

FROM THE STRICT CHEBYSHEV APPROXIMANT OF A VECTOR TO THE STRICT SPECTRAL APPROXIMANT OF A MATRIX

KRYSTYNA ZIĘTAK

*Wrocław School of Information Technology
Wejherowska 28, 54-239 Wrocław, Poland
and*

*Wrocław University of Science and Technology, Faculty of Fundamental Problems of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
E-mail: krystyna.zietak@pwr.edu.pl*

Abstract. The main aim of this review paper is approximation of a complex rectangular matrix, with respect to the unitarily invariant norms, by matrices from a linear subspace or from a convex closed subset of matrices. In particular, we focus on the properties and characterizations of the strict spectral approximant which is in some sense the best among all approximants to a given matrix with respect to the spectral norm. We formulate a conjecture that the strict spectral approximant to a matrix by matrices from a linear subspace is the limit of approximants with respect to the c_p norms of Schatten when p tends to infinity. This conjecture corresponds to the conjecture stated by Rogers and Ward on approximation by positive semidefinite matrices and to the known analogous property of the strict Chebyshev approximation of a vector, proved by Descloux. We describe some special cases of an approximation of a matrix that confirm our conjecture and we discuss an attempt to prove the conjecture.

Additionally, we focus on orthogonality of matrices in the sense of Birkhoff and James and approximation by matrices whose spectrum is in a strip. For the last case we recall a conjecture on characterization of the best approximation of a matrix in the spectral norm by matrices whose spectrum is in a strip. This kind of approximation is a generalization of the famous concept of Halmos of approximation by positive operators.

2010 *Mathematics Subject Classification*: Primary: 15A60, 47A58, 65F20, 65F30; Secondary: 15A18, 15A57, 41A65, 41A36, 49J52, 65D10, 65F35.

Key words and phrases: approximation of matrices, unitarily invariant norm, c_p norm of Schatten, lexicographic ordering, subdifferential of norm, dual matrix, strict spectral approximant, c_p -minimal approximation, Birkhoff–James orthogonality, zero trace matrix, positive definite matrix, restricted spectrum, overdetermined system of linear equations, l_p -solution, Chebyshev solution, strict Chebyshev solution, Pólya algorithm.

The paper is in final form and no version of it will be published elsewhere.

For the convenience of the reader we recall properties of the Chebyshev solutions of overdetermined system of linear equations and the solutions with respect to the l_p norm, and the properties of dual matrices, related to subdifferentials of norm of a matrix, defined for the unitarily invariant norms.

The paper is based on a series of papers by the author and on many papers of other authors, related to the considered problems.

1. Introduction. A matrix nearness problem, i.e. finding a matrix nearest to a given matrix from some given class of matrices where a distance is measured in a norm, is a subject of many papers. A survey of nearness classical problems is given in [42] with particular emphasis on the fundamental properties of symmetry, positive definiteness, orthogonality, rank-deficiency and instability. For example, one considers an approximation of a given matrix by symmetric, unitary, normal, lower rank, positive semi-definite or instable matrices with respect to the unitarily invariant norms. These kinds of nearness problems were studied in [27, 31, 50, 81, 93]; for generalization of some of these problems see, for example, [51, 56, 57].

Let \mathbb{M} be a nonempty closed convex subset of the normed linear space $\mathbb{C}^{m \times n}$ of complex $m \times n$ matrices. For a given $A \in \mathbb{C}^{m \times n}$ we consider the problem of finding a matrix $B \in \mathbb{M}$ such that

$$\|A - B\| = \min_{X \in \mathbb{M}} \|A - X\|. \quad (1.1)$$

We will consider unitarily invariant norms. A norm $\|\cdot\|$ is unitarily invariant if

$$\|WA\| = \|AZ\| = \|A\|$$

for all unitary matrices W and Z .

Let

$$A = U\Sigma V^H \quad (1.2)$$

be the singular value decomposition of a matrix $A \in \mathbb{C}^{m \times n}$, i.e. U and V are unitary, $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal with the ordered singular values

$$\sigma_1(A) \geq \sigma_2(A) \geq \dots \geq 0$$

on the main diagonal. We denote the singular value decomposition by SVD for short. The c_p (Schatten) norm of A , $1 \leq p \leq \infty$, is equal to the l_p norm of the vector $\sigma(A) = [\sigma_1(A), \sigma_2(A), \dots, \sigma_t(A)]^T$, $t = \min\{m, n\}$,

$$\|A\|_p = \left(\sum_{j=1}^t \sigma_j^p(A) \right)^{1/p}. \quad (1.3)$$

We have assumed that the singular values are ordered decreasingly, hence the spectral norm $\|A\|_\infty$ is equal to $\sigma_1(A)$ (the largest singular value of A). For $p = 1$ we have the trace (nuclear) norm, i.e. the sum of all singular values. For $p = 2$ we obtain the Frobenius norm.

If $1 < p < \infty$, then the c_p norm is strictly convex (see [102]). Thus for $1 < p < \infty$ the c_p -approximant $B = B_p \in \mathbb{M}$ to A satisfying (1.1) is unique. The spectral norm and the trace norm are not strictly convex. Therefore the spectral approximant and the trace approximant may be not unique in the general case. We select among all approximants

to A an approximant which can be considered as the best from the best. For the spectral norm such an approximant, called a *strict spectral approximant*, was introduced in [106]. For the trace norm Legg and Ward [60] have introduced the *canonical trace approximant*. Both approximants are determined uniquely.

A matrix $B^{(\text{st})} \in \mathbb{M}$ is a *strict spectral approximant to a matrix A* if the vector $\sigma(R^{(\text{st})})$ of the ordered singular values of the residual matrix $R^{(\text{st})} = A - B^{(\text{st})}$ is minimal with respect to the lexicographic ordering in the set of vectors $\{\sigma \in \mathbb{R}^t : \sigma = \sigma(A - X), X \in \mathbb{M}\}$ (see [106]). Examples of the strict spectral approximants are given in [108, 110] and they are recalled in Section 9.

The concept of the strict spectral approximation rests on the *strict Chebyshev approximation of a vector* by vectors from a linear subspace of \mathbb{R}^m , introduced by Rice [78] (see also [24, 94]). The definition of $B^{(\text{st})}$ follows from a more general definition of a *strict approximation* introduced in [106], where we have proved that there exists exactly one strict spectral approximant to any given complex matrix A by elements from a closed convex set. The strict approximation includes the strict Chebyshev approximation of a vector and the strict spectral approximation of a matrix as particular cases.

The lexicographic ordering was used also by Young [100] to distinguish certain solution of the Nevalinna–Pick problem for matrix-valued functions, called *superoptimal approximation* (compare [23, 99]).

Let us recall that the c_p norms are unitarily invariant and depend only on the singular values of a matrix. They correspond to the l_p norm of a vector. Therefore there are some analogies between properties of solutions of problem (1.1) for the c_p norms, $1 \leq p \leq \infty$, and an approximation of a vector in l_p norms. Some of such analogies are mentioned in [110].

In this paper we discuss an attempt to prove the following conjecture, stated in the report [103].

CONJECTURE 1.1. *Let \mathbb{M} be a linear subspace of $\mathbb{C}^{m \times n}$ and let $B_p \in \mathbb{M}$ be the c_p -approximant to $A \in \mathbb{C}^{m \times n}$. As p tends to ∞ , the matrix $R_p = A - B_p$ converges to $A - B^{(\text{st})}$, where $B^{(\text{st})}$ is the strict spectral approximant to A .*

Conjecture 1.1 is motivated by the well known result of Descloux [24] that the l_p -approximants of a real vector by elements from a linear real subspace of \mathbb{R}^m converge to the strict Chebyshev approximant as $p \rightarrow \infty$ (see the next section). This property is called the Pólya algorithm for computing the strict Chebyshev approximant.

Conjecture 1.1 is analogous to the conjecture of Rogers and Ward for the approximation by positive operators, stated in [80] in 1981. We recall their conjecture in Section 7. Our Conjecture 1.1 was formulated independently of Rogers and Ward, and it was presented by the author, for example, at the XII Householder Symposium (Lake Arrowhead, USA, 1993), the SIAM Conference on Linear Algebra in Signals, Systems and Control (Boston, USA, 2001), the Conference on Numerical Matrix Analysis and Operator Theory (Helsinki, Finland, 2008) and the Householder Symposium (Zeuthen, Germany, 2008).

Legg and Ward [60] have proved that the c_p -approximants of a matrix by matrices from a convex subset converge to the canonical trace approximant when p tends to 1

(the result of Legg and Ward is formulated for the linear bounded operators in a finite-dimensional Hilbert space). The canonical trace approximant corresponds to the natural l_1 -approximant to a vector by elements from a linear subspace (a specific solution of an overdetermined system of linear equations in the l_1 norm), studied by Fischer [33] (for the case of L_1 approximation see [55]). Fischer has proved that the natural l_1 -approximant is the limit of the l_p -approximants of a vector when $p \rightarrow 1$ (see Section 7). Thus there is an analogy between the vector case and the matrix case for $p \rightarrow 1$. This fact is an argument that Conjecture 1.1 is worth to be considered because for the vector case an analogous property has been proved by Descloux [24] for $p \rightarrow \infty$ (see (3.1) in Section 3). In the paper we present an idea of incomplete proof of Conjecture 1.1. For this reason we review the properties of the strict spectral approximants.

We also discuss the approximation of a given matrix by zero-trace matrices and by matrices with prescribed spectrum (see [11, 52, 57, 104, 107]). We focus on the orthogonality in the Birkhoff and James sense (see [12, 13, 14, 38, 49, 64]).

The paper is organized as follows. In Section 2 we recall properties of solutions of overdetermined systems of real linear equations with respect to the l_p norm and the Chebyshev norm ($p = \infty$). It is continued in Section 3, where one considers the Pólya algorithm and one formulates some doubts about some known proofs of its convergence. An introduction to approximation of a matrix by positive matrices is presented in Section 4. An extension of this kind of approximation to the case of matrices with spectra in a strip is recalled in Section 5. Properties of dual matrices, i.e. subdifferentials of norm of a matrix, for unitarily invariant norms are recalled in Section 6. This notion can be applied to characterize approximants of a matrix with respect to the unitarily invariant norms as it is done in Section 7 concerning an approximation by matrices from a linear subspace. Particular cases of such problems, related to a concept of orthogonality of matrices in the Birkhoff and James sense, are presented in Section 8. We give a new proof of one of the results of Grover [38], concerning the Birkhoff and James orthogonality of a rectangular complex matrix to any subspace of matrices (see Theorem 8.4).

Sections 9 and 10 are devoted to the strict spectral approximation, which is the main topic of the paper. In Section 9 we give examples confirming Conjecture 1.1. An attempt to prove the conjecture is developed in Section 10.

In Section 11 we consider [45, Lemma 2.5] and [30, Lemma 5]. We show that the proof of [30, Lemma 5] is false. This lemma is crucial in an elementary proof of the result of Descloux, proposed by Egger and Huotari in [30] (for further analysis of [30, Lemma 5] see [111]). Therefore the proof of Egger and Huotari needs improvements. It is out of the scope of the paper. In the paper [45] one proposes a generalization of the result of Descloux [24] about the convergence of the Pólya algorithm to the case of cylindrical sets. Unfortunately, it is known that [45, Lemma 2.5] is false (see [30]). In Section 11 we give an example which confirm it. Therefore it is not clear whether this generalization of Huotari, Legg and Townsend is valid.

2. The Chebyshev and l_p -solutions of an overdetermined system of linear equations. Let us consider an overdetermined system of real linear equations

$$Gx = b, \quad G \in \mathbb{R}^{m \times n}, \quad b \in \mathbb{R}^m, \quad m > n. \tag{2.1}$$

We assume that G is of full rank, i.e. $\text{rank}(G) = n$. In this section we focus on solutions of (2.1) with respect to the l_p norms ($r = [r_1, \dots, r_m]^T \in \mathbb{R}^m$)

$$\|r\|_p = \left(\sum_{j=1}^m |r_j|^p \right)^{1/p}, \quad 1 \leq p \leq \infty.$$

For $p = \infty$ we obtain the Chebyshev norm (l_∞ norm)

$$\|r\|_\infty = \max_{1 \leq j \leq m} |r_j|.$$

The k -major l_p norms are defined for $1 \leq p < \infty$ as follows (see [70, 97])

$$\|r\|_{k,p} = \left(\max_{0 \leq \alpha_i \leq 1} \sum_{i=1}^m \alpha_i |r_i|^p \right)^{1/p}, \tag{2.2}$$

where $\sum_{i=1}^m \alpha_i = k$. Notice that the expression on the right hand side in brackets just picks out the leading k terms (in modulus) of the vector r .

A vector $x^{(\text{Ch})} \in \mathbb{R}^n$ is a Chebyshev solution of (2.1) if

$$\|Gx^{(\text{Ch})} - b\|_\infty = \delta = \min_{x \in \mathbb{R}^n} \|Gx - b\|_\infty. \tag{2.3}$$

The problem (2.3) is connected with a discrete Chebyshev approximation of a function whose values are given on a finite set of points. The vector $Gx^{(\text{Ch})}$ can be interpreted as an (linear) approximant to b by vectors from the linear subspace spanned by columns of the matrix G .

The problem (2.3) with the l_∞ norm replaced by the l_p norm, $1 < p < \infty$, has a unique solution $x^{(p)}$ because the l_p norm is strictly convex in this case and G is of full rank. For $p = \infty$ and $p = 1$ the solution can be not unique in the general case. Therefore for $p = \infty$ and $p = 1$ one selects special solutions: the strict Chebyshev solution (see [78]) and the natural l_1 -solution (see [33]), respectively. We now recall the definition of the natural l_1 -solution of $Gx = b$. The natural l_1 -solution is the unique vector $x^{(\text{nat})}$ that minimizes the expression (see, for example, [76])

$$\sum_{j=1}^m |(Gx - b)_j| \ln(|(Gx - b)_j|) \tag{2.4}$$

over all l_1 -solutions x of $Gx = b$, where $(y)_j$ denotes the j -th component of a vector y . Here we put $\ln(0) = 0$. Fischer [33] has proved that the natural l_1 -solution is the limit of the l_p -solutions as $p \rightarrow 1$. The definition of the strict Chebyshev solution will be recalled later.

The characterization of the solution of the overdetermined system $Gx = b$ with respect to the l_p norm, $1 < p < \infty$, follows from the general Theorem 1.7 in [94], characterizing the linear approximation. This characterization specialized to our case $Gx = b$ can be formulated as follows (see, for example, [69, formula (3)], [71, Theorem 2.1], [101], [109, page 217]). The vector $x^{(p)}$ is the l_p -solution of $Gx = b$ if and only if

$$G^T v = 0 \quad \text{for} \quad v = [v_1, \dots, v_m]^T, \quad v_j = \text{sign}(r_j) |r_j|^{p-1},$$

where $r = b - Gx^{(p)}$. The characterization of the solutions with respect to the k -major l_p norms is presented in [97]. In fact all these characterizations apply the concept of dual vectors (subdifferentials of a norm), see Section 6, [87, 88, 109]. The properties of the approximation with respect to the l_p norm are investigated, for example, in [32, 75, 94].

Let us consider problem (2.3). Let $\mathbb{J}(x)$ denote the set of indices i of the components of $r(x) = [r_1(x), \dots, r_m(x)]^T = b - Gx$ that (in modulus) equal $\|r(x)\|_\infty$. A vector $x^{(\text{Ch})} \in \mathbb{R}^n$ is a Chebyshev solution of (2.1) if and only if there exists a subset \mathbb{I} of $\mathbb{J}(x^{(\text{Ch})})$, containing at most $(n + 1)$ indices, and a nontrivial vector $\lambda = [\lambda_1, \dots, \lambda_m]^T \in \mathbb{R}^m$ such that (see [94, p. 27])

$$G^T \lambda = 0, \quad \lambda_i = 0 \text{ for } i \notin \mathbb{I}, \quad \lambda_i \text{sign}(r_i(x^{(\text{Ch})})) \geq 0 \text{ for } i \in \mathbb{I},$$

where we put $\text{sign}(0) = 0$.

It is known that if G satisfies the Haar condition, i.e. every $n \times n$ submatrix of G is nonsingular, the Chebyshev solution of (2.1) is unique (see, for example, [94, Section 2.3]). Let \mathbb{S}_G be a set of all Chebyshev solutions of $Gx = b$ and let δ be the approximation error in (2.3). We define

$$\mathbb{J}_G = \{i : |r_i(x)| = \delta \text{ for all } x \in \mathbb{S}_G\}.$$

This set is called the characteristic set of b relative to G (see [24, 26]). It is known that for every Chebyshev solution $x^{(\text{Ch})}$ of (2.1) we have

$$r_j(x^{(\text{Ch})}) = \varepsilon_j \delta \quad \text{for } j \in \mathbb{J}_G, \tag{2.5}$$

where $\varepsilon_j = \pm 1$ independently on $x^{(\text{Ch})}$ (see [24, 78, 94]). Therefore the signs of some components of the residual vector $b - Gx^{(\text{Ch})}$ whose absolute values are equal to the error δ are common for all Chebyshev solutions $x^{(\text{Ch})}$ of (2.1) (for an extension of this property to the case of overdetermined system of complex linear equations see [15]). We will show later what in the case of the spectral approximation of a matrix corresponds to this property of the Chebyshev approximation of a vector (see Theorems 7.3 and 10.2). The above property (2.5) of the Chebyshev solutions is a base of an algorithm for computing the strict Chebyshev solution (see [26]).

We now recall the definition of a strict Chebyshev approximant to b by vectors of the form Gx . We use a version of the definition proposed in [30, 46], adapting it to (2.3). Let $|r(x)|$ denote a vector whose components are moduli of components of $r(x) = b - Gx$. Let $\Psi(|r(x)|)$ be a vector whose components are components of $|r(x)|$, arranged in non-increasing order. Let $x^{(\text{st})}$ be such that $\Psi(|r(x^{(\text{st})}|)$ is minimal in the lexicographic ordering on $\{\Psi(|r(x)|) : x \in \mathbb{R}^n\}$. Then $Gx^{(\text{st})}$ is the unique strict Chebyshev approximant to b . Huotari and Li [46] have shown that this definition is equivalent the more complicated original definition of strict Chebyshev approximants introduced by Rice in [78].

