

KOROVKIN TESTS, APPROXIMATION, AND ERGODIC THEORY

STEFANO SERRA CAPIZZANO

ABSTRACT. We consider sequences of $s \cdot k(n) \times t \cdot k(n)$ matrices $\{A_n(f)\}$ with a block structure spectrally distributed as an L_1 p -variate $s \times t$ matrix-valued function f , and, for any n , we suppose that $A_n(\cdot)$ is a linear and positive operator. For every fixed n we approximate the matrix $A_n(f)$ in a suitable linear space \mathcal{M}_n of $s \cdot k(n) \times t \cdot k(n)$ matrices by minimizing the Frobenius norm of $A_n(f) - X_n$ when X_n ranges over \mathcal{M}_n . The minimizer \tilde{X}_n is denoted by $\mathcal{P}_{k(n)}(A_n(f))$. We show that only a simple Korovkin test over a *finite* number of polynomial test functions has to be performed in order to prove the following general facts:

1. the sequence $\{\mathcal{P}_{k(n)}(A_n(f))\}$ is distributed as f ,
2. the sequence $\{A_n(f) - \mathcal{P}_{k(n)}(A_n(f))\}$ is distributed as the constant function 0 (i.e. is spectrally clustered at zero).

The first result is an ergodic one which can be used for solving numerical approximation theory problems. The second has a natural interpretation in the theory of the preconditioning associated to cg-like algorithms.

1. INTRODUCTION

In this paper we deal with the approximation of a given sequence of matrices $\{A_n(f)\}$ by another sequence $\mathcal{P}_{k(n)}(A_n(f))$ of matrices obtained from the first one by projecting (via the operators $\{\mathcal{P}_{k(n)}(\cdot)\}$) each matrix $A_n(f)$ over a linear space of matrices with regard to the scalar product associated with the Frobenius norm. In the theory of preconditioning this kind of matrix approximation is called optimal [6].

Here we assume that each $A_n(\cdot)$ is a linear and positive operator, that each sequence $\{A_n(f)\}$ is distributed as f (in the Weyl-Tyrtysnikov sense [31] as reported in equation (4)), and that $\|A_n(f)\|_{\text{sp}} \leq \|f\|_{L_\infty}$ when $f \in L_\infty$ with $\|\cdot\|_{\text{sp}}$ being the spectral norm [2]. We firstly deduce that $\mathcal{P}_{k(n)}(\cdot)$ is a linear and positive operator. This preliminary fact is then used to prove a powerful and somewhat surprising result: if $\{\mathcal{P}_{k(n)}(A_n(g))\}$ approximates $\{A_n(g)\}$ for a *finite* set of test polynomials g in the sense that the sequence $\{\mathcal{P}_{k(n)}(A_n(g)) - A_n(g)\}$ is clustered at zero, then each sequence $\{\mathcal{P}_{k(n)}(A_n(f))\}$ approximates the sequence $\{A_n(f)\}$ for any integrable function f .

This result is in the style of Korovkin theory [17, 22]. However, while the classical Korovkin-style theorems are used to deduce convergence over the class of continuous

Received by the editor February 2, 1998 and, in revised form, November 20, 1998.

1991 *Mathematics Subject Classification*. Primary 65F10, 65D15, 15A60, 47B65, 28Dxx.

Key words and phrases. Distributions and ergodic theory, Toeplitz matrices, Korovkin theorem, circulants and τ matrices, discrete transforms.

©2000 American Mathematical Society

functions, here the result is much more powerful since we are able to treat the L_1 case.

We stress that we use the matrix version of the Korovkin theorems developed in [20, 22] and the simple but powerful technique used by Tilli in the Toeplitz context [29]. Indeed, as a special case of the general result, we consider also the optimal approximation of the multilevel Toeplitz matrices with $s \times t$ unstructured blocks by algebras/matrix spaces: we obtain that the optimal approximation also distributes as f for all the known trigonometric algebras [10, 16]. This approximation result has two aspects: the first is related to the use of cg-like algorithms [5, 21, 20, 31]. The second is not so classical, because we propose an algorithm for approximating a multivariate Lebesgue integrable function f . We use its Fourier coefficients, the optimal approximation of the Toeplitz matrices $\{A_n(f)\}$ in the circulant algebra [7], and the ergodic theorem obtained as a byproduct of the main analysis: the given procedure is cheap enough because its arithmetic and parallel cost is asymptotically that of the well known fast Fourier transform (FFT) [3].

The paper is organized as follows. In Section 2 we introduce some notation and some preliminary results. In Section 3 we give the basic properties of the optimal approximation, and in Section 4 we describe the linear spaces of matrices in which we approximate the sequence $\{A_n(f)\}$. Section 5 is devoted to the Korovkin-style theorems, and in Section 6 we prove the main results. Section 7 is then addressed to some applications, and Section 8 to a few related numerical experiments. The final Section 9 of conclusions ends the paper.

2. NOTATIONS AND PRELIMINARIES

We denote by M_s the linear space $\mathbf{C}^{s \times s}$ of all square matrices of order s with complex entries. If $A \in M_s$, then the symbol $\sigma_j(A)$ denotes the j -th singular value of A , and, if A is Hermitian, then the symbol $\lambda_j(A)$ denotes the j -th eigenvalue of A , where both sets are in nonincreasing order. The space M_s is equipped with the *trace norm* defined as [2]

$$\|A\|_{\text{tr}} = \sum_{j=1}^s \sigma_j(A), \quad A \in M_s.$$

It is interesting to observe that the trace norm coincides with the ordinary trace when A is Hermitian nonnegative definite. Moreover, each square matrix can be looked at as a special linear combination of nonnegative definite matrices. In particular, we have

$$(1) \quad A = (\text{Re}A)^+ - (\text{Re}A)^- + \mathbf{i}(\text{Im}A)^+ - \mathbf{i}(\text{Im}A)^-, \quad \mathbf{i}^2 = -1,$$

where $\text{Re}A = (A + A^*)/2$ and $\text{Im}A = (A - A^*)/(2\mathbf{i})$ are Hermitian matrices and X^* denotes the transpose conjugate of X . In addition, for any Hermitian matrix A we have $A = U \text{diag}(d_1, \dots, d_s) U^*$, where the values $\{d_j\}$ are real and U is unitary. Finally, the matrices A^+ and A^- are defined as $A^+ = U \text{diag}(d_1^+, \dots, d_s^+) U^*$ and $A^- = U \text{diag}(d_1^-, \dots, d_s^-) U^*$ with $d^+ = \max\{0, d\}$ and $d^- = \max\{0, -d\}$.

The following basic result will be used later.

Lemma 2.1 ([2, Bhatia]). *For any pair of matrices $A, B \in M_s$, we have*

$$(2) \quad \sum_{j=1}^s |\sigma_j(A) - \sigma_j(B)| \leq \|A - B\|_{\text{tr}}.$$

If A and B are Hermitian, then we also have

$$(3) \quad \sum_{j=1}^s |\lambda_j(A) - \lambda_j(B)| \leq \|A - B\|_{\text{tr}}.$$

Now we consider a class $\{A_n(f)\}_{n \in \mathbf{N}}$ of $k(n) \times k(n)$ block matrices with $s \times s$ blocks depending on the function $f : \Omega \rightarrow M_s$, where $\Omega \subset \mathbf{R}^p$, $p \geq 1$, is a compact set with nonempty interior equipped with a probability measure μ so that $\mu(\Omega) = 1$. Without loss of generality and in order to apply the Korovkin-type results of Section 5, we will suppose that an ϵ extension of Ω is a subset of I^p with $I = [-\pi, \pi]$: if it is not the case, then we can make a linear change of variables so that an ϵ extension of the new definition set is contained in I^p . Here $\{k(n)\}$ is a monotone increasing sequence of positive integers, and n can be thought as a positive integer or as a p -dimensional vector of integers as in multilevel structures [31, 29]. For instance, when we consider p -level Toeplitz matrices [31, 29, 23] or discretizations by equispaced finite differences over rectangular p -dimensional domains Ω of boundary value differential problems [14], the quantity $k(n)$ equals $N(n) = n_1 \cdot n_2 \cdots n_p$ with $n = (n_1, n_2, \dots, n_p) \in \mathbf{N}_+^p$. When the domain Ω is not rectangular, it is evident that $k(n)$ is generically an increasing function of n , where n can be, roughly speaking, the inverse of the stepsize of the considered discretization.

More precisely, we deal with μ -integrable matrix-valued functions. So it is natural to use the L_1 -norm induced by the trace norm over M_s , i.e.,

$$\|f\|_{L_1} = \int_{\Omega} \|f(x)\|_{\text{tr}} d\mu(x).$$

In addition, we assume that $A_n(\cdot)$ is *linear and positive* in the sense that for any $f, g \in L_1 = L_1(\Omega, M_s)$ and α, β constants, we require that $A_n(\alpha f + \beta g) = \alpha A_n(f) + \beta A_n(g)$ and that the matrices $\{A_n(f)\}$ are all nonnegative definite if the symbol f is Hermitian and nonnegative definite almost everywhere with regard to the measure μ .

We denote by $\{\sigma_i(A_n(f))\}_{i \leq s \cdot k(n)}$ and $\{\lambda_i(A_n(f))\}_{i \leq s \cdot k(n)}$ the complete set of the singular values and eigenvalues of $A_n(f)$, respectively. We suppose that the singular values of $\{A_n(f)\}$ are distributed as f with regard to the probability space (Ω, μ) [31, 29, 21] in the Weyl-Tyrtysnikov sense: for any $F \in C_0$ (continuous function with bounded support) the ergodic formula

$$(4) \quad \lim_{n \rightarrow \infty} \frac{1}{s \cdot k(n)} \sum_{i=1}^{s \cdot k(n)} F(\sigma_i(A_n(f))) = \int_{\Omega} \sum_{j=1}^s F(\sigma_j(f(x))) d\mu(x)$$

holds true. Notice that the right-hand side of equation (4) makes sense because each function $F(\sigma_j(f(x)))$ is measurable and bounded over the probability space (Ω, μ) . If all the matrices are Hermitian, then we suppose that for any test function $F \in C_0$ the equation

$$(5) \quad \lim_{n \rightarrow \infty} \frac{1}{s \cdot k(n)} \sum_{i=1}^{s \cdot k(n)} F(\lambda_i(A_n(f))) = \int_{\Omega} \sum_{j=1}^s F(\lambda_j(f(x))) d\mu(x)$$

holds. We also assume the following:

$$(6) \quad \lim_{n \rightarrow \infty} \frac{1}{s \cdot k(n)} \|A_n(f)\|_{\text{tr}} = \int_{\Omega} \|f(x)\|_{\text{tr}} d\mu(x) = \|f\|_{L_1}.$$

Finally, as emphasized in the next lemma, certain relations in the style of equation (6) are a simple consequence of assumption (4) in the case where f is in $L_\infty(\Omega, M_s)$ with respect to μ .

Lemma 2.2. *Let $\{\|\cdot\|_p\}_{\infty \geq p \geq 1}$ be the class of the Schatten p -norms [2, Bhatia, p. 92] over M_s , with*

$$\begin{aligned} \|A\|_p^p &= \sum_{j=1}^s \sigma_j^p(A), & p < \infty, \\ \|A\|_\infty &= \sigma_1(A), & A \in M_s, \end{aligned}$$

and let $\{\|\cdot\|_{L_p}\}_{\infty \geq p \geq 1}$ be a class of norms [2] over $L_\infty(\Omega, M_s)$, with

$$\begin{aligned} \|f\|_{L_p}^p &= \int_\Omega \|f(x)\|_p^p d\mu(x), & p < \infty, \\ \|f\|_{L_\infty} &= \text{esssup}_x \|f(x)\|_\infty, & f \in L_\infty(\Omega, M_s). \end{aligned}$$

Then

$$\lim_{n \rightarrow \infty} \frac{1}{s \cdot k(n)} \|A_n(f)\|_p = \|f\|_{L_p}, \quad p < \infty,$$

and

$$\lim_{n \rightarrow \infty} \|A_n(f)\|_\infty = \|f\|_{L_\infty}.$$

Proof. It is enough to set F continuous, with bounded support, and such that it coincides with $F(z) = z^p$ over $[0, \|f\|_{L_\infty}]$, and to recall that

$$\|A_n(f)\|_{\text{sp}} = \|A_n(f)\|_\infty \leq \|f\|_{L_\infty}.$$

The application of relation (4) concludes the proof of the first part.

