The theory of Generalized Locally Toeplitz sequences: a review, an extension, and a few representative applications

Carlo Garoni[a,b] and Stefano Serra-Capizzano[b,c]

[a]University of Roma 'Tor Vergata', Department of Mathematics, Via della Ricerca Scientifica, 00133 Roma, Italy.
Email: garoni@mat.uniroma2.it.

[b]University of Insubria, Department of Science and High Technology, Via Valleggio 11, 22100 Como, Italy.
Email: carlo.garoni@uninsubria.it, stefano.serrac@uninsubria.it.

[c]Uppsala University, Department of Information Technology, Division of Scientific Computing,
Box 337, SE-751 05 Uppsala, Sweden. Email: stefano.serra@it.uu.se.

August 14, 2015

Abstract

We review and extend the theory of Generalized Locally Toeplitz (GLT) sequences, which goes back to the pioneering work by Tilli on Locally Toeplitz (LT) sequences and was developed by the second author during the last decade. Informally speaking, a GLT sequence $\{A_n\}_n$ is a sequence of matrices with increasing size, equipped with a Lebesgue-measurable function κ (the so-called symbol). This function characterizes the asymptotic singular value distribution of $\{A_n\}_n$; in the case where the matrices A_n are Hermitian, it also characterizes the asymptotic eigenvalue distribution of $\{A_n\}_n$. Three fundamental examples of GLT sequences are: (i) the sequence of Toeplitz matrices generated by a function f in L^1; (ii) the sequence of diagonal sampling matrices containing the evaluations of a Riemann-integrable function a over a uniform grid; (iii) any zero-distributed sequence, i.e., any sequence of matrices possessing an asymptotic singular value distribution characterized by the identically zero function. The symbol of the GLT sequence (i) is f, the symbol of the GLT sequence (ii) is a, and the symbol of any GLT sequence of the form (iii) is 0. The set of GLT sequences is a *-algebra. More precisely, suppose that $\{A_n^{(1)}\}_n, \ldots, \{A_n^{(r)}\}_n$ are GLT sequences with symbols $\kappa_1, \ldots, \kappa_r$, and let $A_n = \text{ops}(A_n^{(1)}, \ldots, A_n^{(r)})$ be a matrix obtained from $A_n^{(1)}, \ldots, A_n^{(r)}$ by means of certain algebraic operations 'ops', such as linear combinations, products, inversions and Hermitian transpositions; then, $\{A_n\}_n$ is a GLT sequence with symbol $\kappa = \text{ops}(\kappa_1, \ldots, \kappa_r)$.

As already proved in several contexts, the theory of GLT sequences is a powerful apparatus for computing the asymptotic singular value and eigenvalue distribution of the discretization matrices A_n arising from the numerical approximation of continuous problems, such as integral equations and, especially, partial differential equations. Indeed, when the discretization parameter n tends to infinity, the discretization matrices A_n give rise to a sequence $\{A_n\}_n$, which often turns out to be a GLT sequence.

However, in this work we are not concerned with the applicative interest of the theory of GLT sequences, for which we limit to outline some of the numerous applications and to refer the reader to the available literature. On the contrary, we focus on the mathematical foundations. We propose slight (but relevant) modifications of the original definitions, and we introduce for the first time the concept of LT sequences in the multivariate/multilevel setting. With the new definitions, based on the notion of approximating class of sequences, we are able to enlarge the applicability of the theory, by generalizing and/or simplifying a lot of key results. In particular, we remove a technical hypothesis concerning the Riemann-integrability of the so-called 'weight functions', which appeared in the statement of many spectral distribution and algebraic results for GLT sequences. Moreover, we provide a formal and detailed proof of the fact that the sequences of matrices mentioned in items (i)–(iii) fall in the class of LT sequences. Several versions of this result were already present in previous papers, but only partial proofs were given.

As a final step, we extend the theory of GLT sequences. We first prove an approximation result, which is particularly useful to show that a given sequence of matrices is a GLT sequence. By using this result, we provide a new and easier proof of the fact that $\{A_n^{-1}\}_n$ is a GLT sequence with symbol κ^{-1} whenever $\{A_n\}_n$ is a GLT sequence of invertible matrices with symbol κ and $\kappa \neq 0$ almost everywhere. Finally, using again the approximation result, we prove that $\{f(A_n)\}_n$ is a GLT sequence with symbol $f(\kappa)$, as long as $f : \mathbb{R} \to \mathbb{R}$ is continuous and $\{A_n\}_n$ is a GLT sequence of Hermitian matrices with symbol κ. This has important implications, e.g., in proving that the geometric mean of two GLT sequences is still a GLT sequence, with symbol given by the the geometric mean of the symbols.

Contents

1 Introduction

We review and extend the theory of Generalized Locally Toeplitz (GLT) sequences, which stems from Tilli's work on Locally Toeplitz (LT) sequences [52] and from the theory of classical Toeplitz operators [14, 31, 51], and was developed by the second author in [44, 45]. In Section 1.1, we mention some of the main applications of this theory. In Section 1.2, we summarize its main features. In Section 1.3, we describe the main contributions of this work.

1.1 Applications of the theory of GLT sequences

As already proved in several contexts, the theory of GLT sequences is a powerful apparatus for computing/analyzing the asymptotic spectral distribution of the discretization matrices arising from the numerical approximation of continuous problems, such as Integral Equations (IEs) and, especially, Partial Differential Equations (PDEs). Let us explain this point in more detail. When discretizing a linear PDE by means of a linear numerical method, the actual computation of the numerical solution u_n reduces to solving a linear system $A_n\mathbf{u}_n = \mathbf{b}_n$. The size d_n of this linear system increases when the discretization parameter n tends to infinity. Hence, what we actually have is not just a single linear system, but an whole sequence of linear systems with increasing size; and what is often verified in practice is that the sequence of discretization matrices $\{A_n\}_n$ enjoys an asymptotic spectral distribution in the Weyl sense, which is somehow related to the spectrum of the differential operator associated with the considered PDE. More precisely, it often happens that, for a large set of test functions F (usually, for all continuous functions F with bounded support), the following limit relation holds:

$$\lim_{n\to\infty}\frac{1}{d_n}\sum_{j=1}^{d_n}F(\lambda_j(A_n))=\frac{1}{\mu_k(D)}\int_D F(f(\mathbf{x}))d\mathbf{x}, \tag{1.1}$$

where $\lambda_j(A_n),\ j=1,\ldots,d_n$, are the eigenvalues of A_n, μ_k is the Lebesgue measure in $\mathbb{R}^k$, and $f:D\subset\mathbb{R}^k\to\mathbb{C}$ is a measurable function. In this situation, f is referred to as the spectral symbol of the sequence $\{A_n\}_n$. The spectral symbol provides a 'compact' and accurate description of the asymptotic spectral distribution of the discretization matrices A_n. Indeed, the informal (but important) meaning behind (1.1) can be summarized as follows: assuming that f is at least Riemann-integrable and n is large enough, a suitable ordering of the eigenvalues $\lambda_j(A_n)$, $j=1,\ldots,d_n$, assigned in correspondence with a uniform grid on D, reconstructs approximately the surface $\mathbf{x}\to f(\mathbf{x})$ (the graph of f). In other words, the spectrum of A_n 'behaves' (asymptotically) like a uniform sampling of f over D. The theory of GLT sequences, in combination with the results of [27, 30] concerning the asymptotic spectral distribution of perturbed sequences of matrices, is one of the most powerful and successful tools for computing the spectral symbol f. Indeed, the sequence of discretization matrices $\{A_n\}_n$ turns out to be a GLT sequence, or a perturbation of a GLT sequence, for many classes of PDEs and numerical methods. We refer the reader to [44, 45] for applications of the theory of GLT sequences in the context of Finite Difference discretizations of PDEs; to [6, 23, 45, 47] for the Finite Element and collocation settings; and to [19, 25, 26] for recent applications to the case of B-spline Isogeometric Analysis (IgA) approximations of PDEs, both in the collocation and Galerkin frameworks.[1] We also refer the reader to [1, 41] for a look at the GLT approach to deal with sequences of matrices coming from the approximation of IEs.

At this point, it is worth emphasizing that the discretization matrices A_n arising from the numerical approximation of PDEs are often ill-conditioned for large n. In fact, their condition number diverges when $n\to\infty$. The knowledge of the spectral symbol f, which can be attained through the theory of GLT sequences, is not only interesting in itself, but could also be exploited in two different ways: (a) to analyze/predict the convergence rate of known iterative methods, such as preconditioned Krylov and multigrid methods, when they are applied to the ill-conditioned linear systems with coefficient matrix A_n; (b) to design effective preconditioners and iterative solvers for these linear systems. The reason is clear: the convergence properties of general purpose iterative methods depend on the spectral features of the matrix to which they are applied. Hence, the spectral information provided by f can be conveniently used for designing fast iterative solvers and/or analyzing their convergence properties. In this respect, we recall that recent estimates on the superlinear convergence of the Conjugate Gradient method are strictly related to the asymptotic spectral distribution of the matrices to which the method is applied; see [5]. We also refer the reader to [16, 17, 18] for recent developments in the IgA framework, where the spectral symbol was exploited to design ad hoc iterative solvers for the IgA discretization matrices.

Despite the numerous applications of the theory of GLT sequences, we should say now that the focus of this work is not on the applications. On the contrary, we will be interested only in the mathematical foundations of the theory, as explained in Section 1.3. However, in order to give an idea of the applicative interest, we will outline some applications at the end, in Section 6.

1.2 Summary of the theory of GLT sequences

Informally speaking, a GLT sequence $\{A_n\}_n$ is a sequence of matrices with increasing size, equipped with a Lebesgue-measurable (complex-valued) function κ. This function is referred to as the symbol of $\{A_n\}_n$ and it is defined over a domain D of the form $[0,1]^d\times[-\pi,\pi]^d,\ d\geq 1$. Due to the experience coming from the applications and to a detected analogy between the theory of GLT sequences and the Fourier Analysis,[2] a point of $D=[0,1]^d\times[-\pi,\pi]^d$ is usually denoted by $(\mathbf{x},\boldsymbol{\theta})$, where $\mathbf{x}=(x_1,\ldots,x_d)$ are the so-called 'physical variables', while $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_d)$ are the 'Fourier variables'.

The main theoretical properties of GLT sequences are summarized in items **GLT 1 – GLT 9**, which will be made more precise in Section 5.6, once the theory of GLT sequences has been developed. In the following, we write $\{A_n\}_n\sim_{\rm GLT}\kappa$ to indicate that $\{A_n\}_n$ is a GLT sequence with symbol κ.

