Linear Algebra and its Applications 456 (2014) 174–198

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Decay properties for functions of matrices over C^*-algebras

Michele Benzi [a,*,1], Paola Boito [b]

[a] Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA
[b] Equipe Calcul Formel, DMI-XLIM UMR 7252 Université de Limoges – CNRS, 123 avenue Albert Thomas, 87060 Limoges Cedex, France

ARTICLE INFO

Article history:
Received 26 July 2013
Accepted 17 November 2013
Available online 7 December 2013
Submitted by L. Rodman

MSC:
primary 47A60, 15A16
secondary 65F60

Keywords:
Matrix functions
C^*-algebras
Exponential decay
Sparse matrices
Graphs
Functional calculus

ABSTRACT

We extend previous results on the exponential off-diagonal decay of the entries of analytic functions of banded and sparse matrices to the case where the matrix entries are elements of a C^*-algebra.

1. Introduction

Decay properties of inverses, exponentials and other functions of band or sparse matrices over $\mathbb{R}$ or $\mathbb{C}$ have been investigated by several authors in recent years [3,4,7,8,12,27,36,37,39–41,44,49]. Such properties play an important role in various applications including electronic structure computations in quantum chemistry [6,15], quantum information theory [19,20,29,56], computational harmonic analysis [37,41], high-dimensional statistics [2], random matrix theory [51] and numerical analysis [13,61], to name a few.

* Corresponding author.
E-mail addresses: benzi@mathcs.emory.edu (M. Benzi), paola.boito@unilim.fr (P. Boito).
[1] Work supported in part by NSF Grant DMS 1115692.

0024-3795/$ – see front matter
http://dx.doi.org/10.1016/j.laa.2013.11.027

Answering a question posed by P.-L. Giscard and coworkers [35], we consider generalizations of existing decay estimates to functions of matrices with entries in more general algebraic structures than the familiar fields $\mathbb{R}$ or $\mathbb{C}$. In particular, we propose extensions to functions of matrices with entries from the following algebras:

1. Commutative algebras of complex-valued continuous functions.
2. Noncommutative algebras of bounded operators on a complex Hilbert space.
3. The real division algebra $\mathbb{H}$ of quaternions.

The theory of complex C^*-algebras provides the natural abstract setting for the desired extensions [43,47,55]. Matrices over such algebras arise naturally in various application areas, including parametrized linear systems and eigenproblems [18,59], differential equations [25], generalized moment problems [52], control theory [16,22,23], and quantum physics [10,11,28]. The study of matrices over C^*-algebras is also of independent mathematical interest; see, e.g., [38,42].

Using the holomorphic functional calculus, we establish exponential off-diagonal decay results for analytic functions of banded $n \times n$ matrices over C^*-algebras, both commutative and noncommutative. Our decay estimates are expressed in the form of computable bounds on the norms of the entries of $f(A)$ where $A = [a_{ij}]$ is an $n \times n$ matrix with entries $a_{ij} = a_{ji}^*$ in a C^*-algebra $\mathcal{A}_0$ and f is an analytic function defined on a suitable open subset of $\mathbb{C}$ containing the spectrum of A, viewed as an element of the C^*-algebra $\mathcal{M}_n(\mathcal{A}_0)$ $(= \mathcal{A}_0^{n\times n})$. The interesting case is when the constants in the bounds do not depend on n. Functions of more general sparse matrices over $\mathcal{A}_0$ will also be discussed.

For the case of functions of $n \times n$ quaternion matrices, we identify the set of such matrices with a (real) subalgebra of $\mathbb{C}^{2n\times 2n}$ and treat them as a special type of complex block matrices; as we will see, this will impose some restrictions on the type of functions that we are allowed to consider.

2. Background on C^*-algebras

In this section we provide definitions and notations used throughout the remainder of the paper, and recall some of the fundamental results from the theory of C^*-algebras. The reader is referred to [47,55] for concise introductions to this theory and to [43] for a more systematic treatment. See also [1,14] for applications of C^*-algebras to numerical analysis.

Recall that a *Banach algebra* is a complex algebra $\mathcal{A}_0$ with a norm making $\mathcal{A}_0$ into a Banach space and satisfying

$$\|ab\| \leqslant \|a\|\|b\|$$

for all $a, b \in \mathcal{A}_0$. In this paper we consider only *unital* Banach algebras, i.e., algebras with a multiplicative unit I with $\|I\| = 1$.

An *involution* on a Banach algebra $\mathcal{A}_0$ is a map $a \mapsto a^*$ of $\mathcal{A}_0$ into itself satisfying

(i) $(a^*)^* = a$,
(ii) $(ab)^* = b^*a^*$,
(iii) $(\lambda a + b)^* = \bar{\lambda}a^* + b^*$

for all $a, b \in \mathcal{A}_0$ and $\lambda \in \mathbb{C}$. A *C^*-algebra* is a Banach algebra with an involution such that the *C^*-identity*

$$\left\|a^*a\right\| = \|a\|^2$$

holds for all $a \in \mathcal{A}_0$. Note that we do not make any assumption on whether $\mathcal{A}_0$ is commutative or not.

Basic examples of C^*-algebras are:

1. The commutative algebra $C(\mathcal{X})$ of all continuous complex-valued functions on a compact Hausdorff space $\mathcal{X}$. Here the addition and multiplication operations are defined pointwise, and the

norm is given by $\|f\|_\infty = \max_{t\in\mathcal{X}} |f(t)|$. The involution on $C(\mathcal{X})$ maps each function f to its complex conjugate f^*, defined by $f^*(t) = \overline{f(t)}$ for all $t \in \mathcal{X}$.

2. The algebra $\mathcal{B}(\mathcal{H})$ of all bounded linear operators on a complex Hilbert space $\mathcal{H}$, with the operator norm $\|T\|_{op} = \sup \|Tx\|_{\mathcal{H}}/\|x\|_{\mathcal{H}}$, where the supremum is taken over all nonzero $x \in \mathcal{H}$. The involution on $\mathcal{B}(\mathcal{H})$ maps each bounded linear operator T on $\mathcal{H}$ to its adjoint, T^*.

Note that the second example contains as a special case the algebra $\mathcal{M}_n(\mathbb{C})$ $(= \mathbb{C}^{k\times k})$ of all $k \times k$ matrices with complex entries, with the norm being the usual spectral norm and the involution mapping each matrix $A = [a_{ij}]$ to its Hermitian conjugate $A^* = [\overline{a}_{ji}]$. This algebra is noncommutative for $k \geqslant 2$.

Examples 1 and 2 above provide, in a precise sense, the "only" examples of C^*-algebras. Indeed, every (unital) commutative C^*-algebra admits a faithful representation onto an algebra of the form $C(\mathcal{X})$ for a suitable (and essentially unique) compact Hausdorff space $\mathcal{X}$; and, similarly, every unital (possibly noncommutative) C^*-algebra can be faithfully represented as a norm-closed subalgebra of $\mathcal{B}(\mathcal{H})$ for a suitable complex Hilbert space $\mathcal{H}$.

More precisely, a map ϕ between two C^*-algebras is a $*$-*homomorphism* if ϕ is linear, multiplicative, and such that $\phi(a^*) = \phi(a)^*$; a $*$-*isomorphism* is a bijective $*$-homomorphism. Two C^*-algebras are said to be *isometrically* $*$-*isomorphic* if there is a norm-preserving $*$-isomorphism between them, in which case they are indistinguishable as C^*-algebras. A $*$-*subalgebra* $\mathcal{B}_0$ of a C^*-algebra is a subalgebra that is $*$-closed, i.e., $a \in \mathcal{B}_0$ implies $a^* \in \mathcal{B}_0$. Finally, a C^*-subalgebra is a norm-closed $*$-subalgebra of a C^*-algebra. The following two results are classical [32,33].

Theorem 1 *(Gelfand). Let $\mathcal{A}_0$ be a commutative C^*-algebra. Then there is a compact Hausdorff space $\mathcal{X}$ such that $\mathcal{A}_0$ is isometrically $*$-isomorphic to $C(\mathcal{X})$. If $\mathcal{Y}$ is another compact Hausdorff space such that $\mathcal{A}_0$ is isometrically $*$-isomorphic to $C(\mathcal{Y})$, then $\mathcal{X}$ and $\mathcal{Y}$ are necessarily homeomorphic.*

Theorem 2 *(Gelfand–Naimark). Let $\mathcal{A}_0$ be a C^*-algebra. Then there is a complex Hilbert space $\mathcal{H}$ such that $\mathcal{A}_0$ is isometrically $*$-isomorphic to a C^*-subalgebra of $\mathcal{B}(\mathcal{H})$.*

We will also need the following definitions and facts. An element $a \in \mathcal{A}_0$ of a C^*-algebra is *unitary* if $aa^* = a^*a = I$, *Hermitian* (or *self-adjoint*) if $a^* = a$, *skew-Hermitian* if $a^* = -a$, *normal* if $aa^* = a^*a$. Clearly, unitary, Hermitian and skew-Hermitian elements are all normal. Any element $a \in \mathcal{A}_0$ can be written uniquely as $a = h_1 + \mathrm{i}h_2$ with h_1, h_2 Hermitian and $\mathrm{i} = \sqrt{-1}$.

For any (complex) Banach algebra $\mathcal{A}_0$, the *spectrum* of an element $a \in \mathcal{A}_0$ is the set of all $\lambda \in \mathbb{C}$ such that $\lambda I - a$ is not invertible in $\mathcal{A}_0$. We denote the spectrum of a by $\sigma(a)$. For any $a \in \mathcal{A}_0$, the spectrum $\sigma(a)$ is a nonempty compact subset of $\mathbb{C}$ contained in the closed disk of radius $r = \|a\|$ centered at 0. The complement $r(a) = \mathbb{C}\backslash\sigma(a)$ of the spectrum of an element a of a C^*-algebra is called the *resolvent set* of a. The *spectral radius* of a is defined as $\rho(a) = \max\{|\lambda|;\ \lambda \in \sigma(A)\}$. *Gelfand's formula* for the spectral radius [32] states that

$$\rho(a) = \lim_{m\to\infty} \left\|a^m\right\|^{\frac{1}{m}}. \tag{1}$$

Note that this identity contains the statement that the above limit exists.

If $a \in \mathcal{A}_0$ (a C^*-algebra) is Hermitian, $\sigma(a)$ is a subset of $\mathbb{R}$. If $a \in \mathcal{A}_0$ is normal (in particular, Hermitian), then $\rho(a) = \|a\|$. This implies that if a is Hermitian, then either $-\|a\| \in \sigma(a)$ or $\|a\| \in \sigma(a)$. The spectrum of a skew-Hermitian element is purely imaginary, and the spectrum of a unitary element is contained in the unit circle $S^1 = \{z \in \mathbb{C};\ |z| = 1\}$.

An element $a \in \mathcal{A}_0$ is *nonnegative* if $a = a^*$ and the spectrum of a is contained in $\mathbb{R}_+$, the nonnegative real axis. Any linear combination with real nonnegative coefficients of nonnegative elements of a C^*-algebra is nonnegative; in other words, the set of all nonnegative elements in a C^*-algebra $\mathcal{A}_0$ form a (nonnegative) *cone* in $\mathcal{A}_0$. For any $a \in \mathcal{A}_0$, aa^* is nonnegative, and $I + aa^*$ is invertible in $\mathcal{A}_0$. Furthermore, $\|a\| = \sqrt{\rho(a^*a)} = \sqrt{\rho(aa^*)}$, for any $a \in \mathcal{A}_0$.

Finally, we note that if $\|\cdot\|_*$ and $\|\cdot\|_{**}$ are two norms with respect to which $\mathcal{A}_0$ is a C^*-algebra, then $\|\cdot\|_* = \|\cdot\|_{**}$.

3. Matrices over a C^*-algebra

Let $\mathcal{A}_0$ be a C^*-algebra. Given a positive integer n, let $\mathcal{A} = \mathcal{M}_n(\mathcal{A}_0)$ be the set of $n \times n$ matrices with entries in $\mathcal{A}_0$. Observe that $\mathcal{A}$ has a natural C^*-algebra structure, with matrix addition and multiplication defined in the usual way (in terms, of course, of the corresponding operations on $\mathcal{A}_0$). The involution is naturally defined as follows: given a matrix $A = [a_{ij}] \in \mathcal{A}$, the adjoint of A is given by $A^* = [a_{ji}^*]$. The algebra $\mathcal{A}$ is obviously unital, with unit

$$I_n = \begin{bmatrix} I & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & I \end{bmatrix}$$

where I is the unit of $\mathcal{A}_0$. The definition of unitary, Hermitian, skew-Hermitian and normal matrix are the obvious ones.