For the second part, we first remark that $\limsup_{n \rightarrow \infty} \|A_n(f)\|_\infty \leq \|f\|_{L_\infty}$. Now, by contradiction, we suppose that

$$\liminf_{n \rightarrow \infty} \|A_n(f)\|_\infty = \alpha < \|f\|_{L_\infty}.$$

Consider a value $\epsilon > 0$ so that $\alpha + 2\epsilon < \|f\|_{L_\infty}$. There exists a subsequence $\{n_q\}$ so that for any $q \geq \bar{q}$ we have

$$\|A_{n_q}(f)\|_\infty = \alpha + \epsilon < \|f\|_{L_\infty}.$$

Take a globally continuous function F , piecewise linear and with bounded support, satisfying the following relations:

$$(7) \quad F(z) = \begin{cases} 0 & \text{if } \{z \leq \alpha + \epsilon\} \cup \{z \geq \|f\|_{L_\infty} + \epsilon\}, \\ \text{linear} & \text{over } (\alpha + \epsilon, \alpha + 2\epsilon) \cup (\|f\|_{L_\infty}, \|f\|_{L_\infty} + \epsilon), \\ 1 & \text{over } [\alpha + 2\epsilon, \|f\|_{L_\infty}]. \end{cases}$$

Set $\Omega_\epsilon = \{x \in \Omega : \sigma_1(f(x)) \in [\alpha + 2\epsilon, \|f\|_{L_\infty}]\}$. From the definition of $\|f\|_{L_\infty}$ it follows that $\mu(\Omega_\epsilon) > 0$. Therefore for $n = n_q$, q large enough, and F as in (7), we find that

$$(8) \quad \frac{1}{s \cdot k(n_q)} \sum_{i=1}^{s \cdot k(n_q)} F(\sigma_i(A_{n_q}(f))) \equiv 0$$

and

$$(9) \quad \int_\Omega \sum_{j=1}^s F(\sigma_j(f(x))) d\mu(x) \geq \int_\Omega F(\sigma_1(f(x))) d\mu(x) \geq \int_{\Omega_\epsilon} 1 d\mu(x) = \mu(\Omega_\epsilon) > 0.$$

Finally, for q large enough, we observe that (8) and (9) contradict (4). □

It should be noticed that the Frobenius norm is the norm denoted by $\|\cdot\|_2$, the trace norm is the one denoted by $\|\cdot\|_1$, and the spectral norm coincides with $\|\cdot\|_\infty$.

Remark 2.1. All the results of this paper apply to class of matrices $A_n(f)$ such that $A_n(\cdot)$ is a linear positive operator and under the assumption that (4) holds. We observe that there exist several important classes of matrices that satisfy these hypotheses: multilevel block Toeplitz matrices [32, 29], finite difference and finite element matrices discretizing elliptic problems over bounded domains of \mathbf{R}^p and with Dirichlet boundary conditions [30, 25, 27].

3. THE OPTIMAL FROBENIUS APPROXIMATION

In this section we allow the function f to be valued in the space $\mathbf{C}^{s \times t}$. Then we suppose to have the block matrices $\{A_n(f)\}_{n \in \mathbf{N}}$ of dimension $k(n) \times k(n)$ with $s \times t$ blocks. Now, in order to make sense of the notion of positivity, we assume the following *locality property*. Let $f : \Omega \rightarrow \mathbf{C}^{s \times t}$ and $g : \Omega \rightarrow \mathbf{C}^{s_1 \times t_1}$, with $s \geq s_1$ and $t \geq t_1$, and let us suppose that $g(\cdot)$ is a submatrix of $f(\cdot)$ almost everywhere. With these hypotheses we require that $A_n(g)$ is a submatrix of $A_n(f)$. In this way, we say that the operator $A_n(\cdot)$ is positive if it is positive over the $k \times k$ matrix-valued functions with $k = \min\{s, t\}$.

Remark 3.1. It should be observed that, under this *locality property*, relation (4) still holds for $\{A_n(f)\}$ when s is replaced by $\min\{s, t\}$.

3.1. Matrix spaces and optimal approximation. Let U be a unitary complex $K \times K$ matrix; then, for any fixed K, s, t , by $\mathcal{M}(U)$ we denote the vector space of all the matrices simultaneously block diagonalized by the unitary transforms $U(s) = U \otimes I_s$ and $U^*(t) = U^* \otimes I_t$.

More precisely,

$$(10) \quad \mathcal{M}(U) = \{X = U(s)\Delta U^*(t) : \Delta = \text{diag}_{j=1, \dots, K}(\Delta_j)\},$$

Δ_j being $s \times t$ complex matrices.

Some generalizations have been considered [4, 10], by modifying the canonical form Δ which is assumed to be diagonal in the definition of $\mathcal{M}(U)$: here, for the sake of simplicity, we restrict our analysis to the case where Δ is diagonal.

The *optimal preconditioning operator* $\mathcal{P}_K = \mathcal{P}_K[U]$ is defined on $\mathbf{C}^{K \cdot s \times K \cdot t}$ and takes values in $\mathcal{M}(U)$, where both vector spaces are equipped with the Frobenius norm $\|X\|_F^2 = \sum_{i,j} |x_{i,j}|^2 = \|X\|_2^2$. Then

$$(11) \quad \mathcal{P}_K(A) = \arg \min_{X \in \mathcal{M}(U)} \|A - X\|_F,$$

where the minimum exists and is unique since $\mathcal{M}(U)$ is a linear finite dimensional space and $(\mathbf{C}^{K \cdot s \times K \cdot t}, \langle \cdot, \cdot \rangle_F)$ is a Hilbert space with $\langle A, B \rangle_F = \text{tr}(A^*B)$, $\|A\|_F^2 = \langle A, A \rangle_F$ and where $\text{tr}(\cdot)$ is the ordinary trace.

The following two results have been proved in a more general setting in [10]: the first theorem concerns the algebraic properties of the optimal operator, the second stresses some interesting geometric features.

Theorem 3.1. *Let $A, B \in \mathbf{C}^{K \cdot s \times K \cdot t}$, and let $\mathcal{P}_K = \mathcal{P}_K[U, s, t]$ be defined as in (11). The following algebraic properties hold:*

1. $\mathcal{P}_K(A)$ satisfies the explicit formula

$$(12) \quad \mathcal{P}_K(A) = U(s)\delta(U^*(s)AU(t))U^*(t),$$

where $\delta(U^*(s)AU(t))$ is the block diagonal matrix with $s \times t$ blocks such that $(\delta(U^*(s)AU(t)))_{i,i} = (U^*(s)AU(t))_{i,i}$, for $i = 1, \dots, K$;

2. $\mathcal{P}_K(\alpha A + \beta B) = \alpha \mathcal{P}_K(A) + \beta \mathcal{P}_K(B)$ with $\alpha, \beta \in \mathbf{C}$;
3. $\mathcal{P}_K(A^*) = (\mathcal{P}_K^*(A))^*$, where $\mathcal{P}_K^* = \mathcal{P}_K[U^*, t, s]$ and s and t interchange their roles;
4. if $s = t$, $\mathcal{P}_K(A)$ is Hermitian if A is;
5. $\text{Trace}(\mathcal{P}_K(A)) = \text{Trace}(A)$ (in the $s \times t$ block sense).

For $X \in \mathbf{C}^{K \cdot s \times K \cdot t}$, we set

$$\text{Trace}(X) = \sum_{j=1}^K X[(j-1)s + 1 : js, (j-1)t + 1 : jt] \in \mathbf{C}^{s \times t},$$

where $X[\alpha : \beta, \gamma : \delta]$ is the submatrix of A obtained by deleting each row whose index does not belong to $[\alpha, \beta]$ and each column whose index does not belong to $[\gamma, \delta]$. It should be remarked that the invariance of the trace in the block sense plainly implies the invariance of the ordinary trace when $s = t$.

Theorem 3.2. *Let $A \in \mathbf{C}^{K \cdot s \times K \cdot t}$ and $\mathcal{P}_K = \mathcal{P}_K[U]$. The following geometric properties hold:*

1. $\sup_{X: \|X\|_* = 1} \|\mathcal{P}_K(X)\|_* = 1$, with $\|\cdot\|_*$ being the spectral norm or the Frobenius norm,
2. $\|A - \mathcal{P}_K(A)\|_F^2 = \|A\|_F^2 - \|\mathcal{P}_K(A)\|_F^2$.

Now the following result is a simple consequence of Theorem 3.1.

Theorem 3.3. *If $s = t$ and A is Hermitian, then the eigenvalues of $\mathcal{P}_K(A)$ are contained in the closed real interval $[\lambda_{\min}(A), \lambda_{\max}(A)]$ containing all the eigenvalues of A . Moreover, when A is positive definite, so is $\mathcal{P}_K(A)$.*

In short, this operator is always linear, and preserves the Hermitianity, the positivity and the block trace; therefore it is a *linear and positive* operator in the sense stated in [24], namely, it is linear and, if $s = t$, it maps nonnegative definite matrices into nonnegative definite matrices.

Therefore the composition $\mathcal{P}_K(A_n(f))$ with $K = k(n)$ is another linear positive operator in a somewhat different sense [21], because, under the assumption that $s = t$, it maps nonnegative definite matrix-valued functions into nonnegative definite matrices.

The following stability result is a consequence of relation (6) and of the trace preserving property of the operator \mathcal{P}_K stated in Theorem 3.1. The style of the proof follows that of Lemma 3.1 in [29].

Lemma 3.1. *If $f \in L_1(\Omega, M_s)$, then, for n large enough,*

$$(13) \quad \frac{1}{s \cdot k(n)} \|\mathcal{P}_{k(n)}(A_n(f))\|_{\text{tr}} \leq 5 \|f\|_{L_1}.$$

Proof. Consider the Hermitian matrix-valued functions

$$f_1 = (\text{Re}f)^+, \quad f_2 = (\text{Re}f)^-, \quad f_3 = (\text{Im}f)^+, \quad f_4 = (\text{Im}f)^-.$$

From (1), we have $f(x) = f_1(x) - f_2(x) + \mathbf{i}f_3(x) - \mathbf{i}f_4(x)$ for all $x \in \Omega$, where each matrix f_i is positive semidefinite. Since the operator $\mathcal{P}_{k(n)}(A_n(\cdot))$ is linear, we have

$$(14) \quad \|\mathcal{P}_{k(n)}(A_n(f))\|_{\text{tr}} \leq \sum_{i=1}^4 \|\mathcal{P}_{k(n)}(A_n(f_i))\|_{\text{tr}},$$

and by the positivity of the operator, each $\mathcal{P}_{k(n)}(A_n(f_i))$ is positive semidefinite and therefore the trace norm coincides with the usual trace. But the operator $\mathcal{P}_{k(n)}(\cdot)$ preserves the trace, and therefore

$$\|\mathcal{P}_{k(n)}(A_n(f_i))\|_{\text{tr}} = \text{tr}(A_n(f_i)), \quad i = 1, 2, 3, 4.$$

So in light of equation (6), there exists an integer \bar{n} so that for $n \geq \bar{n}$ we find that

$$\frac{1}{s \cdot k(n)} \text{tr}(A_n(f_i)) \leq (1.25) \int_{\Omega} \|f_i(x)\|_{\text{tr}} d\mu(x) = (1.25) \|f_i\|_{L_1},$$

and since $\|f_i\|_{L_1} \leq \|f\|_{L_1}$ the claimed thesis follows. □

Notice that the estimate of the constant 5 appearing in (13) can be improved, although for our purposes it is good enough.