GLT 1. The symbol $\kappa:[0,1]^d\times[-\pi,\pi]^d\to\mathbb{C}$ of a GLT sequence $\{A_n\}_n$ characterizes the asymptotic singular value distribution of $\{A_n\}_n$. This means that, for all continuous functions F with bounded support, we have

$$\lim_{n\to\infty}\frac{1}{d_n}\sum_{j=1}^{d_n}F(\sigma_j(A_n))=\frac{1}{(2\pi)^d}\int_{[0,1]^d\times[-\pi,\pi]^d}F(|\kappa(\mathbf{x},\boldsymbol{\theta})|)d\mathbf{x}d\boldsymbol{\theta}, \tag{1.2}$$

[1] IgA is a modern and successful paradigm introduced in [15, 33] for analyzing problems governed by PDEs. Its goal is to improve the connection between numerical simulation and Computer-Aided Design (CAD) systems. The main idea in IgA is to use directly the geometry provided by CAD systems and to approximate the unknown solutions of PDEs by the same type of functions (usually, B-splines or NURBS). In this way, it is possible to save about 80% of the CPU time, which is normally employed in the translation between two different languages (e.g., between Finite Elements and CAD or between Finite Differences and CAD).

[2] It was noted in [45] that the theory of GLT sequences can be seen as a generalized Fourier Analysis.

where d_n is the size of A_n and $\sigma_j(A_n),\ j=1,\ldots,d_n$, are the singular values of A_n. If moreover the matrices A_n are Hermitian, then the symbol κ also characterizes the asymptotic spectral distribution of $\{A_n\}_n$. This means that, for all continuous functions F with bounded support, we have

$$\lim_{n\to\infty}\frac{1}{d_n}\sum_{j=1}^{d_n}F(\lambda_j(A_n))=\frac{1}{(2\pi)^d}\int_{[0,1]^d\times[-\pi,\pi]^d}F(\kappa(\mathbf{x},\boldsymbol{\theta}))d\mathbf{x}d\boldsymbol{\theta}, \tag{1.3}$$

where $\lambda_j(A_n),\ j=1,\ldots,d_n$, are the singular values of A_n.

GLT 2. Any sequence of (multilevel) Toeplitz matrices $\{T_n(f)\}_n$ generated by a function f in $L^1([-\pi,\pi]^d)$ is a GLT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta})=f(\boldsymbol{\theta})$.

GLT 3. Any sequence of (multilevel) diagonal sampling matrices $\{D_n(a)\}_n$ containing the evaluations of a Riemann-integrable function $a:[0,1]^d\to\mathbb{C}$ over a uniform grid is a GLT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta})=a(\mathbf{x})$.

GLT 4. Any zero-distributed sequence $\{Z_n\}_n$, i.e., any sequence of matrices possessing an asymptotic singular value distribution characterized by the identically zero function, in the sense of eq. (1.2), is a GLT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta})=0$ (identically).

GLT 5. If $\{A_n\}_n\sim_{\rm GLT}\kappa$ then $\{A_n^*\}_n\sim_{\rm GLT}\overline{\kappa}$, where A_n^* is the Hermitian transpose of A_n.

GLT 6. If $A_n=\sum_{i=1}^r\alpha_i\prod_{j=1}^{q_i}A_n^{(i,j)}$, where $r,q_1,\ldots,q_r\in\mathbb{N}$, $\alpha_1,\ldots,\alpha_r\in\mathbb{C}$ and $\{A_n^{(i,j)}\}_n\sim_{\rm GLT}\kappa_{ij}$, then $\{A_n\}_n\sim_{\rm GLT}\kappa=\sum_{i=1}^r\alpha_i\prod_{j=1}^{q_i}\kappa_{ij}$.

GLT 7. If $\{A_n\}_n\sim_{\rm GLT}\kappa$ and $\kappa\neq 0$ a.e., then $\{A_n^\dagger\}_n\sim_{\rm GLT}\kappa^{-1}$, where $A_n^\dagger$ is the (Moore–Penrose) pseudoinverse of A_n.

GLT 8. If $\{A_n\}_n\sim_{\rm GLT}\kappa$ and each A_n is Hermitian, then $\{f(A_n)\}_n\sim_{\rm GLT}f(\kappa)$ for all continuous functions $f:\mathbb{R}\to\mathbb{R}$.

GLT 9. $\{A_n\}_n\sim_{\rm GLT}\kappa$ if and only if there exist GLT sequences $\{B_{n,m}\}_n\sim_{\rm GLT}\kappa_m$ such that κ_m converges to κ in measure and $\{B_{n,m}\}_n$ 'converges' to $\{A_n\}_n$.

At this stage, we cannot be more precise about the 'convergence' of $\{B_{n,m}\}_n$ to $\{A_n\}_n$. We only anticipate that, by saying « $\{B_{n,m}\}_n$ 'converges' to $\{A_n\}_n$ », we mean that « $\{\{B_{n,m}\}_n\}_m$ is an approximating class of sequences for $\{A_n\}_n$ (as $m\to\infty$) ». Things will become more clear in Section 3, where the notion of approximating class of sequences is introduced. We note that items **GLT 5 – GLT 7** can be summarized by saying that the set of GLT sequences is a *-algebra. Informally speaking, we may rephrase them as follows: suppose that $\{A_n^{(1)}\}_n,\ldots,\{A_n^{(r)}\}_n$ are GLT sequences with symbols $\kappa_1,\ldots,\kappa_r$, and let $A_n={\rm ops}(A_n^{(1)},\ldots,A_n^{(r)})$ be a matrix obtained from $A_n^{(1)},\ldots,A_n^{(r)}$ by means of certain algebraic operations 'ops', such as linear combinations, products, (pseudo)inversions and Hermitian transpositions; then, $\{A_n\}_n$ is a GLT sequence with symbol $\kappa={\rm ops}(\kappa_1,\ldots,\kappa_r)$.

1.3 Contributions of this work

As already pointed out, in this work we are not concerned with the applicative interest of the theory of GLT sequences. On the contrary, we focus on the mathematical foundations of the theory.

We first propose slight (but relevant) modifications of the original definitions of separable Locally Toeplitz (sLT) sequences and GLT sequences appeared in [44, 45]. Moreover, we introduce for the first time the notion of LT sequences in the multivariate/multilevel setting. We stress that sLT and LT sequences are nothing else than specific examples of GLT sequences. However, their role in developing the theory of GLT sequences is so fundamental that they deserve a special name and also a dedicated section (Section 4). With the new definitions, which are based on the notion of approximating class of sequences [29, 42], we are able to enlarge the applicability of the theory of GLT sequences, by generalizing and/or simplifying a lot of key results. In particular, we remove a technical hypothesis concerning the Riemann-integrability of the so-called 'weight functions', which appeared in the statement of fundamental spectral distribution and algebraic results for GLT sequences, such as [44, Theorems 4.5 and 4.8] and [45, Theorem 2.2]. We also show that the product of two LT sequences with symbols $\kappa(\mathbf{x},\boldsymbol{\theta})=a(\mathbf{x})f(\boldsymbol{\theta})$ and $\tilde\kappa(\mathbf{x},\boldsymbol{\theta})=\tilde a(\mathbf{x})\tilde f(\boldsymbol{\theta})$ is a LT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta})\tilde\kappa(\mathbf{x},\boldsymbol{\theta})=a(\mathbf{x})\tilde a(\mathbf{x})f(\boldsymbol{\theta})\tilde f(\boldsymbol{\theta})$, under the only assumption that the two 'generating' functions $f,\tilde f$ are conjugate (i.e., one in L^p and the other in L^q, being p,q conjugate exponents). In this way, we remove from [44, Theorem 5.3] both the assumption that $f,\tilde f$ are in L^∞ and the technical condition in [44, eq. (41)]. In addition, we provide a formal proof of the fact that the sequences of matrices mentioned in items **GLT 2 – GLT 4** fall in the class of (multilevel) LT sequences; different versions of this result appeared in many previous papers, but only partial proofs were given. The latter results allow us to show that the sequence $\{D_n(a)T_n(f)\}_n$, obtained as the product of the diagonal sampling matrix $D_n(a)$ associated with the function a and the Toeplitz matrix $T_n(f)$ generated by f, is a LT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta})=a(\mathbf{x})f(\boldsymbol{\theta})$, under the only assumptions that a is Riemann-integrable and f is in L^1.

As a final step, we also extend the theory of GLT sequences. This is the completely new part of this work. We first provide an approximation result in Section 5.3, which is essentially the content of item **GLT 9** and is particularly useful to show that a given sequence of matrices is a GLT sequence. By using this result, we provide in Section 5.5 a new and easier proof of item **GLT 7**. Finally, using again the approximation result of item **GLT 9**, we prove item **GLT 8**.

