A weak-type (∞, ∞) inequality for the triangular projection

Tomasz Gałązka[a], Adam Osękowski[a,*], Yahui Zuo[b]

[a] *Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland*

[b] *School of Mathematics and Statistics, Central South University, Changsha 410085, People's Republic of China*

Abstract

The paper contains the proof of the weak-type (∞, ∞) estimate for the triangular projection on Schatten classes. The argument rests on a number of trace inequalities for matrices, which are of independent interest and connections.

Keywords: Triangular projection, matrix, singular value, trace inequality.
2010 MSC: 15A18, 15A42, 47B06.

1. Introduction

The purpose of this paper is to establish a new weak-type estimate for the so-called triangular projection, an important object arising in the operator theory. Assume that $\mathcal{H}$ is a separable complex Hilbert space and let $B(\mathcal{H})$ denote the algebra of all bounded linear operators on $\mathcal{H}$. For any compact operator A, we define its modulus by $|A| = (A^*A)^{1/2}$ and let $s_1(A)$, $s_2(A)$, $\ldots$ be the singular values of A, i.e., the eigenvalues of $|A|$ in decreasing order and repeated according to multiplicity. A compact operator A belongs to the Schatten p-class C_p (here $1 \leq p < \infty$), if

$$\|A\|_{C_p} = \left(\sum_{n \geq 1} s_n(A)^p \right)^{1/p} < \infty.$$

The space C_p, equipped with the Schatten p-th norm $\|\cdot\|_p$, is a Banach space; in the special case $p = 1$, it is sometimes referred to as the trace class, while C_2 is called the Hilbert-Schmidt class. We let C_∞ be the ideal of compact operators and endow it with the usual operator norm $\|\cdot\|_{C_\infty} = \|\cdot\|$. If the Hilbert space $\mathcal{H}$ is finite-dimensional, i.e., $\mathcal{H} = \mathbb{C}^N$ for some N, we will denote the corresponding

*Corresponding author

URL: T.Galazka@mimuw.edu.pl (Tomasz Gałązka), A.Osekowski@mimuw.edu.pl (Adam Osękowski), zuoyahui@csu.edu.cn (Yahui Zuo)

Schatten classes by C_p^N, $1 \le p \le \infty$. Given a self-adjoint operator $A \in C_p$, let $A = \int_{-\infty}^{\infty} \lambda dE_\lambda$ stand for its spectral decomposition. This allows us to define $f(A)$ for any Borel function f on $\mathbb{R}$, by $f(A) = \int_{-\infty}^{\infty} f(\lambda) \mathrm{d}E_\lambda$.

Suppose further that $\xi_1, \xi_2, \ldots$ is an orthonormal basis of $\mathcal{H}$. Given an operator $A \in C_1$, we define its trace by $\mathrm{Tr}(A) = \sum_{n=1}^{\infty} a_{n,n}$, where $(a_{i,j})_{i,j}$ is the infinite matrix representation of A, with the complex coefficients $a_{i,j} = \langle A\xi_i, \xi_j \rangle$ for all i, j. This definition of the trace is independent of a particular choice of the basis. It is not difficult to check that $\|A\|_{C_p} = \big(\mathrm{Tr}(|A|^p)\big)^{1/p}$ for any finite p. See [11, 13] for the more systematic presentation of the subject.

The central role in the paper will be played by a certain important operator on matrices, the so-called lower triangular projection. If $\xi = (\xi_n)_{n \ge 1}$ is a given orthonormal basis of $\mathcal{H}$ and A is an element of $B(\mathcal{H})$ with the matrix representation $(a_{i,j})_{i,j}$, then we define $T(A) = (T(A)_{i,j})_{i,j}$ by setting

$$(T(A))_{i,j} = \begin{cases} a_{i,j} & \text{if } i \ge j, \\ 0 & \text{otherwise.} \end{cases}$$

This operator plays an important role in analysis and operator theory. First, it serves as a convenient tool for providing efficient lower bounds for a number of estimates in the theory of noncommutative martingales and Schur multipliers: see e.g. [7, 8, 10]. It applies to the theory of (p,q)-summing operators and properties of unconditional bases of $B(\mathcal{H})$ (see Section 4 in [9]) and gives a characterization of optimal constants in Menshov-Rademacher inequality [3]. The triangular projection can be also regarded as the matrix analogue of the analytic projection P_a acting on trigonometric series by the formula

$$P_a\left(\sum_{n=-\infty}^{\infty} a_n e^{int}\right) = \sum_{n=0}^{\infty} a_n e^{int}.$$

See [6, 12] for more on this subject and consult the references therein.

Our motivation comes from the question about the boundedness properties of the triangular projection on Schatten classes (see Chapters II and III in [6] and Section 1 in [9] for a detailed presentation). It is well-known that T is bounded on C_p when $1 < p < \infty$. More specifically, we have $\|T\|_{C_p \to C_p} = O(p)$ as $p \to \infty$ and $\|T\|_{C_p \to C_p} = O((p-1)^{-1})$ as $p \to 1$. Since T is self-adjoint on C_2, we also have $\|T\|_{C_p \to C_p} = \|T\|_{C_{p'} \to C_{p'}}$ for all $1 < p < \infty$, where $p' = p/(p-1)$ is the harmonic conjugate to p. On the other hand, the triangular projection is not bounded on C_1 or C_∞, actually, we have the more precise result: $\|T\|_{C_1^N \to C_1^N} = \|T\|_{C_\infty^N \to C_\infty^N} = O(\log N)$ as $N \to \infty$. However, as a substitute in the case $p = 1$, T satisfies the weak-type bound