It follows from the Gelfand–Naimark representation theorem (Theorem 2 above) that each $A \in \mathcal{A}$ can be represented as a matrix T_A of bounded linear operators, where T_A acts on the direct sum $\mathscr{H} = \mathcal{H} \oplus \cdots \oplus \mathcal{H}$ of n copies of a suitable complex Hilbert space $\mathcal{H}$. This fact allows us to introduce an operator norm on $\mathcal{A}$, defined as follows:

$$\|A\| := \sup_{\|x\|_{\mathscr{H}}=1} \|T_A x\|_{\mathscr{H}}, \tag{2}$$

where

$$\|x\|_{\mathscr{H}} := \sqrt{\|x_1\|_{\mathcal{H}}^2 + \cdots + \|x_n\|_{\mathcal{H}}^2}$$

is the norm of an element $x = (x_1, \dots, x_n) \in \mathscr{H}$. Relative to this norm, $A \in \mathcal{A}$ is a C^*-algebra. Note that $\mathcal{A}$ can also be identified with the tensor product of C^*-algebras $\mathcal{A}_0 \otimes \mathcal{M}_n(\mathbb{C})$.

Similarly, Gelfand's Theorem (Theorem 1 above) implies that if $\mathcal{A}_0$ is commutative, there is a compact Hausdorff space $\mathcal{X}$ such that any $A \in \mathcal{A}$ can be identified with a continuous matrix-valued function

$$A : \mathcal{X} \longrightarrow \mathcal{M}_n(\mathbb{C}).$$

In other words, A can be represented as an $n \times n$ matrix of continuous, complex-valued functions: $A = [a_{ij}(t)]$, with domain $\mathcal{X}$. The natural C^*-algebra norm on $\mathcal{A}$, which can be identified with $C(\mathcal{X}) \otimes \mathcal{M}_n(\mathbb{C})$, is now the operator norm

$$\|A\| := \sup_{\|x\|=1} \|Ax\|, \tag{3}$$

where $x = (x_1, \dots, x_n) \in [C(\mathcal{X})]^n$ has norm $\|x\| = \sqrt{\|x_1\|_\infty^2 + \cdots + \|x_n\|_\infty^2}$ with $\|x_i\|_\infty = \max_{t \in \mathcal{X}} |x_i(t)|$, for $1 \leqslant i \leqslant n$.

Since $\mathcal{A}$ is a C^*-algebra, all the definitions and basic facts about the spectrum remain valid for any matrix A with entries in $\mathcal{A}_0$. Thus, the spectrum $\sigma(A)$ of $A \in \mathcal{A}$ is the set of all $\lambda \in \mathbb{C}$ such that $\lambda I_n - A$ is not invertible in $\mathcal{A}$. If $0 \in \sigma(A)$, we will also say that A is *singular*. The set $\sigma(A)$ is a nonempty compact subset of $\mathbb{C}$ completely contained in the disk of radius $\|A\|$ centered at 0. The definition of spectral radius and Gelfand's formula (1) remain valid. Hermitian matrices have real spectra, skew-Hermitian matrices have purely imaginary spectra, unitary matrices have spectra contained in S^1, and so forth. Note, however, that it is not true in general that a normal matrix A over a C^*-algebra can be unitarily diagonalized [42].

In general, it is difficult to estimate the spectrum of a matrix $A = [a_{ij}]$ over a C^*-algebra. It will be useful for what follows to introduce the *matricial norm* of A [53,54], which is defined as the $n \times n$ real nonnegative matrix

$$\hat{A} = \begin{bmatrix} \|a_{11}\| & \|a_{12}\| & \dots & \|a_{1n}\| \\ \|a_{21}\| & \|a_{22}\| & \dots & \|a_{nn}\| \\ \vdots & \ddots & \ddots & \vdots \\ \|a_{n1}\| & \|a_{n2}\| & \dots & \|a_{nn}\| \end{bmatrix}.$$

The following result shows that we can obtain upper bounds on the spectral radius and operator norm of a matrix A over a C^*-algebra in terms of the more easily computed corresponding quantities for $\hat{A}$. As usual, the symbol $\|\cdot\|_2$ denotes the spectral norm of a matrix with real or complex entries.

Theorem 3. *For any $A \in \mathcal{A}$, the following inequalities hold*:

1. $\|A\| \leqslant \|\hat{A}\|_2$;
2. $\rho(A) \leqslant \rho(\hat{A})$.

Proof. To prove the first item, observe that

$$\|A\| = \sup \|T_A x\|_{\mathscr{H}} = \sup \left[\sum_{i=1}^{n} \left\| \sum_{j=1}^{n} \pi(a_{ij}) x_j \right\|^2 \right]^{\frac{1}{2}},$$

where $\pi(a_{ij})$ is the Gelfand–Naimark representation of $a_{ij} \in \mathcal{A}_0$ as a bounded operator on $\mathcal{H}$, and the sup is taken over all n-tuples $(x_1, x_2, \dots, x_n) \in \mathscr{H}$ with $\|x_1\|_{\mathcal{H}}^2 + \|x_2\|_{\mathcal{H}}^2 + \cdots + \|x_n\|_{\mathcal{H}}^2 = 1$.

Using the triangle inequality and the fact that the Gelfand–Naimark map is an isometry we get

$$\|A\| \leqslant \sup \left[\sum_{i=1}^{n} \left(\sum_{j=1}^{n} \|a_{ij}\| \|x_j\|_{\mathcal{H}} \right)^2 \right]^{\frac{1}{2}},$$

or, equivalently,

$$\|A\| \leqslant \sup_{(\xi_1,\dots,\xi_n) \in \Xi^n} \left[\sum_{i=1}^{n} \left(\sum_{j=1}^{n} \|a_{ij}\| \xi_j \right)^2 \right]^{\frac{1}{2}},$$

where $\Xi^n := \{(\xi_1, \dots, \xi_n) \mid \xi_i \in \mathbb{R}_+ \ \forall i \text{ and } \sum_{i=1}^{n} \xi_i^2 = 1\}$. On the other hand,

$$\|\hat{A}\|_2 = \sup_{(\xi_1,\dots,\xi_n) \in S^n} \left[\sum_{i=1}^{n} \left| \sum_{j=1}^{n} \|a_{ij}\| \xi_j \right|^2 \right]^{\frac{1}{2}},$$

where S^n denotes the unit sphere in $\mathbb{C}^n$. Observing that $\Xi^n \subset S^n$, we conclude that $\|A\| \leqslant \|\hat{A}\|_2$.

To prove the second item we use the characterizations $\rho(A) = \lim_{m\to\infty} \|A^m\|^{\frac{1}{m}}$, $\rho(\hat{A}) = \lim_{m\to\infty} \|\hat{A}^m\|_2^{\frac{1}{m}}$ and the fact that $\|A^m\| \leqslant \|\widehat{A^m}\|_2$, which we just proved. A simple inductive argument shows that $\|\widehat{A^m}\|_2 \leqslant \|\hat{A}^m\|_2$ for all $m = 1, 2, \dots$, thus yielding the desired result. □

Remark 1. A version of item 1 of the previous theorem was proved by A. Ostrowski in [53], in the context of matrices of linear operators on normed vector spaces. Related results can also be found in [34] and [60].

Remark 2. If A is Hermitian or (shifted) skew-Hermitian, then $\hat{A}$ is real symmetric. In this case $\|A\| = \rho(A)$, $\|\hat{A}\|_2 = \rho(\hat{A})$ and item 2 reduces to item 1. On the other hand, in the more general case where A is normal, the matrix $\hat{A}$ is not necessarily symmetric or even normal and we obtain the bound

$$\|A\| = \rho(A) \leqslant \rho(\hat{A}),$$

which is generally better than $\|A\| \leqslant \|\hat{A}\|_2$.

Next, we prove a simple invertibility condition for matrices over the commutative C^*-algebra $C(\mathcal{X})$.

Theorem 4. *A matrix A over $\mathcal{A}_0 = C(\mathcal{X})$ is invertible in $\mathcal{A} = \mathcal{M}_n(\mathcal{A}_0)$ if and only if for each $t \in \mathcal{X}$ the $n \times n$ matrix $A(t) = [a_{ij}(t)]$ is invertible in $\mathcal{M}_n(\mathbb{C})$.*

Proof. The theorem will be proved if we show that

$$\sigma(A) = \bigcup\nolimits_{t\in\mathcal{X}} \sigma\big(A(t)\big). \tag{4}$$

Assume first that $\lambda \in \sigma(A(t_0))$ for some $t_0 \in \mathcal{X}$. Then $\lambda I_n - A(t_0)$ is not invertible, therefore $\lambda I_n - A(t)$ is not invertible for all $t \in \mathcal{X}$, and $\lambda I_n - A$ fails to be invertible as an element of $\mathcal{A} = \mathcal{M}_n(\mathcal{A}_0)$. This shows that

$$\bigcup\nolimits_{t\in\mathcal{X}} \sigma\big(A(t)\big) \subseteq \sigma(A).$$

To prove the reverse inclusion, we show that the resolvent sets satisfy

$$\bigcap\nolimits_{t\in\mathcal{X}} r\big(A(t)\big) \subseteq r(A).$$

Indeed, if $z \in r(A(t))$ for all $t \in \mathcal{X}$, then the matrix-valued function $t \mapsto (zI_n - A(t))^{-1}$ is well defined and necessarily continuous on $\mathcal{X}$. Hence, $z \in r(A)$. This completes the proof. □

Clearly, the set $\mathcal{K} = \{A(t);\ t \in \mathcal{X}\}$ is compact in $\mathcal{M}_n(\mathbb{C})$. The spectral radius, as a function on $\mathcal{M}_n(\mathbb{C})$, is continuous and thus the function $t \mapsto \rho(A(t))$ is continuous on the compact set $\mathcal{X}$. It follows that this function attains its maximum value for some $t_0 \in \mathcal{X}$. The equality (4) above then implies that

$$\rho(A) = \rho\big(A(t_0)\big) = \max_{t\in\mathcal{X}} \rho\big(A(t)\big).$$

If A is normal, we obviously have

$$\|A\| = \max_{t\in\mathcal{X}} \big\|A(t)\big\|_2. \tag{5}$$

Recalling that for any $A \in \mathcal{A}$ the norm satisfies $\|A\| = \sqrt{\rho(AA^*)} = \sqrt{\rho(A^*A)}$, we conclude that the identity (5) holds for any matrix A over the C^*-algebra $C(\mathcal{X})$.

As a special case, consider an $n \times n$ matrix with entries in $C(\mathcal{X})$, where $\mathcal{X} = [0, 1]$. Each entry $a_{ij} = a_{ij}(t)$ of A is a continuous complex-valued function of t. One can think of such an A in different ways. As a mapping of $[0, 1]$ into $\mathcal{M}_n(\mathbb{C})$, A can be regarded as a continuous curve in the space of all $n \times n$ complex matrices. On the other hand, A is also a point in the C^*-algebra of matrices over $C(\mathcal{X})$.

Theorem 4 then states that A is invertible if and only if the corresponding curve $\mathcal{K} = \{A(t);\ t \in [0, 1]\}$ does not intersect the set of singular $n \times n$ complex matrices, i.e., if and only if $\mathcal{K}$ is entirely contained in the group $\mathcal{GL}_n(\mathbb{C})$ of invertible $n \times n$ complex matrices.

Example 1. As a simple illustration of Theorem 3, consider the 2×2 Hermitian matrix over $C([0, 1])$:

$$A = A(t) = \begin{bmatrix} \mathrm{e}^{-t} & t^2 + 1 \\ t^2 + 1 & \mathrm{e}^{t} \end{bmatrix}.$$

Observing now that

$$\hat{A} = \begin{bmatrix} 1 & 2 \\ 2 & \mathrm{e} \end{bmatrix},$$

we obtain the bound $\rho(A) = \|A\| \leqslant \|\hat{A}\|_2 \approx 4.03586$.

By direct computation we find that $\det(\lambda I_2 - A(t)) = \lambda^2 - 2(\cosh t)\lambda - t^2(t^2+2)$ and thus the spectrum of $A(t)$ (as an element of the C^*-algebra of matrices over $C([0,1])$) consists of all the numbers of the form

$$\lambda_{\pm}(t) = \cosh t \pm \sqrt{\cosh^2 t + t^2\left(t^2+2\right)}, \quad 0 \leqslant t \leqslant 1$$

(a compact subset of $\mathbb{R}$). Also note that $\det(A(t)) = -t^2(t^2+2)$ vanishes for $t = 0$, showing that A is *not* invertible in $\mathcal{M}_n(C([0,1]))$.

Finding the maxima and minima over $[0,1]$ of the continuous functions $\lambda_-(t)$ and $\lambda_+(t)$ we easily find that the spectrum of $A(t)$ is given by

$$\sigma\left(A(t)\right) = [-0.77664, 0] \cup [2, 3.86280],$$

where the results have been rounded to five decimal digits. Thus, in this simple example $\|\hat{A}\|_2 = 4.03586$ gives a pretty good upper bound for the true value $\|A\| = 3.86280$.