2 Mathematical background

2.1 Notation and terminology

- $\mathbb{R}^{m\times n}$ (resp. $\mathbb{C}^{m\times n}$) is the space of real (resp. complex) $m \times n$ matrices.
- O_m and I_m denote, respectively, the $m\times m$ zero matrix and the $m\times m$ identity matrix. Sometimes, when the dimension m can be inferred from the context, O and I are used instead of O_m and I_m.
- If $\mathbf{x}$ is a vector and X is a matrix, then $\mathbf{x}^T$ and $\mathbf{x}^*$ (resp. X^T and X^*) are the transpose and the transpose conjugate of $\mathbf{x}$ (resp. X).
- Given $X\in\mathbb{C}^{m\times m}$, $\Lambda(X)$ is the spectrum of X and $\rho(X)$ is the spectral radius of X, i.e., $\rho(X)=\max_{\lambda\in\Lambda(X)}|\lambda|$. The eigenvalues of X are denoted by $\lambda_j(X),\ j=1,\ldots,m$.
- Let $X\in\mathbb{C}^{m\times m}$ be a matrix with only real eigenvalues (e.g., a Hermitian matrix). Unless otherwise stated, it is understood that the eigenvalues of X are labeled in non-increasing order: $\lambda_{\max}(X)=\lambda_1(X)\ge\ldots\ge\lambda_m(X)=\lambda_{\min}(X)$. In addition, we set $\lambda_j(X)=+\infty$ if $j<1$ and $\lambda_j(X)=-\infty$ if $j>m$.
- If $X\in\mathbb{C}^{m\times m}$, we denote by $\sigma_j(X),\ j=1,\ldots,m$, the singular values of X labeled, as usual, in non-increasing order: $\sigma_{\max}(X)=\sigma_1(X)\ge\ldots\ge\sigma_m(X)=\sigma_{\min}(X)$. In addition, we set $\sigma_j(X)=+\infty$ if $j<1$ and $\sigma_j(X)=-\infty$ if $j>m$.
- Given $X\in\mathbb{C}^{m\times m}$ and $1\le p\le\infty$, $\|X\|_p$ denotes the Schatten p-norm of X, which is defined as the p-norm of the vector $(\sigma_1(X),\ldots,\sigma_m(X))$ formed by the singular values of X; see [8]. The Schatten 1-norm is also called the trace-norm. The Schatten ∞-norm $\|X\|_\infty=\sigma_{\max}(X)$ coincides with the spectral (Euclidean) norm of X; it will be preferably denoted by $\|X\|$. Note that the Schatten norms are unitarily invariant, i.e., $\|UXV\|_p=\|X\|_p$ for all $p\in[1,\infty]$, all $X\in\mathbb{C}^{m\times m}$ and all unitary matrices $U,V\in\mathbb{C}^{m\times m}$. This follows from the fact that X and UXV have the same singular values.
- If $1\le p\le\infty$, the symbol $|\cdot|_p$ denotes the p-norm of both vectors and matrices:
$$|\mathbf{x}|_p=\begin{cases}\left(\sum_{i=1}^m|x_i|^p\right)^{1/p} & \text{if } 1\le p<\infty,\\ \max_{i=1,\ldots,m}|x_i| & \text{if } p=\infty,\end{cases}\qquad \mathbf{x}\in\mathbb{C}^m,$$
$$|X|_p=\max_{\substack{\mathbf{x}\in\mathbb{C}^m\\ \mathbf{x}\ne\mathbf{0}}}\frac{|X\mathbf{x}|_p}{|\mathbf{x}|_p},\qquad X\in\mathbb{C}^{m\times m}.$$
Note that $|X|_2=\|X\|$.
- $\Re(X)$ and $\Im(X)$ are, respectively, the real and the imaginary part of the (square) matrix X:
$$\Re(X)=\frac{X+X^*}{2},\qquad \Im(X)=\frac{X-X^*}{2\mathrm{i}},$$
where i is the imaginary unit ($\mathrm{i}^2=-1$).
- If $z\in\mathbb{C}$ and $\epsilon>0$, we denote by $D(z,\epsilon)$ the disk centered at z and with radius ϵ, i.e., $D(z,\epsilon)=\{w\in\mathbb{C}:\ |w-z|<\epsilon\}$.
- $\zeta_n\overset{n\to\infty}{\longrightarrow}\zeta$ means that $\zeta_n\to\zeta$ as $n\to\infty$.
- $C_c(\mathbb{C})$ (resp. $C_c(\mathbb{R})$) is the space of complex-valued continuous functions defined on $\mathbb{C}$ (resp. $\mathbb{R}$) and with bounded support. Moreover, $C_c^1(\mathbb{R})=C_c(\mathbb{R})\cap C^1(\mathbb{R})$, where $C^1(\mathbb{R})$ is the space of complex-valued functions F defined on $\mathbb{R}$ whose real and imaginary parts $\Re(F)$, $\Im(F)$ are of class C^1 over $\mathbb{R}$ in the classical sense.
- A 'functional' ϕ is any function defined on some vector space (such as, for example, $C_c(\mathbb{C})$ or $C_c(\mathbb{R})$) and taking values in $\mathbb{C}$.
- If $H:\mathbb{R}\to\mathbb{R}$ and the limit $\lim_{x\to\infty}H(x)$ exists, we denote it by $H(\infty)$. Similarly, $H(-\infty)=\lim_{x\to-\infty}H(x)$.
- If $w_i:D_i\to\mathbb{C},\ i=1,\ldots,d$, are arbitrary functions, $w_1\otimes\cdots\otimes w_d:D_1\times\cdots\times D_d\to\mathbb{C}$ denotes the tensor-product function
$$(w_1\otimes\cdots\otimes w_d)(\xi_1,\ldots,\xi_d)=w_1(\xi_1)\cdots w_d(\xi_d),\qquad (\xi_1,\ldots,\xi_d)\in D_1\times\cdots\times D_d.$$
- A function $a:[0,1]^d\to\mathbb{C}$ is said to be Riemann-integrable if its real and imaginary parts $\Re(a),\Im(a):[0,1]^d\to\mathbb{R}$ are Riemann-integrable in the classical sense. Recall that *any Riemann-integrable function is bounded.*
- If $g:D\to\mathbb{C}$ and $E\subseteq D$, we set $\|g\|_{\infty,E}=\sup_{\xi\in E}|g(\xi)|$. If $E=D$ and D can be inferred from the context, we often write $\|g\|_\infty$ instead of $\|g\|_{\infty,D}$. Clearly, $\|g\|_\infty<\infty$ if and only if g is bounded over D.

- χ_E is the characteristic (or indicator) function of the set E,

$$\chi_E(\xi) = \begin{cases} 1, & \text{if } \xi \in E, \\ 0, & \text{otherwise.} \end{cases}$$

- μ_k denotes the Lebesgue measure in $\mathbb{R}^k$. Throughout this work, all the terminology coming from measure theory (such as 'measure', 'measurable', 'a.e.', 'in L^p', etc.) is always referred to the Lebesgue measure.

- If $f : D \subseteq \mathbb{R}^k \to \mathbb{C}$ is in $L^p(D)$ and the domain D is clear from the context, we write $\|f\|_{L^p}$ instead of $\|f\|_{L^p(D)}$ to indicate the L^p-norm of f. Recall that $\|f\|_{L^p} = (\int_D |f|^p)^{1/p}$ for $1 \le p < \infty$, and $\|f\|_{L^\infty} = \operatorname{ess\,sup}_D |f|$ for $p = \infty$.

- We use a notation borrowed from probability theory to indicate sets. For example, if $f, g : D \subseteq \mathbb{R}^k \to \mathbb{C}$, then $\{f \neq 1\} = \{\mathbf{x} \in D : f(\mathbf{x}) \neq 1\}$, $\{f \in D(z, \epsilon)\} = \{\mathbf{x} \in D : f(\mathbf{x}) \in D(z, \epsilon)\}$, $\{0 \le f \le 1,\ g > 2\} = \{\mathbf{x} \in D : 0 \le f(\mathbf{x}) \le 1,\ g(\mathbf{x}) > 2\}$, $\mu_k\{f > 0,\ g < 0\}$ is the measure of the set $\{\mathbf{x} \in D : f(\mathbf{x}) > 0,\ g(\mathbf{x}) < 0\}$, $\chi_{\{f=0\}}$ is the characteristic function of the set where f vanishes, and so on.

2.1.1 Multi-index notation

Throughout this work, we will systematically use the multi-index notation. A multi-index $\boldsymbol{i} \in \mathbb{Z}^d$, also called a d-index, is simply a vector in $\mathbb{Z}^d$; its components are denoted by $i_1, \ldots, i_d$.

- $\mathbf{0}, \mathbf{1}, \mathbf{2}, \ldots$ are the vectors of all zeros, all ones, all twos, ... (their size will be clear from the context).

- For any d-index $\boldsymbol{m}$, $N(\boldsymbol{m}) = \prod_{j=1}^d m_j$ and $\boldsymbol{m} \to \infty$ means that $\min(\boldsymbol{m}) = \min_{j=1,\ldots,d} m_j \to \infty$.

- If $\boldsymbol{h}, \boldsymbol{k}$ are d-indices, $\boldsymbol{h} \le \boldsymbol{k}$ means that $h_r \le k_r$ for all $r = 1, \ldots, d$, while $\boldsymbol{h} \not\le \boldsymbol{k}$ means that $h_r > k_r$ for at least one $r \in \{1, \ldots, d\}$.

- If $\boldsymbol{h}, \boldsymbol{k}$ are d-indices such that $\boldsymbol{h} \le \boldsymbol{k}$, the multi-index range $\boldsymbol{h}, \ldots, \boldsymbol{k}$ is the set $\{\boldsymbol{j} \in \mathbb{Z}^d : \boldsymbol{h} \le \boldsymbol{j} \le \boldsymbol{k}\}$. We assume for the multi-index range $\boldsymbol{h}, \ldots, \boldsymbol{k}$ the standard lexicographic ordering:

$$\Big[\ldots \Big[\big[(j_1, \ldots, j_d) \big]_{j_d = h_d, \ldots, k_d} \Big]_{j_{d-1} = h_{d-1}, \ldots, k_{d-1}} \ldots \Big]_{j_1 = h_1, \ldots, k_1}. \tag{2.1}$$

  For instance, in the case $d = 2$ the ordering is

$$\begin{gathered}(h_1, h_2),\ (h_1, h_2+1),\ \ldots,\ (h_1, k_2),\ (h_1+1, h_2),\ (h_1+1, h_2+1),\ \ldots,\ (h_1+1, k_2),\\ \ldots\ldots,\ (k_1, h_2),\ (k_1, h_2+1),\ \ldots,\ (k_1, k_2).\end{gathered}$$

- When a d-index $\boldsymbol{j}$ varies over a multi-index range $\boldsymbol{h}, \ldots, \boldsymbol{k}$ (this is sometimes written as $\boldsymbol{j} = \boldsymbol{h}, \ldots, \boldsymbol{k}$), it is understood that $\boldsymbol{j}$ varies from $\boldsymbol{h}$ to $\boldsymbol{k}$ following the specific ordering (2.1). For instance, if $\boldsymbol{m} \in \mathbb{N}^d$ and if we write $\mathbf{x} = [x_{\boldsymbol{i}}]_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{m}}$, then $\mathbf{x}$ is a vector of size $N(\boldsymbol{m})$ whose components $x_{\boldsymbol{i}}$, $\boldsymbol{i} = \mathbf{1}, \ldots, \boldsymbol{m}$, are ordered in accordance with (2.1): the first component is $x_{\mathbf{1}} = x_{(1,\ldots,1,1)}$, the second component is $x_{(1,\ldots,1,2)}$, and so on until the last component, which is $x_{\boldsymbol{m}} = x_{(m_1,\ldots,m_d)}$. Similarly, if

$$X = [x_{\boldsymbol{ij}}]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{m}}, \tag{2.2}$$

  then X is a $N(\boldsymbol{m}) \times N(\boldsymbol{m})$ matrix whose components are indexed by two d-indices $\boldsymbol{i}, \boldsymbol{j}$, both varying from $\mathbf{1}$ to $\boldsymbol{m}$ according to the lexicographic ordering (2.1).

- When a multi-index appears as subscript or superscript, we often suppress the parentheses to simplify the notation. For instance, the component of the vector $\mathbf{x} = [x_{\boldsymbol{i}}]_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{m}}$ corresponding to the multi-index $\boldsymbol{i}$ is denoted by $x_{\boldsymbol{i}}$ or by $x_{i_1,\ldots,i_d}$, and we preferably avoid the heavy notation $x_{(i_1,\ldots,i_d)}$.

- Given $\boldsymbol{h}, \boldsymbol{k} \in \mathbb{Z}^d$ with $\boldsymbol{h} \le \boldsymbol{k}$, the notation $\sum_{\boldsymbol{j}=\boldsymbol{h}}^{\boldsymbol{k}}$ indicates the summation over all $\boldsymbol{j}$ in $\boldsymbol{h}, \ldots, \boldsymbol{k}$.