$$\|T(A)\|_{C_{1,\infty}} \le K\|A\|_{C_1}, \tag{1}$$

where $\|A\|_{C_{1,\infty}} = \sup_{n>0} \big(n s_n(A)\big)$ is the standard quasinorm in the weak L^1 (cf. [6, 12]). There is a natural question about the analogue of this weak-type bound in the case $p = \infty$. This in particular leads to the problem of defining an appropriate version of the weak Schatten norm (or rather functional) $\|\cdot\|_{C_{\infty,\infty}}$. Let us briefly discuss the classical, measure-theoretic approach to this problem, due to Bennett, DeVore and Sharpley [1]. Suppose that $(X, \mathcal{F}, \mu)$ is a measure space and let $f : X \to \mathbb{C}$ be a measurable function. The decreasing rearrangement $f^* : (0, \infty) \to [0, \infty)$ is given by

$$f^*(t) = \inf \Big\{ \lambda \geq 0 : \mu(\{x \in X : |f(x)| > \lambda\}) \leq t \Big\}$$

and the maximal function of f is defined as $f^{**}(t) = \frac{1}{t} \int_0^t f^*(s) \mathrm{d}s$ for $t > 0$. Equivalently,

$$f^{**}(t) = \frac{1}{t} \sup \left\{ \int_E |f| \mathrm{d}\mu : E \in \mathcal{F}, \mu(E) \leq t \right\}$$

for all t. Following Bennett, DeVore and Sharpley, we set

$$\|f\|_{weak(L^\infty)} = \sup_{t>0} (f^{**}(t) - f^*(t))$$

and define the weak-L^∞ as the collection of all (equivalent classes of) functions f with $\|f\|_{weak(L^\infty)} < \infty$. The main motivation behind the introduction of this space comes from interpolation theory. Obviously, weak-L^∞ contains $L^\infty(X, \mathcal{F}, \mu)$. The first important property is that if an operator S is bounded from L^1 to $L^{1,\infty}$ and from L^∞ to $weak(L^\infty)$, then it can be extended to a bounded operator on all L^p spaces, $1 < p < \infty$. In other words, thanks to the above definition, we have a substitute of Marcinkiewicz interpolation theorem for operators which are unbounded on L^∞. There is a further complementary explanation. Namely, the Peetre K-functional for the pair (L^1, L^∞) (consult [2, p.184]) can be expressed explicitly in the form

$$K(f, t; L^1, L^\infty) = \int_0^t f^*(s) \mathrm{d}s = t f^{**}(t), \qquad t > 0,$$

and the weak-L^1 quasinorm is given by the formula $\|f\|_{1,\infty} = \sup_{t>0} t f^*(t) = \sup_{t>0} t \frac{d}{dt} K(f, t; L^1, L^\infty)$. It seems plausible to define the weak-L^∞ functional by interchanging the roles of L^1 and L^∞ in the latter expression. Because of the identity $K(f, t; L^\infty, L^1) = tK(f, t^{-1}; L^1, L^\infty)$, we compute that

$$\sup_{t>0} t \frac{d}{dt} K(f, t; L^\infty, L^1) = \sup_{t>0} \big(f^{**}(t) - f^*(t) \big) = \|f\|_{weak(L^\infty)}.$$

It should be emphasized that there are some drawbacks: in general, the weak-L^∞ is not a linear space and the functional $\|\cdot\|_{weak(L^\infty)}$ is not a norm, it is not even a quasinorm. Nevertheless, these objects are of importance in the study of many classical operators, e.g. the maximal functions of singular integrals. See [1] for the more detailed discussion of this topic.

We easily extend the above concept of weak-L^∞ to the operator context. Define the space $C_{\infty,\infty}$ as the collection of all compact operators A for which

$$\|A\|_{C_{\infty,\infty}} = \sup_{n \geq 1} \left(\frac{1}{n} \sum_{k=1}^{n} s_k(A) - s_n(A) \right) < \infty. \tag{2}$$

So, roughly speaking, $A \in C_{\infty,\infty}$ if its singular values decrease in a moderate manner.

Let us discuss some basic properties of the space $C_{\infty,\infty}$ and the functional $\|\cdot\|_{C_{\infty,\infty}}$. It is clear that the spaces $C_{\infty,\infty}$ and C_∞ are equal as sets, furthermore, it follows directly from (2) that $\|A\|_{C_{\infty,\infty}} \leq \|A\|_{C_\infty}$ for all operators A. On the other hand, the inclusion $C_{\infty,\infty} \hookrightarrow C_\infty$ is not bounded: there is no finite universal constant K such $\|A\|_{C_\infty} \leq K\|A\|_{C_{\infty,\infty}}$ for all A. To see the latter, consider the finite-rank operators with singular values equal to $\ln N$, $\ln(N-1)$, $\dots$, $\ln 2$, 0, 0, $\dots$, and let $N \to \infty$. It should also be emphasized that the functional $\|\cdot\|_{C_{\infty,\infty}}$ is not even a quasinorm: there is no universal constant K such that $\|A+B\|_{C_{\infty,\infty}} \leq K(\|A\|_{C_{\infty,\infty}} + \|B\|_{C_{\infty,\infty}})$. The example is similar to that above: fix a positive integer N and consider the $N \times N$ diagonal matrices with the entries $\ln 1$, $\ln 2$, $\dots$, $\ln N$ and $-\ln 1$, $-\ln 2$, $\dots$, $-\ln(N-1)$, $\ln N$ on the main diagonals, respectively. Then $\|A\|_{C_{\infty,\infty}} = \|B\|_{C_{\infty,\infty}} = 1$, while $\|A+B\|_{C_{\infty,\infty}} = 2\ln N$ can be arbitrarily large. The same calculation shows that there is no nontrivial upper bound for the ratio $\|A_\pm\|_{C_{\infty,\infty}}/\|A\|_{C_{\infty,\infty}}$, where A is self-adjoint and $A_\pm$ are the negative/positive parts of A, defined spectrally.