4. The holomorphic functional calculus

The standard way to define the notion of an analytic function $f(a)$ of an element a of a C^*-algebra $\mathcal{A}_0$ is via contour integration. In particular, we can use this approach to define functions of a matrix A with elements in $\mathcal{A}_0$.

Let $f(z)$ be a complex function which is analytic in an open neighborhood U of $\sigma(a)$. Since $\sigma(a)$ is compact, we can always find a finite collection $\Gamma = \bigcup_{j=1}^{\ell} \gamma_j$ of smooth simple closed curves whose interior parts contain $\sigma(a)$ and entirely contained in U. The curves γ_j are assumed to be oriented counterclockwise.

Then $f(a) \in \mathcal{A}_0$ can be defined as

$$f(a) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} f(z)(zI - a)^{-1}\, dz, \tag{6}$$

where the line integral of a Banach-space-valued function $g(z)$ defined on a smooth curve $\gamma : t \mapsto z(t)$ for $t \in [0,1]$ is given by the norm limit of Riemann sums of the form

$$\sum_{j=1}^{\nu} g\big(z(\theta_j)\big)\big[z(t_j) - z(t_{j-1})\big], \quad t_{j-1} \leqslant \theta_j \leqslant t_j,$$

where $0 = t_0 < t_1 < \cdots < t_{\nu-1} < t_{\nu} = 1$.

Denote by $\mathcal{H}(a)$ the algebra of analytic functions whose domain contains an open neighborhood of $\sigma(a)$. The following well-known result is the basis for the *holomorphic functional calculus*; see, e.g., [43, page 206].

Theorem 5. *The mapping $\mathcal{H}(a) \to \mathcal{A}_0$ defined by $f \mapsto f(a)$ is an algebra homomorphism, which maps the constant function 1 to $I \in \mathcal{A}_0$ and maps the identity function to a. If $f(z) = \sum_{j=0}^{\infty} c_j z^j$ is the power series representation of $f \in \mathcal{H}(a)$ over an open neighborhood of $\sigma(a)$, then we have*

$$f(a) = \sum_{j=0}^{\infty} c_j a^j.$$

Moreover, the following version of the *spectral theorem* holds:

$$\sigma\big(f(a)\big) = f\big(\sigma(a)\big). \tag{7}$$

If a is normal, the following properties also hold:

- $\|f(a)\| = \|f\|_{\infty,\sigma(a)} := \max_{\lambda\in\sigma(a)} |f(\lambda)|$;
- $\overline{f(a)} = [f(a)]^*$; in particular, if a is Hermitian then $f(a)$ is also Hermitian if and only if $f(\sigma(a)) \subset \mathbb{R}$;
- $f(a)$ is normal;
- $f(a)b = bf(a)$ whenever $b \in \mathcal{A}_0$ and $ab = ba$.

Obviously, these definitions and results apply in the case where a is a matrix A with entries in a C^*-algebra $\mathcal{A}_0$. In particular, if $f(z)$ is analytic on a neighborhood of $\sigma(A)$, we define $f(A)$ via

$$f(A) = \frac{1}{2\pi \mathrm{i}} \int_{\Gamma} f(z)(zI_n - A)^{-1}\, dz, \tag{8}$$

with the obvious meaning of Γ.

5. Bounds for the Hermitian case

In this paper we will be concerned mostly with banded and sparse matrices. A matrix $A \in \mathcal{A}$ is *banded with bandwidth* m if a_{ij} is the zero element of $\mathcal{A}_0$ whenever $|i-j| > m$. Banded matrices over a C^*-algebra arise in several contexts; see, e.g., [10,11,25,52] and references therein.

Let $A \in \mathcal{A}$ be a banded Hermitian matrix with bandwidth m. In this section we provide exponentially decaying bounds on the norm of the entries $[f(A)]_{ij}$ $(1 \leqslant i, j \leqslant n)$ of $f(A)$, where f is analytic on a neighborhood of $\sigma(A)$, and we discuss the important case where the bounds do not depend on the order n of the matrix. The results in this section extend to the C^*-algebra setting analogous results for matrices over $\mathbb{R}$ or $\mathbb{C}$ found in [7,8,6]. Functions of non-Hermitian (and nonnormal) matrices are studied in the following sections.

Hermitian matrices have a real spectrum. If $\sigma(A) \subseteq [\alpha, \beta] \subset \mathbb{R}$ then, by replacing A (if necessary) with the shifted and scaled matrix $\frac{2}{\beta-\alpha}A - \frac{\beta+\alpha}{\beta-\alpha}I_n$, we can assume that the spectrum is contained in the interval $\mathcal{I} = [-1, 1]$. We also assume that $f(z)$ is real for real z, so that f maps Hermitian matrices to Hermitian matrices.

Let $\mathbb{P}_k$ denote the set of all complex polynomials of degree at most k on $\mathcal{I}$. Given $p \in \mathbb{P}_k$, the matrix $p(A) \in \mathcal{A}$ is well defined and it is banded with bandwidth at most km. So for any polynomial $p \in \mathbb{P}_k$ and any pair of indices i, j such that $|i-j| > km$ we have

$$\begin{aligned}
\big\|\big[f(A)\big]_{ij}\big\| &= \big\|\big[f(A) - p(A)\big]_{ij}\big\| && (9)\\
&\leqslant \big\|f(A) - p(A)\big\| && (10)\\
&= \rho\big(f(A) - p(A)\big) && (11)\\
&= \sup\big(\sigma\big(f(A) - p(A)\big)\big) = \sup\big(\sigma\big((f-p)(A)\big)\big) && (12)\\
&= \sup\big((f-p)\big(\sigma(A)\big)\big) \leqslant E_k(f), && (13)
\end{aligned}$$

where $E_k(f)$ is the best uniform approximation error for the function f on the interval $\mathcal{I}$ using polynomials of degree at most k:

$$E_k(f) := \min_{p\in\mathbb{P}_k} \max_{t\in\mathcal{I}} \big|f(t) - p(t)\big|.$$

In the above computation:

- (10) follows from (9) as a consequence of the definition of operator norm,
- (11) follows from (10) because A is Hermitian, so $f(A) - p(A)$ is also Hermitian,
- the spectral theorem (7) allows us to obtain (13) from (12).

Next, we recall the classical Bernstein's Theorem concerning the asymptotic behavior of $E_k(f)$ for $k \to \infty$; see, e.g., [50, page 91]. This theorem states that there exist constants $c_0 > 0$ and $0 < \xi < 1$ such that $E_k(f) \leqslant c_0 \xi^{k+1}$. From this we can deduce exponentially decaying bounds for $\|[f(A)]_{ij}\|$ with respect to $|i-j|$, by observing that $|i-j| > km$ implies $k+1 < \frac{|i-j|}{m} + 1$ and therefore

$$\left\|\left[f(A)\right]_{ij}\right\| \leqslant c_0 \xi^{\frac{|i-j|}{m}+1} = c\zeta^{|i-j|}, \qquad c = c_0\xi, \qquad \zeta = \xi^{\frac{1}{m}} \in (0,1). \tag{14}$$

The above bound warrants further discussion. Indeed, as it is stated it is a trivial bound, in the sense that for any *fixed* matrix A and function f such that $f(A)$ is defined one can always find constants $c_0 > 0$ and $0 < \xi < 1$ such that (14) holds for all $i, j = 1, \ldots, n$; all one has to do is pick c_0 large enough. Thus, the entries of $f(A)$ may exhibit no actual decay behavior! However, what is important here is that the constants c_0 and ξ (or at least bounds for them) can be given explicitly in terms of properties of f and, indirectly, in terms of the bounds α and β on the spectrum of A. If we have a sequence $\{A_n\}$ of $n \times n$ matrices such that

- the A_n are banded with bounded bandwidth (independent of n);
- the spectra $\sigma(A_n)$ are all contained in a common interval $\mathcal{I}$ (independent of n), say $\mathcal{I} = [-1, 1]$,

then the bound (14) holds independent of n. In particular, the entries of $f(A_n)$ will actually decay to zero away from the main diagonal as $|i-j|$ and n tend to infinity, at a rate that is uniformly bounded below by a positive constant independent of n.

More specifically, Bernstein's Theorem yields the values $c_0 = \frac{2\chi M(f)}{\chi - 1}$ and $\xi = 1/\chi$, where χ is the sum of the semi-axes of an ellipse $\mathcal{E}_\chi$ with foci in 1 and -1, such that $f(z)$ is continuous on $\mathcal{E}_\chi$ and analytic in the interior of $\mathcal{E}_\chi$ (and $f(z) \in \mathbb{R}$ whenever $z \in \mathbb{R}$); furthermore, we have set $M(f) = \max_{z \in \mathcal{E}_\chi} |f(z)|$.

Summarizing, we have established the following results:

Theorem 6. *Let $\mathcal{A} = \mathcal{A}_0^{n\times n}$ where $\mathcal{A}_0$ is a C*-algebra and let $A \in \mathcal{A}$ be Hermitian with bandwidth m and spectrum contained in $[-1, 1]$. Let the complex function $f(z)$ be continuous on a Bernstein ellipse $\mathcal{E}_\chi$ and analytic in the interior of $\mathcal{E}_\chi$, and assume $f(z) \in \mathbb{R}$ for $z \in \mathbb{R}$. Then there exist constants $c > 0$ and $0 < \zeta < 1$ such that*

$$\left\|\left[f(A)\right]_{ij}\right\| \leqslant c\zeta^{|i-j|}$$

for all $1 \leqslant i, j \leqslant n$. Moreover, one can choose $c = \max\{\|f(A)\|, \frac{2M(f)}{\chi-1}\}$ and $\zeta = (\frac{1}{\chi})^{1/m}$.

Remark 3. Letting $\theta := -\ln\zeta > 0$ the decay bound can be rewritten in the form $\|[f(A)]_{ij}\| \leqslant ce^{-\theta|i-j|}$, which is sometimes more convenient.

Theorem 7. *Let $\mathcal{A}_0$ be a C*-algebra and let $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}_0^{n\times n}$ be a sequence of Hermitian matrices of increasing size, with bandwidths uniformly bounded by $m \in \mathbb{N}$ and spectra all contained in $[-1, 1]$. Let the complex function $f(z)$ be continuous on a Bernstein ellipse $\mathcal{E}_\chi$ and analytic in the interior of $\mathcal{E}_\chi$, and assume $f(z) \in \mathbb{R}$ for $z \in \mathbb{R}$. Then there exist constants $c > 0$ and $0 < \zeta < 1$, independent of n, such that*

$$\left\|\left[f(A_n)\right]_{ij}\right\| \leqslant c\zeta^{|i-j|} = ce^{-\theta|i-j|}, \qquad \theta = -\ln\zeta,$$

for all indices i, j. Moreover, one can choose $c = \max\{\|f(A)\|, \frac{2M(f)}{\chi-1}\}$ and $\zeta = (\frac{1}{\chi})^{1/m}$.

Remark 4. It is worth noting that the decay bounds in the above results are actually *families* of bounds; different choices of the ellipse $\mathcal{E}_\chi$ will result in different bounds. If χ and χ', with $\chi < \chi'$, are both admissible values, choosing χ' will result in a smaller value of ζ, thus yielding a faster asymptotic decay rate, but possibly a larger value of the prefactor c; in general, tighter bounds may be obtained by varying χ for different values of i and j. See [6] for examples and additional discussion of this issue.

Remark 5. The bounds in Theorem 7 essentially state that as long as the possible singularities of f remain bounded away from the interval $[-1, 1]$ or, slightly more generally, from a (fixed) interval $[\alpha, \beta]$ containing the union of all the spectra $\sigma(A_n)$, $n \in \mathbb{N}$, then the entries of $f(A_n)$ decay exponentially fast away from the main diagonal, at a rate that is bounded below by a positive constant that does not depend on n. As a consequence, for every $\varepsilon > 0$ one can determine a bandwidth $M = M(\varepsilon)$ (independent of n) such that

$$\left\| f(A_n) - \left[f(A_n) \right]_M \right\| < \varepsilon$$

holds for all n, where $[B]_M$ denotes the matrix with entries b_{ij} equal to those of B for $|i-j| \leqslant M$, zero otherwise. It is precisely this fact that makes exponential decay an important property in applications; see, e.g., [6]. As a rule, the closer the possibile singularities of f are to the spectral interval $[\alpha, \beta]$, the slower the decay is (that is, the larger is c and the closer ζ is to the upper bound 1, or θ to 0).

Remark 6. For entire functions, such as the exponential function $f(z) = \mathrm{e}^z$, the above exponential decay results are not optimal; indeed, in such cases *superexponential* decay bounds can be established, exploiting the fact that the coefficients in the Chebyshev expansion of f decay superexponentially. For an example of this type of result, see [40]; see also Example 5 below. Also, in some cases improved decay bounds can be obtained by using different tools from polynomial approximation theory, or exploiting additional structure in f or in the spectra $\sigma(A_n)$; see [6].