4. GRAM MATRIX ALGEBRAS AND SPACES

Let $\mathbf{W} = \{W_n\}_{n \in \mathbb{N}}$, with $W_n = \{x_i^{(n)}\}$, be a sequence of sets of $k(n)$ distinct points belonging to Ω . By means of this sequence of sets we define a family of semidefinite scalar products $\{[\cdot, \cdot]_n\}$ over the space of the complex-valued functions defined on Ω : for h_1 and h_2 complex-valued functions we define

$$[h_1, h_2]_n = \sum_{i=1}^{k(n)} (\bar{h}_1 h_2) (x_i^{(n)}).$$

Now, let $F_n = \{f_i^{(n)}\}_{1 \leq i \leq k(n)}$ be a sequence of orthogonal normalized continuous functions with respect to the inner product $[\cdot, \cdot]_n$; therefore we have

$$(15) \quad [f_i^{(n)}, f_j^{(n)}]_n = \delta_{i,j}, \quad i, j \in \{1, \dots, k(n)\},$$

where $\delta_{i,j}$ is the Kronecker delta. If the functions of F_n are polynomials, then we find the classical Gram polynomials. In the general case, adopting the same terminology, we call the functions of F_n Gram functions.

Definition 4.1. A space of matrices of the form $(U_n \otimes I_s) \Delta(U_n^* \otimes I_t)$ is a *Gram space* if and only if the matrix U_n is the generalized Vandermonde [12] matrix $(f_j^{(n)}(x_i^{(n)}))_{i,j=1}^{k(n)}$. Here the set $\{f_j^{(n)}\}_{1 \leq j \leq k(n)}$ is the finite Gram sequence F_n associated with the inner product related to the grid $W_n = \{x_i^{(n)}\}$. If $s = t$, then we have a *Gram algebra*.

Therefore, in the following, by $\{\mathcal{M}(U_n)\}$ we denote the sequence of Gram spaces (or algebras) related to the sets $\mathbf{W} = \{W_n\}$ and $\mathbf{F} = \{F_n\}$.

We notice that the orthogonality condition $U_n^* U_n = I$ is exactly inherited from the discrete orthogonality relations (15) concerning the Gram functions.

The last requirement is that the sequence \mathbf{W} is distributed as Ω with regard to the probability measure μ :

Definition 4.2. A sequence of sets $\mathbf{W} = \{W_n\}_n$ with $W_n = \{x_i^{(n)}\}$ and $\#(W_n) = k(n)$ is called quasi-uniform with regard to (Ω, μ) if for any ball $B(x_0, \rho) \subset \Omega$, $B(x_0, \rho) = \{x \in \mathbf{R}^p : \|x - x_0\|_{\infty} \leq \rho\}$, $\rho > 0$, we have

$$\#\{x_i^{(n)} \in B(x_0, \rho)\} = k(n) \cdot \mu(B(x_0, \rho)) + o(k(n)).$$

We say that \mathbf{W} is distributed as (Ω, μ) , if \mathbf{W} is quasi-uniform and for any positive ϵ there exist $M = M_\epsilon > 0$, $\delta = \delta_\epsilon > 0$ and points $\{v_j\}$, $j = 1, \dots, M$ such that

$$B(v_j, \delta) \subset \Omega, \quad B^\circ(v_j, \delta) \cap B^\circ(v_k, \delta) = \emptyset, \quad j \neq k,$$

with X° denoting the interior part of X and

$$\mu\left(\Omega \setminus \left(\bigcup_{j=1}^M B(v_j, \delta)\right)\right) < \epsilon.$$

Some remarks concerning the preceding definition are needed.

Remark 4.1. In the preceding definition we have introduced two notions, *quasi-uniformity with regard to* (Ω, μ) and *distribution as* (Ω, μ) . These notions often coincide, but they are not equivalent. By definition it is evident that the second implies the first. However it is possible to construct a compact set $\Omega \subset \mathbf{R}$ with nonempty interior endowed with a probability measure, and a grid sequence \mathbf{W} whose points belong to Ω such that

- $\mathbf{W} = \{W_n\}_n$ is quasi-uniform with regard to (Ω, μ) and
- \mathbf{W} is not distributed as (Ω, μ) .

Example. Let $\epsilon < 10^{-1}$, let $\{q_i\}_{i \in \mathbf{N}}$ be the sequence of all the rational numbers belonging to the real interval $[1, 2]$, with $q_0 = 3/2$. Let $S = \bigcup_{i \in \mathbf{N}} B^\circ(q_i, \epsilon 2^{-1})$, $K = [1, 2] \setminus S$ and let $\Omega = [0, 1] \cup K$. The following facts hold:

1. Ω is compact (since S is open).
2. S is dense in $[1, 2]$ and K is totally not connected.
3. The Lebesgue measure $m(K)$ of K is t , and $1 - 2/5 \leq t \leq 1 - 1/5$.

Therefore, setting $\mu = \frac{m}{(1+t)}$, we have that (Ω, μ) is a probability space, Ω is compact, and its interior is nonempty since $\Omega^\circ = (0, 1)$.

Now define $\mathbf{W} = \{W_n\}_n$ with

$$W_n = \{x_i^{(n)} = i/n, i = 0, \dots, n - 1, x_i^{(n)} \in K, i = n, \dots, n + \lceil tn \rceil\}$$

and $k(n) = n + \lceil tn \rceil$. We claim that \mathbf{W} is quasi-uniform with regard to (Ω, μ) : in fact, by item 2 it follows that all the balls of positive radius of Ω are the nontrivial subintervals $[\alpha, \beta]$ of $[0, 1]$. Therefore

$$\begin{aligned} \#\{x_i^{(n)} \in B(x_0, \rho) \subset \Omega\} &= \#\{x_i^{(n)} \in [\alpha, \beta] \subset [0, 1]\} \\ &= \#\{i/n \in [\alpha, \beta] \subset [0, 1], i = 0, \dots, n - 1\} \\ &= n(\beta - \alpha) + O(1) = (n + tn) \frac{(\beta - \alpha)}{1 + t} + O(1) \\ &= k(n) \cdot \mu(B(x_0, \rho)) + o(k(n)). \end{aligned}$$

Nevertheless \mathbf{W} is not distributed as (Ω, μ) . In fact the set Ω cannot be approximated in measure μ by finite unions of internal intervals, since K is totally not connected and has positive μ measure equal to $t/(t + 1)$.

Remark 4.2. If μ is constructed as before (that is as a normalized Lebesgue measure over \mathbf{R}^p), then “ \mathbf{W} is distributed as (Ω, μ) ” implies that Ω is measurable in the Peano-Jordan sense [15] (the multidimensional version of the Riemann measure). In fact Ω is compact and consequently Lebesgue measurable; therefore it can be approximated in measure m by plurirectangles containing Ω . Since \mathbf{W} is distributed

as (Ω, μ) , it follows that it can be approximated in measure m by plurirectangles contained in Ω . Therefore Ω is measurable in the Peano-Jordan sense.

It is worth noticing that if \mathbf{W} is distributed as (Ω, μ) , then, in the light of Remark 4.2, for any continuous function f defined over Ω , we can use the values $\{f(x_i^{(n)})\}$ in Riemann sums in order to approximate $\int_{\Omega} f d\mu$. More precisely, the following statement holds true:

Lemma 4.1. *Suppose that \mathbf{W} is distributed as (Ω, μ) and f is continuous. Then*

$$\sum_{i=1}^{k(n)} f(x_i^{(n)}) = k(n) \cdot \int_{\Omega} f d\mu + o(k(n)).$$

Proof. For any fixed $\epsilon > 0$, there exist a positive δ and an integer $M = M_{\epsilon}$ so that we find values $\{\alpha_j\}_{j \leq M}$ and points $\{v_j\}_{j \leq M}$ of Ω for which

1. $B(v_j, \delta) \subset \Omega$, $B^{\circ}(v_j, \delta) \cap B^{\circ}(v_k, \delta) = \emptyset$, $j \neq k$,
2. $\mu(\Omega \setminus (\bigcup_j B(v_j, \delta))) < \epsilon$,
3. $|f(x) - \sum_{j=1}^M \alpha_j Ch_{B(v_j, \delta)}(x)| < \epsilon$ if $x \in \bigcup_j B(v_j, \delta)$.

Here Ch_X denotes the characteristic function of a set X . From this we have

$$\begin{aligned} \sum_{i=1}^{k(n)} f(x_i^{(n)}) &= \sum_{i=1}^{k(n)} \sum_{j=1}^M \alpha_j Ch_{B(v_j, \delta)}(x_i^{(n)}) + A_{n, \epsilon} + B_{n, \epsilon} \\ &= \sum_{j=1}^M \alpha_j \sum_{i=1}^{k(n)} Ch_{B(v_j, \delta)}(x_i^{(n)}) + A_{n, \epsilon} + B_{n, \epsilon} \\ &= \sum_{j=1}^M \alpha_j \# \{x_i^{(n)} \in B(v_j, \delta)\} + A_{n, \epsilon} + B_{n, \epsilon} \end{aligned}$$

where, by virtue of items 1, 2, and 3, and of the quasi-uniform distribution of \mathbf{W} , we find that

$$|A_{n, \epsilon}| < \sum_{x_i^{(n)} \in \bigcup_j B(v_j, \delta)} \epsilon \leq \epsilon k(n),$$

$$\begin{aligned} |B_{n, \epsilon}| &< \sum_{x_i^{(n)} \in \Omega \setminus \bigcup_j B(v_j, \delta)} \|f\|_{L_{\infty}} \leq \mu\{\Omega \setminus \bigcup_j B(v_j, \delta)\} k(n) \|f\|_{L_{\infty}} + o(k(n)) \\ &< \epsilon k(n) \|f\|_{L_{\infty}} + o(k(n)), \end{aligned}$$

$$\sum_{j=1}^M \alpha_j \# \{x_i^{(n)} \in B(v_j, \delta)\} = k(n) \int_{\Omega} \sum_{j=1}^M \alpha_j Ch_{B(v_j, \delta)}(x) d\mu + o(k(n)),$$

and

$$\int_{\Omega} |f(x) - \sum_{j=1}^M \alpha_j Ch_{B(v_j, \delta)}(x)| d\mu < \epsilon(1 + \|f\|_{L_{\infty}}).$$

Finally, by joining all the previous 5 relations and using the arbitrariness of ϵ , the proof is concluded. \square

5. SOME KOROVKIN-TYPE THEOREMS

In this section we recall some results [20, 10] of approximation theory based on the concepts of linear positive operators and in the spirit of the Korovkin theorems [17].

5.1. Approximation theory premises. Since the goodness of the optimal preconditioners in the case of Gram algebras is roughly speaking decided by the behaviour of an operator [22] over the grid points of W_n , we are motivated to introduce the notions of convergence on discrete sets.

Definition 5.1. Let $f : \Omega \rightarrow X$, where $X = (X, \|\cdot\|)$ is a normed vector space. Let $\{f_n\}_n$ be a sequence of functions belonging to $C(\Omega, X)$ and $\{W_n\}_n \subset \Omega$ be a sequence of sets of pairwise distinct points with $\#(W_n) = k(n)$.

- We say that f_n uniformly converges to f on $\{W_n\}$ if

$$\lim_{n \rightarrow \infty} \sup_{x \in W_n} \|f_n(x) - f(x)\| = 0.$$

- The convergence is $O(1)$ -uniform (on $\{W_n\}$) if

$$\lim_{n \rightarrow \infty} \sup_{x \in W_n \setminus W(J_n)} \|f_n(x) - f(x)\| = 0,$$

where $W(J_n)$ is a set of points of W_n associated with the set of indices J_n of cardinality bounded by an absolute constant.