One of the main goals of the paper is to show that the triangular projection is bounded as an operator from C_∞ to $C_{\infty,\infty}$, under the additional compactness assumption.

Theorem 1.1. *For any N and any operator $A \in C_\infty^N$ we have the estimate*

$$\|T(A)\|_{C_{\infty,\infty}^N} \leq 32\|A\|_{C_\infty^N}. \tag{3}$$

In addition, if $A \in C_\infty$ and $T(A)$ is compact, then

$$\|T(A)\|_{C_{\infty,\infty}} \leq 32\|A\|_{C_\infty}. \tag{4}$$

As we will see, the main difficulty lies in proving (3); the passage to the infinite-dimensional case involves standard limiting arguments only. This observation reduces the problem to the analysis of appropriate trace inequalities for finite-dimensional matrices, which are of their own interest.

As an exemplary application of the above theorem, consider the Toeplitz matrices

$$A_n = \begin{bmatrix} 0 & -1 & -\frac{1}{2} & -\frac{1}{3} & \dots & -\frac{1}{n} \\ 1 & 0 & -1 & -\frac{1}{2} & \dots & -\frac{1}{n-1} \\ \frac{1}{2} & 1 & 0 & -1 & \dots & -\frac{1}{n-2} \\ \frac{1}{3} & \frac{1}{2} & 1 & 0 & \dots & -\frac{1}{n-3} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{1}{n} & \frac{1}{n-1} & \frac{1}{n-2} & \frac{1}{n-3} & \dots & 0 \end{bmatrix}, \qquad n = 1, 2, \dots.$$

which are closely related to the so-called Hilbert matrix, playing an important role in harmonic analysis (see e.g. Chapter II in [6] or Section 1 in [9]). It is well-known that $\sup_{n\geq 1}\|A_n\| = \pi$ and $\sup_{n\geq 1}\|T(A_n)\| = \infty$: a convenient reference is [4, p. 306]. Theorem 1.1 implies that $\|T(A_n)\|_{C_{\infty,\infty}} \leq 32\pi$, $n = 1, 2, \dots$.

The second application is related to the interpolation results obtained in [1]. Let S denote the Calderón maximal operator which acts on nonnegative functions on $(0, \infty)$ by

$$S(f)(t) = \frac{1}{t}\int_0^t f(s)\mathrm{d}s + \int_t^\infty f(s)\frac{\mathrm{d}s}{s}, \qquad t > 0.$$

Theorem 2.2 of [1], combined with (1) and Theorem 1.1 above yields the following fact. With a slight abuse of notation, the symbol A^{**} stands for the maximal function of A, that is, $A^{**}(t) = \frac{1}{t}\sum_{1\leq k\leq t} s_k(A)$. This should not lead to any confusion: the notation will be used just once, in the statement below, and it is consistent with the above measure-theoretic context.

Theorem 1.2. *Let A be a compact operator whose projection $T(A)$ is also compact. Then for any $t > 0$ we have*

$$T(A)^{**}(t) \leq CS(A^{**})(t),$$

where C is a universal constant.

Finally, we would like to comment on the compactness assumption appearing in the second part of Theorem 1.1. It plays the role of an appropriate localization which is very natural if we compare the result to similar statements in the functional context. For example, for the analytic projection P_a on the real line (cf. Chapter III in [6]), one proves easily that the estimate

$$\|P_a f\|_{weak(L^\infty(\mathbb{R}))} \leq K\|f\|_{L^\infty(\mathbb{R})} \tag{5}$$

does not hold in general with any finite constant K. To fix the inequality, one can use one of the following options: (i) consider its periodic version; (ii) investigate (5) for compactly supported functions f (which leads to the so-called restricted weak-type estimates [1]); (iii) localize the norm on the left, replacing it by $\|P_a f\|_{weak(L^\infty(0,1))}$. All these modifications, on the level of matrices, correspond to imposing the compactness assumption on the function on the left. Furthermore, note that (i) suggests a simple sufficient condition on A which guarantees the compactness of $T(A)$: this is the case if A is of finite rank.

The rest of the paper is organized as follows. The next section contains some preliminary material: we introduce some basic matrices and study their properties. Section 3 is devoted to the proof of Theorem 1.1.

2. Some special matrices and their properties

We start with a simple fact which follows directly from the definition of matrix multiplication. Here and below, the symbol '0' will denote a square or rectangular matrix (the dimension will be always specified or clear from the context), with all entries equal to zero.

Lemma 2.1. *Suppose that A, B are square matrices of the block form*

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & B_{12} \\ B_{21} & 0 \end{bmatrix},$$

such that the corresponding blocks in A and B have the same dimension. Then

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA) = 0. \tag{6}$$

From now on, we assume that $A = (a_{i,j})_{1 \le i,j \le N}$ is a fixed matrix of dimension $N \times N$. For a given $1 \le k \le N$, define

$$C_k = \begin{bmatrix} 0 & 0 & \dots & 0 & a_{1,k} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & a_{2,k} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & a_{k-1,k} & 0 & \dots & 0 \\ a_{k,1} & a_{k,2} & \dots & a_{k,k-1} & a_{k,k} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

and

$$D_k = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ a_{k,1} & a_{k,2} & \dots & a_{k,k-1} & a_{k,k} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

The following properties of these matrices will be needed later.