Methods for computing the strict Chebyshev solution $x^{(\text{st})}$ of overdetermined system $Gx = b$ are proposed, for example, in [1, 18, 26, 92, 94].

In the next section we recall the result of Descloux [24] that the l_p -solutions $x^{(p)}$ of $Gx = b$, $\text{rank}(G) = n$, converge to the strict Chebyshev solution when $p \rightarrow \infty$. This property is used in the process called the Pólya algorithm, where the Chebyshev solution is found via the l_p -solutions (see [74]).

3. The strict Chebyshev solution and the Pólya algorithm. Let real functions $f(\xi), g_1(\xi), \dots, g_n(\xi)$ be determined on a finite set of points $\Xi = \{\xi_1, \dots, \xi_m\}$ with $m \geq n$. We assume that the functions $g_1(\xi), \dots, g_n(\xi)$ are linearly independent. For example, we can select $g_j(\xi) = \xi^{j-1}$. The problem of an approximation of $f(\xi)$ by linear combinations of the functions $g_1(\xi), \dots, g_n(\xi)$ over Ξ can be formulated as an approximation of a given vector $b = [f(\xi_1), \dots, f(\xi_m)]^T \in \mathbb{R}^m$ by vectors from a linear subspace $\mathbb{V} = \text{span}\{g^{(1)}, \dots, g^{(n)}\}$ of \mathbb{R}^m , where $g^{(j)} = [g_j(\xi_1), \dots, g_j(\xi_m)]^T$ are linearly independent. Let the columns of a matrix $G \in \mathbb{R}^{m \times n}$ be formed by the vectors $g^{(1)}, \dots, g^{(n)}$. Then we approximate the vector b by vectors from a linear subspace $\mathbb{V} = \{Gx : x \in \mathbb{R}^n\}$ with respect to a norm $\|\cdot\|$. It is equivalent to solving an overdetermined system of linear equations $Gx = b$ with respect to the norm $\|\cdot\|$. In particular, we can select the l_p norm for $1 \leq p \leq \infty$ (see Section 2).

It is obvious that for any $x \in \mathbb{R}^n$

$$\lim_{p \rightarrow \infty} \|x\|_p = \|x\|_\infty.$$

Therefore it is not surprising that there is a link between the Pólya algorithm and a strict Chebyshev approximation. In 1913 Pólya [74] proposed an algorithm to obtain the best uniform-norm (Chebyshev) approximants to continuous functions by polynomials via l_p -approximations. Descloux has proved in [24] that the Pólya algorithm converges in the case of a linear approximation of a function on a finite set of points. For problem (2.3) this well-known result of Descloux can be formulated as follows

$$\lim_{p \rightarrow \infty} (Gx^{(p)} - b) = Gx^{(\text{st})} - b, \tag{3.1}$$

where $x^{(p)}$ is the l_p -solution of $Gx = b$ and $x^{(\text{st})}$ is the strict Chebyshev solution. We formulate this property in the form (3.1) to underline that $Gx^{(p)}$ is the l_p -approximation to b by vectors from a linear subspace \mathbb{V} .

The properties of the Pólya algorithm and the strict Chebyshev approximations are the subject of a series of papers by Huotari and his coauthors (see, for example, [28, 29, 30, 46, 47, 48]). In [28] Egger and Huotari give two important examples for the approximation over closed convex sets in \mathbb{R}^m . In the first example the sequence of the l_p -approximants converges when $p \rightarrow \infty$, however, the limit is not the strict approximant. In the second example they show that there exists a closed, convex set in \mathbb{R}^n for which the l_p -approximants to a given vector of \mathbb{R}^m fail to converge as $p \rightarrow \infty$. Some characterizations of sets on which the Pólya algorithm converges to the strict approximants are given in [46]. Future results, presented in the series of papers by Marano and his coauthors (see [48, 69, 77]), concern a rate of convergence of the Pólya algorithm.

There are different proofs of the result of Descloux. These proofs concern the original problem from [24], i.e. a convergence of the Pólya algorithm (3.1) (see [30, 69]) and some generalizations of this problem (see [45, 68]). The idea of the proof of [30, Theorem 7] of Egger and Huotari about the convergence of the Pólya algorithm is different from the original proof of Descloux [24]. Both proofs are a contrario, which leads to a construction of an approximant which would be a better approximant in the l_p norm than the best one. Using our formulation (3.1) of the problem, one can observe that in the proof of

Egger and Huotari one constructs for a given vector x such its modification δx that does not change certain elements of the residual vector (residuum) $b - G(x + \delta x)$ in comparison with the residuum $b - Gx$. Unfortunately, this construction is based on Lemma 5 in [30]. In Section 11 we show that the proof of Lemma 5 of Egger and Huotari is false (see [111] for future analysis of Lemma 5). Therefore the proof of Egger and Huotari of the result of Descloux is not correct. Lemma 5 is part (ii) of Lemma 3.1 formulated below.

We would like to emphasize that Lemma 5, concerning properties of linear subspaces of \mathbb{R}^m , follows as a particular case of Lemma 2.5 in [45], which is for E -cylindrical subsets of \mathbb{R}^m . Unfortunately, Egger and Huotari [30] state without any comments that Lemma 2.5 in [45] is false. In Section 11 we analyze the original proofs of Lemma 5 and Lemma 2.5, presented in [30, 45], respectively.

We now formulate the lemma based on [30, Lemma 5] and [45, Lemma 2.5] for linear subspaces. Let $y = [y_1, \dots, y_m]^T \in \mathbb{R}^m$, $1 \leq k \leq m$. We introduce the auxiliary notation

$$\Pi_k(y) = [y_1, \dots, y_k, 0, \dots, 0]^T. \tag{3.2}$$

LEMMA 3.1. *Let $\mathbb{V} = \{Gu : u \in \mathbb{R}^n\}$ with $G \in \mathbb{R}^{m \times n}$ of rank n . Let $1 \leq k \leq m$ and $\varepsilon > 0$. Then*

- (i) *if $x \in \mathbb{V}$ then there exists $\eta > 0$ such that if $y \in \mathbb{V}$ and $\|\Pi_k(x - y)\|_\infty < \eta$, then there exists a vector $z \in \mathbb{V}$ such that $\Pi_k(z) = \Pi_k(y)$ and $\|x - z\|_\infty < \varepsilon$;*
- (ii) *there exists $\eta > 0$ such that if $y \in \mathbb{V}$ and $\|\Pi_k(y)\|_\infty < \eta$, then there exists a vector $z \in \mathbb{V}$ such that $\Pi_k(z) = \Pi_k(y)$ and $\|z\|_\infty < \varepsilon$.*

The parts (i) and (ii) are equivalent. Part (ii) follows from part (i) by putting $x = 0$. On the other hand, part (i) follows from part (ii) by replacing y and z by $y - x$ and $z - x$, respectively. The first part of Lemma 3.1 is Lemma 2.5 from [45], restricted to subspaces. The second part is Lemma 5 from [30] with the condition $\|z\|_1 < \varepsilon$ replaced by the weaker condition $\|z\|_\infty < \varepsilon$.

Unfortunately, there are also some doubts about the proof of Descloux [24]. Borowsky [15] states that the formulation of [24, Part (c)] is incorrect (modulus is missing) and he proves a stronger version of [24, Part (c)] (see [15, Lemma 4]), because the original version of Descloux was not sufficient for the proof of [15, Theorem 6]. In fact the situation is more complicated than it was noticed by Borowsky. The set of approximations satisfying condition (1) in [24] is not a linear subspace because due to the condition [24, (1)] there is a dependence between the coefficients, defined in the new approximation problem following formula [24, (1)]. Moreover, the proof of a crucial inequality involving a constant w on top of page 1022 is not presented. Therefore some people question the completeness of the proof of Descloux.

4. The c_p -minimal positive and trace class positive approximants of a matrix. Halmos [39] initiated the study of positive operator approximation. Operators are bounded linear transformations on a complex Hilbert space. We formulate his result and further results of Legg, Rogers, and Ward only for matrices. Let $A \in \mathbb{C}^{n \times n}$ have

the Cartesian representation $A = B + \iota C$, where $\iota = \sqrt{-1}$ and the matrices B, C are Hermitian. Then (see [3, 16, 17, 39, 80])

$$\inf\{\|A - P\|_\infty : P \geq 0\} = \inf\{r : r \geq \|C\|_\infty, B + (r^2I - C^2)^{1/2} \geq 0\}, \tag{4.1}$$

where $X^{1/2}$ denotes the principal square root of X . The inequality $P \geq 0$ denotes that the matrix P is Hermitian positive semidefinite. We will say for short that P is positive. From (4.1) it follows that the positive number r of which $\delta(A)$ is the infimum must satisfy two conditions

$$r^2I - C^2 \geq 0, \quad B + (r^2I - C^2)^{1/2} \geq 0. \tag{4.2}$$

Let the first infimum in (4.1) be denoted by $\delta_H(A)$. Then

$$P_H = B + ((\delta_H(A))^2I - C^2)^{1/2} \tag{4.3}$$

is a positive approximant to A (see [39]). Higham [41] has proposed an algorithm for computing the Halmos approximant P_H .

Rogers and Ward [80] have proved that each complex matrix A has a positive approximant P_{\min} solving (4.1) such that $A - P_{\min}$ is normal and such that for each positive matrix P , $P \neq P_{\min}$,

$$\|A - P\|_p > \|A - P_{\min}\|_p \tag{4.4}$$

for all finite p sufficiently large. The matrix P_{\min} is called the c_p -minimal positive approximant to A or a canonical minimal positive approximant. It is a positive approximant with respect to the spectral norm. The construction of the c_p -minimal positive approximant is based on the concept of the strict Chebyshev approximant to a vector (see the proof of Theorem 1.1 in [80]). We would like to note that P_{\min} is minimal in the sense given by (4.4) and in general case it is not minimal in the sense of the partial Loewner order. We say that Y is below X , $X \geq Y$, with respect the partial Loewner order on the set of Hermitian matrices if and only if $X - Y$ is positive semidefinite. Bouldin [16] has proved that the Halmos approximant P_H (see (4.3)) is maximal in the Loewner ordering, i.e. $P_H \geq P$ for all positive approximants P , and there need not be a positive approximant minimal in this sense. The Halmos approximant P_H is maximal also in the strict spectral ordering because every singular value of $A - P_H$ is equal to $\delta_H(A)$ (see [108]). The strict spectral ordering is defined in the following way (see [108]): we say that $X \geq_{st} Y$ if the vector $\sigma(X)$ of the ordered singular values of X is bigger than $\sigma(Y)$ in the lexicographic order.

Let P_p be the c_p -positive approximant of A with respect to the c_p norm. Rogers and Ward [80] have formulated the following conjecture.

CONJECTURE 4.1. *The c_p -positive approximant P_p of A converges to the c_p -minimal approximant P_{\min} when $p \rightarrow \infty$.*

In 1985 Rogers and Ward [80] have formulated also a question concerning the case $p \rightarrow 1$: does $\lim_{p \rightarrow 1}(A - P_p)$ exist? If the answer is yes, then can the limit be identified by any characteristics? Affirmative answers are given by Legg and Ward [60]. We now recall shortly their result. Let \mathbb{M} be a convex subset of matrices of order n . Then the best c_p -approximant from \mathbb{M} to A converges to a selected trace class approximant $A^{(\text{tr})}$

as $p \rightarrow 1$. This approximant is called the canonical trace class approximant. Furthermore, $A^{(\text{tr})}$ is the unique trace class approximant minimizing (compare (2.4) and [33])

$$\sum_{i=1}^n \sigma_i(A - X) \ln \sigma_i(A - X)$$

over the set of the c_1 -approximants X to A from \mathbb{M} . This result of Legg and Ward [60] has been obtained for matrices independently of an analogous result of Fischer for vectors and $p \rightarrow 1$, obtained in 1983 in [33] (see Section 2).

Properties of approximation by positive matrices in the unitarily invariant norms are investigated, for example, in [11] (for the c_1 norm see [2]).

5. Approximation with restricted spectra. The idea of approximation by positive matrices has been extended to other cases of approximation of matrices (see, for example, [40, 84, 91]). We now consider approximation with restricted spectra (see, for example, [22, 51]). Let $\mathbb{X}(\mathbb{S})$ denote the set of all complex matrices X of order n , whose spectra $\text{spect}(X)$ are in a convex closed subset \mathbb{S} of a complex plane \mathbb{C} . In [51] one considers the following problem:

$$\min_{X \in \mathbb{X}(\mathbb{S}_a)} \|A - X\|_\infty \tag{5.1}$$

for $\mathbb{S}_a = [0, \infty) \times [0, a] = \{x + yn : 0 \leq x, 0 \leq y \leq a\}$. Unfortunately, the characterization of approximants solving (5.1), given in [51], is not quite correct. Khalil and Maher have not realized that their Theorem 6.2 is not valid for the problem (5.1), but for the following problem (see [57])

$$\min_{X \in \mathbb{Y}(\mathbb{S}_a)} \|A - X\|_\infty, \tag{5.2}$$

where $\mathbb{Y}(\mathbb{S}_a)$ denotes the set all matrices of the form $X = X_1 + iX_2$, where X_1 and X_2 are Hermitian, and $\text{spect}(X_1) \subseteq [0, \infty)$ and $\text{spect}(X_2) \subseteq [0, a]$. If we choose $a = 0$ then the problem (5.2) with \mathbb{S}_0 is the problem of Halmos, i.e. approximation by positive matrices. The following corrected version of the theorem of Khalil and Maher is proved in [57].

THEOREM 5.1. *Let $A = B + iC \in \mathbb{C}^{n \times n}$ for B, C Hermitian. Let*

$$\delta(A) = \inf\{\|A - X\|_\infty : X \in \mathbb{Y}(\mathbb{S}_a)\}.$$

Then

$$\delta(A) = \inf\{r : B + (r^2 I - (C - \tilde{C})^2)^{1/2} \geq 0 \text{ for some Hermitian } \tilde{C}\},$$

where $\text{spect}(\tilde{C}) \subseteq [0, a]$, and for some Hermitian \hat{C} with the spectrum in $[0, a]$ the matrix

$$\hat{X} = B + (\delta(A)^2 I - (C - \hat{C})^2)^{1/2} + i\hat{C} \tag{5.3}$$

is a solution of (5.2).

As in (4.2), the positive numbers r of which $\delta(A)$ is the infimum must satisfy

$$r^2 I - (C - \tilde{C})^2 \geq 0.$$

Theorem 5.1 is formulated in [57] in a bit more general setting covering also the approximation by matrices whose spectrum lies in a quarter.

Unfortunately, the matrix \widehat{C} in (5.3) is not determined explicitly in the general case. However, there are two special cases in which the problem is solved completely. For this purpose one uses retractions. Let F_a be a retraction of real numbers onto the interval $[0, a]$. Then $F_a(x) = x$ for $x \in [0, a]$, $F_a(x) = a$ for $x > a$ and $F_a(x) = 0$ for $x < 0$. It is known that every Hermitian matrix X has the following spectral decomposition

$$X = Q \operatorname{diag}(\lambda_j)Q^H,$$

where Q is unitary and eigenvalues λ_j are real. We define

$$F_a(X) = Q \operatorname{diag}(F_a(\lambda_j))Q^H.$$

Then from Theorem 5.1 we obtain the following corollary (see [57, Corollary 3.3]).

COROLLARY 5.2. *If in Theorem 5.1 B is positive, then $\delta(A) = \|C - F_a(C)\|_\infty$ and*

$$\widehat{X} = B + (\delta(A)^2 I - (C - F_a(C))^2)^{1/2}$$

is a solution of (5.2). If B is not positive and C has the spectrum in $[0, a]$, then $\delta(A)$ is equal to the modulus of the minimal eigenvalue of B and

$$\widehat{X} = B + \delta(A)I.$$

In both cases described in Corollary 5.2 the matrix \widehat{C} in (5.3) has a common form $\widehat{C} = F_a(C)$. It is known that $F_a(C)$ is the strict spectral approximation of C by Hermitian matrices with the spectra in $[0, a]$ (see [57, Theorem 2.1]). Extensive numerical experiments and some theoretical considerations lead to the following conjecture, formulated in [57].

CONJECTURE 5.3. *Let the assumptions of Theorem 5.1 be satisfied. Then*

$$\delta(A) = \inf \{ r : B + (r^2 I - (C - F_a(C))^2)^{1/2} \geq 0, r^2 I - (C - F_a(C))^2 \geq 0 \}$$

and in (5.3) we can put $\widehat{C} = F_a(C)$.

Numerical experiments show that in the general case the matrix \widehat{C} in (5.3) cannot be replaced by any approximant to C by Hermitian matrices with the spectrum in $[0, a]$. It is also confirmed by Example 3.7 presented in [57].

6. Dual matrices and subdifferentials of norm of a matrix for unitarily invariant norms.

We now present some concepts which are used in convex analysis of real vector spaces. It will be adopted to the space of rectangular complex matrices. Let \mathbb{V} be an n -dimensional real vector space, endowed with an inner product $\langle \cdot, \cdot \rangle$ and let ν be a norm on \mathbb{V} . By definition, the polar (the dual norm) of ν is the functional $\nu^* : \mathbb{V} \rightarrow \mathbb{R}$ given by

$$\nu^*(x) = \sup \{ \langle y, x \rangle : y \in \mathbb{B}_\nu \}, \tag{6.1}$$

where \mathbb{B}_ν denotes the unit ball $\{x \in \mathbb{V} : \nu(x) \leq 1\}$. Then ν^* is also a norm and generalized Hölder inequality

$$\langle y, x \rangle \leq \nu(x)\nu^*(y) \tag{6.2}$$

holds for all $x, y \in \mathbb{V}$. Of course, $(\nu^*)^* = \nu$ and for any $x \in \mathbb{V}$ there exists a nonzero $y \in \mathbb{V}$ for which (6.2) becomes equality. The subdifferential of ν at $a \in \mathbb{V}$ is the set (see [79, Section 13], [83])

$$\partial\nu(a) = \{ y \in \mathbb{V} : \langle y, x - a \rangle \leq \nu(x) - \nu(a) \text{ for all } x \in \mathbb{V} \}.$$

If $\nu(x) = \nu^*(y) = 1$, then

$$y \in \partial\nu(x) \quad \text{if and only if} \quad x \in \partial\nu^*(y).$$

Moreover, we have the identity (see, for example, [83, formula (9)])

$$\partial\nu(a) = \{y \in \mathbb{V} : \langle a, y \rangle = \nu(a)\nu^*(y), \nu^*(y) = 1\}. \tag{6.3}$$

Thus the concept of subdifferentials is closely related to dual vectors. If $\nu^*(y) = 1$ and in (6.2) the equality holds, then y is said to be dual to x with respect to the norm ν (see [83]). Other authors adopt slightly different definition of dual vectors (see e.g. [6, 37, 73, 87, 88, 102, 109]) which will be applied in this paper (see (6.4) and (6.6) below).