- The convergence is $o(k(n))$ -uniform (on $\{W_n\}$) if

$$\lim_{n \rightarrow \infty} \sup_{x \in W_n \setminus W(J_n)} \|f_n(x) - f(x)\| = 0,$$

where $W(J_n)$ is a set of points of W_n whose cardinality is $o(k(n))$.

Now let us introduce the following definition.

Definition 5.2 ([22]). Let \mathcal{G} be the linear space

$$(C(I^p, \mathbf{C}^{s \times t}), \|\cdot\|_\infty), \quad I = [-\pi, \pi]$$

of continuous (periodic) functions defined on I^p , and let $\{\Phi_n\}$ be a sequence of linear operators on \mathcal{G} . If $\{q_i\}_{i=1}^3$ is the set of the three test functions $1, \sin(x), \cos(x)$, and $E_{j,k}$ is the matrix of the canonical basis of $\mathbf{C}^{s \times t}$ having 1 in the position (j, k) and zero elsewhere, let $\hat{q}_{i,j,k,l}(x) = E_{j,k}q_i(x_l)$ for $i = 1, 2, 3, (j, k) \in \{1, \dots, s\} \times \{1, \dots, t\}, l \in \{1, \dots, p\}$.

We say that “ $\{\Phi_n\}_n$ satisfies the Korovkin test^p” if $\Phi_n(\hat{q}_{i,j,k,l})$ uniformly converges to $\hat{q}_{i,j,k,l}$, according to one of the notions given in Definition 5.1.

We notice that for $p = s = t = 1$, the latter definition reduces to the classical Korovkin test [17] in the trigonometric scalar case.

In the following we will consider the linear operator

$$(16) \quad L_n[U](f) : x \in I^p \mapsto ([\mathbf{u}(n)](x) \otimes I_s) A_n(f) ([\mathbf{u}(n)](x) \otimes I_t)^* \in \mathbf{C}^{s \times t}.$$

Here $[\mathbf{u}(n)](x)$ is the generic row of U^* , where the grid points have been replaced by the continuous variable $x = (x_1, \dots, x_p)$. Notice that, in light of part 1 of Theorem 3.1, we obtain that the diagonal form of $\mathcal{P}_{k(n)}(A_n(f)) = \mathcal{P}_{k(n)}[U](A_n(f))$ is given by the matrix

$$\text{diag} \left(L_n[U](f)(x_1^{(n)}), \dots, L_n[U](f)(x_{k(n)}^{(n)}) \right).$$

Remark 5.1. It can be seen that $L_n[U](f)$ is the continuous expression of the $s \times t$ diagonal blocks of $U^*(s)A_n(f)U(t)$; hence, the singular values of $\mathcal{P}_{k(n)}(A_n(f))$ are given by those of $L_n[U](f)$ evaluated on the points of W_n .

Now we analyze the problem of the convergence with regard to the concepts of $O(1)$ -uniform convergence or $o(k(n))$ -uniform convergence, by stating a specialized version of the Korovkin theorem.

Theorem 5.1 ([10]). *Let \mathcal{G} be the function space of Definition 5.2, and suppose that the linear operators $L_n[U] : \mathcal{G} \rightarrow \mathcal{G}$, given by (16), satisfy the Korovkin test on a given sequence of grids $\{W_n\} \subset I^p$, with the exception of the points belonging to $W(J_n)$. Then, for any continuous function $f \in \mathcal{G}$, $L_n[U](f)$ uniformly converges to f on the same sequence of grids $\{W_n\}$, with the exception of the points of $W(J_n)$.*

We remark that the related convergence is $O(1)$ -uniform or $o(k(n))$ -uniform according to the cardinality of $W(J_n)$: as usual in the Korovkin results, the convergence only needs to be checked on the test functions. Notice that in Section 2, we supposed that an ϵ extension of Ω is a subset of I^p ; this assumption was considered in order to link Theorem 5.1 with our class of matrices $\{A_n(f)\}$ in which the symbol f is defined over Ω

Indeed, by invoking an application of Tietze’s extension theorem ([18], pp. 422–423), for any continuous function f defined over Ω , there exists a continuous 2π -periodic function \hat{f} defined over I^p so that its trace over Ω is f (see also Theorem 3.2 in [8], pp. 10–11).

6. MAIN THEOREMS

We first present a lemma due to Tyrtysnikov which is one of the tools that are going to be used. Then we give an application of the Korovkin-style theory which allows one to prove the main ergodic and approximation results.

Definition 6.1 ([31]). Consider a sequence of matrices $\{S_n\} \in \mathbf{C}^{n \times m}$ with $m \geq n$ and a set M in the nonnegative real line. Denote by M_ϵ the ϵ -extension of M , which is the union of all balls of radius ϵ centered at points of M . For any n , let $\gamma_n(\epsilon)$ count those singular values of S_n that do not belong to M_ϵ . Assume that, for any $\epsilon > 0$,

$$\gamma_n(\epsilon) = o(n), \quad n \rightarrow \infty.$$

Then M is called a *general cluster*. If $M = \{0\}$, then we say that $\{S_n\}$ is *clustered at zero*.

Lemma 6.1 ([31, 32]). *Let $\{A_n\}$ and $\{B_n\}$ be two sequences of $n \times m$ complex matrices with $m \geq n$. If $\|A_n - B_n\|_F^2 = o(n)$, then the sequence $\{A_n - B_n\}$ is clustered at zero, and if in addition the sequence $\{A_n\}$ has a distribution, then the sequence $\{B_n\}$ has the same distribution. The same statement holds true if, for any $\epsilon > 0$, there exists a new sequence $\{D_n(\epsilon)\}$ so that $\text{rank}(D_n(\epsilon)) \leq \epsilon n$ and $\|A_n - B_n - D_n(\epsilon)\|_F^2 \leq \epsilon n$.*

Theorem 6.1. *Let $f \in C(\Omega, \mathbf{C}^{s \times t})$ be a continuous function. Let us suppose that $L_n[U](q) = q + \epsilon_n(q)$ for each one of the test functions q and with ϵ_n going uniformly to zero on the meshes \mathbf{W} related to U (with at most $o(k(n))$ exceptional points). If \mathbf{W} and Ω are as in Lemma 4.1, then the sequences $\{\mathcal{P}_{k(n)}(A_n(f))\}$ and $\{A_n(f)\}$ have the same distribution, and the sequence $\{\mathcal{P}_{k(n)}(A_n(f)) - A_n(f)\}$ is clustered at zero.*

Proof. From the assumption and Theorem 5.1, we have that $L_n[U](f) = f + \epsilon_n(f)$, where the quantity $\epsilon_n(f)$ is infinitesimal as n for $k(n) - o(k(n))$ grid points. Moreover, by virtue of item 1 of Theorem 3.2, $\epsilon_n(f)$ is uniformly bounded in norm at every grid point.

But from the second part of Theorem 3.2 we have

$$0 \leq \|A_n(f) - \mathcal{P}_{k(n)}(A_n(f))\|_F^2 = \|A_n(f)\|_F^2 - \|\mathcal{P}_{k(n)}(A_n(f))\|_F^2.$$

Therefore

$$\|A_n(f) - \mathcal{P}_{k(n)}(A_n(f))\|_F^2 = \|A_n(f)\|_F^2 - \sum_{i=1}^{k(n)} \|\sigma_i(L_n[U](f)(x_i^{(n)}))\|_F^2,$$

which coincides with

$$\|A_n(f)\|_F^2 - \sum_{i=1}^{k(n)} \|f(x_i^{(n)}) + \epsilon_n(f)(x_i^{(n)})\|_F^2.$$

Now, from the definition of the Frobenius norm, we find that

$$\|A_n(f)\|_F^2 = \sum_{i=1}^{k(n) \min\{s,t\}} \sigma_i^2(A_n(f)).$$

The preceding relation is very interesting because, after division by $\min\{s, t\}k(n)$, it is exactly the sum appearing in the left-hand side of (4), where F has bounded support and over $[0, \|f\|_{L_\infty}]$ it coincides with $F(z) = z^2$ (see also Remark 3.1 and Lemma 2.2). Then, by applying Lemma 2.2, we find that

$$(17) \quad \|A_n(f)\|_F^2 = k(n) \cdot \int_{\Omega} \sum_{j=1}^{\min\{s,t\}} \sigma_j^2(f(x)) d\mu(x) + o(k(n)).$$

In addition, from the structure of $\mathcal{P}_{k(n)}(A_n(f))$ and by exploiting the convergence of $L_n[U](f)$ to f , we may conclude that

$$(18) \quad \|\mathcal{P}_{k(n)}(A_n(f))\|_F^2 = \sum_{i=1}^{k(n)} \|f(x_i^{(n)}) + \epsilon_n(f)(x_i^{(n)})\|_F^2 = \sum_{i=1}^{k(n)} \|f(x_i^{(n)})\|_F^2 + o(k(n)).$$

So, by virtue of the quasi-uniform distribution of points $\{x_i^{(n)}\}$ and by Lemma 4.1, we arrive at

$$(19) \quad \sum_{i=1}^{k(n)} \|f(x_i^{(n)}) + \epsilon_n(f)(x_i^{(n)})\|_F^2 = k(n) \cdot \int_{\Omega} \|f(x)\|_F^2 d\mu(x) + o(k(n)),$$

where

$$\|f(x)\|_F^2 = \sum_{j=1}^{\min\{s,t\}} \sigma_j^2(f(x)).$$

The combination of equations (17) and (19), in light of the powerful Lemma 6.1, allows one to state that $\{\mathcal{P}_{k(n)}(A_n(f))\}$ and $\{A_n(f)\}$ are both distributed as f and that their difference is clustered at zero. \square

Therefore the main result holds for the continuous functions. Now we use the latter theorem as an intermediate step in order to prove the main statement with f merely μ -integrable. Of course, we have to require that the continuous functions are dense in the μ -integrable functions, and this is true under very mild assumptions on Ω and μ : more specifically, it is enough to impose the hypotheses of the Lusin theorem and in particular that μ is a measure on a locally compact Hausdorff space Ω (for more details see Rudin [18, p. 55 and pp. 40–41]). For instance, the assumption concerning the set Ω considered in Lemma 4.1 plainly satisfies this requirement. Moreover, we want to stress that the proof of the first part of the next result is substantially identical to the one followed by Tilli in the case of multilevel matrix-valued Toeplitz structures [29], while in the proof of the second part we use arguments and tools introduced by Tyrtysnikov and Zamarashkin [32].

For notational simplicity we present it for the case $s = t$.