- Operations involving multi-indices that do not have a meaning when considering multi-indices as normal vectors must always be interpreted in the componentwise sense. For instance, $\boldsymbol{np} = (n_1p_1, \ldots, n_dp_d)$, $\alpha\boldsymbol{i}/\boldsymbol{j} = (\alpha i_1/j_1, \ldots, \alpha i_d/j_d)$ for all $\alpha \in \mathbb{C}$ (of course, the division is defined when $j_1, \ldots, j_d \neq 0$), $\boldsymbol{i}^2 = (i_1^2, \ldots, i_d^2)$, $\max(\boldsymbol{i}, \boldsymbol{j}) = (\max(i_1, j_1), \ldots, \max(i_d, j_d))$, $\boldsymbol{i} \bmod \boldsymbol{m} = (i_1 \bmod m_1, \ldots, i_d \bmod m_d)$, and so on.

2.1.2 Matrix-sequences and multilevel diagonal sampling matrices

In all this work, by sequence of matrices (or matrix-sequence) we mean a sequence of the form $\{A_{\boldsymbol{n}}\}_n$, where:

- $\boldsymbol{n} = \boldsymbol{n}(n)$ is a multi-index in $\mathbb{N}^d$ which depends on n;
- n varies in some infinite subset of $\mathbb{N}$, and $\boldsymbol{n} \to \infty$ when $n \to \infty$;
- $A_{\boldsymbol{n}}$ is a square matrix of size $N(\boldsymbol{n})$.

Recall from the previous section that $\boldsymbol{n} \to \infty$ means $\min_{j=1,\ldots,d} n_j \to \infty$. Unless otherwise stated, the multi-index that parameterizes a matrix-sequence is always assumed to be a d-index. We will avoid to repeat this every time, so as to simplify the presentation.

Three classes of matrix-sequences, which can be regarded as the building blocks of the theory of GLT sequences, will be of particular interest in the following: zero-distributed sequences, sequences of multilevel diagonal sampling matrices and sequences of multilevel Toeplitz matrices. Here, we introduce the multilevel diagonal sampling matrices, while zero-distributed sequences and multilevel Toeplitz matrices will be considered in more detail in Sections 2.5.1 and 2.6, respectively. For $\boldsymbol{n} \in \mathbb{N}^d$ and $a : [0,1]^d \to \mathbb{C}$, we define the d-level diagonal sampling matrix $D_{\boldsymbol{n}}(a)$ as the following diagonal matrix of size $N(\boldsymbol{n})$:

$$D_{\boldsymbol{n}}(a) = \operatorname*{diag}_{\boldsymbol{i}=\boldsymbol{1},\ldots,\boldsymbol{n}} a\Big(\frac{\boldsymbol{i}}{\boldsymbol{n}}\Big),$$

where we recall that $\boldsymbol{i}$ varies from $\boldsymbol{1}$ to $\boldsymbol{n}$ according to the lexicographic ordering (2.1). For example, if $d = 2$ then

$$D_{\boldsymbol{n}}(a) = \operatorname*{diag}_{i_1=1,\ldots,n_1} \left[\operatorname*{diag}_{i_2=1,\ldots,n_2} a\Big(\frac{i_1}{n_1}, \frac{i_2}{n_2}\Big) \right].$$

Note that $D_{\boldsymbol{n}}(a)$ can also be defined through a recursive formula: if $d = 1$, then

$$D_n(a) = \operatorname*{diag}_{i=1,\ldots,n} a\Big(\frac{i}{n}\Big);$$

if $d > 1$, then

$$D_{\boldsymbol{n}}(a) = D_{n_1,\ldots,n_d}(a) = \operatorname*{diag}_{i_1=1,\ldots,n_1} D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1}, \cdot\Big)\Big), \tag{2.3}$$

where $a(i_1/n_1, \cdot) : [0,1]^{d-1} \to \mathbb{C}$ is defined by $(x_2, \ldots, x_d) \mapsto a(i_1/n_1, x_2, \ldots, x_d)$.

2.2 Separable functions and multivariate trigonometric polynomials

Let $I_1, \ldots, I_d \subseteq \mathbb{R}$ be measurable sets and let $f : I_1 \times \cdots \times I_d \to \mathbb{C}$. We say that f is separable if there exist measurable functions $f_i : I_i \to \mathbb{C},\ i = 1, \ldots, d$, such that $f = f_1 \otimes \cdots \otimes f_d$. In this case, the functions $f_1, \ldots, f_d$ are called factors of f and $f_1 \otimes \cdots \otimes f_d$ is said to be a factorization of f. Note that the factorization is not unique: it suffices to choose d constants $c_1, \ldots, c_d$ such that $c_1 \cdots c_d = 1$ in order to obtain another factorization $f = c_1 f_1 \otimes \cdots \otimes c_d f_d$. Note also that any separable function is measurable.

Let $f : I_1 \times \cdots \times I_d \to \mathbb{C}$ be separable and take a factorization $f = f_1 \otimes \cdots \otimes f_d$. If $f \in L^p(I_1 \times \cdots \times I_d)$ and f is not a.e. equal to 0, then $f_i \in L^p(I_i)$ for all $i = 1, \ldots, d$. Indeed, for $p < \infty$ we have

$$\int_{I_1 \times \cdots \times I_d} |f|^p = \prod_{i=1}^{d} \int_{I_i} |f_i|^p.$$

Since $\int_{I_i} |f_i|^p \neq 0$ for all i (otherwise $f = 0$ a.e., contrary to the assumption), it follows that $f \in L^p(I_1 \times \cdots \times I_d)$ if and only if $f_i \in L^p(I_i)$ for all i. For the case $p = \infty$, we only prove that $f_1 \in L^\infty(I_1)$, because the proof for the other factors is similar. Since f is not a.e. equal to 0, in particular $f_2 \otimes \cdots \otimes f_d$ is not a.e. equal to 0. Hence, $\mu_{d-1}\{|f_2 \otimes \cdots \otimes f_d| \geq \epsilon\} > 0$ for some $\epsilon > 0$. If we assume by contradiction that $f_1 \notin L^\infty(I_1)$, then $\mu_1\{|f_1| \geq \alpha\} > 0$ for all $\alpha > 0$. This implies that

$$\mu_d\{|f| \geq \alpha\} \geq \mu_d(\{|f_1| \geq \alpha/\epsilon\} \cap \{|f_2 \otimes \cdots \otimes f_d| \geq \epsilon\}) = \mu_1\{|f_1| \geq \alpha/\epsilon\}\, \mu_{d-1}\{|f_2 \otimes \cdots \otimes f_d| \geq \epsilon\} > 0,$$

for all $\alpha > 0$, which is a contradiction to the assumption that $f \in L^\infty(I_1 \times \cdots \times I_d)$. In conclusion, we have proved that, for any $1 \leq p \leq \infty$, the factors $f_1, \ldots, f_d$ appearing in any factorization of a separable function $f \in L^p(I_1 \times \cdots \times I_d)$ are themselves in L^p, provided that f is not a.e. equal to 0. In particular, the following lemma holds.

Lemma 2.1. *Let $f : I_1 \times \cdots \times I_d \to \mathbb{C}$ be a separable function in $L^p(I_1 \times \cdots \times I_d)$, $1 \leq p \leq \infty$. Then, there exist functions $f_i : I_i \to \mathbb{C}$, $i = 1, \ldots, d$, such that $f = f_1 \otimes \cdots \otimes f_d$ a.e. and $f_i \in L^p(I_i)$ for all $i = 1, \ldots, d$.*

A d-variate trigonometric polynomial is a finite linear combination of the Fourier frequencies $e^{i\boldsymbol{j}\cdot\boldsymbol{\theta}}$, $\boldsymbol{j} \in \mathbb{Z}^d$ (here, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d) \in \mathbb{R}^d$ and $\boldsymbol{j} \cdot \boldsymbol{\theta} = j_1\theta_1 + \ldots + j_d\theta_d$). Let $f : \mathbb{R}^d \to \mathbb{C}$ be a separable d-variate trigonometric polynomial, i.e., a d-variate trigonometric polynomial which is separable in the sense specified above. Let $f = f_1 \otimes \cdots \otimes f_d$ be a factorization of f. If f is not identically 0, then $f_1, \ldots, f_d$ are (univariate) trigonometric polynomials. Indeed, since $f_2, \ldots, f_d$ are not identically 0, there exists $(\vartheta_2, \ldots, \vartheta_d)$ such that $f_2(\vartheta_2) \cdots f_d(\vartheta_d) \neq 0$. From the definition of d-variate trigonometric polynomials, we see that $\theta_1 \mapsto f(\theta_1, \vartheta_2, \ldots, \vartheta_d) = f_1(\theta_1)f_2(\vartheta_2) \cdots f_d(\vartheta_d)$ is a (univariate) trigonometric polynomial, and this means that f_1 is a trigonometric polynomial. With the same argument, one can show that $f_2, \ldots, f_d$ are trigonometric polynomials as well.

Lemma 2.2. *Let $f : \mathbb{R}^d \to \mathbb{C}$ be a separable d-variate trigonometric polynomial. Then, there exist trigonometric polynomials $f_1, \ldots, f_d : \mathbb{R} \to \mathbb{C}$ such that $f = f_1 \otimes \cdots \otimes f_d$.*

2.3 Convergence in measure

The convergence in measure is of particular interest in probability theory, and it plays an important role also in the study of GLT sequences. In this section, we recall the definition and provide some basic properties of this convergence.

Definition 2.1 (convergence in measure). Let $f_m, f : D \subseteq \mathbb{R}^k \to \mathbb{C}$ be measurable functions. We say that $f_m \to f$ in measure if, for every $\epsilon > 0$,

$$\lim_{m\to\infty} \mu_k\{|f_m - f| > \epsilon\} = 0.$$

We recall that, according to our notation, $\{|f_m - f| > \epsilon\} = \{\mathbf{x} \in D : |f_m(\mathbf{x}) - f(\mathbf{x})| > \epsilon\}$; see Section 2.1.

Basic results about the convergence in measure are collected in the next lemmas. Since these results are not so popular, we include the details of the proofs for the reader's convenience.

Lemma 2.3. *Let $f_m, g_m, f, g : D \subseteq \mathbb{R}^k \to \mathbb{C}$ be measurable functions.*

(i) If $f_m \to f$ in measure, then $|f_m| \to |f|$ in measure.

(ii) If $f_m \to f$ in measure and $g_m \to g$ in measure, then $\alpha f_m + \beta g_m \to \alpha f + \beta g$ in measure for all $\alpha, \beta \in \mathbb{C}$.

(iii) If $f_m \to f$ in measure, $g_m \to g$ in measure, and $\mu_k(D) < \infty$, then $f_m g_m \to fg$ in measure.

Proof. *(i)* We have $\{||f_m| - |f|| > \epsilon\} \subseteq \{|f_m - f| > \epsilon\}$ and $\mu_k\{||f_m| - |f|| > \epsilon\} \leq \mu_k\{|f_m - f| > \epsilon\}$. Hence, if $f_m \to f$ in measure, $|f_m| \to |f|$ in measure.