Lemma 2.2. *We have $A = \sum_{k=1}^N C_k$, $\sum_{k=1}^N |D_k|^2 = |T(A)|^2$ and $|D_k|^2 \le |C_k|^2$.*

Proof. The first identity is clear. To check the second equality, observe that $(|D_k|^2)_{i,j}$, i.e., the entry of $|D_k|^2$ which lies in the i-th row and j-th column, is equal to $\overline{a_{k,i}} a_{k,j}$ if both i and j do not exceed k, and zero otherwise. Consequently, for any $1 \le i, j \le N$,

$$\left(\sum_{k=1}^N |D_k|^2 \right)_{i,j} = \sum_{k \ge \max\{i,j\}} \overline{a_{k,i}} a_{k,j} = (|T(A)|^2)_{i,j}.$$

To establish the estimate $|D_k|^2 \le |C_k|^2$, set $B_k = C_k - D_k$ and observe that

$$|C_k|^2 = (B_k^* + D_k^*)(B_k + D_k) = |B_k|^2 + B_k^* D_k + D_k^* B_k + |D_k|^2 = |B_k|^2 + |D_k|^2.$$

This completes the proof. □

In our considerations below, it will also be convenient for us to work with a larger algebra of matrices of dimension $N(N+1) \times N(N+1)$. In what follows, the symbol I will stand for the identity matrix in this larger dimension. Furthermore, for any $1 \leq k \leq N$, let S_k be given, in the block form, by

$$S_k = \begin{bmatrix} 0 & D_1^* & D_2^* & \dots & D_k^* & 0 & \dots & 0 \\ D_1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ D_2 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ D_k & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Here each block is a matrix of dimension $N \times N$, there are $N+1$ block-rows and $N+1$ block-columns. The relation between S_N and the triangular projection $T(A)$ is studied in a lemma below.

Lemma 2.3. *For any function $f : \mathbb{R} \to \mathbb{R}$ we have the identity* $\mathrm{Tr}(f(|S_N|)) = 2\,\mathrm{Tr}(f(|T(A)|))$.

Proof. We have

$$|S_N|^2 = \begin{bmatrix} \sum_{k=1}^N |D_k|^2 & 0 \\ 0 & \left[D_i D_j^*\right]_{i,j} \end{bmatrix} = \begin{bmatrix} |T(A)|^2 & 0 \\ 0 & \left[D_i D_j^*\right]_{i,j} \end{bmatrix}$$

and hence for any positive integer m,

$$|S_N|^{2m} = \begin{bmatrix} |T(A)|^{2m} & 0 \\ 0 & \left(\left[D_i D_j^*\right]_{i,j}\right)^m \end{bmatrix}.$$

This implies

$$\mathrm{Tr}(|S_N|^{2m}) = \mathrm{Tr}(|T(A)|^{2m}) + \mathrm{Tr}\left(\left(\left[D_i D_j^*\right]_{i,j}\right)^m\right)$$

and by the tracial property, we have

$$\begin{aligned} \mathrm{Tr}\left(\left(\left[D_i D_j^*\right]_{i,j}\right)^m\right) &= \mathrm{Tr}\left(\left((D_1, D_2, \dots, D_N)^T (D_1^*, D_2^*, \dots, D_N^*)\right)^m\right) \\ &= \mathrm{Tr}\left(\left((D_1^*, D_2^*, \dots, D_N^*)(D_1, D_2, \dots, D_N)^T\right)^m\right) \\ &= \mathrm{Tr}\left(\left(\sum_{k=1}^N |D_k|^2\right)^m\right) = \mathrm{Tr}(|T(A)|^{2m}). \end{aligned}$$

Hence, for any polynomial W on $\mathbb{R}$ we have $\mathrm{Tr}(W(S_N^2)) = 2\,\mathrm{Tr}(W(|T(A)|^2))$ and the claim follows, since the spectra of $|S_N|$ and $|T(A)|$ are finite. $\square$

The final step of this section is the introduction of a certain class $(e_k)_{0\le k\le N}$ of projections on $\mathbb{C}^{N(N+1)}$, motivated by the construction due to Cuculescu [5]. Assume that λ is a fixed positive parameter. Let $e_0 = I$ be the identity matrix (of dimension $N(N+1)\times N(N+1)$) and define, inductively,

$$e_k = e_{k-1}\chi_{(-\infty,\lambda)}(e_{k-1}S_k e_{k-1}) \qquad k = 1,\, 2,\, \ldots,\, N. \tag{7}$$

To gain some intuition about these operators, observe that e_k is a projection onto the subspace on which all the matrices S_1, S_2, ..., S_k are smaller than λI. In other words, e_k can be regarded as a formal version of the projection $\chi_{(-\infty,\lambda)}(\max_{1\le j\le k} S_j)$ (of course, the latter symbol makes no sense: the maximum of matrices is not well-defined). In addition, note that due to the appearance of e_{k-1} in front of the right-hand side of (7), the sequence $(e_k)_{0\le k\le N}$ is nonincreasing.

Further properties of these projections are studied in a lemma below.

Lemma 2.4. *(i) For any k the projection e_k has the block form* $\begin{bmatrix} b & 0 \\ 0 & I_{(N-k)N}\end{bmatrix}$, *where b is a matrix of dimension $(k+1)N\times(k+1)N$ and $I_{(N-k)N}$ is the identity matrix. Furthermore, the upper-left block of b of dimension $N\times N$ is of the form* $\begin{bmatrix} c & 0 \\ 0 & I_{N-k}\end{bmatrix}$.

(ii) For all $1\le j\le k\le N$ we have

$$e_k S_j e_k \le \lambda e_k \quad \text{and} \quad (e_{k-1}-e_k)S_k(e_{k-1}-e_k) \ge \lambda(e_{k-1}-e_k).$$

(iii) For all $1\le k\le j\le N$, we have $\operatorname{Tr}(e_k S_j e_k) = \operatorname{Tr}(e_k S_k e_k)$.

