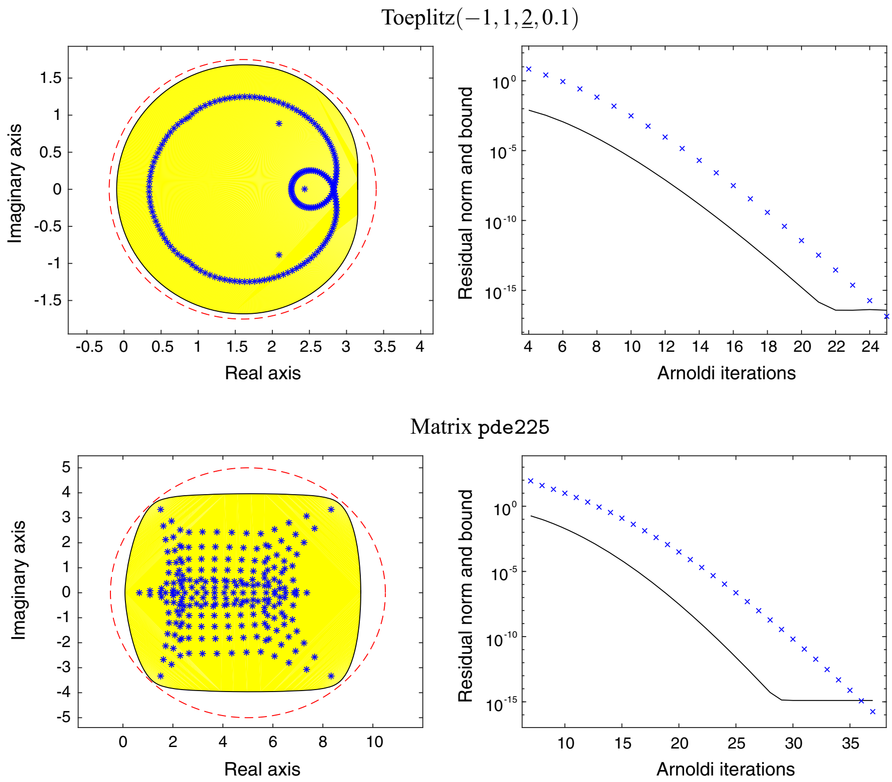

Fig. 2 Example 3.1. Approximation of $e^{-A}\mathbf{v}$, with $\mathbf{v} = (1, \ldots, 1)^T/\sqrt{n}$. Top: A = Toeplitz$(-1, 1, \underline{2}, 0.1) \in \mathscr{B}_{200}(1, 2)$. Bottom: matrix `pde225`. Left: $W(A)$ (yellow area), eigenvalues of A (blue asteriks), and enclosing ellipse E (red dashed line). Right: residual norm as the Arnoldi iteration proceeds in the approximation (black solid line), and residual bound in (3.1) (blue $\times$).

In [42,43], a similar bound is proposed, where, however, a continuum E with rectangular shape is considered, instead of the elliptical one we take in Corollary 2.4.

Example 3.1 Figure 2 shows the behavior of the bound in (3.1) for the residual of the Arnoldi approximation of $e^{-A}\mathbf{v}$ with $\mathbf{v} = (1, \ldots, 1)^T/\sqrt{n}$. The top plots refer to $A \in \mathscr{B}_{200}(1, 2)$ with Toeplitz structure, A = Toeplitz$(-1, 1, \underline{2}, 0.1)$, where the underlined element is on the diagonal, while the previous (resp. subsequent) values denote the lower (resp. upper) diagonal entries. The bottom plots refer to the matrix pde225 of the Matrix Market repository [32]. The left figure shows the field of values of the matrix A (yellow area), its eigenvalues (blue asteriks), and the horizontal ellipse used in the bound (red dashed line). On the right, we plot the residual associated with the Arnoldi approximation as the iteration proceeds (black solid line), and the corresponding values of the bound (blue "$\times$"). Matrix exponentials were computed by the `expm` Matlab function.