Let now $v = [v_1, \dots, v_t]^T \in \mathbb{V} = \mathbb{R}^t$ and let $\langle v, y \rangle = y^T v$. Then the dual norm $\|\cdot\|^*$ to the norm $\|\cdot\|$ of a vector v is determined in the following way (see (6.1)):

$$\|v\|^* = \max_{\|y\| \leq 1} y^T v. \tag{6.4}$$

It is known that the l_q norm is dual to the l_p norm where p and q satisfy

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{6.5}$$

Now a vector v^* is dual to v with respect to the norm $\|\cdot\|$ (v^* is the $\|\cdot\|$ -dual vector to v) if the maximum (6.4) is reached by v^* , i.e. (see (6.2))

$$v^T v^* = \|v\|^*, \quad \|v^*\| = 1. \tag{6.6}$$

The l_∞ -dual vector v^* to v has the following elements

$$v_j^* = \begin{cases} \text{sign}(v_j) & \text{if } v_j \neq 0, \\ h_j & \text{if } v_j = 0, \end{cases} \tag{6.7}$$

where h_j are arbitrary real numbers such that $|h_j| \leq 1$. The l_1 -dual vector has the elements

$$v_j^* = \begin{cases} \text{sign}(v_j)g_j & \text{if } |v_j| = \|v\|_\infty, \\ 0 & \text{if } |v_j| < \|v\|_\infty, \end{cases} \tag{6.8}$$

where $g_j \geq 0$ and $\sum_j g_j = 1$. For $1 < p < \infty$ we have

$$v_j^* = \text{sign}(v_j) \left(\frac{|v_j|}{\|v\|_q} \right)^{q-1}, \tag{6.9}$$

where p and q satisfy (6.5). If $p \rightarrow \infty$, then the l_p -dual vector v^* tends to the following l_∞ -dual vector v (see (6.7))

$$v_j^* = \begin{cases} \text{sign}(v_j) & \text{if } v_j \neq 0, \\ 0 & \text{if } v_j = 0. \end{cases} \tag{6.10}$$

Properties of dual vectors with respect to orthant-monotonic norms were studied in [109]. We now show how concepts of dual vectors and subdifferentials can be extended to the space of complex rectangular matrices endowed with unitarily invariant norms (see [83, 95, 102, 105, 107]).

Let now $\mathbb{V} = \mathbb{K}^{m \times n}$ with $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. In the real linear space $\mathbb{R}^{m \times n}$ of dimension mn , the usual inner product $\langle A, B \rangle$ is defined as $\text{trace}(B^T A)$. The $\mathbb{C}^{m \times n}$ treated

as real space of dimension $2mn$ is endowed with the following real inner product (see [83, 107]; compare [61]):

$$\langle A, B \rangle = \operatorname{Re} \operatorname{trace}(B^H A).$$

Thus in both cases the inner product is given by $\langle A, B \rangle = \operatorname{Re} \operatorname{trace}(B^H A)$. Therefore we can use results from convex analysis in real vector spaces also to the space of complex matrices over \mathbb{R} , for example subdifferentials (subgradients) of $\|A\|$ for unitarily invariant norms. For this purpose we now recall properties of the singular value decomposition and unitarily invariant norms of a matrix.

Let the vector $\sigma(A)$ of ordered singular values of A have the form

$$\sigma(A) = [\mu_1, \dots, \mu_1, \mu_2, \dots, \mu_2, \dots, \mu_\ell, \dots, \mu_\ell]^T, \tag{6.11}$$

where $\mu_1 > \mu_2 > \dots > \mu_\ell \geq 0$. The multiplicity of μ_j is denoted by s_j . We divide the unitary matrices U and V , from the singular value decomposition (1.2) of A , into submatrices according to the form (6.11) of $\sigma(A)$

$$U = [U_1, \dots, U_\ell], \quad V = [V_1, \dots, V_\ell], \quad U_j \in \mathbb{C}^{m \times s_j}, \quad V_j \in \mathbb{C}^{n \times s_j}.$$

Then

$$A = \mu_1 U_1 V_1^H + \dots + \mu_\ell U_\ell V_\ell^H. \tag{6.12}$$

The matrices U and V from the SVD of A are not unique in general case, but the matrices $U_j V_j^H$ in (6.12) corresponding to the nonzero μ_j are determined uniquely (see [9]). The diagonal matrix of ordered singular values Σ can be written as $\Sigma = W \Sigma T^H$ with $W = \operatorname{diag}(W_1, \dots, W_{\ell-1}, W_0)$, $T = \operatorname{diag}(W_1, \dots, W_{\ell-1}, T_0)$, where the blocks $W_0, T_0, W_1, \dots, W_{\ell-1}$ are arbitrary unitary matrices of appropriate orders according to (6.11), and $W_0 = T_0$, if $\mu_\ell \neq 0$. This implies the relation between unitary matrices U and V from (1.2) and unitary matrices X and Y from any singular decomposition $A = X \Sigma Y^H$ (see [9]):

$$X = UW = U \operatorname{diag}(W_1, \dots, W_{\ell-1}, W_0), \quad Y = VT = V \operatorname{diag}(W_1, \dots, W_{\ell-1}, T_0). \tag{6.13}$$

For example, if $m = n$ and A is nonsingular, then $UV^H = XY^H$.

We recall some useful properties of ordered decreasingly singular values and unitarily invariant norms. Namely, singular values are continuous functions of elements of a matrix, i.e. if $\lim_{i \rightarrow \infty} X_i = X_*$, then

$$\lim_{i \rightarrow \infty} \sigma_k(X_i) = \sigma_k(X_*)$$

and for every j we have (see [44, p. 178])

$$|\sigma_j(G + H) - \sigma_j(G)| \leq \sigma_1(H). \tag{6.14}$$

Moreover, if a matrix G is divided into submatrices

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix},$$

then for arbitrary unitarily invariant norm $\|\cdot\|$ we have (see [90, p. 88])

$$\|\operatorname{diag}(G_{11}, G_{22})\| \leq \|G\|. \tag{6.15}$$

Unitarily invariant norms of matrices can be determined by symmetric gauge functions (see [7, Chapter 3], [10, Chapter IV], [44, Chapter 3], [90, Chapter II]). Let $t = \min\{m, n\}$. A map

$$\Phi : \mathbb{R}^t \rightarrow \mathbb{R}_+$$

is called a *symmetric gauge function*, if Φ is a norm and its value does not depend on the order and signs of the components of a vector (see, for example, [10, p. 86] and [44, p. 209]; for the definition of a gauge function defined on \mathbb{C}^t see, for example, [7, p. 138]). Thus, a symmetric gauge function on \mathbb{R}^t is an absolute permutation-invariant norm, which implies that it is a monotone norm. It has been shown by von Neumann [72] that if $\|\cdot\|$ is a given unitarily invariant norm on $\mathbb{C}^{m \times n}$, then there is a symmetric gauge function Φ such that $\|A\| = \Phi(\sigma(A))$ for all $A \in \mathbb{C}^{m \times n}$ (see, for example, [44], [90, p. 78]). Conversely, if $\Phi(\cdot)$ is a given symmetric gauge function on \mathbb{R}^t , then $\|A\| = \Phi(\sigma(A))$ is unitarily invariant on $\mathbb{C}^{m \times n}$. We will denote this norm by $\|\cdot\|_\Phi$.

The l_p norm in \mathbb{R}^t , $1 \leq p \leq \infty$, is a symmetric gauge function and the corresponding unitarily invariant norm of a matrix is the c_p norm of Schatten (see (1.3)). For $p = 1$ we obtain the trace (nuclear) norm. For $p = \infty$ it is the spectral norm. The norm $\|\cdot\|_\Phi$ is strictly convex if and only if Φ is strictly convex (see [102]). Therefore the c_p norm of Schatten is strictly convex if $1 < p < \infty$.

If Φ is a symmetric gauge function, then its dual (polar) Φ^*

$$\Phi^*(x) = \sup_{\Phi(y) \leq 1} y^T x$$

is also a symmetric gauge function (see [10, p. 90]). Duality in unitarily invariant norms is defined via the inner product $\langle A, B \rangle = \text{trace}(A^H B)$ (see [10, p. 92 and 96]). The dual norm to the unitarily invariant norm $\|\cdot\|_\Phi$ is given by

$$\|A\|_\Phi^* = \sup_{\|X\|_\Phi \leq 1} |\langle A, X \rangle| = \sup_{\|X\|_\Phi \leq 1} |\text{trace}(X^H A)| = \|A\|_\Phi^*$$

and it is also unitarily invariant. It is known that the c_q norm is dual to the c_p norm if p and q satisfy (6.5). Thus the trace norm (c_1 norm) is dual to the spectral norm (c_∞ norm).

It turns out that $\|A\|_\Phi$ can be characterized in terms of the dual norm (see [90, Lemma 3.5])

$$\|A\|_\Phi = \max_{\|X\|_{\Phi^*} \leq 1} \text{Re trace}(X^H A).$$

Since $(\Phi^*)^* = \Phi$, we consequently obtain

$$\|A\|_{\Phi^*} = \max_{\|X\|_\Phi \leq 1} |\text{trace}(X^H A)| = \max_{\|X\|_\Phi \leq 1} \text{Re trace}(X^H A). \tag{6.16}$$

A matrix X_* for which the second maximum in (6.16) is reached is called a $\|\cdot\|_\Phi$ -dual matrix to the matrix A . Hence we have for $A \neq 0$

$$\text{Re trace}(X_*^H A) = \|A\|_{\Phi^*}, \quad \|X_*\|_\Phi = 1. \tag{6.17}$$

Then we say that X_* is dual to A with respect to the norm $\|\cdot\|_\Phi$ and we denote the dual matrix X_* by A^* . We have introduced in [102] the notion “dual matrix” because this concept has been used in vector spaces (see, for example, [87, 88, 109]).

Let X_* be a $\|\cdot\|_\Phi$ -dual matrix to A and $\gamma = \text{trace}(X_*^H A)$. Then from (6.16) we obtain $\text{Re } \gamma \geq |\gamma|$, which implies that γ is a real positive number. Thus in (6.17) we have

$$\text{trace}(X_*^H A) = \|A\|_{\Phi^*}, \quad \|X_*\|_\Phi = 1 \tag{6.18}$$

for every $\|\cdot\|_\Phi$ -dual matrix $X_* = A^*$. This property is confirmed by Theorem 6.1, formulated below, because of (6.13) and (6.21). Thus instead of (6.17) we apply the relations (6.18) for dual matrices.

The set of all dual matrices for the unitarily invariant norm $\|\cdot\|_\Phi$ is described in [105] (see also [95] for real case; compare [83]).

THEOREM 6.1. *Let $A \in \mathbb{C}^{m \times n}$ have the SVD $A = U\Sigma V^H$. Then the set of all $\|\cdot\|_\Phi$ -dual matrices to A for an arbitrary unitarily invariant norm $\|\cdot\|_\Phi$ is equal to*

$$\mathbb{G}(A, \Phi) = \{X \in \mathbb{C}^{m \times n} : X = UDV^H, D \text{ is } \|\cdot\|_\Phi\text{-dual to } \Sigma\}, \tag{6.19}$$

where the set of all matrices $D \|\cdot\|_\Phi$ -dual to Σ is equal to

$$\mathbb{G}(\Sigma, \Phi) = \{D = W \text{diag}(\tau_j)T^H : \Sigma = W\Sigma T^H \text{ is a SVD of } \Sigma\} \tag{6.20}$$

and the vector $\tau = [\tau_1, \dots, \tau_t]^T$ is Φ -dual to the vector $\sigma = \sigma(A)$ of singular values of A , i.e. $\langle \sigma, \tau \rangle = \tau^T \sigma = \Phi^*(\sigma)$, $\Phi(\tau) = 1$.

In (6.20) we take all SVD of Σ . Therefore we have in (6.19)

$$\mathbb{G}(A, \Phi) = \{X \in \mathbb{C}^{m \times n} : X = U \text{diag}(d_j)V^H, A = U\Sigma V^H \text{ is any SVD of } A\}, \tag{6.21}$$

where $d = [d_1, \dots, d_t]^T$ is a Φ -dual vector to $\sigma(A)$. A dual matrix to A with respect to the unitarily invariant norm $\|\cdot\|$ is unique if the norm is strictly convex.

Using the characterization of dual matrices, presented in Theorem 6.1, we can specialize it for particular norms (see [102]). The matrix A_∞^* is a c_∞ -dual matrix to A , $\text{rank}(A) = r$, if and only if (see [102, Theorem 4.2] for A real)

$$A_\infty^* = USV^H, \quad \|S\|_\infty = 1, \tag{6.22}$$

where

$$S = \begin{bmatrix} I_r & 0 \\ 0 & Z \end{bmatrix}, \quad \text{if } r < t, \quad S = \begin{bmatrix} I_r \\ 0 \end{bmatrix} \quad \text{if } r = n; \quad S = [I_r \quad 0] \quad \text{if } r = m.$$

The singular values of S form a vector that is l_∞ -dual to $\sigma(A)$ (see (6.7)). The c_1 -dual matrices to A have the form (see [102, Theorem 4.3] for A real)

$$A_1^* = U \begin{bmatrix} BD^*B^H & 0 \\ 0 & 0 \end{bmatrix} V^H, \tag{6.23}$$

where B is unitary of order s , s is the number of singular values of A equal to $\sigma_1(A)$, D^* is diagonal with elements of the l_1 -dual vector to $\sigma(A)$ on the main diagonal (see (6.8)). Thus $D^* = \text{diag}(\mu_j)$ for some $\mu_j \geq 0$, $\sum_j \mu_j = 1$. Of course, for $1 < p < \infty$ the c_p -dual matrix A_p^* is unique and it has the form

$$A_p^* = U \text{diag}(\nu_j)V^H, \quad 1 < p < \infty,$$

where the elements ν_1, \dots, ν_t form the l_p -dual vector to $\sigma(A)$. If $p \rightarrow \infty$, then the c_p -dual matrix A_p^* tends to a specific c_∞ -dual matrix (6.22) with $Z = 0$

$$\lim_{p \rightarrow \infty} A_p^* = U \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} V^H. \tag{6.24}$$

From (6.3) and (6.18) it follows that the set $\mathbb{G}(A, \Phi^*)$ of all $\|\cdot\|_{\Phi^*}$ -dual matrices to A is the subdifferential of $\|A\|_{\Phi}$ (see [107]; compare [61, 83, 95]). Thus the subdifferential of $\|A\|$ is equal to

$$\partial\|A\|_{\Phi} = \{Y : \text{trace}(Y^H A) = \|A\|_{\Phi}, \|Y\|_{\Phi^*} = 1\}.$$

The dual matrices (subdifferential of $\|A\|$) can be used to characterize approximants of a given matrix by matrices from a linear subspace endowed with the unitarily invariant norm $\|A\|_{\Phi}$. We apply it in the next section.

The facial structure of the unitarily invariant norms and extreme points of the unit ball were investigated in [83, 86, 102, 105] (for convex analysis in spaces of matrices see also [61]–[63]). We now recall a characterization of extreme points of unit sphere of matrices for a unitarily invariant norm (see [102]). Let $\mathbb{B}_{\Phi} = \{F \in \mathbb{C}^{m \times n} : \|F\|_{\Phi} = 1\}$. The matrix $F \in \mathbb{B}_{\Phi}$ is the extreme point of \mathbb{B}_{Φ} if and only if the relation

$$F = \frac{1}{2}(F_1 + F_2), \quad F_1, F_2 \in \mathbb{B}_{\Phi},$$

implies $F_1 = F_2 = F$. The matrix F is an extreme point of \mathbb{B}_{Φ} if and only if the vector $\sigma(F)$ is an extreme point of the unit sphere of the linear space \mathbb{R}^t , $t = \min\{m, n\}$, endowed with the norm Φ . In particular, the extreme points of the unit sphere \mathbb{B}_1 for the nuclear (trace) norm are precisely the rank-one matrices of the form $F = uv^H$, $u^H u = v^H v = 1$. The extreme points of the unit sphere \mathbb{B}_{∞} for the spectral norm are matrices whose all singular values are equal to 1.

7. Approximation of a matrix by matrices from a linear subspace. Now we focus on a characterization of the approximants B to A in (1.1) for a unitarily invariant norm. For this purpose we can apply a general theorem of Singer [85, p. 56], based on extreme points of unit ball of conjugate space (see, for example, [58, 64] and the next section) or an approach of Watson [94, pp. 15–18]), based on subdifferentials of norm (dual matrices) (see [104, 107]). We apply the latter approach. The approximants $B \in \mathbb{M}$ to A , i.e. solutions of the problem (1.1), are characterized by the following theorem for linear subspaces (see [107, Corollary 1], compare [104]).

THEOREM 7.1. *Let \mathbb{M} be a linear subspace of $\mathbb{C}^{m \times n}$ endowed with a unitarily invariant norm $\|\cdot\|$. Then $B \in \mathbb{M}$ is a solution of*

$$\min_{X \in \mathbb{M}} \|A - X\| \tag{7.1}$$

if and only if there exists a matrix F such that

$$\text{trace}(F^H(A - B)) = \|A - B\|, \quad \|F\|^* = 1, \quad F \in \mathbb{M}^{\perp}, \tag{7.2}$$

where

$$\mathbb{M}^{\perp} = \{F \in \mathbb{C}^{m \times n} : \text{trace}(F^H X) = 0 \text{ for all } X \in \mathbb{M}\}$$

and $\|\cdot\|^$ is a dual norm to the norm $\|\cdot\|$.*

Corollary 1 in [107] is formulated only for square matrices, however its extension to rectangular matrices is straightforward and obvious.