(iv) For all $1\le k\le N$, we have

$$\operatorname{Tr}((e_{k-1}-e_k)(S_N-S_k)e_k(S_k-\lambda I)) = \operatorname{Tr}((e_{k-1}-e_k)(S_k-\lambda I)e_k(S_N-S_k)) = 0.$$

(v) If $\|A\|\le 1$, then for all $1\le k\le N$, we have

$$\operatorname{Tr}((e_{k-1}-e_k)(S_N-S_k)^2) \le 4\operatorname{Tr}(e_{k-1}-e_k).$$

(vi) For any $\lambda\ge a>0$ we have $\operatorname{Tr}(I-e_N)\le a^{-1}\operatorname{Tr}\big((S_N-(\lambda-a)I)_+\big)$.

Proof. (i) We will only prove the first half, the second is established analogously. We proceed by induction. The claim is clear for $k=0$; assuming its validity for $k-1$, we see that all entries of $e_{k-1}S_k e_{k-1}$, which lie outside the upper-left corner of dimension $(k+1)N\times(k+1)N$, are equal to zero. This proves that e_k is of the desired form.

(ii) This follows directly from the very definition of e_k.

(iii) We rewrite the identity in the form $\operatorname{Tr}(e_k(S_j-S_k)e_k)=0$ or $\operatorname{Tr}(e_k(S_j-S_k))=0$, by the tracial property. The latter equality follows from Lemma 2.1 and part (i).

(iv) We proceed as in the proof of the previous part. By the tracial property, we have

$$\operatorname{Tr}((e_{k-1}-e_k)(S_N-S_k)e_k(S_k-\lambda I)) = \operatorname{Tr}\Big(e_k(S_k-\lambda I)(e_{k-1}-e_k)(S_N-S_k)\Big)$$

and it remains to note that the matrices $e_k(S_k - \lambda I)(e_{k-1} - e_k)$, $S_N - S_k$ have appropriate block structure as in Lemma 2.1. The second trace in the assertion is handled by passing to the adjoint matrix, showing that both traces are actually equal.

(v) Here the argument will be a little more involved. Observe that

$$\begin{aligned}\mathrm{Tr}((e_{k-1} - e_k)(S_N - S_k)^2) &= \sum_{k+1 \le i,j \le N} \mathrm{Tr}((e_{k-1} - e_k)(S_i - S_{i-1})(S_j - S_{j-1})) \\ &= \sum_{j=k+1}^{N} \mathrm{Tr}((e_{k-1} - e_k)(S_j - S_{j-1})^2),\end{aligned}$$

since for $i \neq j$ we have $\mathrm{Tr}((e_{k-1} - e_k)(S_i - S_{i-1})(S_j - S_{j-1})) = 0$ (which can be proved exactly in the same manner as parts (iii) and (iv) above). Next, by part (i), we see that all the entries of the projection $e_{k-1} - e_k$ which lie outside the upper-left corner of dimension $(k+1)N \times (k+1)N$, are equal to zero. Consequently, for $j \ge k+1$,

$$(e_{k-1} - e_k)(S_j - S_{j-1})^2 = (e_{k-1} - e_k) \begin{bmatrix} |D_j|^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Indeed, in the formula for $(S_j - S_{j-1})^2$ there would be an additional entry $|D_j^*|^2$ in the j-th row and j-th column, but it can be omitted due to the block form of $e_{k-1} - e_k$ discussed above. Summing over j and applying Lemma 2.2, we obtain

$$\begin{aligned}&\sum_{j=k+1}^{N} \mathrm{Tr}((e_{k-1} - e_k)(S_j - S_{j-1})^2) \\ &\quad = \mathrm{Tr}\left((e_{k-1} - e_k) \begin{bmatrix} \sum_{j=k+1}^{N} |D_j|^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} \right) \\ &\quad \le \mathrm{Tr}\left((e_{k-1} - e_k) \begin{bmatrix} \sum_{j=k+1}^{N} |C_j|^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} \right).\end{aligned}$$

Now we will prove the identity

$$\mathrm{Tr}\left((e_{k-1} - e_k) \begin{bmatrix} C_i^* C_j & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} \right) = 0, \qquad i,\, j \ge k+1,\, i \neq j. \quad (8)$$

Before we do this, let us see how it shows the claim. Combining the last two identities yields

$$\begin{aligned}
&\operatorname{Tr}((e_{k-1}-e_k)(S_N-S_k)^2) \\
&\leq \operatorname{Tr}\left((e_{k-1}-e_k)\begin{bmatrix} \sum_{i,\,j\geq k+1} C_i^* C_j & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}\right) \\
&= \operatorname{Tr}\left((e_{k-1}-e_k)\begin{bmatrix} |\sum_{j=k+1}^{N} C_j|^2 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}\right).
\end{aligned}$$

But $\sum_{j=k+1}^{N} C_j = A - P_k(A)$, where $P_k(A)$ is obtained from A by leaving the upper-left corner of dimension $k \times k$ unchanged, and changing all the remaining entries to zero. Since $\|P_k(A)\| \leq \|A\| \leq 1$, we get $\|A - P_k(A)\| \leq 2$ and the desired estimate follows.

It remains to show (8). Let us look at the product of the matrices under the trace. We may restrict ourselves to the upper-left corners of dimension $N \times N$, the other parts do not contribute to the trace. By the second half of part (i), the corresponding corner for the projection $e_{k-1} - e_k$ is of the form $\begin{bmatrix} c & 0 \\ 0 & 0 \end{bmatrix}$ for some matrix c of dimension $k \times k$; by the very definition of C_j, the corresponding block form of $C_i^* C_j$ is $\begin{bmatrix} 0 & \mathtt{d} \\ \mathtt{e} & \mathtt{f} \end{bmatrix}$. The product of these two matrices, and hence also the product under the trace in (8), has only zeros on the main diagonal. This gives the claim.