3.2 The inexact Arnoldi method

In an inexact Arnoldi procedure, A is not known exactly (we consider inexactness under the assumptions and in the context of [40]). This may be due for instance to the fact that A is only implicitly available via functional operations with a vector, which can be approximated at some accuracy. To proceed with our analysis, we can formalize this inexactness at each iteration k as

$$\widetilde{\mathbf{v}}_{k+1} = A\mathbf{v}_k + \mathbf{w}_k \approx A\mathbf{v}_k. \tag{3.2}$$

Typically, some form of accuracy criterion is implemented, so that $\|\mathbf{w}_k\| < \epsilon$ for some ϵ. In practice, a different value of this tolerance may be used at each iteration k, i.e., $\epsilon = \epsilon_k$; for this reason, in the following we assume that this tolerance depends on the iteration. The new vector $\widetilde{\mathbf{v}}_{k+1}$ is then orthonormalized with respect to the previous basis vectors to obtain $\mathbf{v}_{k+1}$. In compact form, the original Arnoldi relation becomes

$$(A + \mathscr{E}_m)V_m = V_m H_m + h_{m+1,m}\mathbf{v}_{m+1}\mathbf{e}_m^T, \quad \mathscr{E}_m = [\mathbf{w}_1, \ldots, \mathbf{w}_m]V_m^*.$$

Here H_m is again upper Hessenberg; however, $H_m = V_m^*(A + \mathscr{E}_m)V_m$. Moreover, $\mathscr{E}_m$ changes as m grows.

The quantities $\mathbf{y}_m = V_m f(H_m)\mathbf{e}_1$ and $\mathbf{r}_m = A\mathbf{y}_m - \mathbf{y}_m^{(d)}$ can still be defined as in the exact case; however the inexact Arnoldi relation should be considered to proceed further. Indeed,

$$\begin{aligned}
\mathbf{r}_m &= A\mathbf{y}_m - \mathbf{y}_m^{(d)} = AV_m f(H_m)\mathbf{e}_1 - \mathbf{y}_m^{(d)} && (3.3)\\
&= -\mathscr{E}_m V_m f(H_m)\mathbf{e}_1 + V_m H_m f(H_m)\mathbf{e}_1 - \mathbf{y}_m^{(d)} + \mathbf{v}_{m+1}h_{m+1,m}\mathbf{e}_m^T f(H_m)\mathbf{e}_1 \\
&= -[\mathbf{w}_1, \ldots, \mathbf{w}_m]f(H_m)\mathbf{e}_1 + \mathbf{v}_{m+1}h_{m+1,m}\mathbf{e}_m^T f(H_m)\mathbf{e}_1. && (3.4)
\end{aligned}$$

We can still define $r_m = |h_{m+1,m}\mathbf{e}_m^T f(H_m)\mathbf{e}_1|$, but we observe that now $r_m \neq \|\mathbf{r}_m\|$. Moreover, while r_m is computable, the quantity $\|\mathbf{r}_m\|$ is not available, since A is not known exactly. With the previous notation we can write $\|\mathbf{r}_m\| \leq |\|\mathbf{r}_m\| - r_m| + r_m$ where

$$|\|\mathbf{r}_m\| - r_m| \leq \|[\mathbf{w}_1, \ldots, \mathbf{w}_m]f(H_m)\mathbf{e}_1\|.$$

Therefore, if $\|[\mathbf{w}_1, \ldots, \mathbf{w}_m]f(H_m)\mathbf{e}_1\|$ is smaller than the tolerance for the final requested accuracy, then r_m provides a good measure in a computable stopping criterion.

Following a similar discussion in [39,40], we write

$$\|[\mathbf{w}_1, \ldots, \mathbf{w}_m]f(H_m)\mathbf{e}_1\| = \|\sum_{j=1}^{m} \mathbf{w}_j\mathbf{e}_j^T f(H_m)\mathbf{e}_1\| \leq \sum_{j=1}^{m} \|\mathbf{w}_j\|\,|\mathbf{e}_j^T f(H_m)\mathbf{e}_1|,$$