Theorem 6.2. *Let $f \in L_1(\Omega, M_s)$. Let us assume that $L_n[U](q) = q + \epsilon_n(q)$ for each one of the test functions q and with ϵ_n going uniformly to zero on the meshes \mathbf{W} related to U (with at most $o(k(n))$ exceptional points). If \mathbf{W} and Ω are as in Lemma 4.1, then*

1. *the sequences $\{\mathcal{P}_{k(n)}(A_n(f))\}$ and $\{A_n(f)\}$ have the same distribution f , and*
2. *the sequence $\{\mathcal{P}_{k(n)}(A_n(f)) - A_n(f)\}$ is clustered at zero.*

Proof of the first part. What we want to prove is equivalent to the following:

$$\lim_{n \rightarrow \infty} \frac{1}{s \cdot k(n)} \sum_{i=1}^{s \cdot k(n)} F(\sigma_i(\mathcal{P}_{k(n)}(A_n(f)))) = \int_{\Omega} \sum_{j=1}^s F(\sigma_j(f(x))) d\mu(x),$$

where $\{\sigma_i(\mathcal{P}_{k(n)}(A_n(f)))\}$ are the singular values of $\mathcal{P}_{k(n)}(A_n(f))$ and where F is a generic uniformly continuous bounded function. Following Tilli [29], we reduce the claim to the case where F is continuously differentiable with $\sup |F'| < \infty$, since this set is dense in the continuous functions. Second, we choose a sequence $\{f_m\}$ of functions belonging to $C(\Omega, M_s)$ so that

$$\|f - f_m\|_{L_1} \rightarrow 0.$$

Let us denote D_n as

$$\begin{aligned} & \left| \frac{1}{s \cdot k(n)} \sum_{i=1}^{s \cdot k(n)} F(\sigma_i(\mathcal{P}_{k(n)}(A_n(f)))) - \int_{\Omega} \sum_{j=1}^s F(\sigma_j(f(x))) d\mu(x) \right| \\ & \leq E_{n,m} + F_{n,m} + G_{n,m}. \end{aligned}$$

To handle $E_{n,m}$ we use Lemma 3.1 and Lemma 2.1:

$$\begin{aligned} E_{n,m} &= \left| \frac{1}{s \cdot k(n)} \sum_{i=1}^{s \cdot k(n)} F(\sigma_i(\mathcal{P}_{k(n)}(A_n(f)))) - F(\sigma_i(\mathcal{P}_{k(n)}(A_n(f_m)))) \right| \\ &\leq \frac{\|F'\|_{\infty}}{s \cdot k(n)} \sum_{i=1}^{s \cdot k(n)} |\sigma_i(\mathcal{P}_{k(n)}(A_n(f))) - \sigma_i(\mathcal{P}_{k(n)}(A_n(f_m)))| \\ &\leq \frac{\|F'\|_{\infty}}{s \cdot k(n)} \|\mathcal{P}_{k(n)}(A_n(f)) - \mathcal{P}_{k(n)}(A_n(f_m))\|_{\text{tr}} \\ &= \frac{\|F'\|_{\infty}}{s \cdot k(n)} \|\mathcal{P}_{k(n)}(A_n(f - f_m))\|_{\text{tr}} \\ &\leq 5 \frac{\|F'\|_{\infty}}{s \cdot k(n)} \|f - f_m\|_{L_1}. \end{aligned}$$

The contribution of $F_{n,m}$, owing to Theorem 6.1, is infinitesimal as n , for any m . In fact, for any m the function f_m is continuous, and then

$$F_{n,m} = \left| \frac{1}{s \cdot k(n)} \sum_{i=1}^{s \cdot k(n)} F(\sigma_i(\mathcal{P}_{k(n)}(A_n(f_m)))) - \int_{\Omega} \sum_{j=1}^s F(\sigma_j(f_m(x))) d\mu(x) \right| \rightarrow 0$$

as n tends to ∞ . Finally, the contribution given by $G_{n,m}$ with

$$G_{n,m} = \int_{\Omega} \sum_{j=1}^s |F(\sigma_j(f_m(x))) - F(\sigma_j(f(x)))| d\mu(x)$$

does not depend on n and, in light of Lemma 2.1, goes to zero as m tends to infinity. With this, owing to the arbitrariness of m , the proof of part 1 is concluded.

Proof of part 2. Here we have to analyze the difference $\{\mathcal{P}_{k(n)}(A_n(f)) - A_n(f)\}$. We use an argument which is substantially the one of Tyrtysnikov and Zamarashkin [32]. Let

$$f_M = \begin{cases} f(x) & \text{if } \|f(x)\|_{\text{sp}} \leq M, \\ 0 & \text{otherwise.} \end{cases}$$

Then, for any fixed M , by exploiting the linearity of $A_n(\cdot)$ and of $\mathcal{P}_{k(n)}(\cdot)$, we have

$$\begin{aligned} A_n(f) - \mathcal{P}_{k(n)}(A_n(f)) &= A_n(f - f_M) + [A_n(f_M) - \mathcal{P}_{k(n)}(A_n(f_M))] + \mathcal{P}_{k(n)}(A_n(f_M - f)). \end{aligned}$$

For any M , from the construction of f_M , we have that $f_M \in L_{\infty} \subset L_2$, and therefore, from part 2 of Theorem 3.2, from the first part of this theorem, and from Lemma 2.2, it follows that

$$\|A_n(f_M) - \mathcal{P}_{k(n)}(A_n(f_M))\|_F^2 = \|A_n(f_M)\|_F^2 - \|\mathcal{P}_{k(n)}(A_n(f_M))\|_F^2 = o(k(n)).$$

Therefore, in order to conclude the proof by the application of Lemma 6.1, we have to show that for any positive ϵ , there exists a suitable M so that $A_n(f - f_M)$ and $\mathcal{P}_{k(n)}(A_n(f_M - f))$ can be written as the sum of a term of spectral norm bounded by ϵ and a second term of rank bounded by a quantity $\epsilon k(n)$.

Since $f \in L_1$, from standard measure theory it is clear that

$$(20) \quad \lim_{M \rightarrow \infty} \mu\{x \in \Omega : \sigma_{\max}((f - f_M)(x)) > 0\} = 0.$$

In addition, from the first part we have that the sequences $\{\mathcal{P}_{k(n)}(A_n(f - f_M))\}$ and $\{A_n(f - f_M)\}$ are distributed as $f - f_M$. Therefore, for any positive M and for any positive ϵ (with the exception, at most, of a countable set of values of ϵ [21, 33]), we find that

$$\#\{i : \sigma_i(X_n) > \epsilon\} \leq s^2 \mu\{x \in \Omega : \sigma_{\max}((f - f_M)(x)) > \epsilon\} k(n) + o(k(n)),$$

where $X_n = \mathcal{P}_{k(n)}(A_n(f - f_M))$ or $X_n = A_n(f - f_M)$. Therefore, by the singular value decomposition, for n large enough, the matrices X_n can be written as terms of spectral norm bounded by ϵ and a term whose rank can be bounded as follows:

$$\begin{aligned} \mu\{x \in \Omega : \sigma_{\max}((f - f_M)(x)) > \epsilon\} k(n) \\ \leq \mu\{x \in \Omega : \sigma_{\max}((f - f_M)(x)) > 0\} k(n) \leq \epsilon k(n). \end{aligned}$$

We note that in the latter inequality we have invoked equation (20) with a choice of $M = M(\epsilon)$ large enough. □

The case where $s \neq t$ can be manipulated by using the locality assumption. Indeed, the proof of Theorem 6.2 with $s \neq t$ can be simply done by embedding $f : \Omega \rightarrow \mathbf{C}^{s \times t}$ into a new square matrix-valued function $\hat{f} : \Omega \rightarrow M_k$ with $k = \max\{s, t\}$ and $(\hat{f})_{i,j} = f_{i,j}$ for $1 \leq i \leq s$ and $1 \leq j \leq t$, and $(\hat{f})_{i,j} = 0$ otherwise (we refer also to the last part of the proof of Theorem 1 in [29] for other details).

Remark 6.1. It should be observed that Theorem 6.2 can be read in the context of approximation of functions in the style of the Korovkin results. Indeed, if $L_n[U](q)$ converges to q over \mathbf{W} for the test functions $\{q\}$, then, for any integrable function f , the sequence $\{L_n[U](f)(x)\}$ (with $x \in W_n$) is distributed as f in the sense that

$$\begin{aligned}
 (21) \quad & \lim_{n \rightarrow \infty} \frac{1}{\min\{s, t\} \cdot k(n)} \sum_{i=1}^{k(n)} \sum_{j=1}^{\min\{s, t\}} F(\sigma_j(L_n[U](f)(x_i^{(n)}))) \\
 & = \int_{\Omega} \sum_{j=1}^{\min\{s, t\}} F(\sigma_j(f(x))) d\mu(x),
 \end{aligned}$$

where F is, as usual, continuous with bounded support ($L_n[U](\cdot)$ converges to the identity in a kind of weak* sense). Therefore, with f merely integrable, we cannot claim a kind of pointwise or uniform convergence of the functional operators $\{L_n[U](\cdot)\}$ to the identity, but we can claim a convergence in this distributional sense. In conclusion, we can say that part 1 of Theorem 6.2 can be also considered an extension of the classical Korovkin theorem in the L_1 case. Moreover the optimal Frobenius approximation of matrices $\{A_n(f)\}$, enjoying (4) and (6), is a way of defining new sequences $\{L_n[U](\cdot)\}$ of linear positive operators for which the Korovkin theorem applies in the L_1 case. We notice that the Bernstein polynomial operators [8] are linear and positive and satisfy the Korovkin test, but the approximation does not hold in L_1 [27] in the sense stated in equation (21). On the other hand, the convergence in L_1 is valid for the Cesàro (or Fejér) operators [8] which, indeed, can be viewed as a special case of the operators $\{L_n[U](\cdot)\}$ when the spaces $\{\mathcal{M}_n = \mathcal{M}(U_n)\}$ are the circulants [20] associated to the sequence $\{U = U_n\}$ of the Fourier transforms. Other nonclassical examples [20, 22, 10] of such sequences of operators are obtained by using other spaces, such as those based on sine, cosine, or Hartley transforms [16, 20].

7. APPLICATIONS AND EXAMPLES

In the following we apply the theory developed in the preceding sections to multilevel Toeplitz matrices [5, 13]. Here the Toeplitz matrices $\{A_n(f)\}$ are generated [13] by a Lebesgue-integrable, multivariate, and matrix-valued function f in the sense that the (block) entries of each $A_n(f)$ along the k -th (block) diagonal are given by the k -th Fourier coefficient A_k of the function f acting on I^p , $I = [-\pi, \pi]$, and having values in the space $\mathbf{C}^{s \times t}$ of complex $s \times t$ matrices.

In particular we have p levels of structure whose dimensions are represented by the multi-index $n = (n_1, n_2, \dots, n_p) \in \mathbf{N}_+^p$, so that the inner nonstructured blocks are indexed by subscripts like $k = (k_1, k_2, \dots, k_p)$ with $k_j \in \{-n_j + 1, \dots, n_j - 1\}$. More precisely, setting $x = (x_1, x_2, \dots, x_p)$, $k \cdot x = k_1 x_1 + \dots + k_p x_p$, $i = (i_1, i_2, \dots, i_p)$, and $j = (j_1, j_2, \dots, j_p)$, we have

$$(22) \quad [A_n(f)]_{i,j} = A_{i-j}, \quad A_k = \frac{1}{[2\pi]^p} \int_{I^p} f(x) e^{-i(k \cdot x)} dx.$$

It is understood that the basic blocks of this multilevel structure are given by $\{A_k\}$ with $A_k \in \mathbf{C}^{s \times t}$. To have an idea of the multilevel structure we have to choose an ordering among the indices $\{k_j\}$. The following scheme is the classical one, and is clearly described in [31]. The matrix $A_n(f)$ has dimension $N(n)s \times N(n)t$, where $N(n) = n_1 \cdot n_2 \cdots n_p$, and the symbol $[A_n(f)]_{i,j}$ denotes that we are selecting the block (i_1, j_1) (which is a $(p-1)$ -level Toeplitz matrix); in this block we are selecting the block (i_2, j_2) which is a $(p-2)$ -level Toeplitz matrix; and so on. At the end of this process we find that the basic blocks are given by the $\{A_{i-j}\}$ formally determined in equation (22).

Now with regard to our definitions given in the second section, we have $\Omega = I^p$, $k(n) = N(n)$ and $\mu\{\cdot\} = m\{\cdot\}/[2\pi]^p$ (the Haar measure) where $m\{\cdot\}$ is the usual Lebesgue measure. Moreover, Tilli has shown [29] that the sequence $\{A_n(f)\}$ is distributed as f (for the case where $s = t = 1$, this was proved by Tyrtysnikov and Zamarashkin in a different way [32]).