(ii) We have

$$\{|(\alpha f_m + \beta g_m) - (\alpha f + \beta g)| > \epsilon\} \subseteq \{|\alpha||f_m - f| + |\beta||g_m - g| > \epsilon\} \subseteq \{|\alpha||f_m - f| > \epsilon/2\} \cup \{|\beta||g_m - g| > \epsilon/2\}$$

and

$$\mu_k\{|(\alpha f_m + \beta g_m) - (\alpha f + \beta g)| > \epsilon\} \leq \mu_k\{|\alpha||f_m - f| > \epsilon/2\} + \mu_k\{|\beta||g_m - g| > \epsilon/2\}.$$

Hence, if $f_m \to f$ in measure and $g_m \to g$ in measure, then $\alpha f_m + \beta g_m \to \alpha f + \beta g$ in measure.

(iii) For every m and every $\epsilon, M > 0$, we have

$$\begin{aligned}
\mu_k\{|f_m g_m - fg| > \epsilon\} &\leq \mu_k\{|f_m - f||g_m| + |f||g_m - g| > \epsilon\} \leq \mu_k\{|f_m - f||g_m| > \epsilon/2\} + \mu_k\{|f||g_m - g| > \epsilon/2\} \\
&= \mu_k\{|f_m - f||g_m| > \epsilon/2,\ |g_m| \leq M\} + \mu_k\{|f_m - f||g_m| > \epsilon/2,\ |g_m| > M\} \\
&\quad + \mu_k\{|f||g_m - g| > \epsilon/2,\ |f| \leq M\} + \mu_k\{|f||g_m - g| > \epsilon/2,\ |f| > M\} \\
&\leq \mu_k\{|f_m - f|M > \epsilon/2\} + \mu_k\{|g_m| > M\} \\
&\quad + \mu_k\{M|g_m - g| > \epsilon/2\} + \mu_k\{|f| > M\} \\
&\leq \mu_k\{|f_m - f|M > \epsilon/2\} + \mu_k\{|g_m - g| > M/2\} + \mu_k\{|g| > M/2\} \\
&\quad + \mu_k\{M|g_m - g| > \epsilon/2\} + \mu_k\{|f| > M\}.
\end{aligned} \tag{2.4}$$

Passing to the limit as $m \to \infty$ in (2.4) and using the fact that $f_m \to f$ in measure and $g_m \to g$ in measure, we get

$$\limsup_{m\to\infty} \mu_k\{|f_m g_m - fg| > \epsilon\} \leq \mu_k\{|g| > M/2\} + \mu_k\{|f| > M\}, \tag{2.5}$$

for every $\epsilon, M > 0$. Now we observe that $\mu_k\{|f| > M\} = \int_D \chi_{\{|f|>M\}}$ and $\chi_{\{|f|>M\}} \to 0$ pointwise when $M \to \infty$. Moreover, the convergence of $\chi_{\{|f|>M\}}$ is dominated by the constant 1, which belongs to $L^1(D)$ because $\mu_k(D) < \infty$. Hence, by the dominated convergence theorem [39], $\lim_{M\to\infty} \mu_k\{|f| > M\} = 0$. Similarly, $\lim_{M\to\infty} \mu_k\{|g| > M/2\} = 0$. Passing to the limit as $M \to \infty$ in (2.5), we obtain that $\limsup_{m\to\infty} \mu_k\{|f_m g_m - fg| > \epsilon\} = 0$, for every $\epsilon > 0$. Hence, $f_m g_m \to fg$ in measure. □

Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$ and let $g : D \subset \mathbb{R}^k \to \mathbb{K}$ be a measurable function defined on a set D with $0 < \mu_k(D) < \infty$. Consider the functional

$$\phi_{[g]} : C_c(\mathbb{K}) \to \mathbb{C}, \qquad \phi_{[g]}(F) = \frac{1}{\mu_k(D)} \int_D F(g(\mathbf{x}))d\mathbf{x}. \tag{2.6}$$

$\phi_{[g]}$ is a continuous linear functional on the normed vector space $(C_c(\mathbb{K}), \|\cdot\|_\infty)$, and $\|\phi_{[g]}\| \le 1$. Indeed, the linearity is obvious and the continuity, as well as the bound $\|\phi_{[g]}\| \le 1$, follow from the observation that $|\phi_{[g]}(F)| \le \|F\|_\infty$ for all $F \in C_c(\mathbb{K})$. If g is constant, say $g = \gamma$ a.e., then $\phi_{[g]} = \phi_{[\gamma]}$ is the evaluation functional at γ; that is, $\phi_{[\gamma]}(F) = F(\gamma)$ for every $F \in C_c(\mathbb{K})$.

Lemma 2.4. *Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$, and let $g_m,\ g : D \subset \mathbb{R}^k \to \mathbb{K}$ be measurable functions defined on a set D with $0 < \mu_k(D) < \infty$. If $g_m \to g$ in measure, then $F \circ g_m \to F \circ g$ in $L^1(D)$ for all $F \in C_c(\mathbb{K})$ and $\phi_{[g_m]} \to \phi_{[g]}$ pointwise over $C_c(\mathbb{K})$.*

Proof. Assume that $g_m \to g$ in measure. We show that $F \circ g_m \to F \circ g$ in $L^1(D)$ for all $F \in C_c(\mathbb{K})$; this immediately implies that $\phi_{[g_m]} \to \phi_{[g]}$ pointwise over $C_c(\mathbb{K})$, because

$$|\phi_{[g_m]}(F) - \phi_{[g]}(F)| \le \frac{1}{\mu_k(D)} \|F \circ g_m - F \circ g\|_{L^1}.$$

For every $F \in C_c(\mathbb{K})$, every m and every $\epsilon > 0$,

$$\begin{aligned}\|F \circ g_m - F \circ g\|_{L^1} &= \int_D \big|F(g_m(\mathbf{x})) - F(g(\mathbf{x}))\big| d\mathbf{x} \\ &= \int_{\{|g_m - g| \ge \epsilon\}} \big|F(g_m(\mathbf{x})) - F(g(\mathbf{x}))\big| d\mathbf{x} + \int_{\{|g_m - g| < \epsilon\}} \big|F(g_m(\mathbf{x})) - F(g(\mathbf{x}))\big| d\mathbf{x} \\ &\le 2\|F\|_\infty\, \mu_k\{|g_m - g| \ge \epsilon\} + \omega_F(\epsilon),\end{aligned} \tag{2.7}$$

where ω_F is the modulus of continuity of F,

$$\omega_F(\epsilon) = \sup_{\substack{u,v \in \mathbb{K} \\ |u-v| \le \epsilon}} |F(u) - F(v)|.$$

Since

$$\lim_{m\to\infty} \mu_k\{|g_m - g| \ge \epsilon\} = \lim_{\epsilon \to 0} \omega_F(\epsilon) = 0$$

(because $g_m \to g$ in measure and F is uniformly continuous by the Heine-Cantor theorem), passing first to the $\limsup_{m\to\infty}$ and then to the $\lim_{\epsilon\to 0}$ in (2.7), we conclude that $F \circ g_m \to F \circ g$ in $L^1(D)$. □

Lemma 2.4 admits the following converse.

Lemma 2.5. *Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$, and let $g_m,\ g : D \subset \mathbb{R}^k \to \mathbb{K}$ be measurable functions defined on a set D with $0 < \mu_k(D) < \infty$. If $\phi_{[g_m - g]} \to \phi_{[0]}$ pointwise over $C_c(\mathbb{K})$, then $g_m \to g$ in measure.*

Proof. By hypothesis, for all $F \in C_c(\mathbb{K})$ we have

$$\lim_{m\to\infty} \frac{1}{\mu_k(D)} \int_D F(g_m(\mathbf{x}) - g(\mathbf{x}))d\mathbf{x} = F(0). \tag{2.8}$$

Suppose by contradiction that $g_m \not\to g$ in measure. Then, there exist $\epsilon, \delta > 0$ and a subsequence $\{g_{m_r}\}_r$ such that, for all r,

$$\mu_k\{|g_{m_r} - g| \ge \epsilon\} \ge \delta. \tag{2.9}$$

Take a real function $F \in C_c(\mathbb{K})$ such that $F(0) = 1 = \max_{y\in\mathbb{K}} F(y)$ and $F(y) = 0$ over $\{y \in \mathbb{K} :\ |y| \ge \epsilon\}$. Then, by (2.9), for all r we have

$$\frac{1}{\mu_k(D)} \int_D F(g_{m_r}(\mathbf{x}) - g(\mathbf{x}))d\mathbf{x} = \frac{1}{\mu_k(D)} \int_{\{|g_{m_r} - g| < \epsilon\}} F(g_{m_r}(\mathbf{x}) - g(\mathbf{x}))d\mathbf{x} \le \frac{\mu_k\{|g_{m_r} - g| < \epsilon\}}{\mu_k(D)} \le \frac{\mu_k(D) - \delta}{\mu_k(D)} < 1 = F(0),$$

which is a contradiction to (2.8). □

Remark 2.1. Let $\phi_{[g]}$ be defined as in (2.6) and assume that $\phi_{[g]} = \phi_{[0]}$; then $g = 0$ a.e. Indeed, if $\phi_{[g]} = \phi_{[0]}$, then the constant sequence $\{\phi_{[g]}\}_m$ converges pointwise to $\phi_{[0]}$ over $C_c(\mathbb{K})$. By Lemma 2.5, this implies that $g \to 0$ in measure, i.e., $\mu_k\{|g| \ge \epsilon\} = 0$ for every $\epsilon > 0$. Hence,

$$\mu_k\{g \ne 0\} = \mu_k\{|g| > 0\} = \mu_k\big\{\textstyle\bigcup_{\ell=1}^\infty \{|g| \ge 1/\ell\}\big\} = 0,$$

which means that $g = 0$ a.e.

Lemma 2.6 is the last result we need about the convergence in measure; it will play an important role in the proof of Theorem 5.8. For the proof of Lemma 2.6, we recall that the space generated by the trigonometric monomials

$$\left\{\mathrm{e}^{\mathrm{i}\,(\frac{2\pi}{b_1-a_1}j_1y_1\,+\,\ldots\,+\,\frac{2\pi}{b_k-a_k}j_ky_k)}\,:\,\boldsymbol{j}=(j_1,\ldots,j_k)\in\mathbb{Z}^k\right\}$$

is the set of all finite linear combinations of such monomials; we may call it the space of ‘scaled’ k-variate trigonometric polynomials. This space is dense in $L^1([a_1,b_1]\times\cdots\times[a_k,b_k])$.

Lemma 2.6. *Let $\kappa:[0,1]^d\times[-\pi,\pi]^d\to\mathbb{C}$ be a measurable function. Then, there exists a sequence $\{\kappa_m\}_m$ such that $\kappa_m:[0,1]^d\times[-\pi,\pi]^d\to\mathbb{C}$ is a function of the form*

$$\kappa_m(\mathbf{x},\boldsymbol{\theta})=\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m}a_{\boldsymbol{j}}^{(m)}(\mathbf{x})\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}},\qquad a_{\boldsymbol{j}}\in C^\infty([0,1]^d),\qquad \boldsymbol{N}_m\in\mathbb{N}^d, \tag{2.10}$$

and $\kappa_m\to\kappa$ a.e.