where $\|\mathbf{w}_j\| < \epsilon_j$. As a consequence, $\|[\mathbf{w}_1, \ldots, \mathbf{w}_m]f(H_m)\mathbf{e}_1\|$ is small when either $\|\mathbf{w}_j\|$ or $|\mathbf{e}_j^T f(H_m)\mathbf{e}_1|$ is small, and not necessarily both. By recalling the exponential decay of the entries of $f(H_m)\mathbf{e}_1$, $\|\mathbf{w}_j\|$ is in fact allowed to grow with j, in a way that is inversely proportional to the exponential decay of the corresponding entries of $f(H_m)\mathbf{e}_1$, without affecting the overall accuracy. A priori bounds on $|\mathbf{e}_j^T f(H_m)\mathbf{e}_1|$ can be used to select ϵ_j when estimating $A\mathbf{v}_j$. This relaxed strategy can significantly decrease the computational cost of matrix function evaluations whenever applying A accurately is expensive. However, notice that the field of values of H_m is contained in the field of values of $A+\mathscr{E}_m$. Hence if $W(A)$ is contained in an ellipse ∂E of semi-axes a, b and center c, then $W(A+\mathscr{E}_m) \subset W(A) + W(\mathscr{E}_m)$. Since

$$\sup_{\|z\|=1} |z^* \mathscr{E}_m z| \leq \sup_{\|z\|=1} \|\mathscr{E}_m z\| \leq \sqrt{\sum_{j=1}^{m} \|\mathbf{w}_j\|^2} \leq \sqrt{\sum_{j=1}^{m} \epsilon_j^2} =: \epsilon^{(m)},$$

the set $W(\mathscr{E}_m)$ is contained in the disk centered at the origin and radius $\epsilon^{(m)}$. Therefore $W(A) + W(\mathscr{E}_m)$ is contained in any set whose boundary has minimal distance from ∂E not smaller than $\epsilon^{(m)}$. One such set is contained in the ellipse ∂E_m with semi-axes $a(1+\epsilon^{(m)}/b)$, $b+\epsilon^{(m)}$ and center c. Indeed, $z \in \partial E_m$ can be parameterized as

$$z = \left(1 + \frac{\epsilon^{(m)}}{b}\right) \frac{\rho}{2} \left(Re^{i\theta} + \frac{1}{Re^{i\theta}}\right) + c, \quad 0 \leq \theta \leq 2\pi,$$

with $\rho = \sqrt{a^2 - b^2}$, $R = (a+b)/\rho$. The distance between z and the ellipse ∂E is

$$\left|\frac{\epsilon^{(m)}}{b} \frac{\rho}{2} \left(Re^{i\theta} + \frac{1}{Re^{i\theta}}\right)\right| \geq \left|\frac{\epsilon^{(m)}}{b} \frac{\rho}{2} \left(R - \frac{1}{R}\right)\right| = \epsilon^{(m)}.$$

With these definitions and notations we can introduce the following relaxation strategy for the inexactness in the Arnoldi procedure.

Theorem 3.2 *Let $\mathbf{r}_m$ be the (uncomputable) residual in (3.3) after m steps of the inexact Arnoldi algorithm and associated function f. Let $\epsilon^{(m)} > 0$ be the maximum allowed inexactness tolerance and let $tol > 0$.*

If for every $j \leq m$ we have $\|\mathbf{w}_j\| \leq \overline{\epsilon}_j$ with

$$\overline{\epsilon}_j = \begin{cases} \dfrac{tol}{m} \max\left\{1, \dfrac{1}{s_j}\right\}, & \text{if } \dfrac{tol}{m\, s_j} < \dfrac{1}{m-j+1}\sqrt{(\epsilon^{(m)})^2 - \sum_{k=1}^{j-1} \overline{\epsilon}_k^2} \\ \dfrac{1}{m-j+1}\sqrt{(\epsilon^{(m)})^2 - \sum_{k=1}^{j-1} \overline{\epsilon}_k^2}, & \text{otherwise} \end{cases} \tag{3.5}$$

then

$$|\|\mathbf{r}_m\| - r_m| \leq tol,$$

and $\left(\sum_{j=1}^{m} \overline{\epsilon}_j{}^2\right)^{\frac{1}{2}} \leq \epsilon^{(m)}$. Here s_j is the upper bound for $|\mathbf{e}_j^T f(H_m)\mathbf{e}_1|$ from Theorem 2.3 if j is such that this bound can be determined, otherwise $s_j = 1$; $W(A)$ in

Theorem 2.3 is contained in an ellipse with semiaxes $a \geq b > 0$ and center c, and E is the ellipse with semiaxes $a(1 + \epsilon^{(m)}/b)$, $b + \epsilon^{(m)}$ and center c.