6. Bounds for the normal case

We briefly discuss the case when the banded matrix $A \in \mathcal{A}$ is normal, but not necessarily Hermitian. As usual, we denote by m the bandwidth of A.

The main difference with respect to the previously discussed Hermitian case consists in the fact that $\sigma(A)$ is no longer real. Let $\mathcal{F} \subset \mathbb{C}$ be a compact, connected region containing $\sigma(A)$, and denote by $\mathbb{P}_k$, as before, the set of complex polynomials of degree at most k. Then the argument in (9)–(13) still holds, except that now polynomial approximation is no longer applied on a real interval, but on the complex region $\mathcal{F}$. Therefore, the following bound holds for all indices i, j such that $|i-j| > km$:

$$\left\| \left[f(A) \right]_{ij} \right\| \leqslant \sup \left| (f-p)\big(\sigma(A)\big) \right| \leqslant E_k(f, \mathcal{F}), \tag{15}$$

where

$$E_k(f, \mathcal{F}) := \min_{p \in \mathbb{P}_k} \max_{z \in \mathcal{F}} \left| f(z) - p(z) \right|.$$

Unless more accurate estimates for $\sigma(A)$ are available, a possible choice for $\mathcal{F}$ is the disk of center 0 and radius $\rho(\hat{A})$: see Remark 2.

If f is analytic on $\mathcal{F}$, bounds for $E_k(f, \mathcal{F})$ that decay exponentially with k are available through the use of Faber polynomials: see [8, Theorem 3.3] and the next section for more details. More precisely, there exist constants $\tilde{c} > 0$ and $0 < \tilde{\lambda} < 1$ such that $E_k(f, \mathcal{F}) \leqslant \tilde{c}\tilde{\lambda}^k$ for all $k \in \mathbb{N}$. This result, together with (15), yields for all i and j the bound

$$\left\| \left[f(A) \right]_{ij} \right\| \leqslant c\lambda^{|i-j|} = c\mathrm{e}^{-\theta|i-j|}$$

(where $\theta = -\ln\lambda$) for suitable constants $c > 0$ and $0 < \lambda < 1$, which do not depend on n, although they generally depend on f and $\mathcal{F}$.

Remark 7. For normal elements of a C^*-algebra, one can define a *continuous* functional calculus, which associates to any $a \in \mathcal{A}_0$ a well-defined element $f(a) \in \mathcal{A}_0$, where f is a continuous complex-valued function on $\sigma(a)$; see, e.g., [43, Theorem 4.4.5]. When f is analytic on a neighborhood of $\sigma(a)$, this calculus coincides with the holomorphic functional calculus. Using the continuous calculus one can study decay properties of functions $f(A)$ of a banded normal matrix $A \in \mathcal{A} = \mathcal{A}_0^{n \times n}$ for functions f which have some regularity but are not necessarily analytic. For instance, using Jackson's Theorem

[50, Theorem 45] one can show that the off-diagonal entries of $f(A)$, where A is a banded normal matrix and f has continuous derivatives up to order k with $f^{(k)}$ Lipschitz continuous, are bounded in an algebraically decaying manner away from the off-diagonal; the more derivatives f has, the faster the decay. If $f \in C^\infty(\sigma(A))$, the decay is faster than any power of $(1+|i-j|)^{-1}$ for $|i-j| \to \infty$. It is interesting to observe that these results have no counterpart in the case of functions of $n \times n$ matrices over $\mathbb{R}$ or $\mathbb{C}$, since in this case for any well-defined function $f(A)$ one has $f(A) = p(A)$, where p is a polynomial.

7. Bounds for the general case

If A is not normal, then the equality between (10) and (11) does not hold. We therefore need other explicit bounds on the norm of a function of a matrix.

7.1. The field of values and bounds for complex matrices

Given a matrix $A \in \mathbb{C}^{n\times n}$, the associated *field of values* (or *numerical range*) is defined as

$$W(A) = \left\{ \frac{x^* A x}{x^* x};\ x \in \mathbb{C}^n,\ x \neq 0 \right\}.$$

It is well known that $W(A)$ is a convex and compact subset of the complex plane that contains the eigenvalues of A.

The field of values of a complex matrix appears in the context of bounds for functions of matrices thanks to a result by Crouzeix (see [21]):

Theorem 8 *(Crouzeix). There is a universal constant $2 \leqslant \mathcal{Q} \leqslant 11.08$ such that, given $A \in \mathbb{C}^{n,n}$, $\mathcal{F}$ a convex compact set containing the field of values $W(A)$, a function g continuous on $\mathcal{F}$ and analytic in its interior, then the following inequality holds*:

$$\left\| g(A) \right\|_2 \leqslant \mathcal{Q} \sup_{z\in\mathcal{F}} \left| g(z) \right|.$$

We mention that Crouzeix has conjectured that $\mathcal{Q}$ can be replaced by 2, but so far this has been proved only in some special cases.

Next, we need to review some basic material on polynomial approximation of analytic functions. Our treatment follows the discussion in [8], which in turn is based on [48]; see also [24,58]. In the following, $\mathcal{F}$ denotes a continuum containing more than one point. By a *continuum* we mean a nonempty, compact and connected subset of $\mathbb{C}$. Let G_∞ denote the component of the complement of $\mathcal{F}$ containing the point at infinity. Note that G_∞ is a simply connected domain in the extended complex plane $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. By the Riemann Mapping Theorem there exists a function $w = \Phi(z)$ which maps G_∞ conformally onto a domain of the form $|w| > \rho > 0$ satisfying the normalization conditions

$$\Phi(\infty) = \infty, \qquad \lim_{z\to\infty} \frac{\Phi(z)}{z} = 1; \tag{16}$$

ρ is the *logarithmic capacity* of $\mathcal{F}$. Given any integer $k > 0$, the function $[\Phi(z)]^k$ has a Laurent series expansion of the form

$$\left[\Phi(z)\right]^k = z^k + \alpha_{k-1}^{(k)} z^{k-1} + \cdots + \alpha_0^{(k)} + \frac{\alpha_{-1}^{(k)}}{z} + \cdots \tag{17}$$

at infinity. The polynomials

$$\Phi_k(z) = z^k + \alpha_{k-1}^{(k)} z^{k-1} + \cdots + \alpha_0^{(k)}$$

consisting of the terms with nonnegative powers of z in the expansion (17) are called the *Faber polynomials* generated by the continuum $\mathcal{F}$.

Let Ψ be the inverse of Φ. By C_R we denote the image under Ψ of a circle $|w| = R > \rho$. The (Jordan) region with boundary C_R is denoted by $I(C_R)$. By [48, Theorem 3.17, p. 109], every function $f(z)$ analytic on $I(C_{R_0})$ with $R_0 > \rho$ can be expanded in a series of Faber polynomials:

$$f(z) = \sum_{k=0}^{\infty} \alpha_k \Phi_k(z), \tag{18}$$

where the series converges uniformly inside $I(C_{R_0})$. The coefficients are given by

$$\alpha_k = \frac{1}{2\pi i} \int_{|w|=R} \frac{f(\Psi(w))}{w^{k+1}}\, dw,$$

where $\rho < R < R_0$. We denote the partial sums of the series in (18) by

$$\Pi_k(z) := \sum_{i=0}^{k} \alpha_i \Phi_i(z). \tag{19}$$

Each $\Pi_k(z)$ is a polynomial of degree at most k, since each $\Phi_i(z)$ is of degree i. We now recall a classical result that will be instrumental in our proof of the decay bounds; for its proof see, e.g., [48, Theorem 3.19].

Theorem 9 *(Bernstein). Let f be a function defined on $\mathcal{F}$. Then given any $\varepsilon > 0$ and any integer $k \geqslant 0$, there exists a polynomial Π_k of degree at most k and a positive constant $c(\varepsilon)$ such that*

$$\left| f(z) - \Pi_k(z) \right| \leqslant c(\varepsilon)(q+\varepsilon)^k \quad (0 < q < 1) \tag{20}$$

for all $z \in \mathcal{F}$ if and only if f is analytic on the domain $I(C_{R_0})$, where $R_0 = \rho/q$. In this case, the sequence $\{\Pi_k\}$ converges uniformly to f inside $I(C_{R_0})$ as $k \to \infty$.

Below we will make use of the sufficiency part of Theorem 9. Note that the choice of q (with $0 < q < 1$) depends on the region where the function f is analytic. If f is defined on a continuum $\mathcal{F}$ with logarithmic capacity ρ then we can pick q bounded away from 1 as long as the function is analytic on $I(C_{\rho/q})$. Therefore, the rate of convergence is directly related to the properties of the function f, such as the location of its poles (if there are any). For certain regions, in particular for the case of convex $\mathcal{F}$, it is possible to obtain an explicit value for the constant $c(\varepsilon)$; see [30] and [8, Section 3.7] and [49, Section 2] and the discussion following Theorem 13 below.

We can then formulate the following result on the off-diagonal decay of functions of nonnormal band matrices:

Theorem 10. *Let $A \in \mathbb{C}^{n\times n}$ be m-banded, and let $\mathcal{F}$ be a continuum containing $W(A)$ in its interior. Let f be a function defined on $\mathcal{F}$ and assume that f is analytic on $I(C_{R_0})$ ($\supset W(A)$), with $R_0 = \frac{\rho}{q}$ where $0 < q < 1$ and ρ is the logarithmic capacity of $\mathcal{F}$. Then there are constants $K > 0$ and $0 < \lambda < 1$ such that*

$$\left\| \left[f(A) \right]_{ij} \right\| \leqslant K \lambda^{|i-j|}$$

for all $1 \leqslant i, j \leqslant n$.

Proof. Let $g = f - p_k$ in Theorem 8, where $p_k(z)$ is a polynomial of degree smaller than or equal to k. Then $p_k(A)$ is a banded matrix with bandwidth at most km. Therefore, for all i, j such that $|i-j| > km$ we have

$$\left| \left[f(A) \right]_{ij} \right| = \left| \left[f(A) \right]_{ij} - \left[p_k(A) \right]_{ij} \right| \leqslant \left\| f(A) - p_k(A) \right\|_2 \leqslant \mathcal{Q} \sup_{z\in\mathcal{F}} \left| f(z) - p_k(z) \right|.$$

Now, by Theorem 9 we have that for any $\varepsilon > 0$ there exists a sequence of polynomials Π_k of degree k which satisfy for all $z \in \mathcal{F}$

$$\left| f(z) - \Pi_k(z) \right| \leqslant c(\varepsilon)(q+\varepsilon)^k, \quad \text{where } 0 < q < 1.$$

Therefore, taking $p_k = \Pi_k$ and applying Theorem 9 we obtain

$$\left| \left[f(A) \right]_{ij} \right| \leqslant \mathcal{Q} c(\varepsilon)(q+\varepsilon)^{\frac{|i-j|}{m}}.$$

The thesis follows if we take $\lambda = (q+\varepsilon)^{\frac{1}{m}} < 1$ and $K = \max\{\|f(A)\|_2, \mathcal{Q} c(\varepsilon)\}$. □

We mention that a similar result (for the case of multi-band matrices) can be found in [49, Theorem 2.6].

The assumptions in Theorem 10 are fairly general. In particular, the result applies if $f(z)$ is an entire function; in this case, however, better estimates exist (for instance, in the case of the matrix exponential; see, e.g., [5] and references therein).

The main difficulty in applying the theorem to obtain practical decay bounds is in estimating the constant $c(\varepsilon)$ and the value of q, which requires knowledge of the field of values of A (or an estimate of it) and of the logarithmic capacity of the continuum $\mathcal{F}$ containing $W(A)$. The task is made simpler if we assume (as it is natural) that $\mathcal{F}$ is convex; see the discussion in [8], especially Theorem 3.7. See also [49, Section 2] and the next subsection for further discussion.

The bound in Theorem 10 often improves on previous bounds for diagonalizable matrices in [8] containing the condition number of the eigenvector matrix, especially when the latter is ill-conditioned (these bounds have no analogue in the C^*-algebra setting).

Again, as stated, Theorem 10 is non-trivial only if K and λ are independent of n. We return on this topic in the next subsection.

It is worth noting that since A is not assumed to have symmetric structure, it could have different numbers of nonzero diagonals below and above the main diagonal. Thus, it may be desirable to have bounds that account for the fact that in such cases the rate of decay will be generally different above and below the main diagonal. An extreme case is when A is an upper (lower) Hessenberg matrix, in which case $f(A)$ typically exhibits fast decay below (above) the main diagonal, and generally no decay above (below) it.