Proof. As observed before the statement of the lemma, the space generated by the trigonometric monomials

$$\left\{\mathrm{e}^{\mathrm{i}2\pi\boldsymbol{\ell}\cdot\mathbf{x}}\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}=\mathrm{e}^{\mathrm{i}(2\pi\ell_1x_1+\ldots+2\pi\ell_dx_d+j_1\theta_1+\ldots+j_d\theta_d)}\,:\,\boldsymbol{\ell},\boldsymbol{j}\in\mathbb{Z}^d\right\} \tag{2.11}$$

is dense in $L^1([0,1]^d\times[-\pi,\pi]^d)$. The function $\widetilde{\kappa}_m=\kappa\chi_{\{|\kappa|\le1/m\}}$ belongs to $L^\infty([0,1]^d\times[-\pi,\pi]^d)\subset L^1([0,1]^d\times[-\pi,\pi]^d)$ and converges to κ in measure. Indeed, $\widetilde{\kappa}_m\to\kappa$ pointwise over $[0,1]^d\times[-\pi,\pi]^d$, and it is known that the convergence a.e. on a set of finite measure implies the convergence in measure [39]. Choose a function κ_m belonging to the space generated by the monomials (2.11), such that $\|\kappa_m-\widetilde{\kappa}_m\|_{L^1}\le1/m$. Note that κ_m is a function of the form (2.10). Then, for all $\epsilon>0$,

$$\begin{aligned}\mu_{2d}\{|\kappa_m-\kappa|>\epsilon\}&\le\mu_{2d}\{|\kappa_m-\widetilde{\kappa}_m|>\epsilon/2\}+\mu_{2d}\{|\widetilde{\kappa}_m-\kappa|>\epsilon/2\}=\int_{[0,1]^d\times[-\pi,\pi]^d}\chi_{\{|\kappa_m-\widetilde{\kappa}_m|>\epsilon/2\}}+\mu_{2d}\{|\widetilde{\kappa}_m-\kappa|>\epsilon/2\}\\&\le\int_{[0,1]^d\times[-\pi,\pi]^d}\frac{|\kappa_m-\widetilde{\kappa}_m|}{(\epsilon/2)}+\mu_{2d}\{|\widetilde{\kappa}_m-\kappa|>\epsilon/2\}=\frac{\|\kappa_m-\widetilde{\kappa}_m\|_{L^1}}{(\epsilon/2)}+\mu_{2d}\{|\widetilde{\kappa}_m-\kappa|>\epsilon/2\}\end{aligned}$$

which converges to 0 as $m\to\infty$. Hence, $\kappa_m\to\kappa$ in measure. Since the convergence in measure on a set of finite measure implies the existence of a subsequence that converges a.e. [39], passing to a subsequence of $\{\kappa_m\}_m$ (if necessary) we may assume that $\kappa_m\to\kappa$ a.e. ☐

2.4 Preliminaries on Linear Algebra and Matrix Analysis

Let $X\in\mathbb{C}^{m\times m}$. We know from the Singular Value Decomposition (SVD) that rank(X) is the number of nonzero singular values of X. As a consequence, recalling that $\|X\|=\|X\|_\infty=\sigma_{\max}(X)$, we obtain the following inequality for the trace-norm of X:

$$\|X\|_1=\sum_{i=1}^m\sigma_i(X)\le\mathrm{rank}(X)\|X\|\le m\|X\|,\qquad X\in\mathbb{C}^{m\times m}. \tag{2.12}$$

More generally, let $1<p,q<\infty$ be conjugate exponents ($\frac1p+\frac1q=1$); then, using the Hölder inequality, we have

$$\|X\|_1=\sum_{i=1}^m\sigma_i(X)=\sum_{i=1}^{\mathrm{rank}(X)}\sigma_i(X)\le\left(\sum_{i=1}^{\mathrm{rank}(X)}1^q\right)^{1/q}\left(\sum_{i=1}^{\mathrm{rank}(X)}\sigma_i(X)^p\right)^{1/p}=\mathrm{rank}(X)^{1/q}\|X\|_p\le m^{1/q}\|X\|_p,\qquad X\in\mathbb{C}^{m\times m}. \tag{2.13}$$

Another interesting trace-norm inequality, which provides an upper bound for the trace-norm in terms of the components, is the following:

$$\|X\|_1\le\sum_{i,j=1}^m|x_{ij}|,\qquad X\in\mathbb{C}^{m\times m}. \tag{2.14}$$

The proof is simple. Let $X=U\Sigma V^*$ be an SVD of X. Then, setting $Q=VU^*$, the matrix Q is unitary and we have

$$\|X\|_1=\mathrm{trace}(\Sigma)=\mathrm{trace}(U^*XV)=\mathrm{trace}(XQ)\le\sum_{i=1}^m\sum_{k=1}^m|x_{ik}q_{ki}|\le\sum_{i=1}^m\max_{k=1,\ldots,m}|q_{ki}|\sum_{k=1}^m|x_{ik}|\le\sum_{i=1}^m\sum_{k=1}^m|x_{ik}|.$$

If $1 \le p, q \le \infty$ are conjugate exponents, the following Hölder-type inequality holds for the Schatten norms [8]:

$$\|XY\|_1 \le \|X\|_p \|Y\|_q, \qquad X, Y \in \mathbb{C}^{m\times m}. \tag{2.15}$$

An analogous inequality actually holds for all unitarily invariant norms; see [8, Corollary IV.2.6].

If X is normal, i.e. $XX^* = X^*X$, then X is unitarily diagonalizable, meaning that there exist a unitary matrix U and a diagonal matrix D such that $X = UDU^*$. Using this result, it can be shown that the singular values of X are $|\lambda_j(X)|,\ j = 1, \dots, m$. Consequently, $\|X\| = \rho(X)$ and $\|X\|_p = \left(\sum_{j=1}^m |\lambda_j(X)|^p\right)^{1/p}$ for $1 \le p < \infty$. We observe that, if X is Hermitian ($X^* = X$) or skew-Hermitian ($X^* = -X$), then X is normal. Since $\rho(X) \le |X|_p$ for all $p \in [1, \infty]$ and all matrices $X \in \mathbb{C}^{m\times m}$, it is clear that, whenever X is normal, the inequality $\|X\| \le |X|_p$ holds for all $p \in [1, \infty]$. If X is not normal, an interesting inequality that replaces the previous one is the following:

$$\|X\| \le \sqrt{|X|_1 |X|_\infty}, \qquad X \in \mathbb{C}^{m\times m}. \tag{2.16}$$

This inequality is particularly useful to estimate the spectral norm of a matrix when we have upper bounds for its components. Indeed, we recall that $|X|_1$ and $|X|_\infty$ admit explicit expressions in terms of the components of X. More precisely, $|X|_1$ is the maximum among the 1-norms of the columns of X, and $|X|_\infty$ is the maximum among the 1-norms of the rows of X.

In the next theorems, we recall some important interlacing and perturbation theorems for singular values and eigenvalues. In the statement of Theorems 2.1–2.2, we use the convention introduced in Section 2.1 about the meaning of $\lambda_j(X)$ and $\sigma_j(X)$ when j lies outside the range $1, \dots, m$, being m the size of X.

Theorem 2.1 (**interlacing theorem for singular values**). *Let $Y = X + E$, where $X, E \in \mathbb{C}^{m\times m}$ and* $\mathrm{rank}(E) \le k$*. Then*

$$\sigma_{j-k}(X) \ge \sigma_j(Y) \ge \sigma_{j+k}(X), \qquad j = 1, \dots, m. \tag{2.17}$$

Theorem 2.2 (**interlacing theorem for eigenvalues**). *Let $Y = X + E$, where $X, E \in \mathbb{C}^{m\times m}$ are Hermitian. Let $k^+, k^- \ge 0$ be respectively the number of positive and the number of negative eigenvalues of E:*

$$k^+ = \#\{j \in \{1, \dots, m\} : \lambda_j(E) > 0\}, \qquad k^- = \#\{j \in \{1, \dots, m\} : \lambda_j(E) < 0\}.$$

Then

$$\lambda_{j-k^+}(X) \ge \lambda_j(Y) \ge \lambda_{j+k^-}(X), \qquad j = 1, \dots, m.$$

In particular, if $\mathrm{rank}(E) \le k$ *then*

$$\lambda_{j-k}(X) \ge \lambda_j(Y) \ge \lambda_{j+k}(X), \qquad j = 1, \dots, m. \tag{2.18}$$

Theorem 2.1 can be seen as a corollary of Theorem 2.2. Indeed, for any $A \in \mathbb{C}^{m\times m}$, the eigenvalues of the $(2m) \times (2m)$ Hermitian matrix

$$\widetilde{A} = \begin{bmatrix} O & A \\ A^* & O \end{bmatrix}$$

are $\sigma_j(A),\ -\sigma_j(A),\ j = 1, \dots, m$; see [8, Exercise II.1.15]. Therefore, applying Theorem 2.2 with $\widetilde{Y}$, $\widetilde{X}$, $\widetilde{E}$ in place of Y, X, E, we obtain Theorem 2.1. The proof of Theorem 2.2 can be done by using the result of [8, Exercise III.2.4].

Theorem 2.3 (**perturbation theorem for singular values**). *Let $X, Y \in \mathbb{C}^{m\times m}$, then*

$$|\sigma_j(X) - \sigma_j(Y)| \le \|X - Y\|, \quad j = 1, \dots, m.$$

Theorem 2.4 (**perturbation theorem for eigenvalues**). *Let $X, Y \in \mathbb{C}^{m\times m}$ be Hermitian, then*

$$|\lambda_j(X) - \lambda_j(Y)| \le \|X - Y\|, \quad j = 1, \dots, m.$$

Theorem 2.4 is Weyl's perturbation theorem [8, Corollary III.2.6]. Theorem 2.3 can be seen as a corollary of Theorem 2.4, by considering again the matrices $\widetilde{X}$ and $\widetilde{Y}$. Alternatively, Theorem 2.3 (resp. Theorem 2.4) can be proved by using the minimax principle for singular values [8, Problem III.6.1] (resp. the minimax principle for eigenvalues [8, Corollary III.1.2]). We also refer the reader to [8, Problem II.6.13] for a general perturbation theorem for singular values, which extends Theorem 2.3.

2.4.1 Tensor products and direct sums

If X, Y are matrices of any dimension, say $X \in \mathbb{C}^{m_1\times m_2}$ and $Y \in \mathbb{C}^{\ell_1\times \ell_2}$, the tensor (Kronecker) product of X and Y is the $m_1\ell_1 \times m_2\ell_2$ matrix defined by

$$X \otimes Y = [x_{ij}Y]_{\substack{i=1,\dots,m_1 \\ j=1,\dots,m_2}} = \begin{bmatrix} x_{11}Y & \cdots & x_{1m_2}Y \\ \vdots & & \vdots \\ x_{m_11}Y & \cdots & x_{m_1m_2}Y \end{bmatrix};$$

and the direct sum of X and Y is the $(m_1 + \ell_1) \times (m_2 + \ell_2)$ matrix defined by

$$X \oplus Y = \text{diag}(X, Y) = \begin{bmatrix} X & O \\ O & Y \end{bmatrix}.$$

Tensor products and direct sums possess a lot of nice algebraic properties.