The bound of Theorem 3.2 can be specialized for the functions $f(z) = e^z$ and $f(z) = e^{-\sqrt{z}}$ using respectively Corollaries 2.4 and 2.5.

In the following, we report on some experiments illustrating our findings. We consider the norm of the differential equation residual at time $t = 1$, that is

$$\|A\mathbf{y}_m - \mathbf{y}_m^{(d)}\|, \tag{3.6}$$

where $\mathbf{y}_m = V_m f(H_m)\mathbf{e}_1$ is computed with an inexact Arnoldi procedure. Clearly, the matrices V_m, H_m differ as we allow ϵ_j to vary at each iteration j. Hence, we compared two different strategies for chosing ϵ_j:

(i) A fixed small tolerance $\epsilon_j \equiv tol/m$ for all js, denoting the associated residual norm (3.6) by $||\mathbf{r}_j||$;
(ii) A variable accuracy $\epsilon_j := \bar{\epsilon}_j$ obtained from (3.5), denoting the associated residual norm in (3.6) by $||\bar{\mathbf{r}}_j||$.

We anticipate that our numerical experiments do not emphasize any visible degradation in the differential residual norm, if we relax the accuracy in the construction of the Krylov space as it is done in (ii) above, and the two residual norms stagnate at the same level.

Example 3.3 We consider the approximation of $\exp(-A)\mathbf{v}$ by the inexact Arnoldi procedure. The inexact matrix-vector product is implemented as in (3.2), with $\|\mathbf{w}_j\| = \epsilon_j$. Figure 3 reports our results for $\mathbf{v} = (1, \ldots, 1)^T/\sqrt{n}$ and the same matrices as in Example 3.1: $A = \text{Toeplitz}(-1, 1, \underline{2}, 0.1) \in \mathscr{B}_{200}(1, 2)$ (left), and `pde225` from the Matrix Market repository [32] (right). For this set of experiments, we considered $tol = 10^{-10}$ and $\epsilon^{(m)} = 10^{-1}$. The solid line shows the residual norm $||\mathbf{r}_j||$ as the iteration j proceeds for $\epsilon_j = tol/m$ (dashed line in the plot). The circles display the residual norm $\|\bar{\mathbf{r}}_j\|$ for the variable accuracy $\epsilon_j := \bar{\epsilon}_j$ (increasing asterisk curve in the plot) obtained from (3.5). The maximum number of iterations m was chosen as the smallest value for which the bound (3.1) is lower than tol, respectively $m = 20$ and $m = 31$. A larger, more conservative value could have been considered. The fields of values of the matrices can be obtained starting from those reported in the left plots of Fig. 2, where now the original semi-axes a, b of the elliptical sets considered for the computation of s_j are increased by $\epsilon^{(m)}/b$ and $\epsilon^{(m)}$ respectively. The plots show visually overlapping residual norm histories for the two choices of ϵ_j, illustrating that in practice no loss of information takes place when using the relaxation strategy.

Consider the second order differential equation $\mathbf{y}^{(2)} = A\mathbf{y}$, with $\mathbf{y}(0) = \mathbf{v}$. Its solution can be expressed as $\mathbf{y}(t) = \exp(-t\sqrt{A})\mathbf{v}$, and our results can be applied. This time the upper bound s_j for $|\mathbf{e}_m^T f(H_m)\mathbf{e}_1|$ is obtained from Corollary 2.5.

Example 3.4 For the same experimental setting as in Example 3.3, we consider approximating $\exp(-\sqrt{A})\mathbf{v}$, for the matrix $A = \text{Toeplitz}(-1, 1, \underline{3}, 0.1) \in \mathscr{B}_{200}(1, 2)$, the vector $\mathbf{v} = (1, \ldots, 1)^T/\sqrt{200}$ and $m = 35$ iterations ($W(A)$ is given by translating by 1 the field of values of the Toeplitz matrix in Example 3.1). Figure 4 reports