Therefore, under the assumption that the Korovkin test is satisfied, by virtue of our results, we deduce that the sequence $\{\mathcal{P}_{N(n)}(A_n(f))\}$ is still distributed as f and that it is a good matrix approximation of $\{A_n(f)\}$, since their difference $\{\mathcal{P}_{N(n)}(A_n(f)) - A_n(f)\}$ is clustered at zero. Incidentally this fact establishes that the sequence of spaces $\{\mathcal{M}(U_n)\}$ is asymptotically close in the distribution sense to the sequence of linear spaces $\{\bigcup_{f \in L_1} A_n(f)\}$, and this fact translates to fast convergence when $\mathcal{P}_{N(n)}(A_n(f))$ is used as preconditioner for $A_n(f)$ in pcg-like methods [6, 20, 22, 10].

Moreover, as shown in [20, 10], the Korovkin test is satisfied in the case of all the well known multilevel trigonometric algebras (circulants [7], ω -circulants with $|\omega| = 1$, Hartley and the 8 sine/cosine transforms based algebras described in [16]) and in all the possible combinations of them [20, 10].

This fact has a natural interpretation in the context of the preconditioning problem [5] as discussed in Subsections 8.2 and 8.3. However, this aspect has been deeply analyzed in [6, 31, 5, 19, 10, 20, 22, 23].

Here we want to stress another application in approximation theory.

7.1. Applications to the numerical approximation of functions. We consider some approximation theory problems: let us take a complex-valued p -variate function f defined over I^p and let us suppose that $f \in L_1$. It is natural to suppose that we do not know its analytical expression, or that its analytical expression is not significant for computational purposes. On the other hand, we make the assumption that the Fourier coefficients $\{a_k\}$ are known: this fact is for instance not unusual in physics, where a function can be known only by a formal power series. We consider the following problems.

1. Given a function F continuous with bounded support, compute

$$\int_{I^p} F(|f(x)|)dx \quad \text{or} \quad \int_{I^p} F(f(x))dx$$

if f is either complex or real-valued.

2. Give a “plot” of the essential range of $|f|$ or of f itself if f is real-valued.

The first problem has a trivial but expensive computational solution. Take into account relation (4) with $k(n) = N(n)$ and $A_n(f)$ being Toeplitz as in (22), where $s = t = 1$.

Of course, the quoted formula has only a theoretical interest since the calculation of all the eigenvalues of the multilevel Toeplitz matrices $\{A_n(f)\}$ of large dimension is a very expensive task. However, in the preceding section we have proved that the eigen/singular values of the optimal circulant operator [6] $\mathcal{P}_{N(n)}(A_n(f)) = C_n(f)$ are distributed as $|f|$ or f depending on whether f is complex-valued or real-valued. Then, for any F continuous with bounded support, it follows that [31]

$$(23) \quad \lim_{n \rightarrow \infty} \frac{1}{N(n)} \sum_{i=1}^{N(n)} F(\lambda_i^{(n)}) = \frac{1}{[2\pi]^p} \int_{I^p} F(f(x)) dx$$

if f is real-valued and $\{\lambda_i^{(n)}\}$ is the complete set of the eigenvalues of $C_n(f)$, and

$$(24) \quad \lim_{n \rightarrow \infty} \frac{1}{N(n)} \sum_{i=1}^{N(n)} F(\sigma_i^{(n)}) = \frac{1}{[2\pi]^p} \int_{I^p} F(|f(x)|) dx$$

if f is complex-valued and $\{\sigma_i^{(n)}\}$ denotes the singular values of $C_n(f)$.

We remark that all the eigenvalues of $C_n(f)$ can be computed with total arithmetic cost $O(N(n) \log N(n))$ and $O(\log N(n))$ parallel steps [3], through use of FFTs, and therefore we have a cheap approximate numerical solution of problem 1.

For problem 2, with the help of the preceding ergodic formula, it is easy to see [27] that α belongs to the essential range [18] $\mathcal{ER}(|f|)$ of $|f|$ (or $\alpha \in \mathcal{ER}(f)$) if and only if for any $\epsilon > 0$

$$(25) \quad \lim_{n \rightarrow \infty} \frac{Z(\alpha - \epsilon, \alpha + \epsilon, n)}{Z(\alpha - \epsilon, \alpha + \epsilon, n/2)} = 2^p, \quad n \text{ even,}$$

where $Z(x, y, n)$ denotes the number of singular values or eigenvalues of $C_n(f)$ belonging to (x, y) with $x < y$.

We observe that an evident criticism of (25) is that we do not have information concerning the convergence rate of the limit quantity in the left-hand side of (25) (see also [27]). However, fixing ϵ and computing all the eigenvalues of $C_n(f)$ and $C_{n/2}(f)$ with n even and large enough, still within $O(N(n) \log N(n))$ arithmetic operations and $O(\log N(n))$ parallel steps, it is possible to decide if a value α belongs to the essential range of a given Lebesgue integrable function defined over a bounded domain.

8. NUMERICAL EXPERIMENTS

This section is organized in three parts. In the first we give numerical evidence of the procedure proposed in Section 7.1 to approximate functions. Since the procedure is based on a special instance of the first part of Theorem 6.2, we test the adherence of the numerical experiments with part 1 of Theorem 6.2 too.

In the second part we test the second part of Theorem 6.2 with regard to block Toeplitz matrices $\{A_n(f)\}$: in particular we test the clustering at zero of $\{A_n(f) - \mathcal{P}_{N(n)}(A_n(f))\}$ and of $\{(\mathcal{P}_{N(n)}(A_n(f)))^+ A_n(f) - I\}$ for different Gram matrix spaces. Here we assume $s \geq t$ and, given a matrix X , we denote its Moore-Penrose pseudo-inverse by X^+ .

In the third and last part we use the approximation of Toeplitz matrices by a special Gram matrix space as preconditioner in preconditioned conjugate gradient

(pcg) algorithms [1]. All the numerical experiments are in perfect agreement with the theoretical part.

In all the numerical experiments we use three choices of Gram matrix spaces that are well known in the relevant literature:

- The block circulant class [7] $\mathcal{M}(U_n)$,

$$U_n \equiv F_n = \frac{1}{\sqrt{n}} \left[e^{\frac{i2jk\pi}{n}} \right]_{j,k=0}^{n-1},$$

which is associated to the discrete Fourier transform (DFT). We recall that the DFT has a noteworthy and well recognized range of applications [5].

- The block τ class [10] $\mathcal{M}(U_n)$,

$$U_n \equiv S_n = \sqrt{\frac{2}{n+1}} \left[\sin \left(\frac{(j+1)(k+1)\pi}{n+1} \right) \right]_{j,k=0}^{n-1},$$

which is associated to the discrete sine transform - I (DST-I). This class of matrices is especially suitable in the numerical treatment of differential elliptic problems with Dirichlet boundary conditions [11].

- The block “cosine” class [7] $\mathcal{M}(U_n)$,

$$U_n \equiv T_n = \sqrt{\frac{2}{n}} \left[\eta_j \cos \left(\frac{(2k+1)j\pi}{2n} \right) \right]_{j,k=0}^{n-1},$$

with $\eta_0 = \frac{1}{2}$, $\eta_j = 1$ for $j = 1, 2, \dots, n - 1$. It is worth pointing out that this class, which is associated to the discrete cosine transform - III (DCT-III), has received some attention in the context of image deblurring [5].

8.1. Part 1. The spirit of this subsection is the following. We fix a function $f : I^p \rightarrow \mathbf{C}^{s \times t}$ and a matrix space in which we approximate (the Frobenius optimal approximation) the corresponding Toeplitz matrix $A_n(f)$ for different sizes n . We consider a value α and a tolerance ϵ : the procedure sketched in subsection 7.1 decides if “ α belongs to $\mathcal{ER}(g)$ within a precision ϵ ”. Here g is f or $|f|$ depending on whether f is real valued or not. More specifically, if the answer is negative, i.e., the limit in the left-hand side of (25) is not 2^p , then it means that $\alpha \notin \mathcal{ER}(g)$ (and for all $\beta \in (\alpha - \epsilon, \alpha + \epsilon)$ we have $\beta \notin \mathcal{ER}(g)$). If the answer is positive i.e., (25) holds true, then this means that $\beta \in \mathcal{ER}(g)$ for some $\beta \in (\alpha - \epsilon, \alpha + \epsilon)$.

All the experiments are done in MATLAB on a PC 486, and in all the tables the label * is a “dummy” symbol which is used for denoting badly defined quantities (division by zero).

Example 1. $s = t = 1$ and $f : I \rightarrow \mathbf{R}$ with

$$f(x) = \text{sign}(x).$$

We put $U_n \equiv F_n$ (circulant approximation) and $\epsilon = 10^{-1}$ while $\alpha \in \{-1, 1\}$ (Table 1), $\alpha = 8/9 = 0.8889$ (Table 2), $\alpha = 0.5$ (Table 3), $\alpha = 0$ (Table 4). The value Z denotes $Z(\alpha - \epsilon, \alpha + \epsilon, n)$ (see equation (25)), R denotes

$$\frac{Z(\alpha - \epsilon, \alpha + \epsilon, n)}{Z(\alpha - \epsilon, \alpha + \epsilon, n/2)}$$

and $L = \log_2(R)$. In light of (25) we know that

$$(26) \quad L \equiv L_n \rightarrow p = 1 \quad \text{iff} \quad \alpha \in \mathcal{ER}(f)$$

within a precision ϵ .

TABLE 1. (Example 1): $\alpha = \pm 1, U_n = F_n$

n	16	32	64	128	256	512
Z	7	15	31	63	127	255
R	*	2.14	2.06	2.03	2.01	2.00
L	*	1.09	1.04	1.02	1.01	1.00

TABLE 2. (Example 1): $\alpha = 8/9 = 0.8889, U_n = F_n$

n	16	32	64	128	256	512
Z	7	15	22	18	18	18
R	*	2.14	1.46	0.81	1	1
L	*	1.09	0.55	-0.28	0	0

TABLE 3. (Example 1): $\alpha = 0.5, U_n = F_n$

n	16	32	64	128	256	512
Z	0	0	0	0	0	0
R	*	*	*	*	*	*
L	*	*	*	*	*	*

TABLE 4. (Example 1): $\alpha = 0, U_n = F_n$

n	16	32	64	128	256	512
Z	0	2	2	2	2	2
R	*	*	1	1	1	1
L	*	*	0	0	0	0

Table 1 shows very clearly that ± 1 belongs to $\mathcal{ER}(f)$ while Tables 3 and 4 show that $\{0.5, 0\} \cap \mathcal{ER}(f) = \emptyset$. Table 2 is more interesting: the value $8/9 = 0.8889 \notin \mathcal{ER}(f)$, but the interval $(8/9 - \epsilon, 8/9 + \epsilon) = (0.7889, 0.9889)$ is very close to 1 and $m\{x \in [-\pi, \pi] : f(x) = 1\} = \pi$. This little pathology is evident since for $n = 16, 32$ it seems that $8/9 \in \mathcal{ER}(f)$. However it is enough to inspect Table 2 for $n \geq 64$ to recover the correct negative information.