For diagonalizable matrices over $\mathbb{C}$, such a result can be found in [8, Theorem 3.5]. Here we state an analogous result without the diagonalizability assumption. We say that a matrix A has *lower bandwidth* $p > 0$ if $a_{ij} = 0$ whenever $i - j > p$ and *upper bandwidth* $s > 0$ if $a_{ij} = 0$ whenever $j - i > s$. We note that if A has lower bandwidth p then A^k has lower bandwidth kp for $k = 0, 1, 2, \ldots$, and similarly for the upper bandwidth s. Combining the argument found in the proof of [8, Theorem 3.5] with Theorem 10, we obtain the following result.

Theorem 11. *Let $A \in \mathbb{C}^{n\times n}$ be a matrix with lower bandwidth p and upper bandwidth s, and let the function f satisfy the assumptions of Theorem 10. Then there exist constants $K > 0$ and $0 < \lambda_1, \lambda_2 < 1$ such that for $i \geqslant j$*

$$\left| \left[f(A) \right]_{ij} \right| < K \lambda_1^{i-j} \tag{21}$$

and for $i < j$

$$\left| \left[f(A) \right]_{ij} \right| < K \lambda_2^{j-i}. \tag{22}$$

The constants λ_1 and λ_2 depend on the position of he poles of f relative to the continuum $\mathcal{F}$; they also depend, respectively, on the lower and upper bandwidths p and s of A. For an upper Hessenberg matrix ($p = 1$, $s = n$) only the bound (21) is of interest, particularly in the situation (important in applications) where we consider sequences of matrices of increasing size. Similarly, for a lower Hessenberg matrix ($s = 1$, $p = n$) only (22) is meaningful. More generally, the bounds are of interest when they are applied to sequences of $n \times n$ matrices $\{A_n\}$ for which either p or s (or both) are fixed

as n increases, and such that there is a fixed connected compact set $\mathcal{F} \subset \mathbb{C}$ containing $W(A_n)$ for all n and excluding the singularities of f (if any). In this case the relevant constants in Theorem 11 are independent of n, and we obtain uniform exponential decay bounds.

Next, we seek to generalize Theorem 10 to the C^*-algebra setting. In order to do this, we need to make some preliminary observations. If T is a bounded linear operator on a Hilbert space $\mathcal{H}$, then its numerical range is defined as $W(T) = \{\langle Tx, x\rangle;\ x \in \mathcal{H}, \|x\| = 1\}$. The generalization of the notion of numerical range to C^*-algebras (see [9]) is formulated via the Gelfand–Naimark representation: $a \in \mathcal{A}_0$ is associated with an operator T_a defined on a suitable Hilbert space. Then $\overline{W(T_a)}$, the closure of $W(T_a)$, does not depend on the particular $*$-representation that has been chosen for $\mathcal{A}_0$. In other words, the closure of the numerical range is well defined for elements of C^*-algebras (whereas the numerical range itself, in general, is not). This applies, in particular, to elements of the C^*-algebra $\mathcal{A} = \mathcal{A}_0^{n\times n}$.

Let us now consider a matrix $A \in \mathcal{A}$. In the following, we will need easily computable bounds on $\overline{W(A)}$. Theorem 3 easily implies the following simple result:

Proposition 1. *Let $A \in \mathcal{A}$. Then $\overline{W(A)}$ is contained in the disk of center 0 and radius $\|\hat{A}\|_2$.*

We are now in a position to derive bounds valid in the general, nonnormal case.

7.2. Bounds for the nonnormal case

Our aim is to extend the results in the previous section to the case where A is a matrix over a C^*-algebra. In [21], Crouzeix provides a useful generalization of his result from complex matrices to bounded linear operators on a Hilbert space $\mathcal{H}$. Given a set $E \subset \mathbb{C}$, denote by $\mathcal{H}_b(E)$ the algebra of continuous and bounded functions in $\overline{E}$ which are analytic in the interior of E. Furthermore, for $T \in \mathcal{B}(\mathcal{H})$ let $\|p\|_{\infty,T} := \sup_{z\in\overline{W(T)}} |p(z)|$. Then we have (see [21], Theorem 2):

Theorem 12. *For any bounded linear operator $T \in \mathcal{B}(\mathcal{H})$ the homomorphism $p \mapsto p(T)$ from the algebra $\mathbb{C}[z]$, with norm $\|\cdot\|_{\infty,T}$, into the algebra $\mathcal{B}(\mathcal{H})$, is bounded with constant $\mathcal{Q}$. It admits a unique bounded extension from $\mathcal{H}_b(W(T))$ into $\mathcal{B}(\mathcal{H})$. This extension is bounded with constant $\mathcal{Q}$.*

Using again the Gelfand–Naimark representation together with the notion of numerical range for elements of $\mathcal{A}$, we obtain as a consequence:

Corollary 1. *Given $A \in \mathcal{A}$, the following bound holds for any complex function g analytic on a neighborhood of $\overline{W(A)}$:*

$$\big\|g(A)\big\| \leqslant \mathcal{Q}\|g\|_{\infty,A} = \mathcal{Q} \sup_{z\in\overline{W(A)}} \big|g(z)\big|.$$

Since we wish to obtain bounds on $\|[f(A)]_{ij}\|$, where the function $f(z)$ can be assumed to be analytic on an open set $S \supset \overline{W(A)}$, we can choose $g(z)$ in Corollary 1 as $f(z) - p_k(z)$, where $p_k(z)$ is any complex polynomial of degree bounded by k. The argument in (9)–(13) can then be adapted as follows:

$$\begin{aligned}
\big\|\big[f(A)\big]_{ij}\big\| &= \big\|\big[f(A) - p_k(A)\big]_{ij}\big\| && (23)\\
&\leqslant \big\|f(A) - p_k(A)\big\| && (24)\\
&\leqslant \mathcal{Q}\|f - p_k\|_{\infty,A} && (25)\\
&= \mathcal{Q} \sup_{z\in\overline{W(A)}} \big|f(z) - p_k(z)\big| && (26)\\
&\leqslant \mathcal{Q}E_k\big(f, \overline{W(A)}\big), && (27)
\end{aligned}$$

where $E_k(f, \overline{W(A)})$ is the degree k best approximation error for f on the compact set $\overline{W(A)}$. In order to make explicit computations easier, we may of course replace $\overline{W(A)}$ with a larger but more manageable set in the above argument, as long as the approximation theory results used in the proof of Theorem 10 can be applied.

Theorem 13. *Let $A \in \mathcal{A}$ be an $n \times n$ matrix of bandwidth m and let the function f and the continuum $\mathcal{F} \supset \overline{W(A)}$ satisfy the assumptions of Theorem 10. Then there exist explicitly computable constants $K > 0$ and $0 < \lambda < 1$ such that*

$$\left\| \left[f(A) \right]_{ij} \right\| \leqslant K \lambda^{|i-j|}$$

for all $1 \leqslant i, j \leqslant n$.

A simple approach to the computation of constants K and λ goes as follows. It follows from Proposition 1 that the set $\mathcal{F}$ in Theorem 13 can be chosen as the disk of center 0 and radius $r = \|\hat{A}\|_2$. Assume that $f(z)$ is analytic on an open neighborhood of the disk of center 0 and radius $R > r$. The standard theory of complex Taylor series gives the following estimate for the Taylor approximation error [30, Corollary 2.2]:

$$E_k(f, \mathcal{C}) \leqslant \frac{M(R)}{1 - \frac{r}{R}} \left(\frac{r}{R} \right)^{k+1}, \tag{28}$$

where $M(R) = \max_{|z|=R} |f(z)|$. Therefore we can choose

$$K = \max\left\{ \left\| f(A) \right\|, \mathcal{Q} M(R) \frac{r}{R - r} \right\}, \qquad \lambda = \left(\frac{r}{R} \right)^{1/m}.$$

The choice of the parameter R in (28) is somewhat arbitrary: any value of R will do, as long as $r < R < \min |\zeta|$, where ζ varies over the poles of f (if f is entire, we let $\min |\zeta| = \infty$). Choosing as large a value of R as possible gives a better asymptotic decay rate, but also a potentially large constant K. For practical purposes, one may therefore want to pick a value of R that ensures a good trade-off between the magnitude of K and the decay rate: see the related discussion in [8] and [6].

As in the previous section, we are also interested in the case of a sequence $\{A_n\}_{n \in \mathbb{N}}$ of matrices of increasing size over $\mathcal{A}_0$. In order to obtain a uniform bound, we reformulate Corollary 1 as follows.

Corollary 2. *Let $\{A_n\}_{n \in \mathbb{N}}$ be a sequence of $n \times n$ matrices over a C^*-algebra $\mathcal{A}_0$ such that there exists a connected compact set $\mathcal{C} \subset \mathbb{C}$ that contains $\overline{W(A_n)}$ for all n, and let g be a complex function analytic on a neighborhood of $\mathcal{C}$. The following uniform bound holds*:

$$\left\| g(A_n) \right\| \leqslant \mathcal{Q} \|g\|_{\infty, \mathcal{C}} = \mathcal{Q} \sup_{z \in \mathcal{C}} \left| g(z) \right|.$$

We then have a version of Theorem 13 for matrix sequences having uniformly bounded bandwidths and fields of values:

Theorem 14. *Let $\{A_n\}_{n \in \mathbb{N}} \subset \mathcal{A}$ be a sequence of $n \times n$ matrices over a C^*-algebra $\mathcal{A}_0$ with bandwidths uniformly bounded by m. Let the complex function $f(z)$ be analytic on a neighborhood of a connected compact set $\mathcal{C} \subset \mathbb{C}$ containing $\overline{W(A_n)}$ for all n. Then there exist explicitly computable constants $K > 0$ and $0 < \lambda < 1$, independent of n, such that*

$$\left\| \left[f(A_n) \right]_{ij} \right\| \leqslant K \lambda^{|i-j|}$$

for all indices i, j.

In other words: as long as the singularities of f (if any) stay bounded away from a fixed compact set $\mathcal{C}$ containing the union of all the sets $\overline{W(A_n)}$, and as long as the matrices A_n have bandwidths less than a fixed integer m, the entries of $f(A_n)$ decay exponentially fast away from the main diagonal, at a rate bounded below by a fixed positive constant as $n \to \infty$. The larger the distance between the singularities of f and the compact $\mathcal{C}$, the larger this constant is.

Finally, it is straightforward to generalize Theorem 11 to the case of matrices over a general C^*-algebra. This completes the desired extension to the C^*-algebra setting of the known exponential decay results for analytic functions of banded matrices over $\mathbb{R}$ or $\mathbb{C}$.

8. The case of quaternion matrices

Matrices over the real division algebra $\mathbb{H}$ of quaternions have many interesting properties; see, e.g., [46,62] and the references therein, as well as [45] for a recent application of quaternion matrices to signal processing. There is, however, very little in the literature about functions of matrices over $\mathbb{H}$, except for the very special case of the inverse $f(A) = A^{-1}$ of an invertible quaternion matrix. This is no doubt due to the fact that fundamental difficulties arise even in the scalar ($n = 1$) case when attempting to extend the classical theory of complex analytic functions to functions of a quaternion variable [26,57].

The formal evaluation of functions of matrices with quaternion entries was considered by Giscard and coworkers in [35], the same paper that raised the question that led to the present work. In order to even state a meaningful generalization of our decay results for analytic functions of banded (or sparse) matrices over $\mathbb{R}$ or $\mathbb{C}$ to matrices over $\mathbb{H}$, we first need to restrict the class of functions under consideration to those analytic functions that can be expressed by convergent power series with *real* coefficients. In this case, no ambiguity can arise when considering functions of a matrix of the form

$$f(A) = a_0 I_n + a_1 A + a_2 A^2 + \cdots + a_k A^k + \cdots, \quad A \in \mathbb{H}^{n\times n},$$

since the real field $\mathbb{R}$ is the center of the quaternion algebra $\mathbb{H}$. In contrast, for functions expressed by power series with (say) complex coefficients we would have to distinguish between "left" and "right" power series, since $a_k A^k \neq A^k a_k$ in general. Fortunately, many of the most important functions (like the exponential, the logarithm, the trigonometric and hyperbolic functions and their inverses, etc.) can be represented by power series with real coefficients.

Next, we note that the quaternion algebra $\mathbb{H}$ is a real C^*-algebra [31], not a complex one; note that $\mathbb{H}$ is a noncommutative division algebra, and the Gelfand–Mazur Theorem [47] states that a (complex) C^*-algebra which is a division algebra is $*$-isomorphic to $\mathbb{C}$ and thus it is necessarily commutative. Hence, we cannot immediately apply the results from the previous sections to functions of quaternion matrices.