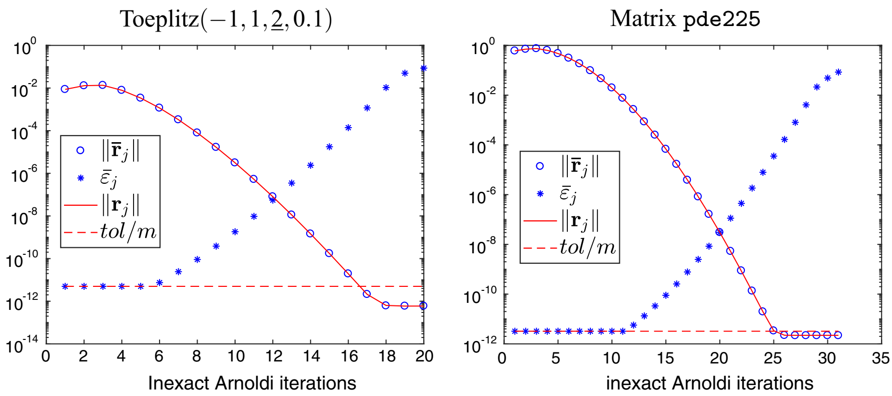

Fig. 3 Example 3.3, approximation of $e^{-A}\mathbf{v}$ with $\mathbf{v} = (1, \ldots, 1)^T/\sqrt{n}$. Residual norm $\|\mathbf{r}_j\|$ with constant accuracy $\epsilon_j = tol/m$, and residual norm $\|\bar{\mathbf{r}}_j\|$ with $\epsilon_j = \bar{\epsilon}_j$ by (3.5) as the inexact Arnoldi method proceeds. Left: For $A = \text{Toeplitz}(-1, 1, \underline{2}, 0.1) \in \mathscr{B}_{200}(1, 2)$. Right: For matrix `pde225` from the Matrix Market repository [32]

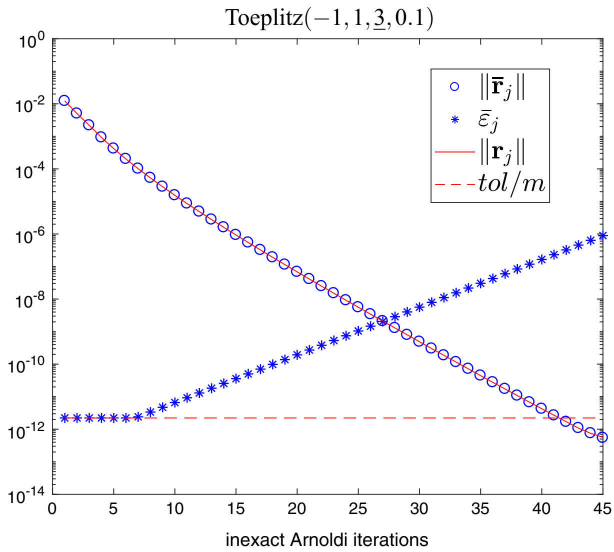

Fig. 4 Example 3.4. Approximation of $\exp(-\sqrt{A})\mathbf{v}$ with $A = \text{Toeplitz}(-1, 1, \underline{3}, 0.1) \in \mathscr{B}_{200}(1, 2)$ and $\mathbf{v} = (1, \ldots, 1)^T/\sqrt{n}$. The residual norm $||\mathbf{r}_j||$ is obtained with constant accuracy $\epsilon_j = tol/m$; the residual norm $\|\bar{\mathbf{r}}_j\|$ is obtained with $\epsilon_j = \bar{\epsilon}_j$ given by (3.5).

on our findings, with the same description as for the previous example. Here s_j in (3.5) is obtained from Corollary 2.5, and it is used to relax the accuracy ϵ_j. Similar considerations apply.

4 Conclusions

We have considered the approximation of $f(A)\mathbf{v}$ by means of the inexact Arnoldi method, in which matrix-vector products with A cannot be computed exactly. We have first derived computable bounds for the off-diagonal decay pattern of functions of non-Hermitian banded matrices. The accuracy of the bounds depends on the quality of the set enclosing and approximating the field of values of A. Then we have used these estimates to devise a new relaxation strategy for inexact matrix-vector operations, that does not influence the convergence of the residual norm in the matrix function approximation, while decreasing the computational cost for the inexact matrix-vector product. Similar results can be obtained for other Krylov-type approximations whose projection and restriction matrix H_m has a semi-banded structure. This is the case for instance of the Extended Krylov subspace approximation; see, e.g., [30] and references therein.