Of course, $\text{trace}(F^H(A - B)) = \text{trace}(F^H A)$ and $\text{trace}(F^H X) = 0$ for all $X \in \mathbb{M}$. Therefore we can choose a common matrix F for all solutions of (7.1). The matrix $Y = (A - B)/\|A - B\|$ satisfies $\text{trace}(F^H Y) = 1 = \|F\|^*$, $\|Y\| = 1$. Thus Y is a $\|\cdot\|$ -dual to F and $Y \in \partial\|F\|^*$. It implies the following corollary (see [107, Corollary 2]).

COROLLARY 7.2. *There exists $F \in \mathbb{M}^\perp$, $\|F\|^* = 1$, such that every solution B of (7.1) has the form*

$$B = A - \|A - B\|Y \tag{7.3}$$

for appropriate $Y \in \partial\|F\|^*$ such that the matrix B in (7.3) belongs to \mathbb{M} . If the subdifferential of $\|F\|^*$ is unique, then the problem (7.1) has a unique solution.

In Corollary 7.2 the approximants B to a complex A are expressed by means of the approximation error $\|A - B\|$ and the subdifferential of $\|F\|^*$. Therefore it can be expressed by the SVD of F (compare [104, Theorem 4.2], [107, Corollary 2]). However, for our future purpose it is better to use the SVD of $R = A - B$ (compare [104, Theorem 4.3]).

The properties of approximants with respect to the Ky Fan k norms and the c_p norms are presented for real matrices in [96, 98]. The case of the spectral norm and real matrices is studied in [104] (see [107] for complex matrices). Since our main goal is proving Conjecture 1.1 for complex matrices, the next theorem is formulated for complex matrices instead of real matrices as it was done in [104, Theorem 4.3]. However, we omit the proof of this extension (compare the proof of Theorem 10.2 in Section 10 which is a further generalization).

THEOREM 7.3. *Let $\|\cdot\|_\infty$ be the spectral norm and let an approximant \hat{B} to A by elements of a linear subspace \mathbb{M} of $\mathbb{C}^{m \times n}$ be such that $\hat{R} = A - \hat{B}$ has the smallest number s of singular values equal to $\delta = \|A - \hat{B}\|_\infty$ over all approximants. Let \hat{R} have the SVD $\hat{R} = \hat{U}\hat{\Sigma}\hat{V}^H$. Then for every approximant B to A there exists a Z such that $\|Z\|_\infty \leq \delta$ and*

$$B = A - \delta\hat{U}_1\hat{V}_1 - \hat{U}_2Z\hat{V}_2^H, \tag{7.4}$$

where $\hat{U} = [\hat{U}_1, \hat{U}_2]$, $\hat{V} = [\hat{V}_1, \hat{V}_2]$, $\hat{U}_1 \in \mathbb{C}^{m \times s}$, $\hat{V}_1 \in \mathbb{C}^{n \times s}$. Moreover, the matrix $\hat{U}_1\hat{V}_1^H$ is determined uniquely, i.e. it does not depend on the choice of \hat{U}, \hat{V} in the SVD of \hat{R} . If $s = \min\{m, n\}$ then the approximant B to A is unique and it has the form

$$B = A - \delta\hat{U}\hat{V}_1^H \quad \text{or} \quad B = A - \delta\hat{U}_1\hat{V}^H,$$

if $s = m$ or $s = n$, respectively.

The following matrix approximation problems in the spectral norm, inspired by some problem introduced by Greenbaum and Trefethen in [36] and connected with the iterative GMRES method for solving system of linear equations, are the goal of [66]:

$$\min_{u \in \mathbb{P}_m} \|A^{m+1} - u(A)\|_\infty, \tag{7.5}$$

$$\min_{u \in \mathbb{P}_m} \|I - Au(A)\|_\infty, \tag{7.6}$$

$$\min_{u \in \mathbb{P}_m} \|v(A) - u(A)\|_\infty, \tag{7.7}$$

where \mathbb{P}_m denotes the set of polynomials of degree at most m and $v(A)$ is a polynomial of the matrix A , $m < \deg v$. In [66] one investigates under which assumptions these problems have unique solutions. Separately the case of A being a Jordan block is considered. In [66] some results from [107] are applied to obtain characterizations of solutions of the problems (7.5)–(7.7) and to prove their uniqueness.

8. Orthogonality of matrices. In this section we present particular cases of the problem (7.1). In the first case, suggested to Friedland by Halmos, one considers the following problem (see [34])

$$\delta_{A,C} \equiv \min_{\alpha \in \mathbb{R}} \|A - \alpha C\|_\infty,$$

where $A, C \in \mathbb{C}^{m \times n}$, but α is a real number. Friedland [34] has given a characterization of real numbers a, b such that for $\beta \in [a, b]$ we have $\|A - \beta C\|_\infty = \delta_{A,C}$. These numbers a, b are solutions of certain polynomial equations. Friedland has mentioned that the special case where A is a real $n \times n$ matrix and $C = I$ is interesting in the theory of partial differential equations [82].

The next approximation problem is related to the notion of orthogonality of matrices in the Birkhoff and James sense [14, 49]. Let A and C be complex $m \times n$ matrices. We say that A is orthogonal to C in the Birkhoff and James sense, if

$$\|A + \mu C\| \geq \|A\| \quad \text{for every complex scalar } \mu. \tag{8.1}$$

Orthogonality is not a symmetric relation. The condition (8.1) can be interpreted as follows. Let A be not in the linear subspace \mathbb{M} spanned by the matrix C . Then the zero matrix is the best approximation to A among all matrices in \mathbb{M} .

For the c_p norms, $1 \leq p < \infty$, square complex matrices and $A = I$, the problem (8.1) has been studied by Kittaneh [52]. The main result of his shows that $\text{trace}(C) = 0$ if and only if

$$\|I + \mu C\|_p \geq \|I\|_p \quad \text{for } \mu \in \mathbb{C}. \tag{8.2}$$

This characterization of zero trace matrices does not hold for $p = \infty$. Ziętak [107] extends the result of Kittaneh to a more general class of unitarily invariant norms. Namely, she has proved that if a norm $\|\cdot\|$ is unitarily invariant and such that the subgradient of $\|I\|$ is unique, then $\text{trace}(C) = 0$ if and only if (see [107, Theorem 5])

$$\min_{z \in \mathbb{C}} \|I + zC\| = \|I\|.$$

It is easy to verify that the subgradient of $\|I\|_p$ is unique for $1 \leq p < \infty$ and it is not unique for $p = \infty$. For an arbitrary unitarily invariant norm it is shown that if $\text{trace}(C) = 0$, then (see [107, Theorem 4])

$$\|I + \mu C\| \geq \|I\|, \quad \mu \in \mathbb{C}.$$

Properties of orthogonality of complex square matrices A and C were investigated also by Bhatia and Šemrl [12]. They characterize matrices A and C which are orthogonal with respect to the c_p norms for $1 \leq p \leq \infty$. We now recall their characterizations (see [12, Theorems 1.1 and 2.1]).

THEOREM 8.1. *Let A and C be square matrices.*

- (i) *A matrix A is orthogonal to C with the spectral norm if and only if there exists a unit vector x such that $\|Ax\|_2 = \|A\|_\infty$ and $\langle Ax, Cx \rangle = 0$.*
- (ii) *Let A have the polar decomposition $A = U|A|$ (U unitary, $|A|$ Hermitian nonnegative definite). If for any $1 \leq p < \infty$ we have*

$$\text{trace}(|A|^{p-1}U^H C) = 0,$$

then A is orthogonal to C in the c_p norm of Schatten. The converse is true for all A , if $1 < p < \infty$, and for all invertible A , if $p = 1$.

Li and Schneider [64] have continued these investigations of Bhatia and Šemrl giving characterizations which can be derived from the following specialized version of the general theorem [85, p. 170] of Singer (see [64, Proposition 2.1])

PROPOSITION 8.2. *Let $\|\cdot\|$ be a norm on $\mathbb{F}^{m \times n}$, where \mathbb{F} is \mathbb{R} or \mathbb{C} , and let $A, C \in \mathbb{F}^{m \times n}$ be such that A is not a multiple of C . Then*

$$\|A + \mu C\| \geq \|A\| \quad \text{for all } \mu \in \mathbb{F}$$

if and only if there exist h extreme points $F_1, \dots, F_h \in \mathbb{F}^{m \times n}$ of the unit ball $\{Y \in \mathbb{F}^{m \times n} : \|\cdot\|^ \leq 1\}$ in the dual (conjugate) space endowed with the dual norm $\|\cdot\|^*$; with $h \leq 3$ in the complex case and $h \leq 2$ in the real case, and positive numbers $\alpha_1, \dots, \alpha_h$, $\alpha_1 + \dots + \alpha_h = 1$ such that*

$$\sum_{j=1}^h \alpha_j \langle F_j, C \rangle = 0, \quad \langle F_j, A \rangle = \|A\|, \quad j = 1, \dots, h.$$

Characterizations of Li and Schneider presented for the rectangular matrices $A, C \in \mathbb{C}^{m \times n}$, $m \leq n$, and the c_p norms, $1 < p \leq \infty$, are similar as the characterizations of Bhatia and Šemrl presented for square matrices in Theorem 8.1. Therefore we omit them (see [64, Theorems 3.1(b) and 3.2]). For the spectral norm Li and Schneider present also a characterization by means of numerical ranges of the matrices $U^H C V$ and $U^H C A^H U$, where columns of U form a basis for the eigenspace of $A A^H$ corresponding to the largest eigenvalue, and $V = A^H U / \sigma_1(A)$ (see [64, Theorem 3.1(c)]): A is orthogonal to C with respect to the spectral norm if and only if zero belongs to the numerical range of $U^H C V$ or $U^H C A^H U$. The characterization of Li and Schneider for trace norm ($p = 1$) is different from the above one of Bhatia and Šemrl. Therefore we recall it (see [64, Theorem 3.3]).

THEOREM 8.3. *Let \mathbb{F} be \mathbb{R} or \mathbb{C} and let $\|\cdot\|_1$ be the trace norm. Let $A, C \in \mathbb{F}^{m \times n}$, $m \leq n$. The following conditions are equivalent:*

- (a) $\|A + \mu C\|_1 \geq \|A\|_1$ for all $\mu \in \mathbb{F}$.
- (b) There exists $F \in \mathbb{F}^{m \times n}$ such that $\|F\|_\infty \leq 1$, and

$$\text{trace}(A F^H) = \|A\|_1, \quad \text{trace}(C F^H) = 0.$$

- (c) For any SVD of A ; $A = U \Sigma V^H$ we have

$$U^H C V = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

where $|\text{trace}(C_{11})| \leq \|C_{22}\|_1$, C_{11} is $k \times k$ with $k = \text{rank}(A)$, and by convention $\|C_{22}\|_1 = 0$ if $m = k$.

The characterization given in Theorem 8.3(b) can be derived also from Theorem 7.1 and the formula for the c_∞ -dual matrices.

Bhatia and Šemrl [12] have formulated the following conjecture. Let ν be any norm on the vector space \mathbb{C}^n and let $\|\cdot\|$ be the matrix norm induced by ν on the space of $n \times n$ complex matrices. Then (8.1) holds if and only if there exists a unit vector x such that $\nu(Ax) = \|A\|$ and $\nu(Ax + \mu Cx) \geq \nu(Ax)$ for all μ . Li and Schneider [64] have given a counterexample to this conjecture for ν being the l_p norm with $p \neq 2$. Benitez and his coauthors [8] have extended this counterexample to finite-dimensional normed spaces whose matrix norm is induced by a vector norm not induced by an inner product. In fact they characterize inner product spaces by means of the notion of orthogonality. Let \mathbb{V} be a real finite-dimensional normed space with unit sphere $\mathbb{B}_\mathbb{V}$ and let $\mathcal{L}(\mathbb{V})$ be the space of linear operators from \mathbb{V} into itself. In [8] it is proved that \mathbb{V} is an inner product space if and only if for $A, C \in \mathcal{L}(\mathbb{V})$ A is orthogonal to C if and only if there exists $x \in \mathbb{B}_\mathbb{V} : \|A\| = \|Ax\|, Ax \perp Cx$, where the concept of Birkhoff’s orthogonality for vectors is defined in the same way as for matrices (see [12, 14, 49]).

Recently Bhattacharyya and Grover [13] have investigated the Birkhoff–James orthogonality when the Banach space has structures. In particular, they consider Hilbert C^* -modules (see also [4]). In [5] one presents some characterizations of the Birkhoff–James orthogonality of a pair of bounded operators and one considers a center of mass of a pair of operators.

Grover [38] obtains a necessary and sufficient condition for a square complex matrix A to be orthogonal in the Birkhoff–James sense to any subspace of square complex matrices. We now present a new proof of Theorem 1 of Grover, extended to the case of rectangular matrices, using Theorem 7.1. The original formulation of [38, Theorem 1] covers also the case of real subspace of complex matrices. We do not consider this case here.

THEOREM 8.4. *Let $A \in \mathbb{C}^{m \times n}$ and let s be the multiplicity of the largest singular value $\sigma_1(A) = \|A\|_\infty$ of A . Let \mathbb{W} be any linear subspace of $\mathbb{C}^{m \times n}$. Then A is orthogonal to \mathbb{W} if and only if there exists a density matrix P of rank at most s such that*

$$A^*AP = \|A\|_\infty^2 P, \quad AP \in \mathbb{W}^\perp, \quad \|P\|_1 = 1,$$

where

$$\mathbb{W}^\perp = \{F \in \mathbb{C}^{m \times n} : \text{trace}(F^H X) = 0 \text{ for all } X \in \mathbb{W}\}.$$

REMARK. Positive semidefinite matrix with trace equal to 1 is called density matrix.

Proof. The assumption that A is orthogonal to \mathbb{W} means that the zero matrix is a best approximation to A by matrices from \mathbb{W} . Let $A = U\Sigma V^H$ be the SVD of A . The zero matrix is the best approximation to A in the spectral norm if and only if there exists a matrix F such that (see Theorem 7.1)

$$\text{trace}(F^H A) = \|A\|_\infty, \quad \|F\|_1 = 1, \quad F \in \mathbb{M}^\perp.$$

This implies that F is a $\|\cdot\|_1$ -dual matrix to A . Therefore F has to have the form (see (6.23))

$$F = U \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} V^H,$$

where S positive semi-definite matrix of the order s and $\|S\|_1 = 1$. Let us define

$$P = V \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}^H.$$

The matrix P is positive semi-definite of rank $\leq s$, $\|P\|_1 = 1$, hence P is a density matrix and it satisfies

$$A^*AP = \|A\|_\infty P, \quad AP = \|A\|_\infty F \in \mathbb{W}^\perp,$$

because the nonzero eigenvalues of S corresponds to the singular values of A equal to $\|A\|_\infty$. This completes the proof. ■

The remaining properties of P , described in [38, Theorem 1], also hold. We omit them.

The concept of orthogonality of matrices can be used for investigations of numerical ranges. The numerical range of rectangular A with respect to B can be expressed in terms of the Birkhoff–James orthogonality (see [21, Theorem 1]). Chmieliński defines and investigates the ε -Birkhoff orthogonality in [19], and considers linear mappings preserving the Birkhoff–James orthogonality in [20].

9. Examples of strict spectral approximants. The notion of c_p -minimal approximant introduced by Rogers and Ward [80] for the approximation by positive operators can be extended to an approximation by elements from a convex subset of $\mathbb{C}^{m \times n}$. What is more, the c_p -minimal approximant is in fact the strict spectral approximant. It follows from the next theorem, proved in [108].

THEOREM 9.1. *Let \mathbb{M} be a closed convex subset of $\mathbb{C}^{m \times n}$. A matrix $\widehat{B} \in \mathbb{M}$ is the strict spectral approximant to A if and only if for every $B \in \mathbb{M}$, $B \neq \widehat{B}$, we have*

$$\|A - B\|_p > \|A - \widehat{B}\|_p \quad \text{for all } p \text{ sufficiently large.}$$

From Theorem 9.1 we obtain the following corollary (see [108, Corollary 2]).

COROLLARY 9.2. *If \widehat{B} is a common c_p -approximant to A for every finite p , then \widehat{B} is the strict spectral approximant. Moreover, if all singular values of $A - \widehat{B}$ are equal to $\delta(A) = \|A - \widehat{B}\|_\infty$, then the spectral approximant is unique.*

We now recall examples of matrices which are the strict spectral approximants because they satisfy the assumptions of Corollary 9.2 (see [108, 110]).

Let \mathbb{Z} be the set of all zero-trace complex matrices of order n . The set \mathbb{Z} is the linear space of dimension $n^2 - 1$. It is known that the matrix (see [107, Theorem 2])

$$B = A - \frac{\text{tr}(A)}{n} I \tag{9.1}$$

is the approximant to A by zero-trace matrices with respect to every unitarily invariant norm. Since each singular value of $A - B$ is equal to $|\text{tr}(A)|/n$, the matrix B is the unique spectral approximant to A , so it is the strict spectral approximant.

If a matrix $C \in \mathbb{C}^{n \times n}$ is zero-trace, then (see [107, Theorem 4])

$$\min_{z \in \mathbb{C}} \|I + zC\|_p = \|I\|_p, \quad 1 \leq p \leq \infty.$$

Therefore the zero matrix is the c_p -approximant (for every p) to the identity matrix by matrices from the linear subspace spanned by C . Thus the strict spectral approximant

$B^{(st)}$ is the zero matrix. This implies the uniqueness of the spectral approximant because all singular values of the residual matrix $I - B^{(st)} = I$ are the same.

In [108] we apply the concept of the strict spectral approximation to certain approximation problems. If $A = B + \imath C$ is normal, B, C Hermitian, then in each unitarily invariant norm the positive part $B^{(+)}$ of B

$$B^{(+)} = \frac{1}{2}((B^H B)^{1/2} + B)$$

is a positive approximant to A (see [11]). Corollary 9.2 implies that $B^{(+)}$ is the strict spectral positive approximant, which confirms some properties of positive approximants (see [108, Section 2]). Inspired by [65], we also consider approximation problems with singular value preserving functions in [108]. Moreover, in [108, Section 4] we have shown that the Moore–Penrose generalized inverse A^\dagger can be interpreted as the strict spectral approximant to the zero matrix by generalized inverses of $A \in \mathbb{C}^{m \times n}$, i.e. by matrices from the set $\{X \in \mathbb{C}^{n \times m} : AXA = A\}$.