To obtain the desired generalization, we make use of the fact that quaternions can be regarded as 2×2 matrices over $\mathbb{C}$ through the following representation:

$$\mathbb{H} = \{q = a + b\mathrm{i} + c\mathrm{j} + d\mathrm{k};\ a, b, c, d \in \mathbb{R}\} \cong \left\{ Q = \begin{pmatrix} a + b\mathrm{i} & c + d\mathrm{i} \\ -c + d\mathrm{i} & a - b\mathrm{i} \end{pmatrix} \right\}.$$

The modulus (or norm) of a quaternion is given by $|q| = \sqrt{a^2 + b^2 + c^2 + d^2} = \|Q\|_2$, where Q is the matrix associated with q.

Thus, we represent matrices over quaternions as complex block matrices with blocks of size 2×2. In this way the real algebra $\mathbb{H}^{n\times n}$ with the natural operator norm

$$\|A\| = \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}, \qquad x = (x_1, \ldots, x_n) \in \mathbb{H}^n, \qquad \|x\| = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}},$$

is isomorphic to a norm-closed real subalgebra $\mathcal{B}$ of the C^*-algebra $\mathcal{A} = \mathbb{C}^{2n\times 2n}$. The operator norm of an $n \times n$ quaternion matrix A turns out to coincide with the spectral norm of the $2n \times 2n$ complex matrix $\varphi(A)$ that corresponds to A in this representation: $\|A\| = \|\varphi(A)\|_2$ (see [46, Theorem 4.1]).

Let now f be a function that can be expressed by a power series $f(z) = \sum_{k=0}^{\infty} a_k z^k$ with $a_k \in \mathbb{R}$, and assume that the power series has radius of convergence $R > \|A\| = \|\varphi(A)\|_2$. Then the function $f(A)$ is well defined,[2] and is given by the convergent power series $f(A) = \sum_{k=0}^{\infty} a_k A^k$.

The theory developed in Sections 5–7 can now be applied to obtain the desired exponential decay bounds for functions of banded quaternion matrices, at least for those analytic functions that can be expressed by convergent power series with real coefficients.

Remark 8. Let f be a function that can be expressed by a convergent power series with real coefficients, as above, and let A be an $n \times n$ matrix with entries in a finite-dimensional real C^*-algebra $\mathcal{A}_{\mathbb{R}}$. Then the decay results presented in Sections 5–8 apply also to $f(A)$.

Indeed, any finite-dimensional real C^*-algebra is isometrically $*$-isomorphic to a direct sum of full matrix algebras over $\mathbb{C}$, $\mathbb{R}$ and $\mathbb{H}$ (see for instance [31, Theorem 5.22]). Given such a representation for the algebra of $n \times n$ matrices over $\mathcal{A}_{\mathbb{R}}$, one may apply the decay bounds of Sections 5–7 to the real and complex components of A, and the results of this section to the quaternionic components, thus obtaining decay bounds for $f(A)$.

9. General sparsity patterns

Following [8] and [6], we sketch an adaptation of Theorems 6 and 13 to the case where the $n \times n$ matrix $A \in \mathcal{A}$ is not necessarily banded, but it has a more general sparsity pattern.

Recall that the graph G_A associated with A is defined as follows:

- G_A has n nodes,
- nodes i and j are connected by an edge if and only if $a_{ij} \neq 0$.

The *geodetic distance* $d(i, j)$ between nodes i and j is the length of the shortest path connecting node i to node j. If there is no path from i to j, then we set $d(i, j) = \infty$. Observe that in general $d(i, j) \neq d(j, i)$, unless the sparsity pattern of A is symmetric.

Also recall that the *degree* of a node i is the number of nodes of G_A that are directly connected by an edge to node i, that is, the number of neighbors of node i. It is equal to the number of nonzero entries in the i-th row of A.

Let $a_{ij}^{(k)}$ be the (i, j)-th entry of the matrix A^k. It can be proved that $a_{ij}^{(k)} = 0$ whenever $d(i, j) > k$, for all positive integers k. In particular, if $d(i, j) > k$ then the (i, j)-th entry of $p_k(A)$ is zero, for any polynomial $p_k(z)$ of degree bounded by k. Therefore, equations (9) and (23) still hold if the condition $|i - j| > km$ is replaced by $d(i, j) > k$. Bounds for $\|[f(A)]_{ij}\|$ are then obtained in a straightforward way: we have

$$\left\|\left[f(A)\right]_{ij}\right\| \leqslant c\xi^{d(i,j)}$$

for the Hermitian case, and

$$\left\|\left[f(A)\right]_{ij}\right\| \leqslant K\lambda^{d(i,j)}$$

for the non-Hermitian case, where the constants c, ξ, K, η are the same as in Theorems 6 and 13 and their proofs.

Results for functions of sequences of matrices (Theorems 7 and 14) can also be adapted to general sparsity patterns in a similar way. Note that the hypothesis that the matrices A_n have uniformly bounded bandwidth should be replaced by the condition that the degree of each node of the graph associated with A_n should be uniformly bounded by a constant independent of n.

[2] Since the subalgebra of the C^*-algebra $\mathbb{C}^{2n\times 2n}$ that corresponds via φ to $\mathbb{H}^{n\times n}$ is closed under linear combinations with real coefficients and norm-closed, the matrix $f(A)$ is a well-defined quaternion matrix that satisfies $\varphi(f(A)) = f(\varphi(A))$.

10. Examples

In this section we show the results of some experiments on the decay behavior of $f(A)$ for various choices of f and A and comparisons with *a priori* decay bounds. We consider matrices over commutative C^*-algebras of continuous functions, block matrices, and matrices over the noncommutative quaternion algebra.

10.1. Matrices over $C([a,b])$

Here we consider simple examples of matrices over $C([a,b])$, the algebra of (complex-valued) continuous functions defined on a closed real interval $[a,b]$.

Let A be such a matrix: each entry of A can be written as $a_{ij} = a_{ij}(t)$, where $a_{ij}(t) \in C([a,b])$. Let $f(z)$ be a complex analytic function such that $f(A)$ is well defined. In order to compute $f(A)$ we consider two approaches.

1. A symbolic (exact) approach, based on the integral definition (6). This approach goes as follows:
   - Assuming $z \notin \sigma(A)$, compute symbolically $M = f(z)(zI - A)^{-1}$. Recall that the entries of M are meromorphic functions of t and z. In particular, if A is invertible the inverse $B = A^{-1}$ can be computed symbolically, and its entries are elements of $C([a,b])$.
   - Compute $\det(M)$ and factorize it as a polynomial in z. The poles of the entries of M are roots of $\det(M)$ with respect to the variable z.
   - Apply the residue theorem: $[f(M)]_{ij}$ is the sum of the residues of M_{ij} at the roots of $\det(M)$. Such residues can be computed via a Laurent series expansion: see for instance the Maple commands `series` and `residue`.

   The norms $\|[f(A)]_{ij}\|_\infty$ can be computed symbolically (see for instance the Maple command `maximize`) or numerically via standard optimization methods. The exact approach is rather time-consuming and can only be applied to moderate-sized matrices.
2. An approximate hybrid (numerical-symbolic) approach, based on polynomial approximation of $f(z)$. In the present work we employ the following technique:
   - Compute the coefficients of the Chebyshev approximating polynomial $p(z)$ that approximates $f(z)$ up to a predetermined degree or tolerance. Here we use the function `chebpoly` of the `chebfun` package [17] for Matlab. If necessary, scale and shift A so that its spectrum is contained in $[-1,1]$.
   - Symbolically compute $f(A) \approx p(A)$.

   This approach gives results that are virtually indistinguishable from the exact (purely symbolic) approach, but it is much more efficient and can be applied to relatively large matrices.

Example 2. Let C be the following bidiagonal Toeplitz matrix of size $n \times n$ over $C([1,2])$:

$$C = \begin{bmatrix} 1 & \mathrm{e}^{-t} & & \\ & 1 & \ddots & \\ & & \ddots & \mathrm{e}^{-t} \\ & & & 1 \end{bmatrix}.$$

Obviously, C has an inverse in $C([1,2])$, which can be expressed as a (finite) Neumann series. We compute C^{-1} symbolically, using the Symbolic Math Toolbox of Matlab, and then we compute the ∞-norms of its elements using the Matlab function `fminbnd`. Fig. 1 shows the corresponding mesh plot of the matrix $[\|b_{ij}\|_\infty]$ with $B = C^{-1}$ for $n = 20$. Note the rapid off-diagonal decay.

Example 3. Let $A = CC^T$, where C is defined as in Example 2. The inverse of A can be computed symbolically as $A^{-1} = C^{-T}C^{-1}$. Fig. 2 shows the mesh plot of the matrix of infinity norms of elements of A^{-1} for $n = 20$.

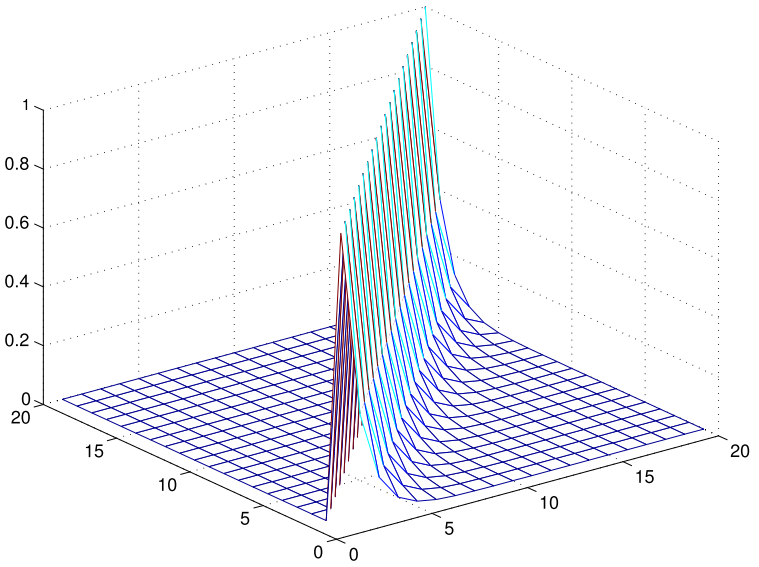

Fig. 1. Decay behavior for the inverse of the bidiagonal matrix in Example 2.

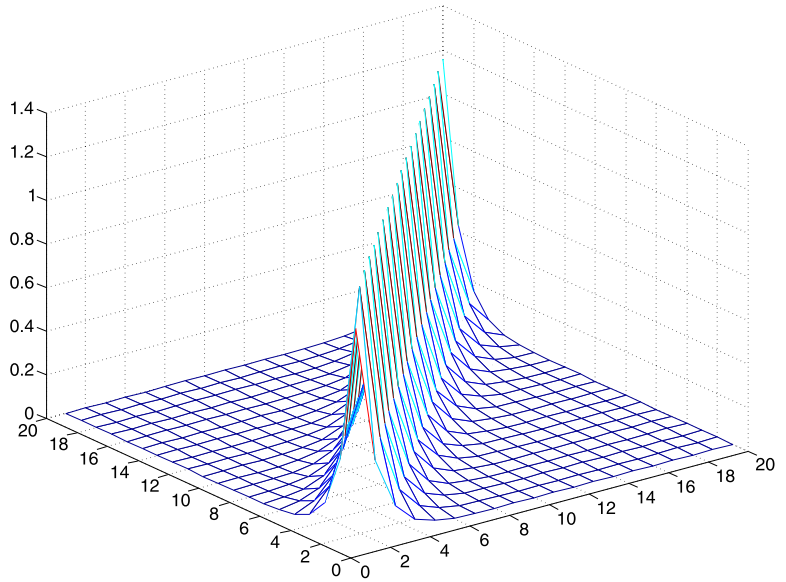

Fig. 2. Decay behavior for the inverse of the tridiagonal matrix in Example 3.

Next we consider the matrix exponential.

Example 4. Let A be a tridiagonal Toeplitz Hermitian matrix as in Example 3. We first scale it so that its spectrum is contained in $[-1, 1]$. This is done by replacing A with $A/\|\hat{A}\|_2$, where $\hat{A}$ is the matrix of infinity norms of the entries of A. Next, we compute an approximation of the exponential of A as

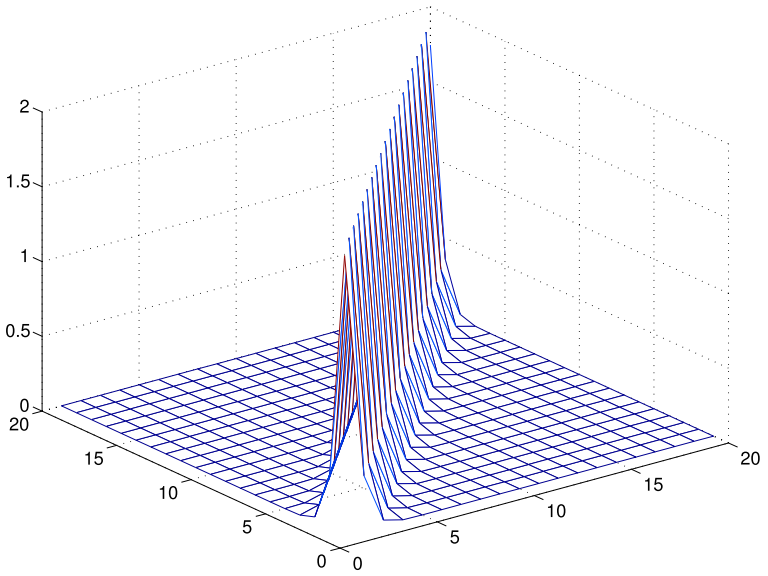

Fig. 3. Decay behavior for the exponential of the scaled tridiagonal matrix in Example 4.

$$e^A \approx \sum_{j=0}^{k} c_j T_j(A),$$

where the coefficients $\{c_j\}_{j=0,\dots,k}$ are computed numerically using the `chebpoly` function of Chebfun [17], and the matrices $T_j(A)$ are computed symbolically using the Chebyshev recurrence relation. Here we choose $n = 20$ and $k = 8$. See Fig. 3 for the mesh plot of the matrix of norms of elements of e^A.