Acknowledgements We are indebted with Leonid Knizhnerman for a careful reading of an earlier version of this manuscript, and for his many insightful remarks which led to great improvements of our results. We also thank Michele Benzi for several suggestions and the two referees whose remarks helped us improve the presentation.

A Technical proofs

Proof of corollary 2.4

Let $\rho = \sqrt{a^2 - b^2}$ be the distance between the foci and the center of the ellipse (i.e., the boundary of E), and let $R = (a+b)/\rho$. Then a conformal map for E is

$$\phi(w) = \frac{w - c - \sqrt{(w-c)^2 - \rho^2}}{\rho R}, \tag{A.1}$$

and its inverse is

$$\psi(z) = \frac{\rho}{2}\left(Rz + \frac{1}{Rz}\right) + c\,; \tag{A.2}$$

see, e.g., [41, chapter II, Example 3]. Notice that

$$\max_{|z|=\tau} |e^{\psi(z)}| = \max_{|z|=\tau} e^{\Re(\psi(z))} = e^{\frac{\rho}{2}\left(R\tau + \frac{1}{R\tau}\right) + c_1}.$$

Hence by Theorem 2.3 we get

$$\left|\left(e^A\right)_{k,\ell}\right| \le 2\frac{\tau}{\tau - 1} e^{c_1} e^{\frac{\rho}{2}\left(R\tau + \frac{1}{R\tau}\right)} \left(\frac{1}{\tau}\right)^{\xi}.$$

The optimal value of $\tau > 1$ that minimizes $e^{\frac{\rho}{2}\left(R\tau+\frac{1}{R\tau}\right)}\left(\frac{1}{\tau}\right)^{\xi}$ is

$$\tau = \frac{\xi + \sqrt{\xi^2 + \rho^2}}{\rho R}.$$

Moreover, the condition $\tau > 1$ is satisfied if and only if $\xi > \frac{\rho}{2}\left(R - \frac{1}{R}\right) = b$. Finally, noticing that

$$\psi\left(\frac{\xi + \sqrt{\xi^2 + \rho^2}}{\rho R}\right) - c_1 = \frac{1}{2}\left(\xi + \sqrt{\xi^2 + \rho^2} + \frac{\rho^2}{\xi + \sqrt{\xi^2 + \rho^2}}\right) = \xi q(\xi),$$

and collecting ξ the proof is completed. □

Proof of corollary 2.5

The function $f(z) = \exp(-\sqrt{z})$ is analytic on $\mathbb{C} \setminus (-\infty, 0)$. Since we consider the principal square root, then $\Re(\sqrt{z}) \geq 0$, and

$$|\exp(-\sqrt{z})| = \exp(-\Re(\sqrt{z})) \leq 1.$$

Hence, by Theorem 2.3 we can determine τ for which

$$\left|\left(e^{-\sqrt{A}}\right)_{k,\ell}\right| \leq 2\frac{\tau}{\tau - 1}\left(\frac{1}{\tau}\right)^{\xi}.$$

For every $\varepsilon > 0$ close enough to zero, we set the parameter

$$\tau_\varepsilon = |\phi(\varepsilon)| = \left|\frac{c - \varepsilon + \sqrt{(c-\varepsilon)^2 - \rho^2}}{\rho R}\right|,$$

with $\phi(w)$ as in (A.1) and $\psi(z)$ its inverse (A.2). Then the ellipse $\{\psi(z),\ |z| = \tau_\varepsilon\}$ is contained in $\mathbb{C}\setminus(-\infty, 0]$. Letting $\varepsilon \to 0$ concludes the proof. □

References

1. Afanasjew, M., Eiermann, M., Ernst, O.G., Güttel, S.: Implementation of a restarted Krylov subspace method for the evaluation of matrix functions. Linear Algebra Appl. **429**(10), 2293–2314 (2008). https://doi.org/10.1016/j.laa.2008.06.029
2. Beckermann, B.: Image numérique, GMRES et polynômes de Faber. C. R. Math. Acad. Sci. Paris **340**(11), 855–860 (2005)
3. Benzi, M., Boito, P.: Decay properties for functions of matrices over C^*-algebras. Linear Algebra Appl. **456**, 174–198 (2014). https://doi.org/10.1016/j.laa.2013.11.027
4. Benzi, M., Boito, P., Razouk, N.: Decay properties of spectral projectors with applications to electronic structure. SIAM Rev. **55**(1), 3–64 (2013). https://doi.org/10.1137/100814019