Example 2. $s = t = 1$ and $f : I \rightarrow \mathbf{R}$ with

$$f(x) = x^2.$$

We use $U_n \equiv S_n$ (τ approximation) and $\epsilon = 10^{-1}$. Moreover, the meaning of the quantities Z, R and L is the same as in the preceding example. In Table 5 we check that $\alpha = 2$ belongs to $\mathcal{ER}(f)$, and the answer is positive for $n \geq 128$. In Table 6 we make a little change: we test that $[\gamma, \delta] = [0.25, 1]$ contains some points of $\mathcal{ER}(f)$ and we use the ergodic formula related to part 1. of Theorem 6.2 to compute $m\{x \in [-\pi, \pi] : f(x) \in [\gamma, \delta] = [0.25, 1]\} = 1$. In fact, if

$$m\{x \in [-\pi, \pi]^p : f(x) = \gamma\} = 0 \quad \text{and} \quad m\{x \in [-\pi, \pi]^p : f(x) = \delta\} = 0,$$

TABLE 5. (Example 2): $\alpha = 2, U_n = S_n$

n	16	32	64	128	256	512
Z	0	1	1	3	6	12
R	*	*	1	3	2	2
L	*	*	0	1.58	1	1

TABLE 6. (Example 2): $[\gamma, \delta] = [0.25, 1], r = 1, U_n = S_n$

n	16	32	64	128	256	512
Z	3	5	10	20	41	82
R	*	1.66	2	2	2.05	2
L	*	0.73	1	1	1.03	1
$s(n)$	1.1781	0.9817	0.9817	0.9817	1.0063	1.0063

TABLE 7. (Example 2): $\alpha = 10, U_n = F_n$

n	16	32	64	128	256
Z	16	32	64	128	256
R	*	2	2	2	2
L	*	1	1	1	1

TABLE 8. (Example 2): $\alpha = 1, U_n = F_n$

n	16	32	64	128	256
Z	0	30	62	126	254
R	*	*	2.06	2.03	2.01
L	*	*	1.04	1.02	1.01

then part 1 of Theorem 6.2 implies that

(27)

$$\lim_{n \rightarrow \infty} \frac{1}{N(n)} \#\{j : \lambda_j(\mathcal{P}_{N(n)}(A_n(f))) \in [\gamma, \delta]\} = \frac{1}{[2\pi]^p} m\{x \in [-\pi, \pi]^p : f(x) \in [\gamma, \delta]\}.$$

In Table 6 the quantity $s(n)$ denotes $\frac{2\pi}{n}Z$. By virtue of (27) with $p = 1$, we know that $s(n)$ must converge to $1 = m\{x \in [-\pi, \pi] : f(x) \in [0.25, 1]\}$, and this forecast is fully confirmed by the last row of Table 6.

Example 3. $s = t = 2$ and $f : I \rightarrow \mathbf{C}^{2 \times 2}$ with

$$f(x) = \frac{\sqrt{2}}{2} \begin{pmatrix} 10 & -\mathbf{i} \cdot \text{sign}(x) \\ 10 & \mathbf{i} \cdot \text{sign}(x) \end{pmatrix}.$$

We set $U_n \equiv F_n$ (circulant approximation) and $\epsilon = 10^{-1}$. The singular values of the function $f(x)$ are two constant functions $\sigma_{\max}(f(x)) = 10$ and $\sigma_{\min}(f(x)) = 1$. From Tables 7 and 8 it is evident that (except for two outlying singular values) one half of the singular values of the $2n \times 2n$ matrix $\mathcal{P}_n(A_n(f))$ lie in $(1 - \epsilon, 1 + \epsilon)$ and the other half lie in $(10 - \epsilon, 10 + \epsilon)$.

Therefore, by virtue of (25) and (26), we deduce that 1 and 10 belong to $\mathcal{ER}(f) = \mathcal{ER}(\sigma_1(f(x))) \cup \mathcal{ER}(\sigma_2(f(x)))$ (see [28]). Indeed, since all the singular values

(except two of them) are in an ϵ -neighbourhood of 1 or 10, we know that $\mathcal{ER}(f)$ exactly coincides with $\{1, 10\}$.

Example 4. $s = t = 2$ and $f : I \rightarrow \mathbf{R}^{2 \times 2}$, with

$$f(x) = \begin{pmatrix} x^2 & 0 \\ 1 & \text{sign}(x) \end{pmatrix}.$$

We use the circulant approximation; that is, $U_n \equiv F_n$. From a direct explicit calculation, we obtain that

$$\sigma_{\max}(f(x)) = \sqrt{\frac{2 + x^4 + \sqrt{4 + x^8}}{2}} \geq \sqrt{2}$$

and

$$\sigma_{\min}(f(x)) = \sqrt{\frac{2 + x^4 - \sqrt{4 + x^8}}{2}} \sim \frac{x^2}{\sqrt{2}}.$$

We apply an instance of the ergodic formula in part 1 of Theorem 6.2 to compute $m\{x \in [-\pi, \pi] : \sigma_{\min}(f(x)) \in [0, 0.1]\} \approx 2\sqrt{\sqrt{2} \cdot 0.1}$. In fact, if $m\{x \in [-\pi, \pi]^p : \sigma_{\min}(f(x)) = \gamma \text{ or } \sigma_{\max}(f(x)) = \gamma\} = 0$ and $m\{x \in [-\pi, \pi]^p : \sigma_{\min}(f(x)) = \gamma \text{ or } \sigma_{\max}(f(x)) = \delta\} = 0$, then part 1 of Theorem 6.2 implies that

$$\begin{aligned} (28) \quad & \lim_{n \rightarrow \infty} \frac{1}{N(n)} \#\{j : \sigma_j(\mathcal{P}_{N(n)}(A_n(f))) \in [\gamma, \delta]\} \\ &= \frac{1}{[2\pi]^p} (m\{x \in [-\pi, \pi]^p : \sigma_{\min}(f(x)) \in [\gamma, \delta]\} \\ & \quad + m\{x \in [-\pi, \pi]^p : \sigma_{\max}(f(x)) \in [\gamma, \delta]\}). \end{aligned}$$

As in Table 6, in Table 9 the quantity $s(n)$ is $\frac{2\pi}{n}Z$. By virtue of (28) with $p = 1$ and since $\sigma_{\max}(f(x)) \geq \sqrt{2}$, we know that $s(n)$ must converge to

$$m\{x \in [-\pi, \pi] : \sigma_{\min}(f(x)) \in [0, 0.1]\} \approx 2\sqrt{\sqrt{2} \cdot 0.1} = 0.7521,$$

and this forecast is fully confirmed by the last row of Table 9.

TABLE 9. (Example 4): $[\gamma, \delta] = [0, 0.1]$, $r \approx 0.7521$, $U_n = F_n$

n	16	32	64	128	256
Z	2	4	8	16	31
R	*	2	2	2	1.93
L	*	1	1	1	0.95
$s(n)$	0.7854	0.7854	0.7854	0.7854	0.7609

8.2. Part 2. From part 2 of Theorem 6.2 we know that the sequence

$$\{\mathcal{P}_{k(n)}(A_n(f)) - A_n(f)\}$$

is clustered at zero. When $s \geq t$ and when $\mu\{x \in \Omega : \sigma_{\min}(f(x)) = 0\} = 0$ it is easy to prove that the new sequence $\{(\mathcal{P}_{k(n)}(A_n(f)))^+ A_n(f) - I\}$ is clustered at zero too. This fact has a practical importance in the context of the preconditioning. Here we report a special and more detailed version of this statement in the case where $A_n(\cdot)$ is the multilevel block Toeplitz operator.

Proposition 8.1 ([10, 23]). *Let us assume that $f \in C(I^p, \mathbf{C}^{s \times t})$ with $s \geq t$.*

p1.: *If $p = 1$ and the minimal singular value $\sigma_{\min}(f(x))$ of f is strictly positive, then for any $\epsilon > 0$*

$$(\mathcal{P}_{N(n)}(A_n(f)))^+ A_n(f)$$

has singular values in $(1 - \epsilon, 1 + \epsilon)$ except, at most, $N_\epsilon = O(1)$ outliers.

p2.: *If $p > 1$ and the minimal singular value $\sigma_{\min}(f(x))$ of f is strictly positive, then for any $\epsilon > 0$*

$$(\mathcal{P}_{N(n)}(A_n(f)))^+ A_n(f)$$

has singular values in $(1 - \epsilon, 1 + \epsilon)$ except, at most,

$$N_\epsilon = O(N(n) \sum_{i=1}^p n_i^{-1})$$

outliers.

p3.: *If the set of the zeros of the minimal singular value $\sigma_{\min}(f(x))$ of f has zero Lebesgue measure, then for any $\epsilon > 0$*

$$(\mathcal{P}_{N(n)}(A_n(f)))^+ A_n(f)$$

has singular values in $(1 - \epsilon, 1 + \epsilon)$ except, at most, $N_\epsilon = o(N(n))$ outliers.

In the following set of examples we test numerically part 2 of Theorem 6.2 and the previous proposition.

Example 5. $s = 2$, $t = 1$ and $f : I \rightarrow \mathbf{R}^{2 \times 1}$ is such that

$$f(x) = \begin{pmatrix} 5 + x^2 \\ \text{sign}(x) \end{pmatrix}.$$

Example 6. $s = 2$, $t = 1$ and $f : I \rightarrow \mathbf{R}^{2 \times 1}$ is such that

$$f(x) = \begin{pmatrix} 2 + 2 \cos(x) \\ \text{sign}(x)(2 + 2 \cos(x)) \end{pmatrix}.$$

Tables 10 and 11 concern Example 5, and Tables 12 and 13 concern Example 6. In particular, in Tables 10 and 12 we report the behaviour of the singular values of $\{\mathcal{P}_{N(n)}(A_n(f)) - A_n(f)\}$, and in the first row we report the behaviour of the “biggest outlier”.

In Tables 11 and 13 we report the behaviour of the singular values of

$$\{(\mathcal{P}_{N(n)}(A_n(f)))^+ A_n(f)\},$$

and in the first two rows we report the behaviour of the “smallest outlier” and the “biggest outlier”. The symbol N denotes the number of singular values belonging to a cluster of radius $\epsilon = 10^{-1}$, while R denotes the related percentage with respect to the total number of singular values.

From Tables 10 and 13, it is evident that the singular values cluster as n increases. However, there are differences between Example 5 and Example 6.

In Table 11 the singular values of $\{(\mathcal{P}_{N(n)}(A_n(f)))^+ A_n(f)\}$ belong to $(0.7, 1, 3)$ for all the considered choices of n , and we have a constant number of outliers with regard to $\epsilon = 10^{-1}$. In Table 13 for $n = 128$ we have $\sigma_n^{(n)} = 0.03$ and $\sigma_1^{(n)} = 14.55$, and it is clear that $\sigma_n^{(n)} \rightarrow 0$ and $\sigma_1^{(n)} \rightarrow \infty$ as n tends to infinity. Moreover, in Table 13 the number of the outliers is not constant even if it grows very slowly.

TABLE 10. (Example 5): $A_n - \mathcal{P}_n, U_n = F_n$

n	16	32	64	128
$\sigma_1^{(n)}$	2.21	2.26	2.28	2.29
ϵ	0.1	0.1	0.1	0.1
N	2	21	52	117
R	12%	65%	81%	91%

TABLE 11. (Example 5): $\mathcal{P}_n^+ A_n, U_n = F_n$

n	16	32	64	128
$\sigma_n^{(n)}$	0.80	0.79	0.79	0.79
$\sigma_1^{(n)}$	1.26	1.27	1.27	1.28
ϵ	0.1	0.1	0.1	0.1
N	14	30	62	126
R	87%	93%	96%	98%

TABLE 12. (Example 6): $A_n - \mathcal{P}_n, U_n = F_n$

n	16	32	64	128
$\sigma_1^{(n)}$	2.50	2.66	2.79	2.91
ϵ	0.1	0.1	0.1	0.1
N	4	19	50	114
R	25%	59%	78%	89%

TABLE 13. (Example 6): $\mathcal{P}_n^+ A_n, U_n = F_n$

n	16	32	64	128
$\sigma_n^{(n)}$	0.20	0.12	0.07	0.03
$\sigma_1^{(n)}$	3.45	5.34	8.72	14.55
ϵ	0.1	0.1	0.1	0.1
N	9	23	51	111
R	56%	71%	79%	86%

The reason for this evident difference is due to the “conditioning” of the generating functions. Table 11 concerns Example 5, where the conditioning of f is given by

$$(29) \quad \frac{\text{ess sup}_x \sigma_{\max}(f(x))}{\text{ess inf}_x \sigma_{\min}(f(x))} = \sqrt{\frac{(5 + \pi^2)^2 + 1}{26}} = 2.9227.$$

Moreover, $\sigma_{\min}(A_n(f)) \geq \sigma_{\min}(A_n(\hat{f}))$, where $\hat{f}(x) = 5 + x^2$ is a 1×1 submatrix of f (see Remark 2.4 in [28]). But $\sigma_{\min}(A_n(\hat{f})) > 5$ (see [26]) and $\sigma_{\max}(A_n(f)) < \text{ess sup}_x \sigma_{\max}(f(x)) = \sqrt{(5 + \pi^2)^2 + 1} = 14.9032$ (see Theorems 2.1 and 2.4 in [28]). In conclusion, the matrices of the sequence $\{A_n(f)\}$ are well conditioned and the

related spectral condition numbers are bounded by

$$(30) \quad \frac{\sqrt{(5 + \pi^2)^2 + 1}}{5} = 2.9806.$$

Notice that in this specific example the condition numbers of f and $A_n(f)$ displayed in (29) and (30) are very close.