Observe that $\| \sum_{j=0}^{k} c_j T_j(A) \|_\infty \leqslant \sum_{j=0}^{k} |c_j|$, so $|c_k|$ gives an estimate of the correction to the approximation due to the highest order term (see also [8, Section 4.1]). If this correction is sufficiently small, we can assume that the Chebyshev approximation is accurate enough for our purposes. In this example we have $c_8 = 1.9921 \cdot 10^{-7}$ and $c_9 = 1.1037 \cdot 10^{-8}$.

Example 5. Consider the tridiagonal Hermitian Toeplitz matrix of size 20×20 over $C([0, 1])$:

$$A = \begin{bmatrix} 1 & e^{-t} & & \\ e^{-t} & 1 & \ddots & \\ & \ddots & \ddots & e^{-t} \\ & & e^{-t} & 1 \end{bmatrix}.$$

We scale A so that $\sigma(A) \subset [-1, 1]$ and then compute the Chebyshev approximation $e^A \approx \sum_{j=0}^{12} c_j A^j$. The approximation error is bounded in norm by $3.9913 \cdot 10^{-14}$. The decay behavior of e^A and the comparison with decay bounds for different choices of χ (cf. Theorem 7) are shown in Figs. 4 and 5. The semi-logarithmic plot clearly shows the superexponential decay in the entries of the first row of e^A, which is to be expected since the coefficients c_k in the Chebyshev expansion of e^z decay faster than exponentially as $k \to \infty$ [50, page 96]. In contrast, our bounds, being based on Bernstein's Theorem, only decay exponentially. Nevertheless, for $\chi = 20$ the exponential bound decays so fast that for large enough column indices (say, $j \approx 5$ or larger) it is very good for all practical purposes.

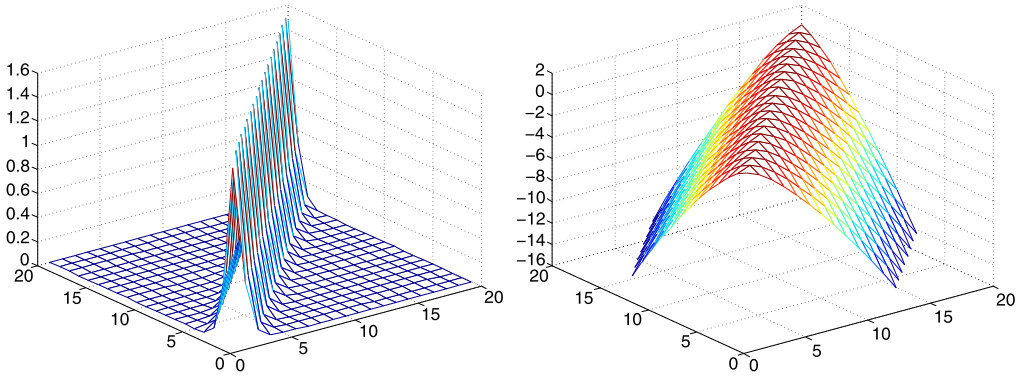

Fig. 4. Linear mesh plot (left) and $\log_{10}$ mesh plot (right) for e^A as in Example 5.

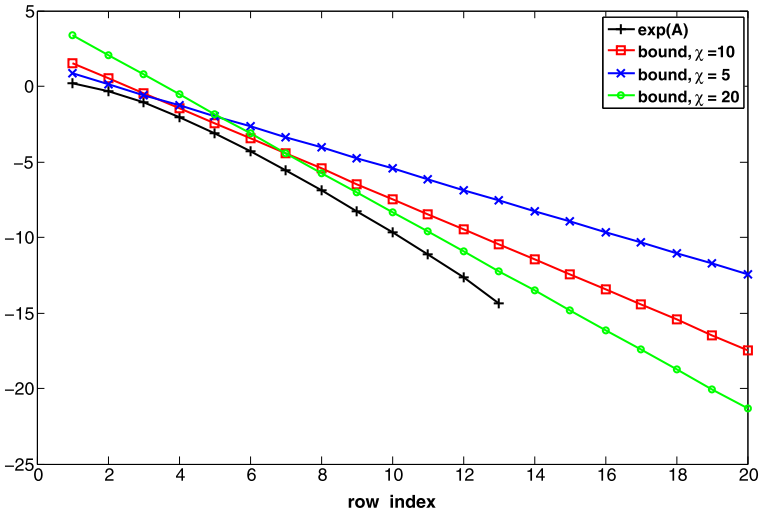

Fig. 5. Comparison between the first row, in norm, of e^A as in Example 5 and theoretical bounds, for several values of χ. The vertical axis is shown in $\log_{10}$ scale.

Example 6. Consider the tridiagonal Hermitian Toeplitz matrix $A = (a_{ij}(t))$ of size 20×20 over $C([0,1])$ defined by

$$a_{jj} = 1, \quad j = 1, \ldots, 20,$$
$$a_{j,j+1} = a_{j+1,j} = 1, \quad j = 2k+1,\ k = 1, \ldots, 9,$$
$$a_{j,j+1} = a_{j+1,j} = t, \quad j = 4k+2,\ k = 0, \ldots, 4,$$
$$a_{j,j+1} = a_{j+1,j} = t^2 - 1, \quad j = 4k,\ k = 0, \ldots, 4.$$

We scale A so that $\sigma(A) \subset [-1,1]$ and then compute the Chebyshev approximation $f(A) \approx \sum_{j=0}^{14} c_j A^j$, where $f(z) = \ln(z+5)$. The approximation error is bounded in norm by $1.7509 \cdot 10^{-14}$. The decay behavior of $f(A)$, compared with decay bounds for different choices of χ (cf. Theorem 7), is shown in Fig. 6. The semi-logarithmic plot clearly shows the exponential decay in the entries of a row of $f(A)$. Note that the decay bounds are somewhat pessimistic in this case.

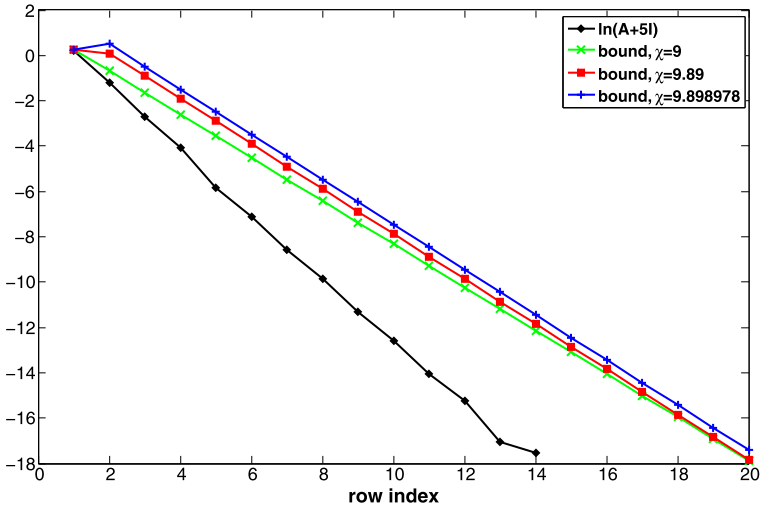

Fig. 6. Comparison between the 7th row, in norm, of $\ln(A+5I)$ as in Example 6 and theoretical bounds, for several values of χ. The vertical axis is shown in $\log_{10}$ scale.

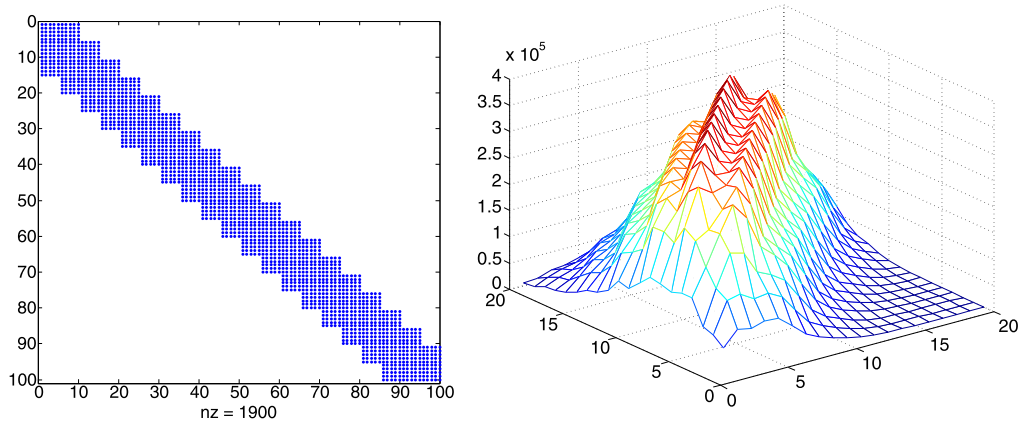

Fig. 7. Sparsity pattern of A in Example 7 (left) and decay behavior in the norms of the blocks of e^A (right).

10.2. Block matrices

If we choose $\mathcal{A}_0$ as the noncommutative C^*-algebra of $k \times k$ complex matrices, then $\mathcal{A} = \mathbb{C}^{nk\times nk}$ can be identified with the C^*-algebra of $n \times n$ matrices with entries in $\mathcal{A}_0$.

Example 7. Let $\mathcal{A}_0 = \mathbb{C}^{5\times 5}$ and consider a banded non-Hermitian matrix A of size 20×20 with entries over $\mathcal{A}_0$. Thus, A is 100×100 as a matrix over $\mathbb{C}$. The entries of each block are chosen at random according to a uniform distribution over $[-1, 1]$. The matrix A has lower bandwidth 2 and upper bandwidth 1. Fig. 7 shows the sparsity pattern of A and the decay behavior of the spectral norms of the blocks of e^A.

10.3. Matrices over quaternions

As discussed in Section 8, we represent matrices over quaternions as complex block matrices with blocks of size 2×2.

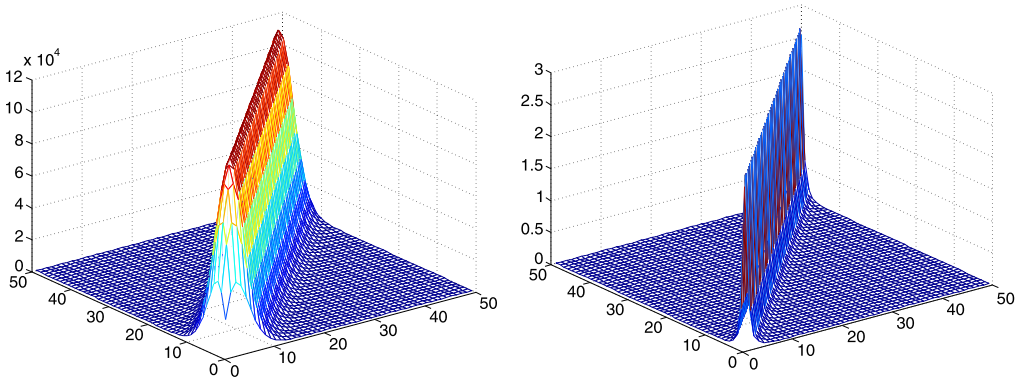

Fig. 8. Decay behavior for the exponential and the logarithm of a tridiagonal matrix over quaternions (Example 8).