5. Benzi, M., Golub, G.H.: Bounds for the entries of matrix functions with applications to preconditioning. BIT **39**(3), 417–438 (1999). https://doi.org/10.1023/A:1022362401426
6. Benzi, M., Razouk, N.: Decay bounds and O(n) algorithms for approximating functions of sparse matrices. Electron. Trans. Numer. Anal. **28**, 16–39 (2007)
7. Benzi, M., Simoncini, V.: Decay bounds for functions of Hermitian matrices with banded or Kronecker structure. SIAM J. Matrix Anal. Appl. **36**(3), 1263–1282 (2015). https://doi.org/10.1137/151006159
8. Botchev, M.A., Grimm, V., Hochbruck, M.: Residual, restarting and Richardson iteration for the matrix exponential. SIAM J. Sci. Comput. **35**(3), A1376–A1397 (2013)
9. Canuto, C., Simoncini, V., Verani, M.: On the decay of the inverse of matrices that are sum of Kronecker products. Linear Algebra Appl. **452**, 21–39 (2014). https://doi.org/10.1016/j.laa.2014.03.029
10. Cowen, C.C., Harel, E.: An Effective Algorithm for Computing the Numerical Range (1995). https://www.math.iupui.edu/ccowen/Downloads/33NumRange.html
11. Del Buono, N., Lopez, L., Peluso, R.: Computation of the exponential of large sparse skew-symmetric matrices. SIAM J. Sci. Comput. **27**(1), 278–293 (2005). https://doi.org/10.1137/030600758
12. Demko, S.G.: Inverses of band matrices and local convergence of spline projections. SIAM J. Numer. Anal. **14**(4), 616–619 (1977). https://doi.org/10.1137/0714041
13. Demko, S.G., Moss, W.F., Smith, P.W.: Decay rates for inverses of band matrices. Math. Comp. **43**(168), 491–499 (1984). https://doi.org/10.2307/2008290
14. Dinh, K.N., Sidje, R.B.: Analysis of inexact Krylov subspace methods for approximating the matrix exponential. Math. Comput. Simul. **138**, 1–13 (2017). https://doi.org/10.1016/j.matcom.2017.01.002. http://www.sciencedirect.com/science/article/pii/S0378475417300034
15. Druskin, V., Knizhnerman, L.: Krylov subspace approximation of eigenpairs and matrix functions in exact and computer arithmetic. Numer. Linear Algebra Appl. **2**(3), 205–217 (1995). https://doi.org/10.1002/nla.1680020303
16. Eiermann, M.: Fields of values and iterative methods. Linear Algebra Appl. **180**, 167–197 (1993). https://doi.org/10.1016/0024-3795(93)90530-2
17. Eiermann, M., Ernst, O.G., Güttel, S.: Deflated restarting for matrix functions. SIAM J. Matrix Anal. Appl. **32**(2), 621–641 (2011)
18. Eijkhout, V., Polman, B.: Decay rates of inverses of banded M-matrices that are near to Toeplitz matrices. Linear Algebra Appl. **109**, 247–277 (1988). https://doi.org/10.1016/0024-3795(88)90211-X. http://www.sciencedirect.com/science/article/pii/002437958890211X
19. Ellacott, S.W.: Computation of Faber series with application to numerical polynomial approximation in the complex plane. Math. Comp. **40**(162), 575–587 (1983)
20. Freund, R.: On polynomial approximations to $f_a(z) = (z-a)^{-1}$ with complex a and some applications to certain non-hermitian matrices. Approx. Theory Appl. **5**, 15–31 (1989)
21. Frommer, A., Güttel, S., Schweitzer, M.: Convergence of restarted Krylov subspace methods for Stieltjes functions of matrices. SIAM J. Matrix Anal. Appl. **35**(4), 1602–1624 (2014)
22. Frommer, A., Güttel, S., Schweitzer, M.: Efficient and stable Arnoldi restarts for matrix functions based on quadrature. SIAM J. Matrix Anal. Appl. **35**(2), 661–683 (2014). https://doi.org/10.1137/13093491X
23. Frommer, A., Schimmel, C., Schweitzer, M.: Bounds for the decay of the entries in inverses and Cauchy-Stieltjes functions of certain sparse, normal matrices. Numer. Linear Algebra Appl. **25**(4), e2131 (2018). https://doi.org/10.1002/nla.2131
24. Giraud, L., Langou, J., Rozložník, M., van den Eshof, J.: Rounding error analysis of the classical Gram-Schmidt orthogonalization process. Numer. Math. **101**(1), 87–100 (2005)
25. Güttel, S.: Rational Krylov approximation of matrix functions: numerical methods and optimal pole selection. GAMM-Mitt. **36**(1), 8–31 (2013). https://doi.org/10.1002/gamm.201310002
26. Güttel, S., Knizhnerman, L.: A black-box rational Arnoldi variant for Cauchy-Stieltjes matrix functions. BIT **53**(3), 595–616 (2013). https://doi.org/10.1007/s10543-013-0420-x
27. Higham, N.J.: Functions of Matrices: Theory and Computation. Society for Industrial and Applied Mathematics, Philadelphia (2008)
28. Hochbruck, M., Lubich, C.: On Krylov subspace approximations to the matrix exponential operator. SIAM J. Numer. Anal. **34**, 1911–1925 (1997)
29. Iserles, A.: How large is the exponential of a banded matrix? New Zealand J. Math. **29**, 177–192 (2000)
30. Knizhnerman, L., Simoncini, V.: A new investigation of the extended Krylov subspace method for matrix function evaluations. Numer. Linear Algebra Appl. **17**(4), 615–638 (2010). https://doi.org/10.1002/nla.652