(i) Associativity: for all matrices X, Y, Z, $(X \otimes Y) \otimes Z = X \otimes (Y \otimes Z)$ and $(X \oplus Y) \oplus Z = X \oplus (Y \oplus Z)$. This means that we can omit parentheses in expressions like $X_1 \otimes X_2 \otimes \cdots \otimes X_d$ or $X_1 \oplus X_2 \oplus \cdots \oplus X_d$.

(ii) The relations $(X_1 \otimes Y_1)(X_2 \otimes Y_2) = (X_1X_2) \otimes (Y_1Y_2)$ and $(X_1 \oplus Y_1)(X_2 \oplus Y_2) = (X_1X_2) \oplus (Y_1Y_2)$ hold whenever X_1, X_2 can be multiplied and Y_1, Y_2 can be multiplied.

(iii) For all matrices X, Y, $(X \otimes Y)^* = X^* \otimes Y^*$, $(X \oplus Y)^* = X^* \oplus Y^*$ and $(X \otimes Y)^T = X^T \otimes Y^T$, $(X \oplus Y)^T = X^T \oplus Y^T$.

(iv) Bilinearity (of tensor products): $(\alpha_1X_1 + \alpha_2X_2) \otimes (\beta_1Y_1 + \beta_2Y_2) = \alpha_1\beta_1(X_1 \otimes Y_1) + \alpha_1\beta_2(X_1 \otimes Y_2) + \alpha_2\beta_1(X_2 \otimes Y_1) + \alpha_2\beta_2(X_2 \otimes Y_2)$ for all $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ and for all matrices X_1, X_2, Y_1, Y_2 such that X_1, X_2 are summable and Y_1, Y_2 are summable.

(v) Multi-index formula (for tensor products): if we have d matrices $X_h \in \mathbb{C}^{m_h \times m_h}$, $h = 1, \ldots, d$, then

$$(X_1 \otimes X_2 \otimes \cdots \otimes X_d)_{\boldsymbol{ij}} = (X_1)_{i_1j_1}(X_2)_{i_2j_2} \cdots (X_d)_{i_dj_d}, \qquad \boldsymbol{i}, \boldsymbol{j} = \boldsymbol{1}, \ldots, \boldsymbol{m}, \tag{2.19}$$

where $\boldsymbol{m} = (m_1, m_2, \ldots, m_d)$. This means that, for all $\boldsymbol{i}, \boldsymbol{j}$ in the multi-index range $\boldsymbol{1}, \ldots, \boldsymbol{m}$, the $(\boldsymbol{i}, \boldsymbol{j})$-th entry of $X_1 \otimes X_2 \otimes \cdots \otimes X_d$ is given by (2.19). Note that it makes sense to talk about the $(\boldsymbol{i}, \boldsymbol{j})$-th entry of $X_1 \otimes X_2 \otimes \cdots \otimes X_d$, because we have fixed for the set $\boldsymbol{1}, \ldots, \boldsymbol{m}$ the lexicographic ordering (2.1). Note also that (2.19) can be rewritten in the form (2.2) as follows:

$$X_1 \otimes \cdots \otimes X_d = [(X_1)_{i_1j_1}(X_2)_{i_2j_2} \cdots (X_d)_{i_dj_d}]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}.$$

Eq. (2.19) is of fundamental importance and, indeed, it motivates the introduction of multi-indices to index the entries of a matrix formed by a sum of one or more tensor products. To understand better the importance of (2.19), try to write the (i, j)-th entry of $X_1 \otimes X_2 \otimes \cdots \otimes X_d$ as a function of two linear indices $i, j = 1, \ldots, N(\boldsymbol{m})$.

From (i)–(v), a lot of other interesting properties follow. For example, if X, Y are invertible, then $X \otimes Y$ is invertible, its inverse being $X^{-1} \otimes Y^{-1}$. If X, Y are normal (resp. Hermitian, symmetric, unitary) then $X \otimes Y$ is also normal (resp. Hermitian, symmetric, unitary). If $X \in \mathbb{C}^{m \times m}$ and $Y \in \mathbb{C}^{\ell \times \ell}$, the eigenvalues and singular values of $X \otimes Y$ (resp. $X \oplus Y$) are $\{\lambda_i(X)\lambda_j(Y) :\ i = 1, \ldots, m,\ j = 1, \ldots, \ell\}$ and $\{\sigma_i(X)\sigma_j(Y) :\ i = 1, \ldots, m,\ j = 1, \ldots, \ell\}$ (resp. $\{\lambda_i(X) :\ i = 1, \ldots, m\} \cup \{\lambda_j(Y) :\ j = 1, \ldots, \ell\}$ and $\{\sigma_i(X) :\ i = 1, \ldots, m\} \cup \{\sigma_j(Y) :\ j = 1, \ldots, \ell\}$). In particular, for all $X \in \mathbb{C}^{m \times m}$ and $Y \in \mathbb{C}^{n \times n}$,

$$\|X \oplus Y\| = \max(\|X\|, \|Y\|), \qquad \|X \otimes Y\| = \|X\|\,\|Y\|, \tag{2.20}$$

$$\|X \oplus Y\|_p = \left(\|X\|_p^p + \|Y\|_p^p\right)^{1/p}, \qquad \|X \otimes Y\|_p = \|X\|_p\,\|Y\|_p, \qquad 1 \le p < \infty, \tag{2.21}$$

$$\text{rank}(X \oplus Y) = \text{rank}(X) + \text{rank}(Y), \qquad \text{rank}(X \otimes Y) = \text{rank}(X)\,\text{rank}(Y). \tag{2.22}$$

A property of tensor products, which is not as popular as the previous ones, is given in Lemma 2.7. For the proof we refer the reader to [24, Lemma 1.2].

Lemma 2.7. *For all $\boldsymbol{m} \in \mathbb{N}^d$ and all permutations σ of the set $\{1, \ldots, d\}$, there exists a permutation matrix $\Pi_{\boldsymbol{m};\sigma}$ of size $N(\boldsymbol{m})$ such that*

$$X_{\sigma(1)} \otimes X_{\sigma(2)} \otimes \cdots \otimes X_{\sigma(d)} = \Pi_{\boldsymbol{m};\sigma}(X_1 \otimes X_2 \otimes \cdots \otimes X_d)\Pi_{\boldsymbol{m};\sigma}^T,$$

for all matrices $X_1 \in \mathbb{C}^{m_1 \times m_1}, X_2 \in \mathbb{C}^{m_2 \times m_2}, \ldots, X_d \in \mathbb{C}^{m_d \times m_d}$.

Lemma 2.7 says that the tensor product is 'almost' commutative. It is important to notice that the permutation matrix $\Pi_{\boldsymbol{m};\sigma}$ depends only on $\boldsymbol{m}$ and σ, and not on the specific matrices $X_1, X_2, \ldots, X_d$.

Concerning the 'distributive properties' of tensor products with respect to direct sums, a result analogous to Lemma 2.7 holds. It shows that these properties hold modulo permutation transformations which only depend on the dimensions of the involved matrices. For the proof of Lemma 2.8, see [24, Lemma 1.4].

Lemma 2.8. *For all $\ell \in \mathbb{N}$ and $\boldsymbol{m} \in \mathbb{N}^d$, there exists a permutation matrix $Q_{\ell,\boldsymbol{m}}$ of size $\ell(m_1 + \ldots + m_d)$ such that*

$$X \otimes (Y_1 \oplus Y_2 \oplus \cdots \oplus Y_d) = Q_{\ell,\boldsymbol{m}}[(X \otimes Y_1) \oplus (X \otimes Y_2) \oplus \cdots \oplus (X \otimes Y_d)]Q_{\ell,\boldsymbol{m}}^T,$$

for all matrices $X \in \mathbb{C}^{\ell \times \ell}$, $Y_1 \in \mathbb{C}^{m_1 \times m_1}, Y_2 \in \mathbb{C}^{m_2 \times m_2}, \ldots, Y_d \in \mathbb{C}^{m_d \times m_d}$.

Lemma 2.8 gives the distributive law on the left; the distributive law on the right holds without permutation transformations. Indeed, it follows directly from the definition of tensor products and direct sums that, for all matrices $X_1, \ldots, X_d, Y$,

$$(X_1 \oplus X_2 \oplus \cdots \oplus X_d) \otimes Y = (X_1 \otimes Y) \oplus (X_2 \otimes Y) \oplus \cdots \oplus (X_d \otimes Y). \tag{2.23}$$

We end this section with a result about direct sums, which is completely analogous to Lemma 2.7: it shows that the direct sum operation is 'almost' commutative.

Lemma 2.9. *For all $\boldsymbol{m} \in \mathbb{N}^d$ and all permutations σ of the set $\{1, \ldots, d\}$, there exists a permutation matrix $V_{\boldsymbol{m};\sigma}$ of size $m_1 + \ldots + m_d$ such that*

$$X_{\sigma(1)} \oplus X_{\sigma(2)} \oplus \cdots \oplus X_{\sigma(d)} = V_{\boldsymbol{m};\sigma}(X_1 \oplus X_2 \oplus \cdots \oplus X_d)V_{\boldsymbol{m};\sigma}^T,$$

for all matrices $X_1 \in \mathbb{C}^{m_1 \times m_1}, X_2 \in \mathbb{C}^{m_2 \times m_2}, \ldots, X_d \in \mathbb{C}^{m_d \times m_d}$.