We have proved that $A \in \mathbb{C}^{n \times n}$ has a unique spectral approximation by Hermitian matrices if and only if all singular values of its skew-Hermitian part $(A - A^H)/2$ are equal. Thus $\widehat{B} = (A + A^H)/2$ is the unique spectral approximant to A and it is the strict spectral approximant (see [110]). If A has not a unique spectral approximant, then \widehat{B} is also the strict spectral approximation, because \widehat{B} is the best approximant to A by Hermitian matrices for every unitarily invariant norm, so also for the c_p norms.

Future examples, in which the solutions of the problem (1.1) with the c_p norms do not depend on p , are given in [59].

In [104] we have considered a special case of the problem (1.1) for the spectral norm and \mathbb{M} being the linear real subspace of $\mathbb{R}^{m \times n}$ of dimension

$$\dim(\mathbb{M}) = mn - 1.$$

Let a matrix E with $\text{rank}(E) = r$ span the orthogonal complement \mathbb{M}^\perp of \mathbb{M} . If $m = n$ and $E = I$, then it is the approximation of a real matrix by zero-trace real matrices, considered earlier (see (9.1)). For the spectral norm every c_∞ -dual matrix E_∞^* to E determines a spectral approximant to A (see [104, (4.32)])

$$B = A - \frac{\text{tr}(E^T A)}{\|E\|_1} E_\infty^*, \tag{9.2}$$

because the conditions (7.2) are satisfied by $F = E/\|E\|^*$, hence

$$\text{trace}(F^T(A - B)) = \text{trace}(F^T A).$$

We recall that the trace norm $\|\cdot\|_1$ is the dual norm to the spectral norm. If E is of full rank, then A has the unique spectral approximant. Let E of rank r have the SVD

$$E = U_e \Sigma_e V_e^T.$$

Then (see [104, (4.33)])

$$B^{(st)} = A - \frac{\text{tr}(E^T A)}{\|E\|_1} U_e \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} V_e^T$$

is the strict spectral approximant to A . If we approximate A with respect to the c_p norm, then the best approximant has the following form (compare (9.2))

$$B_p = A - \frac{\text{tr}(E^T A)}{\|E\|_q} E_p^*,$$

where $1/p + 1/q = 1$ and E_p^* is the c_p -dual matrix to E . It is obvious that

$$\lim_{p \rightarrow \infty} \|E\|_p = \|E\|_\infty, \quad \lim_{p \rightarrow \infty} \|E\|_q = \|E\|_1.$$

The c_p -dual matrix E_p^* tends to the following specific c_∞ -dual matrix E_∞^* as $p \rightarrow \infty$ (see (6.9), (6.22), (6.24)):

$$U_e \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} V_e^T,$$

because we can choose $Z = 0$ in (6.22) and we use (6.10). From (6.7) and (6.9) we deduce that

$$\lim_{p \rightarrow \infty} B_p = B^{(\text{st})},$$

which confirms Conjecture 1.1 in this simple case.

At the end of this section we consider the following trivial cases for which Conjecture 1.1 is also true. Let the linear subspace \mathbb{M} of $\mathbb{R}^{n \times n}$ be spanned by symmetric commuting matrices C_1, \dots, C_k and let real symmetric A commute with every C_j . Then there exists an orthogonal matrix Q such that

$$A = Q \text{diag}(\lambda_\ell) Q^T, \quad C_j = Q \text{diag}(\lambda_\ell^{(j)}) Q^T \quad \text{for } j = 1, \dots, k. \tag{9.3}$$

The eigenvalues of A and C_j are real and their moduli are singular values of A and C_j , respectively. Let

$$b = [\lambda_1(A), \dots, \lambda_n(A)]^T \tag{9.4}$$

and let

$$X = \sum_{j=1}^k x_j C_j.$$

Then

$$A - X = Q \text{diag} \left(\lambda_\ell(A) - \sum_{j=1}^k x_j \lambda_\ell^{(j)} \right) Q^T.$$

We construct a matrix C whose columns are formed by the eigenvalues of C_1, \dots, C_k , respectively, in an order according to the spectral decompositions (9.3)

$$C = \begin{bmatrix} \lambda_1^{(1)} & \dots & \lambda_1^{(k)} \\ \vdots & \ddots & \vdots \\ \lambda_n^{(1)} & \dots & \lambda_n^{(k)} \end{bmatrix}. \tag{9.5}$$

Then for $1 \leq p \leq \infty$ we have

$$\|A - X\|_p = \|b - Cx\|_p,$$

where $x = [x_1, \dots, x_k]^T$. On the left-hand side we have the c_p norm, on the right-hand side there is the l_p norm. Let for $1 < p < \infty$ the vector $x^{(p)} = [x_1^{(p)}, \dots, x_k^{(p)}]^T$ solve in the l_p norm the overdetermined linear system $Cx = b$. Then the limit $\lim_{p \rightarrow \infty} Cx^{(p)}$ is equal

to $Cx^{(st)}$, where $x^{(st)} = [x_1^{(st)}, \dots, x_n^{(st)}]^T$ is the strict Chebyshev solution of $Cx = b$. It follows from the result of Descloux [24]. Then we have

$$\lim_{p \rightarrow \infty} Q \operatorname{diag}(x_j^{(p)})Q^T = Q \operatorname{diag}(x_j^{(st)})Q^T.$$

It is obvious that the matrix

$$X^{(p)} = Q \operatorname{diag}(x_j^{(p)})Q^T$$

is the c_p -approximant to A . From the definitions of the strict spectral approximant and the strict Chebyshev approximant we deduce that the matrix

$$X^{(st)} = Q \operatorname{diag}(x_j^{(st)})Q^T$$

is the strict spectral approximant to A and it is the limit of $X^{(p)}$ as $p \rightarrow \infty$. Thus Conjecture 1.1 is true in this case.

If in the above last example we replace the real pairwise commuting symmetric matrices A, C_1, \dots, C_k by pairwise commuting Hermitian matrices, then the matrix Q in (9.3) is unitary, but eigenvalues are still real and the singular values are moduli of eigenvalues. Therefore Conjecture 1.1 is also true for Hermitian commuting matrices. However, for the case of normal commuting matrices the situation is more complicated because eigenvalues can be complex.

Commuting normal matrices are simultaneously unitary diagonalizable (see [43, Theorem 2.5.5]). Thus the approximation of A by matrices from the linear subspace spanned by pairwise commuting normal matrices C_1, \dots, C_k which commute also with A leads to the overdetermined system of linear equations $Cx = b$ with b and C determined in (9.4) and (9.5), where in this case C and b can be complex. Fortunately, singular values of normal matrices are moduli of eigenvalues and the result of Descloux (3.1) is extended to the overdetermined systems of complex linear equations by Borowsky [15]. Namely, if the vector $x^{(p)}$ is the l_p -solution of the overdetermined system of complex linear equations $Cx = b$, then $\lim_{p \rightarrow \infty} Cx^{(p)} = Cx^{(st)}$ (see [15, Theorem 6]), where the vector $x^{(st)}$ is the strict Chebyshev solution, defined in the complex case analogously as in the real case. Borowsky [15] applies a similar formulation of the definition of the strict Chebyshev solution of the system of complex linear equations as it is in the paper of Descloux [24] for the real case, without the lexicographical ordering which is used in [30] and in [106].

Unfortunately, there are some doubts about correctness of the proof of Borowsky of the generalization of the result of Descloux to complex case in [15] (see [111]). The proof of Borowsky for the complex case is based upon the proof of Descloux [24] for real case (see comments about the proof of [24, Theorem] in Section 3). In [111] we analyze the proof of Borowsky and we propose its improvement. Therefore Conjecture 1.1 is true also for normal commuting matrices.

At the end of this section we would like to mention an interesting property of commuting normal matrices (see [67], for commuting symmetric matrices see [35]). Let A, C_1, \dots, C_k be pairwise commuting normal matrices. Then for the approximation problem (1.1) with the spectral norm and \mathbb{M} being a linear subspace spanned by C_1, \dots, C_k

$$\min_{\alpha_1, \dots, \alpha_k \in \mathbb{C}} \left\| A - \sum_{i=1}^k \alpha_i C_i \right\|_{\infty} = \max_{\substack{x \in \mathbb{C}^n \\ \|x\|_2=1}} \min_{\alpha_1, \dots, \alpha_k \in \mathbb{C}} \left\| Ax - \sum_{i=1}^k \alpha_i C_i x \right\|_2.$$

10. An attempt to prove Conjecture 1.1. Let \mathbb{M} be a linear subspace of $\mathbb{C}^{m \times n}$ and let us consider the problem (1.1) for the spectral norm. Let us denote the approximation error by δ_∞ . Then

$$\delta_\infty = \min_{X \in \mathbb{M}} \|A - X\|_\infty. \tag{10.1}$$

Let $B^{(st)} \in \mathbb{M}$ be the strict spectral approximant to A . This means that the vector $\sigma(A - B^{(st)})$ is minimal with respect to the ordinary lexicographic ordering in the set $\{\sigma(A - X) : X \in \mathbb{M}\}$ (see definition of the strict spectral approximant in Introduction). For example, the vector $[3, 2, 2, 0]^T$ is bigger than $[3, 2, 1, 1]^T$. From the definition of the strict spectral approximation it follows that the matrix $R^{(st)} = A - B^{(st)}$ has the smallest number of the singular values equal to the error of approximation. The existence and uniqueness of the strict spectral approximant is proved in [106] by means of the following sequence of auxiliary norms

$$\|R\|_{k,2} = (\sigma_1^2(R) + \dots + \sigma_k^2(R))^{1/2} \tag{10.2}$$

for $k = 1, \dots, t = \min\{m, n\}$. This is a particular case of the Ky Fan p - k norm for $p = 2$:

$$\|R\|_{k,p} = \left(\sum_{j=1}^k \sigma_j^p(R) \right)^{1/p} \tag{10.3}$$

(compare (1.3) and (2.2)). For $k = t$ and $p = 2$ we obtain the Frobenius norm. We define the nested sequence of minimizing sets \mathbb{S}_k , with

$$\mathbb{S}_0 = \{R = A - X : X \in \mathbb{M}\}, \tag{10.4}$$

and the set \mathbb{S}_k of the minimizers of $\|R\|_{k,2}$ over \mathbb{S}_{k-1} . Then

$$\mathbb{S}_t \subseteq \mathbb{S}_{t-1} \subseteq \dots \subseteq \mathbb{S}_1 \subseteq \mathbb{S}_0.$$

In particular, we have

$$\mathbb{S}_1 = \{R \in \mathbb{C}^{m \times n} : A - R \text{ is spectral approximant to } A \text{ over } \mathbb{M}\}.$$

The values of the first k norms $\|\cdot\|_{j,2}$, $j = 1, \dots, k$, are fixed on the set \mathbb{S}_k . Therefore the set \mathbb{S}_k can alternatively be defined as the set of minimizers of the singular value σ_k of R on \mathbb{S}_{k-1} . Therefore we obtain the same sets of minimizers if we replace the norms $\|\cdot\|_{k,2}$ by the Ky Fan p - k norms $\|\cdot\|_{k,p}$ for arbitrary $p \geq 2$ (see (10.3)). The strict spectral approximant $B^{(st)}$ is the matrix from \mathbb{M} such that $A - B^{(st)}$ is the unique element of \mathbb{S}_t because the Frobenius norm is strictly convex (see [106]). The same matrix will be determined if the norms (10.2) are replaced by the Ky Fan p - k norms because the c_p norm is strictly convex for $1 < p < \infty$. The properties of dual matrices for the Ky Fan p - k norm, $1 < p < \infty$, follow from [97] where one considers solutions of the overdetermined linear system $Ax = b$ with respect to the k -major l_p norms.

Let $R^{(st)}$ have the singular values decomposition

$$R^{(st)} = U \Sigma V^H, \tag{10.5}$$

where U and V are unitary matrices, and $\Sigma = \text{diag}(\sigma_j(R^{(st)}))$. Let the vector $\sigma(R^{(st)})$ have the form

$$\sigma(R^{(st)}) = [\mu_1, \dots, \mu_1, \mu_2, \dots, \mu_2, \dots, \mu_\ell, \dots, \mu_\ell]^T, \quad \mu_1 > \mu_2 > \dots > \mu_\ell \geq 0. \tag{10.6}$$

The multiplicity of μ_j is denoted by s_j . From the properties of the sets \mathbb{S}_k for $k = 0, \dots, \ell - 1$ we obtain

$$\mu_{k+1} = \min_{R \in \mathbb{S}_{t_k}} \sigma_{t_k+1}(R), \tag{10.7}$$

where

$$t_0 = 0, \quad t_k = s_1 + \dots + s_k \quad \text{for } k > 0. \tag{10.8}$$

We divide the unitary matrices U and V from (10.5) into submatrices according to the form (10.6) of $\sigma(R^{(st)})$:

$$\begin{aligned} U &= [U_1, \dots, U_\ell], & V &= [V_1, \dots, V_\ell], \\ U_j &\in \mathbb{C}^{m \times s_j}, & V_j &\in \mathbb{C}^{n \times s_j}. \end{aligned} \tag{10.9}$$

Thus the matrix $R^{(st)}$ can be written in the form (compare Section 6)

$$R^{(st)} = \mu_1 U_1 V_1^H + \dots + \mu_\ell U_\ell V_\ell^H. \tag{10.10}$$

The matrices $U_j V_j^H$ ($j = 1, \dots, \ell$), corresponding to nonzero μ_j , are unique (see, for example, [9] for properties of the SVD).

We will now investigate the additional properties of the sets \mathbb{S}_k . It is noted in [106], without any proof, that the sets \mathbb{S}_k are convex. The set \mathbb{S}_1 is convex because if $0 < \lambda < 1$, $X_1, X_2 \in \mathbb{M}$ and $\|A - X_1\|_\infty = \|A - X_2\|_\infty = \delta_\infty$, then $A - X_1, A - X_2 \in \mathbb{S}_1$ and

$$\|A - (\lambda X_1 + (1 - \lambda)X_2)\|_\infty \leq \lambda \|A - X_1\|_\infty + (1 - \lambda) \|A - X_2\|_\infty = \delta_\infty.$$

Thus $\|A - (\lambda X_1 + (1 - \lambda)X_2)\|_\infty = \delta_\infty$, hence $\lambda(A - X_1) + (1 - \lambda)(A - X_2) = A - (\lambda X_1 + (1 - \lambda)X_2) \in \mathbb{S}_1$. We now prove that also other sets \mathbb{S}_k are convex.

LEMMA 10.1. *Let \mathbb{S}_0 be defined in (10.4) and let \mathbb{S}_k be the set of the minimizers of $\|R\|_{k,2}$ over \mathbb{S}_{k-1} . Then \mathbb{S}_k are convex for $k = 1, \dots, t$.*

Proof. We prove the lemma by induction on the index k . For $k = 1$ it has been already done. Let $2 \leq k < t$ and let \mathbb{S}_{k-1} be convex. Let $A - X_1, A - X_2 \in \mathbb{S}_k$. Then $A - X_1, A - X_2 \in \mathbb{S}_{k-1}$ and for $0 < \lambda < 1$ we obtain $\lambda(A - X_1) + (1 - \lambda)(A - X_2) \in \mathbb{S}_{k-1}$, since $\mathbb{S}_k \subseteq \mathbb{S}_{k-1}$ and \mathbb{S}_{k-1} is convex. Hence (see (10.2))

$$\|A - (\lambda X_1 + (1 - \lambda)X_2)\|_{k,2} \leq \lambda \|A - X_1\|_{k,2} + (1 - \lambda) \|A - X_2\|_{k,2} = \|A - X_1\|_{k,2}.$$

Thus $\|A - (\lambda X_1 + (1 - \lambda)X_2)\|_{k,2} = \|A - X_1\|_{k,2}$, because $\|A - X_1\|_{k,2}$ is the smallest value of the $\|\cdot\|_{k,2}$ norm on \mathbb{S}_k . This implies that $\lambda(A - X_1) + (1 - \lambda)(A - X_2) \in \mathbb{S}_k$, which completes the proof. ■

We know that if we approximate a real matrix A by elements of the real linear subspace \mathbb{M} then for every spectral approximant $B \in \mathbb{M}$ to A the residual matrix $R = A - B$ has at least $t_1 = s_1$ singular values equal to $\sigma_1(A - X) = \mu_1$ (see Theorem 7.3). The same property can be proved for the case of complex matrices (compare the proof of Theorem 10.2 below). Therefore from the way of construction of the sets \mathbb{S}_j ($j = 1, \dots, t$) we conclude that

$$\mathbb{S}_1 = \mathbb{S}_2 = \dots = \mathbb{S}_{t_1}. \tag{10.11}$$

Theorem 10.2 implies that the analogous relations are valid for the rest sets \mathbb{S}_j

$$\mathbb{S}_{t_k+1} = \mathbb{S}_{t_k+2} = \dots = \mathbb{S}_{t_{k+1}}, \quad k = 1, \dots, \ell - 1. \tag{10.12}$$

For $k = \ell$ we have $\mathbb{S}_{t_\ell} = \mathbb{S}_t = \{R^{(st)}\}$.

The matrix $R^{(st)} = A - B^{(st)}$ can be written in the following form for $1 \leq k < \ell$ (see (10.5)–(10.10)):

$$R^{(st)} = \mu_1 U_1 V_1^H + \dots + \mu_k U_k V_k^H + [U_{k+1}, \dots, U_\ell] Z^{(st)} [V_{k+1}, \dots, V_\ell]^H \tag{10.13}$$

with

$$Z^{(st)} = \text{diag}(\mu_{k+1} I_{s_{k+1}}, \mu_{k+2} I_{s_{k+2}}, \dots, \mu_\ell I_{s_\ell}) \in \mathbb{R}^{\ell_k \times \ell_k}, \tag{10.14}$$

where

$$\ell_k = s_{k+1} + \dots + s_\ell, \tag{10.15}$$

I_j denotes the identity matrix of order j and $\|Z^{(st)}\|_\infty = \mu_{k+1}$. We now prove a characterization of the sets \mathbb{S}_{t_k} for $k = 1, \dots, \ell - 1$ (see (10.11) and (10.12)).