31. Kürschner, P., Freitag, M.A.: Inexact methods for the low rank solution to large scale Lyapunov equations. arXiv preprint arXiv:1809.06903 (2018)
32. Matrix Market: A Visual Repository of Test Data for Use in Comparative Studies of Algorithms for Numerical Linear Algebra. Mathematical and Computational Sciences Division, National Institute of Standards and Technology; available online at http://math.nist.gov/MatrixMarket
33. Mastronardi, N., Ng, M., Tyrtyshnikov, E.E.: Decay in functions of multiband matrices. SIAM J. Matrix Anal. Appl. **31**(5), 2721–2737 (2010). https://doi.org/10.1137/090758374
34. The MathWorks, Inc.: MATLAB 7, r2013b edn. (2013)
35. Meurant, G.: A review on the inverse of symmetric tridiagonal and block tridiagonal matrices. SIAM J. Matrix Anal. Appl. **13**(3), 707–728 (1992). https://doi.org/10.1137/0613045
36. Pozza, S., Tudisco, F.: On the stability of network indices defined by means of matrix functions. SIAM J. Matrix Anal. Appl. **39**(4), 1521–1546 (2018)
37. Rinehart, R.F.: The equivalence of definitions of a matric function. Amer. Math. Monthly **62**(6), 395–414 (1955)
38. Saad, Y.: Analysis of some Krylov subspace approximations to the matrix exponential operator. SIAM J. Numer. Anal. **29**(1), 209–228 (1992). https://doi.org/10.1137/0729014
39. Simoncini, V.: Variable accuracy of matrix-vector products in projection methods for eigencomputation. SIAM J. Numer. Anal. **43**(3), 1155–1174 (2005)
40. Simoncini, V., Szyld, D.B.: Theory of inexact Krylov subspace methods and applications to scientific computing. SIAM J. Sci. Comput. **25**(2), 454–477 (2003)
41. Suetin, P.K.: Series of Faber polynomials. Gordon and Breach Science Publishers (1998). Translated from the 1984 Russian original by E. V. Pankratiev [E. V. Pankrat′ev]
42. Wang, H.: The Krylov Subspace Methods for the Computation of Matrix Exponentials. Ph.D. thesis, Department of Mathematics, University of Kentucky (2015)
43. Wang, H., Ye, Q.: Error bounds for the Krylov subspace methods for computations of matrix exponentials. SIAM J. Matrix Anal. Appl. **38**(1), 155–187 (2017). https://doi.org/10.1137/16M1063733
44. Ye, Q.: Error bounds for the Lanczos methods for approximating matrix exponentials. SIAM J. Numer. Anal. **51**(1), 68–87 (2013). https://doi.org/10.1137/11085935x