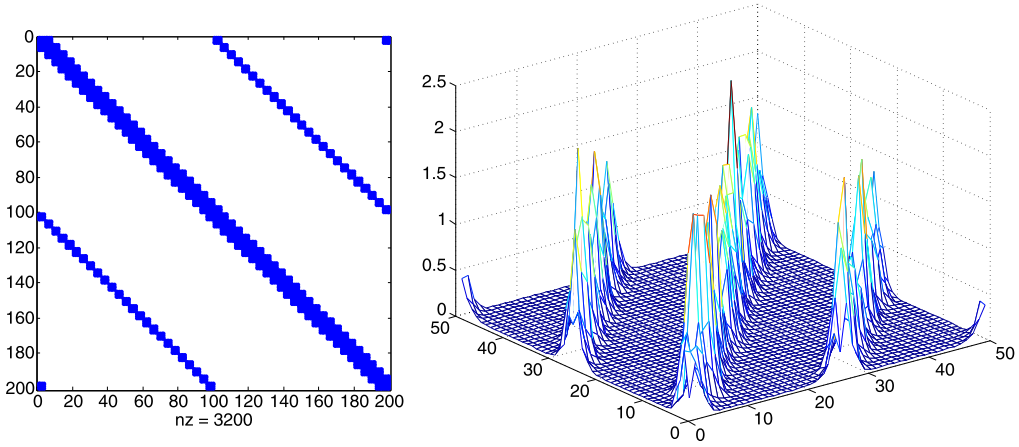

Fig. 9. Sparsity pattern of A as in Example 9 (left) and decay behavior for e^A (right).

Example 8. In this example, $A \in \mathbb{H}^{50\times 50}$ is a Hermitian Toeplitz tridiagonal matrix with random entries, chosen from a uniform distribution over $[-5, 5]$. We form the associated block matrix $\varphi(A) \in \mathbb{C}^{100\times 100}$, compute $f(\varphi(A))$ and convert it back to a matrix in $\mathbb{H}^{50\times 50}$. Fig. 8 shows the mesh plot of the norms of entries of e^A and $\log A$.

Example 9. Here we choose $A \in \mathbb{H}^{50\times 50}$ as a Hermitian matrix with a more general sparsity pattern and random nonzero entries. The decay behavior of e^A is shown in Fig. 9

11. Conclusions and outlook

In this paper we have combined tools from classical approximation theory with basic results from the general theory of C^*-algebras to obtain decay bounds for the entries of analytic functions of banded or sparse matrices with entries of a rather general nature. The theory applies to functions of matrices over the algebra of bounded linear operators on a Hilbert space, over the algebra of continuous functions on a compact Hausdorff space, and over the quaternion algebra, thus achieving a broad generalization of existing exponential decay results for functions of matrices over the real or complex fields. In particular, the theory shows that the exponential decay bounds can be extended *verbatim* to matrices over noncommutative and infinite-dimensional C^*-algebras. In addition, algebraic

decay bounds can be obtained in a similar manner for regular, but not necessarily analytic functions of banded or sparse (normal) matrices over such algebras.

The results in this paper are primarily qualitative in nature, and the bounds can be pessimistic in practice. This is the price one has to pay for the extreme generality of the theory. For entire functions like the matrix exponential, sharper estimates (and superexponential decay bounds) can be obtained by extending known bounds, such as those found in [5] and in [40]. Another avenue for obtaining more quantitative decay results is the Banach algebra approach as found, e.g., in [3,4,12,36,37,41,44]. Although their main focus is on the decay in A^{-1}, together with the representation formula (8) the results found in these papers can be applied to obtain decay results for functions of matrices with various kinds of decay, including algebraic decay.

Future work should address application of the theory to the derivation of specialized bounds for particular functions, such as the matrix exponential, and their use in problems from physics and numerical analysis.

Acknowledgements

We would like to thank Pierre–Louis Giscard (Oxford University) for raising the question that stimulated us to write this paper, Bernd Beckermann (University of Lille I) and Karlheinz Gröchenig (University of Vienna) for relevant bibliographic indications, and Marc Rybowicz (University of Limoges) for useful discussions on symbolic computation of functions of matrices. Thanks also to the handling editor, Leiba Rodman, for carefully reading the manuscript and for helpful comments.

References

[1] W. Arveson, C^*-algebras and numerical linear algebra, J. Funct. Anal. 122 (1994) 333–360.
[2] E. Aune, Computation and modeling for high dimensional Gaussian distributions, PhD thesis, Norwegian University of Science and Technology, Faculty of Information Technology, Mathematics and Electrical Engineering, Department of Mathematical Sciences, 2012.
[3] A.G. Baskakov, Wiener's theorem and the asymptotic estimates of the elements of inverse matrices, Funct. Anal. Appl. 24 (1990) 222–224.
[4] A.G. Baskakov, Estimates for the entries of inverse matrices and the spectral analysis of linear operators, Izv. Math. 61 (1997) 1113–1135.
[5] B. Beckermann, L. Reichel, Error estimation and evaluation of matrix functions via the Faber transform, SIAM J. Numer. Anal. 47 (2009) 3849–3883.
[6] M. Benzi, P. Boito, N. Razouk, Decay properties of spectral projectors with applications to electronic structure, SIAM Rev. 55 (2013) 3–64.
[7] M. Benzi, G.H. Golub, Bounds for the entries of matrix functions with applications to preconditioning, BIT 39 (1999) 417–438.
[8] M. Benzi, N. Razouk, Decay rates and $O(n)$ algorithms for approximating functions of sparse matrices, Electron. Trans. Numer. Anal. 28 (2007) 16–39.
[9] S.K. Berberian, G.H. Orland, On the closure of the numerical range of an operator, Proc. Amer. Math. Soc. 18 (1967) 499–503.
[10] Yu.M. Berezanskii, V.D. Koshmanenko, Axiomatic field theory in terms of operator Jacobi matrices, Teoret. Mat. Fiz. 8 (1971) 175–191.
[11] Yu.M. Berezansky, Commutative Jacobi fields in Fock space, Integral Equations Operator Theory 30 (1998) 163–190.
[12] I.A. Blatov, On algebras and applications of operators with pseudosparse matrices, Sib. Math. J. 37 (1996) 32–52.
[13] N. Bock, M. Challacombe, An optimized sparse approximate matrix multiply for matrices with decay, SIAM J. Sci. Comput. 35 (2013) C72–C98.
[14] A. Böttcher, C^*-algebras in numerical analysis, Irish Math. Soc. Bull. 45 (2000) 57–133.
[15] D.R. Bowler, T. Miyazaki, $O(N)$ methods in electronic structure calculations, Rep. Progr. Phys. 75 (2012) 036503.
[16] J.W. Bunce, Stabilizability of linear systems defined over C^*-algebras, Math. Systems Theory 18 (1985) 237–250.
[17] Chebfun website, http://www2.maths.ox.ac.uk/chebfun/ (accessed July 2013).
[18] P.G. Constantine, D.F. Gleich, G. Iaccarino, Spectral methods for parametrized matrix equations, SIAM J. Matrix Anal. Appl. 31 (2010) 2681–2699.
[19] M. Cramer, J. Eisert, Correlations, spectral gap and entanglement in harmonic quantum systems on generic lattices, New J. Phys. 8 (2006) 71.
[20] M. Cramer, J. Eisert, M.B. Plenio, J. Dreissig, Entanglement-area law for general bosonic harmonic lattice systems, Phys. Rev. A 73 (2006) 012309.
[21] M. Crouzeix, Numerical range and functional calculus in Hilbert space, J. Funct. Anal. 244 (2007) 668–690.
[22] R. Curtain, Riccati equations on noncommutative Banach algebras, SIAM J. Control Optim. 49 (2011) 2542–2557.
[23] R. Curtain, A. Sasane, On Riccati equations in Banach algebras, SIAM J. Control Optim. 49 (2011) 464–475.
[24] J.H. Curtiss, Faber polynomials and Faber series, Amer. Math. Monthly 78 (1971) 577–596.

[25] L. Cvetković, D. Takači, On matrices with operator entries, Numer. Algorithms 42 (2006) 335–344.
[26] C.A. Deavours, The quaternion calculus, Amer. Math. Monthly 80 (1973) 995–1008.
[27] S. Demko, W.S. Moss, P.W. Smith, Decay rates for inverses of band matrices, Math. Comp. 43 (1984) 491–499.
[28] R.S. Doran (Ed.), C^*-Algebras: 1943–1993. A Fifty Year Celebration, Contemp. Math., vol. 167, American Mathematical Society, Providence, RI, 1994.
[29] J. Eisert, M. Cramer, M.B. Plenio, Colloquium: Area laws for the entanglement entropy, Rev. Modern Phys. 82 (2010) 277–306.
[30] S.W. Ellacott, Computation of Faber series with application to numerical polynomial approximation in the complex plane, Math. Comp. 40 (1983) 575–587.
[31] D.R. Farenick, Algebras of Linear Transformations, Springer, New York, 2001.
[32] I.M. Gelfand, Normierte Ringe, Mat. Sb. 9 (1941) 3–23.
[33] I.M. Gelfand, M.A. Neumark, On the imbedding of normed rings in the ring of operators in Hilbert space, Mat. Sb. 12 (1943) 197–213.
[34] M.I. Gil', Operator Functions and Localization of Spectra, Lecture Notes in Math., vol. 1830, Springer, Berlin, 2003.
[35] P.-L. Giscard, S.J. Thwaite, D. Jaksch, Evaluating matrix functions by resummation on graphs: the method of path sums, SIAM J. Matrix Anal. Appl. 34 (2013) 445–469.
[36] K. Gröchenig, Wiener's lemma: Theme and variations. An introduction to spectral invariance and its applications, in: Four Short Courses on Harmonic Analysis, in: Appl. Numer. Harmon. Anal., Birkhäuser, Boston, 2010.
[37] K. Gröchenig, A. Klotz, Noncommutative approximation: inverse-closed subalgebras and off-diagonal decay of matrices, Constr. Approx. 32 (2010) 429–466.
[38] K. Grove, G.K. Pedersen, Diagonalizing matrices over $C(X)$, J. Funct. Anal. 59 (1984) 65–89.
[39] N. Higham, Functions of Matrices. Theory and Computation, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2008.
[40] A. Iserles, How large is the exponential of a banded matrix?, New Zealand J. Math. 29 (2000) 177–192.
[41] S. Jaffard, Propriétés des matrices "bien localisées" près de leur diagonale et quelques applications, Ann. Inst. H. Poincare Anal. Non Lineaire 7 (1990) 461–476.
[42] R. Kadison, Diagonalizing matrices, Amer. J. Math. 106 (1984) 1451–1468.
[43] R. Kadison, J. Ringrose, Fundamentals of the Theory of Operator Algebras. Vol. I, Elementary Theory, Academic Press, Orlando, FL, 1983.
[44] I. Kryshtal, T. Strohmer, T. Wertz, Localization of matrix factorizations, Found. Comput. Math. (2013), in press, arXiv:1305.1618v1.
[45] N. Le Bihan, J. Mars, Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing, Signal Process. 84 (2004) 1177–1199.
[46] T.A. Loring, Factorization of matrices of quaternions, Expo. Math. 30 (2012) 250–267.
[47] B.D. MacCluer, Elementary Functional Analysis, Grad. Texts in Math., Springer, New York, 2009.
[48] A.I. Markushevich, Theory of Functions of a Complex Variable, vol. III, Prentice Hall, Englewood Cliffs, NJ, 1967.
[49] N. Mastronardi, M.K. Ng, E.E. Tyrtyshnikov, Decay in functions of multi-band matrices, SIAM J. Matrix Anal. Appl. 31 (2010) 2721–2737.
[50] G. Meinardus, Approximation of Functions: Theory and Numerical Methods, Springer Tracts Natur. Philos., vol. 13, Springer, New York, 1967.
[51] L. Molinari, Identities and exponential bounds for transfer matrices, J. Phys. A: Math. Theor. 46 (2013) 254004.
[52] A. Osipov, On some issues related to the moment problem for the band matrices with operator elements, J. Math. Anal. Appl. 275 (2002) 657–675.
[53] A.M. Ostrowski, On some metrical properties of operator matrices and matrices partitioned into blocks, J. Math. Anal. Appl. 2 (1961) 161–209.
[54] F. Robert, Normes vectorielles de vecteurs et de matrices, Rev. Fr. Traitement Inform. (Chiffres) 7 (1964) 261–299.
[55] W. Rudin, Functional Analysis, McGraw–Hill, New York, NY, 1973.
[56] N. Schuch, J.I. Cirac, M.M. Wolf, Quantum states on harmonic lattices, Comm. Math. Phys. 267 (2006) 65–92.
[57] A. Sudbery, Quaternionic analysis, Math. Proc. Cambridge Philos. Soc. 85 (1979) 199–225.
[58] P.K. Suetin, Fundamental properties of Faber polynomials, Russian Math. Surveys 19 (1964) 121–149.
[59] C. Tobler, Low-rank tensor methods for linear systems and eigenvalue problems, PhD thesis, ETH, Zürich, 2012.
[60] C. Tretter, Spectral Theory of Block Operator Matrices and Applications, Imperial College Press, London, UK, 2008.
[61] Q. Ye, Error bounds for the Lanczos method for approximating matrix exponentials, SIAM J. Numer. Anal. 51 (2013) 68–87.
[62] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra Appl. 251 (1997) 21–57.