THEOREM 10.2. *Let \mathbb{M} be a linear subspace of $\mathbb{C}^{m \times n}$. Let $\mu_1, \dots, \mu_\ell, t_0, \dots, t_\ell, U_1, \dots, U_\ell$ and V_1, \dots, V_ℓ be determined as in (10.5), (10.6), (10.8) and (10.9), respectively, and let $1 \leq k < \ell$. Then every matrix R from \mathbb{S}_{t_k} has the form*

$$R = \mu_1 U_1 V_1^H + \dots + \mu_k U_k V_k^H + [U_{k+1}, \dots, U_\ell] Z [V_{k+1}, \dots, V_\ell]^H \tag{10.16}$$

for appropriate square matrix Z of order ℓ_k such that

$$B = A - R \in \mathbb{M}, \quad \|Z\|_\infty \leq \mu_k. \tag{10.17}$$

Proof. We prove by induction on the index k of the sets \mathbb{S}_{t_k} that their elements have the form (10.16) and the conditions (10.17) are satisfied. Let $k = 1$. The set \mathbb{S}_1 is determined in (10.11). For $R \in \mathbb{S}_1$ the formula (10.16) follows from Theorem 7.3 where now \widehat{R} is replaced by $R^{(st)}$ and $s = s_1$ because s_1 is the smallest number of singular values of the residual matrix $R = A - B$ equal to δ_∞ over all approximants B to A . Since we have (10.11), the formula (10.16) is valid for $R \in \mathbb{S}_{t_1}$.

In the k -th cycle of induction we assume that every element from \mathbb{S}_{t_k} has the form (10.16). We shall prove that $\mathbb{S}_{t_{k+1}}$ has an analogous form. The idea of the proof is similar to the idea of the proof of Theorem 4.3 in [104], given for real matrices (see also Theorem 7.3). Let $R = A - B \in \mathbb{S}_{t_{k+1}}$, $B \in \mathbb{M}$. Then $R \in \mathbb{S}_{t_k}$ and thus R has the form (10.16) with an appropriate Z such that $\|Z\|_\infty = \mu_{k+1}$ because $R \in \mathbb{S}_{t_{k+1}}$ (see (10.7) and (10.17)). Let us consider the strict spectral approximant $B^{(st)}$ to A . Then $A - B^{(st)}$ has the form (10.13) and $A - B^{(st)} \in \mathbb{S}_{t_{k+1}}$. Let

$$\widehat{B} = (1/2)(B + B^{(st)}).$$

Then $A - \widehat{B} = (1/2)(A - B) + (1/2)(A - B^{(st)}) \in \mathbb{S}_{t_{k+1}}$ because of the convexity of $\mathbb{S}_{t_{k+1}}$. Thus $\widehat{R} = A - \widehat{B}$ has at least s_{k+1} singular values equal to μ_{k+1} . This follows immediately from the construction of the set $\mathbb{S}_{t_{k+1}}$ and the definition of the strict spectral approximation. Let s be the number of the singular values of \widehat{R} equal to μ_{k+1} . Then we have $s \geq s_{k+1}$. We prove that $s = s_{k+1}$.

From the definition of \widehat{B} it follows that

$$A - \widehat{B} = \sum_{j=1}^k \mu_j U_j V_j^H + [U_{k+1}, \dots, U_\ell] \widehat{Z} [V_{k+1}, \dots, V_\ell]^H,$$

where (see (10.14))

$$\widehat{Z} = \frac{1}{2}(Z + Z^{(\text{st})}) = \frac{1}{2}(Z + \text{diag}(\mu_{k+1}I_{s_{k+1}}, \dots, \mu_\ell I_{s_\ell})) \in \mathbb{C}^{\ell_k \times \ell_k},$$

ℓ_k is determined in (10.15) and $\|\widehat{Z}\|_\infty = \mu_{k+1}$ because $A - \widehat{B} \in \mathbb{S}_{t_{k+1}}$ and from the construction of $\mathbb{S}_{t_{k+1}}$ it follows that the t_{k+1} th singular value of $A - \widehat{B}$ is equal to μ_{k+1} . Let \widehat{Z} have the SVD

$$\widehat{Z} = P\widehat{\Sigma}Q^H.$$

Let for $j = 1, \dots, \ell_k$ vectors x_j and y_j denote the j -th columns of the unitary matrices P and Q , respectively. We define

$$\alpha_j = x_j^H Z y_j, \quad \beta_j = x_j^H Z^{(\text{st})} y_j.$$

We notice that $\gamma_j = (1/2)(\alpha_j + \beta_j) = \mu_{k+1}$ for $j = 1, \dots, s$ and obviously $0 \leq \gamma_j \leq (1/2)(|\alpha_j| + |\beta_j|)$. Thus by the following property of the spectral norm (see [89, p. 80])

$$\|X\|_\infty = \max_{\|x\|_2 = \|y\|_2 = 1} |x^H X y|, \quad x \in \mathbb{C}^m, y \in \mathbb{C}^n,$$

we obtain $|\alpha_j| \leq \mu_{k+1}$, $|\beta_j| \leq \mu_{k+1}$ because $\|Z\|_\infty = \|Z^{(\text{st})}\|_\infty = \mu_{k+1}$, and

$$\mu_{k+1} = x_j^H \widehat{Z} y_j = \frac{1}{2}(\alpha_j + \beta_j) \leq \frac{1}{2}(|\alpha_j| + |\beta_j|).$$

This implies that $|\alpha_j| = |\beta_j| = \mu_{k+1}$, i.e.

$$|x_j^H Z y_j| = |x_j^H Z^{(\text{st})} y_j| = \mu_{k+1} \quad (j = 1, \dots, s).$$

Therefore we have equalities in the following Cauchy–Schwarz inequalities

$$\begin{aligned} \mu_{k+1} &= |x_j^H Z y_j| \leq \|x_j\|_2 \|Z y_j\|_2 \leq \mu_{k+1}, \\ \mu_{k+1} &= |x_j^H Z^{(\text{st})} y_j| \leq \|x_j\|_2 \|Z^{(\text{st})} y_j\|_2 \leq \mu_{k+1}. \end{aligned}$$

This implies that $\{Z^{(\text{st})} y_j, x_j\}$, $\{(Z^{(\text{st})})^H x_j, y_j\}$, $\{Z^H x_j, y_j\}$ and $\{Z y_j, x_j\}$ are the pairs of linearly dependent vectors and

$$\|Z y_j\|_2 = \|Z^{(\text{st})} y_j\|_2 = \|Z^H x_j\|_2 = \|(Z^{(\text{st})})^H x_j\|_2 = \mu_{k+1}.$$

Moreover, we have $\alpha_j + \beta_j = |\alpha_j| + |\beta_j|$. Hence α_j and β_j are real and nonnegative. Therefore comparing the norms we conclude that

$$Z^{(\text{st})} y_j = \mu_{k+1} x_j, \quad (Z^{(\text{st})})^H x_j = \mu_{k+1} y_j \quad (j = 1, \dots, s), \tag{10.18}$$

so for $j = 1, \dots, s$ the vectors x_j and y_j are left and right singular vectors of $Z^{(\text{st})}$ corresponding to the singular values equal to μ_{k+1} . This means that $s = s_{k+1}$ because $Z^{(\text{st})}$ has exactly s_{k+1} singular values equal to μ_{k+1} .

Let us partition P and Q into subblocks: $P = [P_1, P_2]$, $Q = [Q_1, Q_2]$, where

$$\begin{aligned} P_1 &= [x_1, \dots, x_{s_{k+1}}] \in \mathbb{C}^{\ell_k \times s_{k+1}}, & P_2 &= [x_{s_{k+1}+1}, \dots, x_{\ell_k}] \in \mathbb{C}^{\ell_k \times (\ell_k - s_{k+1})}, \\ Q_1 &= [y_1, \dots, y_{s_{k+1}}] \in \mathbb{C}^{\ell_k \times s_{k+1}}, & Q_2 &= [y_{s_{k+1}+1}, \dots, y_{\ell_k}] \in \mathbb{C}^{\ell_k \times (\ell_k - s_{k+1})}. \end{aligned}$$

Then (see (10.18))

$$Z^{(\text{st})} Q_1 = \mu_{k+1} P_1, \quad P_1^H Z^{(\text{st})} = \mu_{k+1} Q_1^H,$$

and we conclude that the following relations hold:

$$P_2^H Z^{(\text{st})} Q_1 = 0, \quad P_1^H Z^{(\text{st})} Q_2 = 0, \quad P_1^H Z^{(\text{st})} Q_1 = \mu_{k+1} I_{s_{k+1}}. \tag{10.19}$$

Let $\tilde{Z}_{22}^{(st)} = P_2^H Z^{(st)} Q_2$. The matrix

$$P^H Z^{(st)} Q = \begin{bmatrix} \mu_{k+1} I_{s_{k+1}} & 0 \\ 0 & \tilde{Z}_{22}^{(st)} \end{bmatrix}$$

has the same singular values as $Z^{(st)}$ and

$$Z^{(st)} = \mu_{k+1} P_1 Q_1^H + P_2 \tilde{Z}_{22}^{(st)} Q_2^H.$$

The matrix $Z^{(st)}$ has exactly s_{k+1} singular values equal to μ_{k+1} . Therefore

$$\|P_2 \tilde{Z}_{22}^{(st)} Q_2^H\|_\infty < \mu_{k+1}$$

and (see (10.14))

$$P_1 Q_1^H = \begin{bmatrix} I_{s_{k+1}} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{\ell_k \times \ell_k},$$

because the matrix $P_1 Q_1^H$ corresponding to the singular values equal to μ_{k+1} is unique (compare (6.12)). Thus we obtain

$$P_1 P_1^H = P_1 Q_1^H Q_1 P_1^H = Q_1 Q_1^H = Q_1 P_1^H P_1 Q_1^H = \begin{bmatrix} I_{s_{k+1}} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{\ell_k \times \ell_k},$$

$$P_2 P_2^H = I - P_1 P_1^H = Q_2 Q_2^H = I - Q_1 Q_1^H = \begin{bmatrix} 0 & 0 \\ 0 & I_{\ell_k - s_{k+1}} \end{bmatrix} \in \mathbb{R}^{\ell_k \times \ell_k},$$

$$[U_{k+1}, \dots, U_\ell] P_2 P_2^H = [0, U_{k+2}, \dots, U_\ell], \tag{10.20}$$

$$Q_2 Q_2^H [V_{k+1}, \dots, V_\ell]^H = [0, V_{k+2}, \dots, V_\ell]^H. \tag{10.21}$$

Let us now consider Z determined in (10.16). The vectors x_j and y_j for $j = 1, \dots, s$ are also singular vectors of Z , corresponding to the singular values equal to μ_{k+1} . By the same arguments as for $Z^{(st)}$ it is easy to verify that (see (10.18) and (10.19))

$$\begin{aligned} Z y_j &= \mu_{k+1} x_j, & Z^H x_j &= \mu_{k+1} y_j \quad (j = 1, \dots, s), \\ \text{and } P_1^H Z Q_1 &= \mu_{k+1} I_{s_{k+1}}, & P_2^H Z Q_1 &= 0, & P_1^H Z Q_2 &= 0, \end{aligned}$$

where $s = s_{k+1}$, as we have shown. Thus the last term in (10.16)) can be expressed in the following way (see (10.20) and (10.21)):

$$\begin{aligned} & [U_{k+1}, \dots, U_\ell] Z [V_{k+1}, \dots, V_\ell]^H \\ &= \mu_{k+1} U_{k+1} V_{k+1}^H + [U_{k+1}, \dots, U_\ell] P_2 P_2^H Z Q_2 Q_2^H [V_{k+1}, \dots, V_\ell]^H \\ &= \mu_{k+1} U_{k+1} V_{k+1}^H + [0, U_{k+2}, \dots, U_\ell] Z [0, V_{k+2}, \dots, V_\ell]^H \\ &= \mu_{k+1} U_{k+1} V_{k+1}^H + [U_{k+2}, \dots, U_\ell] Z_{22} [V_{k+2}, \dots, V_\ell]^H, \end{aligned}$$

with the block Z_{22} of Z :

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \in \mathbb{R}^{\ell_k \times \ell_k}, \quad Z_{11} \in \mathbb{R}^{s_{k+1} \times s_{k+1}},$$

and

$$\|[U_{k+1}, \dots, U_\ell] P_2 P_2^H Z Q_2 Q_2^H [V_{k+1}, \dots, V_\ell]^H\|_\infty \leq \mu_{k+1}.$$

Thus $A - B \in \mathbb{S}_{t_{k+1}}$ has the form (10.16) where k is replaced by $k + 1$. This completes the proof. ■

It is known that the signs of some components of the residual vectors $r = b - Gx^{(Ch)}$, whose modules are equal to the approximation error $\|b - Gx^{(Ch)}\|_\infty$, are common for all Chebyshev solutions $x^{(Ch)}$ of the overdetermined system of linear equations $Gx = b$ (see (2.5)). For the problem (10.1) the matrix $U_1V_1^H$ in (10.16), corresponding to the singular values equal to the approximation error $\delta_\infty = \|A - B\|_\infty$ is common for all spectral approximants to A .

Let $X \in \mathbb{M}$, $R = A - X \in \mathbb{S}_0$. If $R \in \mathbb{S}_1$, i.e. X is any spectral approximant to A , then we cannot have $\sigma_i(R) \neq \mu_1$ for some $1 \leq i \leq s_1$ (see Theorem 10.2). If $R \notin \mathbb{S}_1$ then X is not a spectral approximant to A and therefore $\|A - X\|_\infty > \mu_1$. If $R \in \mathbb{S}_{t_k+1}$, then $R \in \mathbb{S}_{t_k}$ and an index i ($t_k + 1 \leq i \leq t_{k+1}$) such that $\sigma_i(R) \neq \mu_{k+1}$ cannot exist because we have to have $\sigma_i(R) = \mu_{k+1}$ for every i , $t_k + 1 \leq i \leq t_{k+1}$ (see Theorem 10.2). If $R \in \mathbb{S}_{t_k}$ and $R \notin \mathbb{S}_{t_{k+1}}$, then by the definition of $\mathbb{S}_{t_{k+1}}$ we have $\sigma_{t_{k+1}}(R) > \mu_{k+1}$. Therefore from Theorem 10.2 we obtain the following corollary.

COROLLARY 10.3. *Let the assumptions of Lemma 10.1 be satisfied. Let s_k be the multiplicity of the singular values of $R^{(st)}$ equal to μ_k defined in (10.6) and let t_k be defined in (10.8). Let $X \in \mathbb{M}$, $R = A - X \in \mathbb{S}_0$.*

- (a) *If $\sigma_i(R) \neq \mu_1$ for some index i satisfying $1 \leq i \leq s_1$, then $\|R\|_\infty > \mu_1$, i.e. $R \notin \mathbb{S}_1$.*
- (b) *If $R \in \mathbb{S}_{t_k}$ and $\sigma_i(R) \neq \mu_{k+1}$ for some index i satisfying $t_k + 1 \leq i \leq t_{k+1}$, then $\sigma_{t_{k+1}}(R) > \mu_{k+1}$, i.e. $R \notin \mathbb{S}_{t_{k+1}}$.*

Let $1 < p < \infty$. The spectral norm corresponds to $p = \infty$. It is obvious that

$$\|Y\|_\infty \leq \|Y\|_p \leq t^{1/p} \|Y\|_\infty \quad (Y \in \mathbb{C}^{m \times n}),$$

where $t = \min\{m, n\}$. This implies that

$$\|Y\|_\infty = \lim_{p \rightarrow \infty} \|Y\|_p.$$

Let $B_p \in \mathbb{M}$ be the c_p -approximant to A and

$$R_p = A - B_p, \quad \delta_p = \|R_p\|_p. \tag{10.22}$$

Let $B \in \mathbb{M}$ be any spectral approximant to A . Then (see (10.1))

$$\delta_\infty = \|A - B\|_\infty \leq \|A - B_p\|_\infty \leq \|A - B_p\|_p \leq \|A - B\|_p \leq t^{1/p} \delta_\infty.$$

Therefore

$$\lim_{p \rightarrow \infty} \delta_p = \delta_\infty. \tag{10.23}$$

The sequence $\{R_p\}_{p=1}^\infty$ is bounded

$$\|R^{(st)}\|_\infty \leq \|R_p\|_\infty \leq \|R_p\|_p \leq \|A\|_p \leq t^{1/p} \|A\|_\infty \leq t \|A\|_\infty$$

because the zero matrix belongs to the linear subspace \mathbb{M} . Thus by the theorem of Bolzano–Weierstrass there exists a subsequence $1 < p_1 < p_2 < \dots$ of integers growing to infinity such that there exists the limit R_∞ of the subsequence $\{R_{p_i}\}_{i=1}^\infty$ with respect to the spectral norm. Therefore the subsequence $\{B_{p_i}\}_{i=1}^\infty$, where $B_{p_i} = A - R_{p_i}$, is convergent to $B_\infty = A - R_\infty$ because

$$\|B_{p_i} - B_\infty\|_\infty = \|R_{p_i} - R_\infty\|_\infty.$$

A sequence of matrices $F_k \in \mathbb{C}^{m \times n}$ converges coordinatewise to F if and only if $\|F_k - F\| \rightarrow 0$ as $k \rightarrow \infty$ in any norm (see [54, Proposition 1, p. 361]). Thus

$$R_\infty = \lim_{i \rightarrow \infty} R_{p_i} \quad \text{and} \quad B_\infty = \lim_{i \rightarrow \infty} B_{p_i}. \tag{10.24}$$

The matrix B_∞ belongs to \mathbb{M} because $B_p \in \mathbb{M}$ and any finite-dimensional metric space is closed. It is easy to verify that B_∞ is a spectral approximant to A . Indeed, let us suppose a contrario that B_∞ is not a spectral approximant to A . Then there exists $\varepsilon > 0$ such that

$$\begin{aligned} \delta_\infty + \varepsilon &\leq \|A - B_\infty\|_\infty \\ &\leq \|A - B_{p_i}\|_\infty + \|B_{p_i} - B_\infty\|_\infty \leq \|A - B_{p_i}\|_{p_i} + \|B_{p_i} - B_\infty\|_\infty. \end{aligned}$$

This implies a contradiction $\delta_\infty + \varepsilon \leq \delta_\infty$ when $i \rightarrow \infty$ because we have (10.23) and (10.24). Therefore B_∞ is a spectral approximant to A and applying the continuity of the singular values (see (6.14)) we obtain the following corollary.