(vi) By the tracial property and parts (ii) and (iii) established above, we have

$$\begin{aligned}
\lambda \operatorname{Tr}(I - e_N) &= \lambda \sum_{k=1}^{N} \operatorname{Tr}(e_{k-1} - e_k) \\
&\leq \sum_{k=1}^{N} \operatorname{Tr}((e_{k-1} - e_k) S_k (e_{k-1} - e_k)) \\
&= \sum_{k=1}^{N} \left[\operatorname{Tr}(e_{k-1} S_k e_{k-1}) - \operatorname{Tr}(e_k S_k e_k) \right] \\
&= \sum_{k=1}^{N} \left[\operatorname{Tr}(e_{k-1} S_N e_{k-1}) - \operatorname{Tr}(e_k S_N e_k) \right] \\
&= \sum_{k=1}^{N} \left[\operatorname{Tr}(S_N) - \operatorname{Tr}(e_N S_N e_N) \right] = \operatorname{Tr}((I - e_N) S_N).
\end{aligned}$$

This implies

$$\begin{aligned}
& a\,\mathrm{Tr}(I-e_N)\\
&\leq \mathrm{Tr}\big((I-e_N)(S_N-(\lambda-a)I)\big)\\
&= \mathrm{Tr}\big((I-e_N)(S_N-(\lambda-a)I)_+\big)+\mathrm{Tr}\big((I-e_N)(S_N-(\lambda-a)I)_-\big)\\
&\leq \mathrm{Tr}\big((I-e_N)(S_N-(\lambda-a)I)_+\big)\leq \mathrm{Tr}\big((S_N-(\lambda-a)I)_+\big),
\end{aligned}$$

where the last two inequalities follow from the fact that the matrix $(S_N-(\lambda-a)I)_-$ is nonpositive, while $I-e_N$ is nonnegative. □

3. Proof of the weak-type estimate

Equipped with all the auxiliary objects, we turn our attention to the weak-type estimate (3). The proof will be split into several intermediate lemmas.

Lemma 3.1. *We have*

$$\mathrm{Tr}\Big((I-e_N)\,(S_N-\lambda I)^2\Big)\leq 2\sum_{k=1}^{N}\mathrm{Tr}\Big((e_{k-1}-e_k)\,(S_N-\lambda I)\,e_{k-1}\,(S_N-\lambda I)\Big). \tag{9}$$

Proof. The argument rests on an appropriate splitting and rearrangement of terms. First, we have $I=\sum_{k=1}^{N}(e_{k-1}-e_k)+e_N$, which implies

$$\begin{aligned}
&\mathrm{Tr}\Big((I-e_N)\,(S_N-\lambda I)^2\Big)\\
&\quad=\sum_{j=1}^{N}\sum_{k=1}^{N}\mathrm{Tr}\left((e_{k-1}-e_k)\,(S_N-\lambda I)\,(e_{j-1}-e_j)\,(S_N-\lambda I)\right)\\
&\qquad+\mathrm{Tr}\left((I-e_N)\,(S_N-\lambda I)\,e_N\,(S_N-\lambda I)\right).
\end{aligned}$$

By the tracial property, all the summands on the right are nonnegative. Consequently, we get

$$\begin{aligned}
&\mathrm{Tr}\Big((I-e_N)\,(S_N-\lambda I)^2\Big)\\
&\leq 2\sum_{k\leq j}\mathrm{Tr}\left((e_{k-1}-e_k)\,(S_N-\lambda I)\,(e_{j-1}-e_j)\,(S_N-\lambda I)\right)\\
&\quad+2\,\mathrm{Tr}\left((I-e_N)\,(S_N-\lambda I)\,e_N\,(S_N-\lambda I)\right)\\
&=2\sum_{k=1}^{N}\mathrm{Tr}\left((e_{k-1}-e_k)\,(S_N-\lambda I)\,(e_{k-1}-e_N)\,(S_N-\lambda I)\right)\\
&\quad+2\sum_{k=1}^{N}\mathrm{Tr}\left((e_{k-1}-e_k)\,(S_N-\lambda I)\,e_N\,(S_N-\lambda I)\right)\\
&=2\sum_{k=1}^{N}\mathrm{Tr}\left((e_{k-1}-e_k)\,(S_N-\lambda I)\,e_{k-1}\,(S_N-\lambda I)\right)
\end{aligned}$$

and we are done. $\square$

The above statement will allow us to obtain the following further bound for the trace $\operatorname{Tr}\Big((I-e_N)\,(S_N-\lambda I)^2\Big)$.