On the other hand, the generating function used in Example 7 is “ill conditioned”, since

$$\frac{\text{ess sup}_x \sigma_{\max}(f(x))}{\text{ess inf}_x \sigma_{\min}(f(x))} = \infty.$$

As a consequence of the fact that $\{A_n(f)\}$ is distributed as f [29] we have that $\{A_n(f)\}$ and $\{\mathcal{P}_{N(n)}(A_n(f))\}$ are asymptotically ill conditioned. In light of the analysis given in [9] we deduce a substantial weakening of the clustering properties of the sequence $\{(\mathcal{P}_{N(n)}(A_n(f)))^+ A_n(f)\}$ and its asymptotic ill-conditioning (compare part p3 of Proposition 8.1 with [9]).

8.3. Part 3. Here we just want to give numerical evidence of the superlinear convergence rate attained by a pcg-like method based on the optimal Toeplitz approximation in a “cosine” space.

Therefore we consider real matrix-valued generating functions $f(x)$ of one variable, leading to one-level block Toeplitz matrices $A_n(f)$ with nonstructured rectangular blocks. For several increasing values of n , we then perform the following calculations.

1. Construction of the preconditioner $\mathcal{P}_n(A_n(f))$, where $U_n = T_n$ is the DCT-III.
2. Computation of the singular values of the preconditioned matrix

$$(31) \quad (\mathcal{P}_n(A_n(f)))^+ A_n(f),$$

which are expected to have a proper cluster at 1 by part p1 of Proposition 8.1. In fact, a neighborhood of center 1 and decreasing radius ϵ_n is detected, containing all the singular values corresponding to the size n with the exception of a constant number of outliers.

3. Solution of the normal equations related to the least-squares problem

$$(32) \quad \min_{\mathbf{x}} \|A_n(f)\mathbf{x} - \mathbf{b}\|_2,$$

where \mathbf{b} is the vector of all ones, by means of a pcg method with starting vector of all zeros and $\mathcal{P}_n(A_n(f))$ as a preconditioner. The iterations stop when the normal equation residual norm is reduced by a factor less than 10^{-12} .

Example 7 ([10]). $s = 2$, $t = 1$ and $f : I \rightarrow \mathbf{R}^{2 \times 1}$ is such that

$$f(x) = \begin{pmatrix} 1 + x^2 \\ |x|(2 - 2 \cos x) \end{pmatrix}.$$

Since the first component of f is strictly positive, the hypothesis of part p1 of Proposition 8.1 ($\sigma_{\min}(f) > 0$) is fulfilled. For every n , the preconditioned matrix (31) has 2 singular values $\sigma_1^{(n)}$ and $\sigma_2^{(n)}$ far from 1, and the remaining ones lie in a ϵ_n -neighborhood of 1, as sketched in Table 14.

The pcg method applied to the least-squares problem (32) achieves the desired precision 10^{-12} after k_n iterations, where k_n tends to decrease with respect to n as displayed in Table 15.

TABLE 14. (Example 7): $\mathcal{P}_n^+ A_n, U_n = T_n$

n	16	32	64	128
$\sigma_1^{(n)}$	2.02	2.16	2.24	2.27
$\sigma_2^{(n)}$	1.99	2.16	2.24	2.27
ϵ_n	0.2964	0.1682	0.0884	0.0575
N	14	30	62	126
R	87%	93%	96%	98%

TABLE 15. (Example 7): pcg iterations, $U_n = T_n$

n	16	32	64	128	256	512
k_n	10	15	13	12	10	9

9. CONCLUSIONS

A general analysis of clustering the sequences $\{\mathcal{P}_{k(n)}(A_n(f))\}$ and $\{A_n(f) - \mathcal{P}_{k(n)}(A_n(f))\}$ is made under very mild assumptions on $\{A_n(f)\}$. We used the LPOs theory, an ergodic assumption (4), and the Korovkin theorem. Our results have applications in constructive approximation and numerical linear algebra. In particular, the Korovkin approach has been very successful for dealing with Toeplitz linear systems. On the other hand, future work should concentrate more attention on the rich variety of matrix-structures coming from the discretization of PDEs with boundary conditions and of elliptic type. In fact, these structures are often related to LPOs [25, 21] and fulfill an assumption in the style of equation (4) [30, 21].

ACKNOWLEDGMENT

Warm thanks go to Milvio Capovani for teaching me the importance of the “structures” in pure and applied computational mathematics.

REFERENCES

- [1] O. Axelsson and G. Lindskög, “The rate of convergence of the preconditioned conjugate gradient method”, *Numer. Math.*, **52** (1986), pp. 499–523. MR **88a**:65037b
- [2] R. Bhatia, *Matrix Analysis*. Springer, New York, 1997. MR **98i**:15003
- [3] D. Bini and V. Pan, *Polynomial and Matrix Computations, Vol. 1: Fundamental Algorithms*. Birkhäuser, Boston, 1994. MR **95k**:65003
- [4] N. Bonanni, *Proprietà spettrali e computazionali di algebre di matrici*. Graduate Thesis in Computer Science, University of Pisa, 1993.
- [5] R.H. Chan and M. Ng, “Conjugate gradient methods for Toeplitz systems”, *SIAM Rev.*, **38** (1996), pp. 427–482. MR **97i**:65048
- [6] T.F. Chan, “An optimal circulant preconditioner for Toeplitz systems”, *SIAM J. Sci. Stat. Comp.*, **9** (1988), pp. 766–771. MR **89e**:65046
- [7] P. Davis, *Circulant Matrices*. John Wiley and Sons, New York, 1979. MR **81a**:15003
- [8] R. De Vore and G. Lorentz, *Constructive Approximation*. Springer-Verlag, Berlin, 1993. MR **95f**:41001
- [9] F. Di Benedetto and S. Serra Capizzano, “A unifying approach to abstract matrix algebra preconditioning”, *Numer. Math.*, **82-1** (1999), pp. 57–90. CMP 99:10
- [10] F. Di Benedetto and S. Serra Capizzano, “Optimal and superoptimal matrix algebra operators”, *TR nr. 360, Dept. of Mathematics - Univ. of Genova* (1997).
- [11] G. Fiorentino and S. Serra, “Fast parallel solvers for elliptic problems”, *Comput. Math. Appl.*, **32** (1996), pp 61–68. MR **97g**:65216

- [12] W. Gautschi, “The condition of Vandermonde-like matrices involving orthogonal polynomials”, *Linear Algebra Appl.*, **52/53** (1983), pp. 293–300. MR **84i**:65043
- [13] U. Grenander and G. Szegő, *Toeplitz Forms and Their Applications*. Second Edition, Chelsea, New York, 1984. MR **88b**:42031
- [14] E. Isaacson and H. Keller, *Analysis of Numerical Methods*. John Wiley and Sons, New York, 1966. MR **34**:924
- [15] C. Jordan, *Cours d’Analyse de l’Ecole Polytechnique: Vol. I*. Gauthier-Villars, Paris, France, 1909. Reprint, CMP 93:03
- [16] T. Kailath and V. Olshevsky, “Displacement structure approach to discrete-trigonometric transform based preconditioners of G. Strang type and T. Chan type”, *Proc. “Workshop on Toeplitz matrices”* Cortona (Italy), September 1996. *Calcolo*, **33** (1996), pp. 191–208. CMP 98:15
- [17] P.P. Korovkin, *Linear Operators and Approximation Theory* (English translation). Hindustan Publishing Co., Delhi, 1960. MR **27**:561
- [18] W. Rudin, *Real and Complex Analysis*. McGraw-Hill, Singapore, 1986. MR **88k**:00002
- [19] S. Serra, “Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems”, *Math. Comp.*, **66** (1997), pp. 651–665. MR **97h**:65056
- [20] S. Serra, “A Korovkin-type Theory for finite Toeplitz operators via matrix algebras”, *Numer. Math.*, **82-1** (1999), pp. 117–142. CMP 99:10
- [21] S. Serra Capizzano, “An ergodic theorem for classes of preconditioned matrices”, *Linear Algebra Appl.*, **282** (1998), pp. 161–183. MR **99k**:65040
- [22] S. Serra Capizzano, “A Korovkin based approximation of multilevel Toeplitz matrices (with rectangular unstructured blocks) via multilevel Trigonometric matrix spaces”, *SIAM J. Numer. Anal.*, **36-6** (1999), pp. 1831–1857. CMP 2000:03
- [23] S. Serra Capizzano, “Approximation of multilevel Toeplitz matrices via multilevel Trigonometric matrix spaces and application to the preconditioning”, *Calcolo*, **36** (1999), pp. 187–213.
- [24] S. Serra Capizzano, “Korovkin Theorems and linear positive Gram matrix algebras approximation of Toeplitz matrices”, *Linear Algebra Appl.*, **284** (1998), pp. 307–334. MR **99h**:65088
- [25] S. Serra Capizzano and C. Tablino Possio, “Spectral and structural analysis of high order finite difference matrices for Elliptic Operators”, *Linear Algebra Appl.* **293** (1999), 85–131. CMP 99:15
- [26] S. Serra, “On the extreme eigenvalues of Hermitian (block) Toeplitz matrices”, *Linear Algebra Appl.*, **270** (1998), pp. 109–129. MR **98k**:15034
- [27] S. Serra Capizzano, “Some theorems on linear positive operators and functionals and their applications”, *TR nr. 26, LAN, Dept. of Mathematics - Univ. of Calabria* (1997). *Comput. Math. Appl.*, in press.
- [28] S. Serra Capizzano and P. Tilli, “Extreme singular values and eigenvalues of non-Hermitian block Toeplitz matrices”, *J. Comput. and Appl. Math.* **108** (1999), 113–130. CMP 99:17
- [29] P. Tilli, “A note on the spectral distribution of Toeplitz matrices”, *Linear Multilin. Algebra*, **45** (1998), pp. 147–159. MR **99j**:65063
- [30] P. Tilli, “Locally Toeplitz sequences: spectral properties and applications”, *Linear Algebra Appl.*, **278** (1998), pp. 91–120. MR **99g**:47057
- [31] E. Tyrtyshnikov, “A unifying approach to some old and new theorems on distribution and clustering”, *Linear Algebra Appl.*, **232** (1996), pp. 1–43. MR **96m**:15018
- [32] E. Tyrtyshnikov and N. Zamarashkin, “Spectra of multilevel Toeplitz matrices: advanced theory via simple matrix relationships”, *Linear Algebra Appl.*, **270** (1998), pp. 15–27. MR **98i**:65030
- [33] H. Widom, *Toeplitz matrices*. In *Studies in real and complex analysis*, I. Hirschman Jr. Ed., Math. Ass. Amer., 1965. MR **32**:1080

DIPARTIMENTO DI ENERGETICA, VIA LOMBROSO 6/17, 50134 FIRENZE, ITALY; DIPARTIMENTO DI INFORMATICA, CORSO ITALIA 40, 56100 PISA, ITALY
E-mail address: `serra@mail.dm.unipi.it`