COROLLARY 10.4. *The limit B_∞ of a convergent subsequence $\{B_{p_i}\}_{i=1}^\infty$ of the c_{p_i} -approximants to A is the spectral approximant to A and for every j*

$$\lim_{i \rightarrow \infty} \sigma_j(A - B_{p_i}) = \sigma_j(A - B_\infty).$$

Let \widehat{R} be the limit of any convergent subsequence of the R_p (see (10.22)). Then Corollary 10.4 implies that $\widehat{R} \in \mathbb{S}_1$. This means that for every convergent subsequence $\{R_{p_i}\}$ of $\{R_p\}$ the singular value $\sigma_j(R_{p_i})$ tends to the common limit μ_1 for $j = 1, \dots, s_1$ (see Theorem 10.2). Therefore we have the following corollary.

COROLLARY 10.5. *Let R_p be determined in (10.22). Then*

$$\lim_{p \rightarrow \infty} \sigma_j(R_p) = \sigma_j(R^{(\text{st})}) = \mu_1 \quad \text{for } j = 1, \dots, s_1.$$

Now, our goal is to prove by induction on k that for $k = 1, \dots, \ell$ we have

$$\lim_{p \rightarrow \infty} \sigma_i(R_p) = \sigma_i(R^{(\text{st})}) = \mu_k \quad \text{for } i = t_{k-1} + 1, \dots, t_k. \tag{10.25}$$

Unfortunately, till now the proof of (10.25) is not complete. For $k = 1$ the limit (10.25) follows from Corollary 10.5 because $t_0 = 0, t_1 = s_1$. We now explain in details a reason why we were not able to prove (10.25).

Let us suppose that for $k, 1 \leq k < \ell$, we have already proved that

$$\lim_{p \rightarrow \infty} \sigma_i(R_p) = \sigma_i(R^{(\text{st})}) \quad \text{for } i = 1, \dots, t_k. \tag{10.26}$$

We should prove that

$$\lim_{p \rightarrow \infty} \sigma_i(R_p) = \sigma_i(R^{(\text{st})}) = \mu_{k+1} \quad \text{for } i = t_k + 1, \dots, t_{k+1}. \tag{10.27}$$

We suppose a contrario that for any index $i_*, t_k + 1 \leq i_* \leq t_{k+1}$,

$$\lim_{p \rightarrow \infty} \sigma_{i_*}(R_p) \neq \sigma_{i_*}(R^{(\text{st})}) = \mu_{k+1} \tag{10.28}$$

(or the limit does not exist). Let us consider a sequence of integers $1 < p_1 < p_2 < \dots$ growing to infinity, such that the subsequence $\{R_{p_i}\}$ of $\{R_p\}$ is convergent to R_∞ . From

(10.26) we deduce that $R_\infty \in \mathbb{S}_{t_k}$ and the assumption (10.28) implies that

$$\sigma_{t_k+1}(R^{(st)}) = \mu_{k+1} < \sigma_{t_k+1}(R_\infty), \tag{10.29}$$

because of Corollary 10.3. Thus, by the continuity of singular values and (10.24), there exist an index i_0 and a number z such that for $i > i_0$ we have

$$\mu_{k+1} < z < \sigma_{t_k+1}(R_{p_i}). \tag{10.30}$$

The existence of z satisfying (10.30) follows from the following considerations. Let

$$\gamma = \sigma_{t_k+1}(R_\infty) - \mu_{k+1}$$

and let ε satisfy $0 \leq \varepsilon < \gamma$ (see (10.29)). Then there exists i_0 such that for $i > i_0$ we have

$$|\sigma_{t_k+1}(R_{p_i}) - \sigma_{t_k+1}(R_\infty)| \leq \varepsilon,$$

and consequently

$$\sigma_{t_k+1}(R_\infty) - \varepsilon \leq \sigma_{t_k+1}(R_{p_i}).$$

Then

$$\mu_{k+1} = \sigma_{t_k+1}(R_\infty) - \gamma < \sigma_{t_k+1}(R_\infty) - \varepsilon \leq \sigma_{t_k+1}(R_{p_i}).$$

Thus (10.30) follows for $z = \sigma_{t_k+1}(R_\infty) - \varepsilon$. The above consideration does not depend on whether $\mu_{k+1} > 0$ or $\mu_{k+1} = 0$. We now would like to select a residual matrix $\tilde{R}_{p_i}^{(k)} = A - \tilde{X}_{p_i}^{(k)}$, $\tilde{X}_{p_i}^{(k)} \in \mathbb{M}$, such that

$$\|\tilde{R}_{p_i}^{(k)}\|_{t_k, p_i} \leq \|R_{p_i}\|_{t_k, p_i}, \tag{10.31}$$

and there exists a constant w such that $0 < w < z$ and

$$\sigma_{t_k+1}(\tilde{R}_{p_i}^{(k)}) < w. \tag{10.32}$$

If such a residual matrix $\tilde{R}_{p_i}^{(k)}$ and a constant w exist then there exists i_1 , $i_1 > i_0$, such that for $i > i_1$ we have

$$(t - t_k)w^{p_i} < z^{p_i},$$

because $w < z$. Then

$$\begin{aligned} \|\tilde{R}_{p_i}^{(k)}\|_{p_i}^{p_i} &= \sum_{q=1}^{t_k} (\sigma_q(\tilde{R}_{p_i}^{(k)}))^{p_i} + \sum_{q=t_k+1}^t (\sigma_q(\tilde{R}_{p_i}^{(k)}))^{p_i} \leq \sum_{q=1}^{t_k} (\sigma_q(R_{p_i}))^{p_i} + (t - t_k)w^{p_i} \\ &< \sum_{q=1}^{t_k} (\sigma_q(R_{p_i}))^{p_i} + z^{p_i} \leq \sum_{q=1}^t (\sigma_q(R_{p_i}))^{p_i} = \|R_{p_i}\|_{p_i}^{p_i}, \end{aligned}$$

which contradicts the fact that $B_{p_i} = A - R_{p_i}$ is the c_{p_i} -approximant to A . Thus (10.28) cannot hold and hence we have (10.27).

It is an open problem whether there exist $\tilde{R}_{p_i}^{(k)}$ and w satisfying (10.31) and (10.32). It is the reason that our proof of (10.25) is not complete. The inequality (10.32), in some sense, corresponds to certain inequality in the proof of Descloux in [24] for the vector case (see the inequality involving a constant w on top of page 1022 in [24]). Unfortunately, the details of the proof of the inequality from [24], involving w , are not given in the paper of Descloux [24].

For $1 < p < \infty$ and $k = 1, \dots, \ell$ we introduce the following auxiliary sets for $j = k, \dots, \ell$: $\mathbb{K}_p^{(j)}$ is the set of residual matrices $R_p^{(j)} = A - B_p^{(j)}$, where $B_p^{(j)} \in \mathbb{M}$ is an

approximant to A with respect to the $\|\cdot\|_{t_j,p}$ norm of Ky Fan (see (10.3)). The indices ℓ and t_j are defined in (10.6) and (10.8), respectively. For $j = \ell$ the set $\mathbb{K}_p^{(\ell)}$ contains the only element $R_p = A - B_p$ where $B_p \in \mathbb{M}$ is the unique c_p -approximant to A . It is obvious that

$$\|R_p\|_{t_j,p} \geq \|R_p^{(j)}\|_{t_j,p}, \quad R_p^{(j)} \in \mathbb{K}_p^{(j)}, \quad j = k, \dots, \ell.$$

The sets $\mathbb{K}_p^{(j)}$ are convex. Therefore we can select in $\mathbb{K}_p^{(j)}$ a matrix $R_{\min}^{(j,p)}$ minimal in lexicographic ordering among all elements of $\mathbb{K}_p^{(j)}$. We conjecture that the matrix $R_{\min}^{(k,p)}$ could be selected as a matrix $\tilde{R}_p^{(k)}$ in (10.31) and (10.32). Moreover we conjecture that for $1 \leq k < \ell$ and $j = k, \dots, \ell$ we have

$$\lim_{p \rightarrow \infty} \sigma_i(R_{\min}^{(j,p)}) = \sigma_i(R^{(\text{st})}) \quad \text{for } i = 1, \dots, t_k \tag{10.33}$$

(compare [24, relation (4)]).

If (10.25) were proved then an attempt to proof of Conjecture 1.1 could be the following. From (10.25) we deduce that all singular values of a limit of every convergent subsequence $\{R_{p_i}\}_{i=1}^\infty$ of the sequence $\{R_p\}_{p=1}^\infty$ are equal to the singular values of $R^{(\text{st})}$. Let a subsequence $\{R_{p_i}\}_{i=1}^\infty$ be convergent to R_∞ (see (10.24)). It is known that R_∞ determines a spectral approximant $B_\infty = A - R_\infty$ to A (see Corollary 10.4). Thus R_∞ has to be equal to $R^{(\text{st})}$ because of the uniqueness of $R^{(\text{st})}$, i.e. because there does not exist any matrix $R \in \mathbb{S}_0$ different from $R^{(\text{st})}$ and having the same singular values as $R^{(\text{st})}$. Therefore the convergent subsequences $\{R_{p_i}\}_{i=1}^\infty$ would have the common limit equal to $R^{(\text{st})}$. This would imply the convergence of R_p to $R^{(\text{st})}$ when $p \rightarrow \infty$, hence it would imply that Conjecture 1.1 is true.

Another approach to prove Conjecture 1.1 could be based upon some idea from the proof of Marana, Ureña [69] for the Pólya algorithm in \mathbb{R}^m . It is out of the scope of this paper.

At the end of this paper we would like to formulate the following open problems:

- Describe dual matrices with respect to the $\|\cdot\|_{k,p}$ norms using the results for the k -major l_p norms of vectors from [97]. Could it help to characterize matrices from $\mathbb{K}_{t_k,p}^{(j)}$ and to solve the open question concerning the inequality (10.32)?
- In [69] one investigates conditions under which the l_p -approximants to a vector coincide with the strict Chebyshev approximant to a vector. How to extend it to the case of approximation of matrices in the c_p norms?

These problems and the proof of the conjecture should be a subject of future investigations.

11. Appendix on lemmas related to the Pólya algorithm. The result of Descloux (3.1) has been extended to an approximation over some more general sets than linear subspaces \mathbb{V} of \mathbb{R}^m (see, for example, [45, 53, 68, 77]). In particular, in [45] one considers an extension to convex E -cylindrical sets. It is known that a linear subspace is a cylindrical set in every direction (see [45, Theorem (2.4)]). As we have written in Section 3, Egger and Huotari [30] say that the crucial Lemma 2.5 in [45] is false, hence it is not clear whether the extension is correct. Egger and Huotari do not write why Lemma 2.5 in [45]

is false. We now analyze the proof of this lemma. Moreover, we show that the proof of [30, Lemma 5], which is a particular case of [45, Lemma 2.5], is not correct.

For the convenience of the reader we present the original proofs of [45, Lemma 2.5] and [30, Lemma 5]. These lemmas are formulated as Lemma 3.1 in Section 3.

We first analyze the following proof of Lemma 5 in [30], to show its weak points: *Let $T : \Pi_k(\mathbb{V}) \rightarrow \mathbb{V}$ be any linear right inverse of Π_k ([54]). Since T is continuous and since the norms l_∞ and l_1 are topologically equivalent, the result follows.* The definition of $\Pi_k(y)$ is given in (3.2). We underline that in [30] one applies row vectors instead of column vectors as in this paper. Therefore in our analysis of this cited proof any linear right inverse has to be replaced by any linear left inverse. We now show that a left inverse needed in the cited proof does not exist.

We now recall the definition of the left inverse of a matrix (see [54, p. 425]). A matrix $F \in \mathbb{R}^{m \times n}$ is said to be left invertible if there exists a matrix $F_L^{(-1)} \in \mathbb{R}^{n \times m}$ such that $F_L^{(-1)}F = I_n$. A matrix $F \in \mathbb{R}^{m \times n}$ is left invertible if and only if $m \geq n$ and $\text{rank}(F) = n$. Analogously, one defines a left inverse of a linear transformation on a space of vectors (see [54, p. 139]). Namely, let \mathbb{V}_1 and \mathbb{V}_2 be linear spaces of dimension n and m , respectively. Let $W : \mathbb{V}_1 \rightarrow \mathbb{V}_2$ be a linear transformation. Then W is said to be left invertible if there exists a linear transformation $W_L : \mathbb{V}_2 \rightarrow \mathbb{V}_1$ such that $W_L W$ the identity transformation on \mathbb{V}_1 . The following conditions are equivalent (see [54, Theorem 2, p. 139]): a linear transformation W is left invertible if and only if every representation of W is an $m \times n$ matrix with $m \geq n$ and of full rank. In our case the linear transformation Π_k can be represented by the $m \times m$ matrix P_k of rank $= k$

$$P_k = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m \times m}, \tag{11.1}$$

where I_k is the identity matrix of order k . The left inverse of P_k exists if and only if $k = m$. Therefore the left inverse of the linear transformation Π_k does not exist in general case. This implies that the proof of [30, Lemma 5] is not correct. Moreover, an analysis of the formulation of Lemma 5 of Egger and Huotari shows in [111] that Lemma 5 does not cover all possible cases which should be considered in the proof of Theorem 7 in [30]. Therefore the proof of Egger and Huotari of the convergence of the Pólya algorithm needs improvements. It is out of the scope of the paper.

We now recall the definition of cylindrical subsets (see [45]). Let $x, v \in \mathbb{R}^m$, $\mathbb{A} \subset \mathbb{R}^m$, and $L(x, v)$ be the straight line in \mathbb{R}^m which contains x and is parallel to the line containing zero vector and v . We define

$$d(x, \mathbb{A}) = \inf\{\|x - y\|_\infty : y \in \mathbb{A}\}.$$

A subset \mathbb{A} of \mathbb{R}^m is v -cylindrical at x , if for any $\varepsilon > 0$ there exists $\delta(x, \varepsilon) > 0$ such that $d(x, L(y, v) \cap \mathbb{A}) < \varepsilon$ whenever $y \in \mathbb{A}$ and $d(y, L(x, v)) < \delta$. The set \mathbb{A} is said to be v -cylindrical if it is v -cylindrical at every x in the closure of \mathbb{A} . Let $E = \{e_1, \dots, e_m\}$ be the standard basis of \mathbb{R}^m . A subset \mathbb{A} of \mathbb{R}^m is said to be E -cylindrical if \mathbb{A} is e_i -cylindrical for $1 \leq i \leq m$.

Let \mathbb{B} be an E -cylindrical subset of \mathbb{R}^m . We now analyze the proof of Lemma 2.5 in [45]: *For each j with $k < j \leq m$, let \mathbb{V}_j be the subspace of \mathbb{R}^m satisfying the equations*

$$x_{k+1} = \dots = x_{j-1} = x_{j+1} = \dots = x_m = 0.$$

Then \mathbb{B}_j , the orthogonal projection of \mathbb{B} onto \mathbb{V}_j , is cylindrical in the direction e_j , so there exists z_j such that

$$(y_1, \dots, y_k, z_j) \in \mathbb{B}_j \tag{11.2}$$

and

$$\max\{|x_j - z_j|, |x_i - y_i| : 1 \leq i \leq k\} < \varepsilon. \tag{11.3}$$

Thus, the coordinates z_j can be chosen independently, so that the vector

$$(y_1, \dots, y_k, z_{k+1}, \dots, z_m) \tag{11.4}$$

satisfies the conclusion of [45, Lemma 2.5]. In this proof for each $k < j \leq m$ one constructs a vector (11.2) satisfying (11.3). Unfortunately, it is not obvious for us that the vector

$$(y_1, \dots, y_k, z_{k+1}, \dots, z_m),$$

defined in (11.4), belongs to \mathbb{B} . We now show that it can be false in our case, i.e. for \mathbb{B} being the linear subspace \mathbb{V} of \mathbb{R}^m .

EXAMPLE 11.1. Let $\mathbb{B} = \mathbb{V}$ be the linear subspace spanned by columns of the following matrix G

$$G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}. \tag{11.5}$$

Let

$$Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Q_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

For $k = 1$ we define the following two subspaces \mathbb{V}_j :

$$\mathbb{V}_2 = \{[x_1, x_2, 0]^T : x_1, x_2 \in \mathbb{R}\} = \{Q_2 x : x \in \mathbb{R}^3\},$$

$$\mathbb{V}_3 = \{[x_1, 0, x_3]^T : x_1, x_3 \in \mathbb{R}\} = \{Q_3 x : x \in \mathbb{R}^3\}.$$

The orthogonal projections of $\mathbb{B} = \mathbb{V}$ onto these two subspaces are the following subspaces

$$\mathbb{B}_2 = \{Q_2 G u : u \in \mathbb{R}\}, \quad \mathbb{B}_3 = \{Q_3 G u : u \in \mathbb{R}\}.$$

Let

$$\begin{aligned} x &= [x_1, x_2, x_3]^T = G[u_1, u_2]^T = [u_1, u_2, u_1 + u_2]^T \in \mathbb{B}, \\ y &= [y_1, y_2, y_3]^T = G[w_1, w_2]^T = [w_1, w_2, w_1 + w_2]^T \in \mathbb{B}. \end{aligned}$$

We consider the vectors

$$Q_2 G[w_1, \widehat{w}_2]^T = [w_1, \widehat{w}_2, 0]^T \in \mathbb{B}_2, \quad Q_3 G[w_1, \widetilde{w}_2]^T = [w_1, 0, w_1 + \widetilde{w}_2]^T \in \mathbb{B}_3$$

and the vector $z = [w_1, z_2, z_3]^T \equiv [w_1, \widehat{w}_2, w_1 + \widetilde{w}_2]^T$ (see (11.4)). According to the assumption of Lemma 3.1 (i) we assume that $\|x - y\|_\infty = |u_1 - w_1| < \eta$. We should take η, \widehat{w}_2 , and \widetilde{w}_2 such that (see (11.3))

$$\max\{|x_2 - z_2|, |x_3 - z_3|, |x_1 - y_1|\} < \varepsilon.$$

Therefore we consider the conditions

$$\begin{aligned} |u_2 - \widehat{w}_2| &< \varepsilon, \\ |u_1 + u_2 - w_1 - \widetilde{w}_2| &\leq |u_1 - w_1| + |u_2 - \widetilde{w}_2| \leq \eta + |u_2 - \widetilde{w}_2| < \varepsilon. \end{aligned}$$

Thus, if, for example, \tilde{w}_2 is such that $|u_2 - \tilde{w}_2| < \eta$ and $\eta < \varepsilon/2$, then

$$\|x - z\|_\infty = \max\{|u_1 - w_1|, |u_2 - \tilde{w}_2|, |u_1 + u_2 - w_1 - \tilde{w}_2|\} < \varepsilon$$

and the vector $z = [w_1, \tilde{w}_2, w_1 + \tilde{w}_2]^T$ is the vector constructed in the proof in [45]. Unfortunately, $z \notin \mathbb{B} = \mathbb{V}$ if $\tilde{w}_2 \neq \hat{w}_2$. Therefore it is not true that z_j in the proof of [45, Lemma 2.5] can be chosen independently, which rises doubts about the proof in [45]. In our case we should put $\tilde{w}_2 = \hat{w}_2$ in order to have $z \in \mathbb{V}$.