Lemma 3.2. *If $\|A\|\le 1$, then we have the estimate*

$$\operatorname{Tr}\Big((I-e_N)\,(S_N-\lambda I)^2\Big)\le 16\operatorname{Tr}(I-e_N). \tag{10}$$

Proof. We need to bound the summands on the right of (9) appropriately. We start with the observation that

$$\begin{aligned}
&\operatorname{Tr}\Big((e_{k-1}-e_k)(S_N-\lambda I)e_{k-1}(S_N-\lambda I)\Big)\\
&=\operatorname{Tr}\Big((e_{k-1}-e_k)(S_N-S_k+S_k-\lambda I)e_{k-1}(S_N-S_k+S_k-\lambda I)\Big)\\
&=\operatorname{Tr}\big((e_{k-1}-e_k)(S_N-S_k)e_{k-1}(S_N-S_k)\big)\\
&\quad+\operatorname{Tr}\big((e_{k-1}-e_k)(S_k-\lambda I)e_{k-1}(S_k-\lambda I)\big)\\
&\le \operatorname{Tr}\big((e_{k-1}-e_k)(S_N-S_k)^2\big)+\operatorname{Tr}\big((e_{k-1}-e_k)(S_k-\lambda I)e_{k-1}(S_k-\lambda I)\big)\\
&=: J_1+J_2.
\end{aligned}$$

Here the second passage follows by Lemma 2.4 (iv) and the inequality is due to the tracial property and the estimate $e_{k-1}\le I$:

$$\begin{aligned}
&\operatorname{Tr}\big((e_{k-1}-e_k)(S_N-S_k)e_{k-1}(S_N-S_k)\big)\\
&=\operatorname{Tr}\big((e_{k-1}-e_k)(S_N-S_k)e_{k-1}(S_N-S_k)(e_{k-1}-e_k)\big)\\
&\le \operatorname{Tr}\big((e_{k-1}-e_k)(S_N-S_k)^2(e_{k-1}-e_k)\big)=\operatorname{Tr}\big((e_{k-1}-e_k)(S_N-S_k)^2\big).
\end{aligned}$$

Now, observe that $J_1\le 4\operatorname{Tr}(e_{k-1}-e_k)$, by Lemma 2.4 (v). Furthermore, by the very definition of e_k, we have $(e_{k-1}-e_k)(S_k-\lambda I)e_{k-1}=(e_{k-1}-e_k)(S_k-\lambda I)(e_{k-1}-e_k)\ge 0$ and since $S_{k-1}\le\lambda I$ on e_{k-1} (and hence also on $e_{k-1}-e_k$), we get

$$(e_{k-1}-e_k)(S_k-\lambda I)(e_{k-1}-e_k)\le (e_{k-1}-e_k)(S_k-S_{k-1})(e_{k-1}-e_k)\le 2(e_{k-1}-e_k).$$

Here in the last passage we used the estimate $\|S_k-S_{k-1}\|\le\|D_k\|\le 2\|A\|\le 2$. Consequently, we obtain $J_2=\operatorname{Tr}\Big(\big((e_{k-1}-e_k)(S_k-\lambda I)(e_{k-1}-e_k)\big)^2\Big)\le 4\operatorname{Tr}(e_{k-1}-e_k)$ and by Lemma 3.1,

$$\operatorname{Tr}\Big((I-e_N)\,(S_N-\lambda I)^2\Big)\le 16\sum_{k=1}^{N}\operatorname{Tr}(e_{k-1}-e_k)=16\operatorname{Tr}(I-e_N).$$

The proof is complete. $\square$

The next statement establishes an important distributional estimate for S_N.

Lemma 3.3. *If $\|A\| \le 1$, then for any $0 < a \le \lambda < \beta$ we have*

$$\operatorname{Tr}(\chi_{[\beta,\infty)}(S_N)) \le \frac{32 \operatorname{Tr}\big((S_N - (\lambda - a))_+\big)}{(\beta-\lambda)^2 a}. \tag{11}$$

Proof. By Lemma 2.4 (ii), we have $e_N S_N e_N \le \lambda e_N$, so

$$\begin{aligned} S_N &= e_N(S_N - \lambda I)e_N + (I - e_N)(S_N - \lambda I)e_N + (S_N - \lambda I)(I - e_N) + \lambda I \\ &\le (I - e_N)(S_N - \lambda I)e_N + (S_N - \lambda I)(I - e_N) + \lambda I. \end{aligned}$$

Therefore, Chebyshev's inequality implies

$$\begin{aligned} &\operatorname{Tr}(\chi_{[\beta,\infty)}(S_N)) \\ &\quad\le \operatorname{Tr}\Big(\chi_{[\beta-\lambda,\infty)}\big((I - e_N)(S_N - \lambda I)e_N + (S_N - \lambda I)(I - e_N)\big)\Big) \\ &\quad\le (\beta-\lambda)^{-2} \operatorname{Tr}\Big(\big((I - e_N)(S_N - \lambda I)e_N + (S_N - \lambda I)(I - e_N)\big)^2\Big) \\ &\quad\le 2(\beta-\lambda)^{-2} \operatorname{Tr}\Big((I - e_N)(S_N - \lambda I)^2\Big). \end{aligned}$$

It suffices to apply the previous lemma and part (vi) of Lemma 2.4. $\square$

If we replace, in all the above considerations, the operators S_n by $-S_n$, $n = 1, 2, \ldots, N$, and repeat the reasoning word-by-word, we obtain the analogous estimate

$$\operatorname{Tr}(\chi_{(-\infty,-\beta]}(S_N)) \le \frac{32 \operatorname{Tr}((-S_N - (\lambda - a)I)_+)}{(\beta-\lambda)^2 a}.$$

Adding this to (11), we get

$$\operatorname{Tr}\big(\chi_{[\beta,\infty)}(|S_N|)\big) \le \frac{32 \operatorname{Tr}\big((|S_N| - (\lambda - a)I)_+\big)}{(\beta-\lambda)^2 a}.$$

Hence by Lemma 2.3, we conclude that

$$\operatorname{Tr}\big(\chi_{[\beta,\infty)}(|T(A)|)\big) \le \frac{32 \operatorname{Tr}\big((|T(A)| - (\lambda - a)I_N)_+\big)}{(\beta-\lambda)^2 a}, \tag{12}$$

where I_N denotes the identity matrix of dimension $N \times N$.

We are ready for the proof of the weak-type estimate.