Acknowledgments. The author would like to thank Jaroslav Zemánek (Institute of Mathematics, Polish Academy of Sciences) for encouraging her to write this paper and the referee for suggesting consideration of normal matrices in Section 9. The author thanks Jean Descloux (École Polytechnique Fédérale de Lausanne) for some comments concerning his proof of [24, Theorem]), and Iwona Wróbel and Paweł Keller (Warsaw University of Technology) for discussions about Lemma 5 of Egger and Huotari.

References

- [1] N. N. Abdelmalek, *Computing the strict Chebyshev solution of overdetermined linear equations*, Math. Comp. 31 (1977), 974–983.
- [2] T. Ando, *Approximation in trace norm by positive semidefinite matrices*, Linear Algebra Appl. 71 (1985), 15–21.
- [3] T. Ando, T. Sekiguchi, T. Suzuki, *Approximation by positive operators*, Math. Z. 131 (1973), 273–282.
- [4] L. Arambašić, R. Rajić, *The Birkhoff–James orthogonality in Hilbert C^* -modules*, Linear Algebra Appl. 437 (2012), 1913–1929.
- [5] M. Barraa, M. Boumazgour, *A note on the orthogonality of bounded linear operators*, Funct. Anal. Approx. Comput. 4 (2012), 65–70.
- [6] F. L. Bauer, *On the field of values subordinate to a norm*, Numer. Math. 4 (1962), 103–113.
- [7] A. Ben-Israel, Th. N. E. Greville, *Generalized Inverses: Theory and Applications*, CMS Books Math., second ed., Springer, Berlin 2003.
- [8] C. Benítez, M. Fernández, M. L. Soriano, *Orthogonality of matrices*, Linear Algebra Appl. 422 (2007), 155–163.
- [9] H. Berens, M. Finzel, *A continuous selection of the metric projection*, in: Numerical Methods of Approximation Theory, vol. 8 (Oberwolfach, 1986), Internat. Schriftenreihe Numer. Math. 81, Birkhäuser, Basel 1987, 21–29.
- [10] R. Bhatia, *Matrix Analysis*, Grad. Texts in Math. 169, Springer, New York 1997.
- [11] R. Bhatia, F. Kittaneh, *Approximation by positive operators*, Linear Algebra Appl. 161 (1992), 1–9.
- [12] R. Bhatia, P. Šemrl, *Orthogonality of matrices and some distance problems*, Linear Algebra Appl. 287 (1999), 77–85.
- [13] T. Bhattacharyya, P. Grover, *Characterization of Birkhoff–James orthogonality*, J. Math. Anal. Appl. 407 (2013), 350–358.
- [14] G. Birkhoff, *Orthogonality in linear metric spaces*, Duke Math. J. 1 (1935), 169–172.
- [15] M. S. Borowsky, *Characterizations of best complex Chebyshev approximate solutions of $Av = b$* , SIAM J. Numer. Anal. 13 (1976), 324–336.
- [16] R. H. Bouldin, *Positive approximants*, Trans. Amer. Math. Soc. 177 (1973), 391–403.

- [17] R. H. Bouldin, *The convex structure of the set of positive approximants for a given operator*, Acta Sci. Math. (Szeged) 37 (1975), 177–190.
- [18] M. Brannigan, *The strict Chebyshev solution of overdetermined systems of equations with rank deficient matrix*, Numer. Math. 40 (1982), 307–318.
- [19] J. Chmieliński, *On an ε -Birkhoff orthogonality*, JIPAM. J. Inequal. Pure Appl. Math. 6 (2005), Art. 79.
- [20] J. Chmieliński, *Remarks on orthogonality preserving mappings in normed spaces and some stability problems*, Banach J. Math. Anal. 1 (2007), 117–124.
- [21] Ch. Chorianopoulos, S. Karanasios, P. Psarrakos, *A definition of numerical range of rectangular matrices*, Linear Multilinear Algebra 57 (2009), 459–475.
- [22] Ch. K. Chui, P. W. Smith, J. D. Ward, *Approximation with restricted spectra*, Math. Z. 144 (1975), 289–297.
- [23] C. Davis, *An extremal problem for extensions of a sesquilinear form*, Linear Algebra Appl. 13 (1976), 91–102.
- [24] J. Descloux, *Approximation in L^p and Chebyshev approximation*, J. Soc. Indust. Appl. Math. 11 (1963), 1017–1026.
- [25] C. S. Duris, V. P. Sreedharan, *Chebyshev and l^1 -solutions of linear equations using least squares solutions*, SIAM J. Numer. Anal. 5 (1968), 491–505.
- [26] C. S. Duris, M. G. Temple, *A finite step algorithm for determining the “strict” Chebyshev solution to $Ax = b$* , SIAM J. Numer. Anal. 10 (1973), 690–699.
- [27] C. Eckart, G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika 1 (1936), 211–218.
- [28] A. Egger, R. Huotari, *The Pólya algorithm on convex sets*, J. Approx. Theory 56 (1989), 212–216.
- [29] A. Egger, R. Huotari, *Rate of convergence of the discrete Pólya algorithm*, J. Approx. Theory 60 (1990), 24–30.
- [30] A. Egger, R. Huotari, *Polya properties in R^n* , Numer. Funct. Anal. Optim. 13 (1992), 123–128.
- [31] K. Fan, A. J. Hoffman, *Some metric inequalities in the space of matrices*, Proc. Amer. Math. Soc. 6 (1955), 111–116.
- [32] J. Fernández-Ochoa, J. Martínez-Moreno, J. M. Quesada, *Asymptotic expression of the linear discrete best l_p -approximation*, J. Approx. Theory 140 (2006), 147–153.
- [33] J. Fischer, *The convergence of the best discrete linear L_p approximation as $p \rightarrow 1$* , J. Approx. Theory 39 (1983), 374–385.
- [34] S. Friedland, *On matrix approximation*, Proc. Amer. Math. Soc. 51 (1975), 41–43.
- [35] A. Greenbaum, L. Gurvits, *Max-min properties of matrix factor norms*, SIAM J. Sci. Comput. 15 (1994), 348–358.
- [36] A. Greenbaum, N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput. 15 (1994), 359–368.
- [37] D. Gries, J. Stoer, *Some results on fields of values of matrices*, SIAM J. Numer. Anal. 4 (1967), 283–300.
- [38] P. Grover, *Orthogonality to matrix subspaces and a distance formula*, Linear Algebra Appl. 445 (2014), 280–288.
- [39] P. R. Halmos, *Positive approximants of operators*, Indiana Univ. Math. J. 21 (1971), 951–960.
- [40] T. L. Hayden, J. Wells, *Approximation by matrices positive semidefinite on a subspace*, Linear Algebra Appl. 109 (1988), 115–130.
- [41] N. J. Higham, *Computing a nearest symmetric positive semidefinite matrix*, Linear Algebra Appl. 103 (1988), 103–118.

- [42] N. J. Higham, *Matrix nearness problems and applications*, in: Applications of Matrix Theory (Bradford, 1988), Inst. Math. Appl. Conf. Ser. New Ser. 22, Oxford Univ. Press, Oxford 1989, 1–27.
- [43] R. A. Horn, Ch. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge 1990.
- [44] R. A. Horn, Ch. R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge 1991.
- [45] R. Huotari, D. Legg, D. Townsend, *The Pólya algorithm on cylindrical sets*, J. Approx. Theory 53 (1988), 335–349.
- [46] R. Huotari, W. Li, *Continuity of metric projection, Pólya algorithm, strict best approximation, and tubularity of convex sets*, J. Math. Anal. Appl. 182 (1994), 834–856.
- [47] R. Huotari, W. Li, *Continuities of metric projection and geometric consequences*, J. Approx. Theory 90 (1997), 319–339.
- [48] R. Huotari, M. Marano, J. Navas, J. M. Quesada, *Rate of convergence of the Pólya algorithm from polyhedral sets*, J. Approx. Theory 135 (2005), 105–113.
- [49] R. C. James, *Orthogonality and linear functionals in normed linear space*, Trans. Amer. Math. Soc. 61 (1947), 265–292.
- [50] W. M. Kahan, *Numerical linear algebra*, Canad. Math. Bull. 9 (1966), 757–801.
- [51] R. Khalil, P. Maher, *Spectral approximation in $L(H)$* , Numer. Funct. Anal. Optim. 21 (2000), 693–713.
- [52] F. Kittaneh, *On zero-trace matrices*, Linear Algebra Appl. 151 (1991), 119–124.
- [53] P. Kosmol, D. Müller-Wichards, *Optimization in Function Space: with Stability Considerations in Orlicz Spaces*, De Gruyter Ser. Nonlinear Anal. Appl. 13, Walter de Gruyter, Berlin 2011.
- [54] P. Lancaster, M. Tismenetsky, *The Theory of Matrices*, second ed., Academic Press, Orlando 1985.
- [55] D. Landers, L. Rogge, *Natural choice of L_1 -approximants*, J. Approx. Theory 33 (1981), 268–280.
- [56] B. Laszkiewicz, K. Ziętak, *Approximation of matrices and a family of Gander methods for polar decomposition*, BIT 46 (2006), 345–366.
- [57] B. Laszkiewicz, K. Ziętak, *Approximation by matrices with restricted spectra*, Linear Algebra Appl. 428 (2008), 1031–1040.
- [58] K. K. Lau, W. O. J. Riha, *Characterization of best approximation in normed linear spaces of matrices by elements of finite-dimensional linear subspaces*, Linear Algebra Appl. 35 (1981), 109–120.
- [59] K. K. Lau, K. Ziętak, *A note on linear approximations of matrices*, Appl. Math. (Warsaw) 21 (1991), 257–263.
- [60] D. A. Legg, J. D. Ward, *A canonical trace class approximant*, Proc. Amer. Math. Soc. 93 (1985), 653–656.
- [61] A. S. Lewis, *The convex analysis of unitarily invariant functions*, J. Convex Anal. 2 (1995), 173–183.
- [62] A. S. Lewis, *Convex analysis on the Hermitian matrices*, SIAM J. Optim. 6 (1996), 164–177.
- [63] A. S. Lewis, M. L. Overton, *Eigenvalue optimization*, in: Acta Numerica 1996, Acta Numer. 5, Cambridge Univ. Press, Cambridge 1996, 149–190.
- [64] C.-K. Li, H. Schneider, *Orthogonality of matrices*, Linear Algebra Appl. 347 (2002), 115–122.
- [65] C.-K. Li, N.-K. Tsing, *On the unitarily invariant norms and some related results*, Linear Multilinear Algebra 20 (1987), 107–119.

- [66] J. Liesen, P. Tichy, *On best approximations of polynomials in matrices in the matrix 2-norm*, SIAM J. Matrix Anal. Appl. 31 (2009), 853–863.
- [67] J. Liesen, P. Tichy, *Max-min and min-max approximation problems for normal matrices revisited*, Electron. Trans. Numer. Anal. 41 (2014), 159–166.
- [68] M. Marano, *Strict approximation on closed convex sets*, Approx. Theory Appl. 6 (1990), 99–109.
- [69] M. Marano, J. N. Ureña, *The linear discrete Pólya algorithm*, Appl. Math. Lett. 8 (1995), no. 6, 25–28.
- [70] J. K. Merikoski, G. De Olivera, *On k -major and k -minor antinorms*, Linear Algebra Appl. 176 (1992), 197–209.
- [71] J. Miao, A. Ben-Israel, *On l_p -approximate solution of linear equations*, Linear Algebra Appl. 199 (1994), 305–327.
- [72] J. von Neumann, *Some matrix inequalities and metrization of metric spaces*, Tomsk Univ. Rev. 1 (1937), 286–300.
- [73] N. Nirschl, H. Schneider, *The Bauer fields of values of a matrix*, Numer. Math. 6 (1964), 355–365.
- [74] G. Pólya, *Sur un algorithm toujours convergent pour obtenir les polynomes de meilleure approximation de Tchebycheff*, C. R. Acad. Sci. Paris 157 (1913), 840–843.
- [75] J. M. Quesada, J. Martínez-Moreno, J. Navas, *Asymptotic behaviour of best l_p -approximations from affine subspaces*, J. Approx. Theory 118 (2002), 275–289.
- [76] J. M. Quesada, J. Martínez-Moreno, J. Navas, J. Fernández-Ochoa, *Rate of convergence of the linear-discrete Pólya 1-algorithm*, J. Approx. Theory 118 (2002), 316–329.
- [77] J. M. Quesada, J. Fernández-Ochoa, J. Martínez-Moreno, J. Bustamante, *The Pólya algorithm in sequence spaces*, J. Approx. Theory 135 (2005), 245–257.
- [78] J. R. Rice, *Tchebycheff approximation in a compact metric space*, Bull. Amer. Math. Soc. 68 (1962), 405–410.
- [79] R. T. Rockafellar, *Convex Analysis*, Princeton Math. Ser. 28, Princeton Univ. Press, Princeton, N.J. 1970.
- [80] D. D. Rogers, J. D. Ward, *C_p -minimal positive approximants*, Acta Sci. Math. (Szeged) 43 (1981), 109–115.
- [81] A. Ruhe, *Closest normal matrix finally found*, BIT 27 (1987), 585–598.
- [82] I. A. Rus, *Un principe du maximum pour les solutions d'un système fortement elliptique*, Glasnik Mat. Ser. III 4 (24) (1969), 75–78.
- [83] E. M. de Sá, *Exposed faces and duality for symmetric and unitarily invariant norms*, Linear Algebra Appl. 197/198 (1994), 429–450.
- [84] A. Shapiro, *Extremal problems on the set of nonnegative definite matrices*, Linear Algebra Appl. 67 (1985), 7–18.
- [85] I. Singer, *Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces*, Grundlehren Math. Wiss. 171, Springer, Berlin 1970.
- [86] W. So, *Facial structures of Schatten p -norms*, Linear Multilinear Algebra 27 (1990), 207–212.
- [87] V. P. Sreedharan, *Solutions of overdetermined linear equations which minimize error in an abstract norm*, Numer. Math. 13 (1969), 146–151.
- [88] V. P. Sreedharan, *Least-squares algorithms for finding solutions of overdetermined linear equations which minimize error in a smooth strictly norm*, J. Approximation Theory 8 (1973), 46–61.
- [89] G. W. Stewart, *Introduction to Matrix Computations*, Computer Sci. Appl. Math., Academic Press, New York 1973.

- [90] G. W. Stewart, J. G. Sun, *Matrix Perturbation Theory*, Computer Sci. Sci. Comput., Academic Press, San Diego 1990.
- [91] T. J. Suffridge, T. L. Hayden, *Approximation by a Hermitian positive semidefinite Toeplitz matrix*, SIAM J. Matrix Anal. Appl. 14 (1993), 721–734.
- [92] J.-P. Thiran, S. Thiry, *Strict Chebyshev approximation for general systems of linear equations*, Numer. Math. 51 (1987), 701–725.
- [93] C. Van Loan, *How near is a stable matrix to an unstable matrix?*, in: Linear Algebra and its Role in Systems Theory (Brunswick, Maine, 1984), Contemp. Math. 47, Amer. Math. Soc., Providence, RI 1985, 465–478.
- [94] G. A. Watson, *Approximation Theory and Numerical Methods*, Wiley, Chichester 1980.
- [95] G. A. Watson, *Characterization of the subdifferential of some matrix norms*, Linear Algebra Appl. 170 (1993), 33–45.
- [96] G. A. Watson, *On matrix approximation problems with Ky Fan k norms*, Numer. Algorithms 5 (1993), 263–272.
- [97] G. A. Watson, *Linear best approximation using a class of k -major l_p norms*, Numer. Algorithms 8 (1994), 135–146.
- [98] G. A. Watson, *The solution of the matrix approximation problems in c_p norms*, manuscript.
- [99] H. J. Woerdeman, *Superoptimal completions of triangular matrices*, Integral Equations Operator Theory 20 (1994), 491–501.
- [100] N. J. Young, *The Nevanlinna–Pick problem for matrix-valued functions*, J. Operator Theory 15 (1986), 239–265.
- [101] K. Ziętak, *Properties of the approximate generalized inverses of a class of matrices*, Linear Algebra Appl. 92 (1987), 259–272.
- [102] K. Ziętak, *On the characterization of the extremal points of the unit sphere of matrices*, Linear Algebra Appl. 106 (1988), 57–75.
- [103] K. Ziętak, *Strict spectral approximation of matrices*, Report N-218, Institute of Computer Science, Univ. of Wrocław, Wrocław, June 1992.
- [104] K. Ziętak, *Properties of linear approximations of matrices in the spectral norm*, Linear Algebra Appl. 183 (1993), 41–60.
- [105] K. Ziętak, *Subdifferentials, faces and dual matrices*, Linear Algebra Appl. 185 (1993), 125–141.
- [106] K. Ziętak, *Strict approximation of matrices*, SIAM J. Matrix Anal. Appl. 16 (1995), 232–234.
- [107] K. Ziętak, *On approximation problems with zero trace matrices*, Linear Algebra Appl. 247 (1996), 169–183.
- [108] K. Ziętak, *Strict spectral approximation of a matrix and some related problems*, Appl. Math. (Warsaw) 24 (1997), 267–280.
- [109] K. Ziętak, *Orthant-monotonic norms and overdetermined linear systems*, J. Approx. Theory 88 (1997), 209–227.
- [110] K. Ziętak, *Properties of the strict spectral approximation of a complex matrix*, Opuscula Math. 22 (2002), 41–46.
- [111] K. Ziętak, *On proof of convergence of the Pólya algorithm for an overdetermined system of complex linear equations*, in preparation.