Proof of (3). Let $\alpha \ge 0$ and $a > 0$. Integrating by parts, we get

$$\operatorname{Tr}\Big(\big(|T(A)| - \alpha I_N\big)_+\Big) = \int_\alpha^\infty \operatorname{Tr}\big(\chi_{[\beta,\infty)}(|T(A)|)\big) \mathrm{d}\beta.$$

Let us bound appropriately the integrand on the right. If $\beta \in [\alpha, \alpha + 2a]$, then we have the obvious estimate $\operatorname{Tr}\big(\chi_{[\beta,\infty)}(|T(A)|)\big) \le \operatorname{Tr}\big(\chi_{[\alpha,\infty)}(|T(A)|)\big)$; for

$\beta > \alpha + 2a$, we apply (12) with $\lambda = \alpha + a$. As the result, we obtain

$$\begin{aligned}\operatorname{Tr}\Big(\big(|T(A)| - \alpha I_N\big)_+\Big) &\le \int_\alpha^{\alpha+2a} \operatorname{Tr}\big(\chi_{[\alpha,\infty)}(|T(A)|)\big)\mathrm{d}\beta \\ &\quad + \int_{\alpha+2a}^{\infty} \frac{32\operatorname{Tr}\big((|T(A)| - \alpha I_N)_+\big)}{(\beta-\alpha-a)^2 a}\mathrm{d}\beta \\ &= 2a\operatorname{Tr}\big(\chi_{[\alpha,\infty)}(|T(A)|)\big) + \frac{32}{a^2}\operatorname{Tr}\big((|T(A)| - \alpha I_N)_+\big).\end{aligned}$$

Putting $a = 8$ and rearranging terms, we get

$$\operatorname{Tr}\Big(\big(|T(A)| - \alpha I_N\big)_+\Big) \le 32\operatorname{Tr}\big(\chi_{[\alpha,\infty)}(|T(A)|)\big),$$

or

$$\frac{\operatorname{Tr}\big(|T(A)|\chi_{[\alpha,\infty)}(|T(A)|)\big)}{\operatorname{Tr}\big(\chi_{[\alpha,\infty)}(|T(A)|)\big)} - \alpha \le 32.$$

Therefore, if we set $\alpha = s_n(T(A))$ (where $1 \le n \le N$ is arbitrary), we obtain

$$\frac{1}{n}\sum_{k=1}^{n} s_k(T(A)) - s_n(T(A)) \le 32.$$

Taking the supremum over n yields $\|T(A)\|_{C_{\infty,\infty}} \le 32\|A\|_{C_\infty}$, which is the claim. □

Proof of (4). Let A be an arbitrary element of C_∞ such that $T(A)$ is compact. Let $P_N : C_\infty \to C_\infty^N$ denote the projection onto the upper-left corner of dimension $N \times N$. Then, by the compactness of $T(A)$, we have the convergence $P_N(T(A)) = T(P_N(A)) \to T(A)$ in norm and $\lim_{N\to\infty} s_n(P_N(T(A)) = s_n(T(A))$ for each n, so

$$\begin{aligned}&\frac{1}{n}\sum_{k=1}^{n} s_k(T(A)) - s_n(T(A)) \\ &= \lim_{N\to\infty}\left(\frac{1}{n}\sum_{k=1}^{n} s_k(T(P_N(A))) - s_n(T(P_N(A)))\right) \le 32,\end{aligned}$$

where the latter estimate follows from (3). This yields the assertion. □

Acknowledgments

The authors would like to thank an anonymous Referee for the careful reading of the first version of the paper and several helpful corrections and suggestions. Tomasz Gałązka and Adam Osękowski are supported by Narodowe Centrum Nauki (Poland), grant number 2018/30/Q/ST1/00072. Yahui Zuo is supported by China Scholarship Council.

References

[1] C. Bennett, R. A. DeVore and R. Sharpley, *Weak-L^∞ and BMO*, Ann. of Math. 113 (1981), 601–611.

[2] P. L. Butzer and H. Berens, Semi-Groups of Operators and Approximation. Springer Verlag, New York, 1967.

[3] S. Chobanyan, S. Levental amd H. Salehi, *On the constant in Menshov-Rademacher inequality*, J. Ineq. Appl. **2006** (2006), Article number: 68969.

[4] M. D. Choi, *Tricks or Treats with the Hilbert Matrix*, Amer. Math. Monthly **90** (1983), 301–312.

[5] I. Cuculescu, *Martingales on von Neumann algebras*, J. Multivariate Anal. **1** (1971), 17–27.

[6] I.C. Gohberg, M.G. Krein, Theory and applications of Volterra operators in Hilbert space, Transl. Math. Monographs, Vol. 24, Amer. Math. Soc. Providence, R.I. 1970.

[7] M. Junge, J. Parcet and Q. Xu, *Rosenthal type inequalities for free chaos*, Ann. Probab. **35** (2007), 1374–1437.

[8] M. Junge and Q. Xu, *On the best constants in some non-commutative martingale inequalities*, Bull. London Math. Soc. **37** (2005), 243–253.

[9] S. Kwapień, A. Pełczyński, *The main triangle projection in matrix spaces and its applications*, Studia Math. **34** (1970), 43–67.

[10] T. Mei, *Notes on Matrix Valued Paraproducts*, Indiana Math. J. **55** (2006), 747–760.

[11] J. R. Ringrose, Compact non-self-adjoint operators (Van Nostrand Reinhold Co., 1971).

[12] A. L. Shields, *An Analogue of a Hardy-Littlewood-Fejer Inequality for Upper Triangular Trace Class Operators*, Math. Z. **182** (1983), 473–484.

[13] B. Simon, Trace ideals and their applications, Cambridge University Press, 1979.