Proof. The proof is done by induction on d. For $d = 1$, the only possible permutation is $\sigma = [1]$ and we can take $V_{m;[1]} = I_m$. For $d = 2$, the only possible permutations are the identity $\sigma = [1, 2]$ and the transposition $\sigma = [2, 1]$, and we can take

$$V_{\boldsymbol{m};[1,2]} = I_{m_1+m_2}, \qquad V_{\boldsymbol{m};[2,1]} = \begin{bmatrix} O & I_{m_2} \\ I_{m_1} & O \end{bmatrix}.$$

For $d \geq 3$, let i be the index for which $\sigma(i) = d$. Define τ to be the permutation of the set $\{1, \ldots, d-1\}$ such that $\tau(j) = \sigma(j)$ for $j = 1, \ldots, i-1$ and $\tau(j) = \sigma(j+1)$ for $j = i, \ldots, d-1$. If $i = d$, then, by induction hypothesis,

$$\begin{aligned} X_{\sigma(1)} \oplus \cdots \oplus X_{\sigma(d)} &= X_{\tau(1)} \oplus \cdots \oplus X_{\tau(d-1)} \oplus X_d \\ &= V_{(m_1,\ldots,m_{d-1});\tau}(X_1 \oplus \cdots \oplus X_{d-1})V_{(m_1,\ldots,m_{d-1});\tau}^T \oplus X_d \\ &= (V_{(m_1,\ldots,m_{d-1});\tau} \oplus I_{m_d})(X_1 \oplus \cdots \oplus X_{d-1} \oplus X_d)(V_{(m_1,\ldots,m_{d-1});\tau} \oplus I_{m_d})^T \end{aligned}$$

and the thesis holds with $V_{\boldsymbol{m};\sigma} = V_{(m_1,\ldots,m_{d-1});\tau} \oplus I_{m_d}$. If $i < d$, then, by induction hypothesis,

$$\begin{aligned} &X_{\sigma(1)} \oplus \cdots \oplus X_{\sigma(d)} = X_{\sigma(1)} \oplus \cdots X_{\sigma(i-1)} \oplus X_d \oplus X_{\sigma(i+1)} \oplus \cdots \oplus X_{\sigma(d)} \\ &= X_{\sigma(1)} \oplus \cdots X_{\sigma(i-1)} \oplus \Big[V_{(m_{\sigma(i+1)}+\ldots+m_{\sigma(d)},m_d);[2,1]}(X_{\sigma(i+1)} \oplus \cdots \oplus X_{\sigma(d)} \oplus X_d)V_{(m_{\sigma(i+1)}+\ldots+m_{\sigma(d)},m_d);[2,1]}^T\Big] \\ &= (I_{m_{\sigma(1)}+\ldots+m_{\sigma(i-1)}} \oplus V_{(m_{\sigma(i+1)}+\ldots+m_{\sigma(d)},m_d);[2,1]}) \big(X_{\sigma(1)} \oplus \cdots \oplus X_{\sigma(i-1)} \oplus X_{\sigma(i+1)} \oplus \cdots \oplus X_{\sigma(d)} \oplus X_d\big) \\ &\quad \cdot (I_{m_{\sigma(1)}+\ldots+m_{\sigma(i-1)}} \oplus V_{(m_{\sigma(i+1)}+\ldots+m_{\sigma(d)},m_d);[2,1]})^T \\ &= U_{\boldsymbol{m};\sigma} \big(X_{\tau(1)} \oplus \cdots \oplus X_{\tau(d-1)} \oplus X_d\big) U_{\boldsymbol{m};\sigma}^T, \end{aligned} \tag{2.24}$$

where $U_{\boldsymbol{m};\sigma} = I_{m_{\sigma(1)}+\ldots+m_{\sigma(i-1)}} \oplus V_{(m_{\sigma(i+1)}+\ldots+m_{\sigma(d)},m_d);[2,1]}$. Using again the induction hypothesis, we obtain

$$X_{\tau(1)} \oplus \cdots \oplus X_{\tau(d-1)} = V_{(m_1,\ldots,m_{d-1});\tau}(X_1 \oplus \cdots \oplus X_{d-1})V_{(m_1,\ldots,m_{d-1});\tau}^T.$$

Substituting this into (2.24), we see that the thesis holds with $V_{\boldsymbol{m};\sigma} = U_{\boldsymbol{m};\sigma}(V_{(m_1,\ldots,m_{d-1});\tau} \oplus I_{m_d})$. □

2.5 Singular value and eigenvalue distribution of a matrix-sequence

Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$. We recall that to any measurable function $g : D \subset \mathbb{R}^k \to \mathbb{K}$ we associate the functional $\phi_{[g]}$ defined over $C_c(\mathbb{K})$ by eq. (2.6). This functional will play an important role in this section and in the remainder of this work.

Definition 2.2 (singular value and eigenvalue distribution of a matrix-sequence, spectral symbol). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence.

- We say that $\{A_{\boldsymbol{n}}\}_n$ has an asymptotic singular value distribution described by a functional $\phi : C_c(\mathbb{R}) \to \mathbb{C}$, and we write $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \phi$, if, for all $F \in C_c(\mathbb{R})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) = \phi(F). \tag{2.25}$$

  In the case where $\phi = \phi_{[|f|]}$ for a measurable function $f : D \subset \mathbb{R}^k \to \mathbb{C}$ defined on a set D with $0 < \mu_k(D) < \infty$, the function f is referred to as the *singular value symbol* of the matrix-sequence $\{A_{\boldsymbol{n}}\}_n$, and we write $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$ instead of $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \phi_{[|f|]}$.

- We say that $\{A_{\boldsymbol{n}}\}_n$ has an asymptotic eigenvalue (or spectral) distribution described by a functional $\phi : C_c(\mathbb{C}) \to \mathbb{C}$, and we write $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \phi$, if, for all $F \in C_c(\mathbb{C})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) = \phi(F). \tag{2.26}$$

  In the case where $\phi = \phi_{[f]}$ for a measurable function $f : D \subset \mathbb{R}^k \to \mathbb{C}$ defined on a set D with $0 < \mu_k(D) < \infty$, the function f is referred to as the *eigenvalue (or spectral) symbol* of the matrix-sequence $\{A_{\boldsymbol{n}}\}_n$, and we write $\{A_{\boldsymbol{n}}\}_n \sim_\lambda f$ instead of $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \phi_{[f]}$.

When we write a relation such as $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \phi$ (resp. $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \phi$), it is understood that ϕ is a functional on $C_c(\mathbb{R})$ (resp. $C_c(\mathbb{C})$), as in Definition 2.2. Similarly, when we write a relation such as $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$ or $\{A_{\boldsymbol{n}}\}_n \sim_\lambda f$, it is understood that f is as in Definition 2.2; that is, f is a measurable function defined on a subset D of some $\mathbb{R}^k$ with $0 < \mu_k(D) < \infty$. Sometimes, for shortness, we will write $\{A_{\boldsymbol{n}}\}_n \sim_{\sigma,\lambda} f$ to indicate that $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$ and $\{A_{\boldsymbol{n}}\}_n \sim_\lambda f$.

Remark 2.2. By definition of $\phi_{[f]}$, see eq. (2.6), the spectral distribution relation $\{A_{\boldsymbol{n}}\}_n \sim_\lambda f$ means that, for all $F \in C_c(\mathbb{C})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) = \frac{1}{\mu_k(D)} \int_D F(f(\mathbf{x}))d\mathbf{x}. \tag{2.27}$$

The informal meaning behind (2.27) is the following. If f is at least Riemann-integrable, n is large enough, and $\{\mathbf{x}_{j,n},\ j = 1, \ldots, N(\boldsymbol{n})\}$ is an equispaced grid on D, then a suitable ordering $\lambda_j(A_{\boldsymbol{n}})$, $j = 1, \ldots, N(\boldsymbol{n})$, of the eigenvalues of $A_{\boldsymbol{n}}$ is such that the pairs $\{(\mathbf{x}_{j,n}, \lambda_j(A_{\boldsymbol{n}})),\ j = 1, \ldots, N(\boldsymbol{n})\}$ reconstruct approximately the hypersurface $\{(\mathbf{x}, f(\mathbf{x})),\ \mathbf{x} \in D\}$. In other words, the spectrum of $A_{\boldsymbol{n}}$ 'behaves' (asymptotically) like a uniform sampling of f over D. For instance, if $k = 1$, $N(\boldsymbol{n}) = n$ and $D = [a, b]$, then the eigenvalues of $A_{\boldsymbol{n}}$ are approximately equal to $f(a + i(b - a)/n)$, $i = 1, \ldots, n$, for n large enough. Similarly, if $k = 2$, $N(\boldsymbol{n}) = n^2$ and $D = [a_1, b_1] \times [a_2, b_2]$, then the eigenvalues of $A_{\boldsymbol{n}}$ are approximately equal to $f(a_1 + i(b_1 - a_1)/n, a_2 + j(b_2 - a_2)/n)$, $i, j = 1, \ldots, n$, for n large enough. A completely analogous meaning can be given also for the singular value distribution relation $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$, which is equivalent to say that, for all $F \in C_c(\mathbb{R})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) = \frac{1}{\mu_k(D)} \int_D F(|f(\mathbf{x})|)d\mathbf{x}. \tag{2.28}$$

Remark 2.3. It is clear from Definition 2.2 that $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$ is equivalent to $\{A_{\boldsymbol{n}}\}_n \sim_\sigma |f|$. Moreover, if every $A_{\boldsymbol{n}}$ is normal and $\{A_{\boldsymbol{n}}\}_n \sim_\lambda f$, then $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$. Indeed, since $A_{\boldsymbol{n}}$ is normal, its singular values coincide with the moduli of the eigenvalues. Therefore, for any fixed $F \in C_c(\mathbb{R})$, by applying the eigenvalue distribution relation (2.27) with the test function $F(|\cdot|) \in C_c(\mathbb{C})$, we get

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) = \lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(|\lambda_j(A_{\boldsymbol{n}})|) = \frac{1}{\mu_k(D)} \int_D F(|f(\mathbf{x})|)d\mathbf{x}.$$

Hence, $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$.

Given a function $f : D \subset \mathbb{R}^k \to \mathbb{C}$, we recall that the essential range of f, denoted by $\mathcal{ER}(f)$, is defined as the set of points $z \in \mathbb{C}$ such that, for every $\epsilon > 0$, the measure of $\{f \in D(z, \epsilon)\}$ is positive. In formulas,

$$\mathcal{ER}(f) = \{z \in \mathbb{C} :\ \mu_k\{f \in D(z, \epsilon)\} > 0 \text{ for all } \epsilon > 0\}.$$

Note that $\mathcal{ER}(f)$ is always closed (the complement is open). Moreover, it can be shown that $f(\mathbf{x}) \in \mathcal{ER}(f)$ for almost every $\mathbf{x} \in D$, i.e., $f \in \mathcal{ER}(f)$ a.e. In addition, whenever f is continuous and D is sufficiently regular (say, D is contained in the closure of its interior), then $\mathcal{ER}(f)$ coincides with the closure of the image of f. The following result is proved in [30, Theorem 2.4].

Proposition 2.1. *If $\{A_{\boldsymbol{n}}\} \sim_\lambda f$, then every point of the essential range $\mathcal{ER}(f)$ strongly attracts the spectrum $\Lambda(A_{\boldsymbol{n}})$ with infinite order. This means that, for every $z \in \mathcal{ER}(f)$, if we order the eigenvalues of $A_{\boldsymbol{n}}$ according to their distance from z,*

$$|\lambda_1(A_{\boldsymbol{n}}) - z| \le |\lambda_2(A_{\boldsymbol{n}}) - z| \le \ldots,$$

then $\lim_{n\to\infty} |\lambda_j(A_{\boldsymbol{n}}) - z| = 0$ for any fixed $j \ge 1$.

2.5.1 Zero-distributed sequences

A class of matrix-sequences that plays a central role in the framework of the theory of GLT sequences is the class of zero-distributed matrix-sequences (or zero-distributed sequences for shortness). A zero-distributed sequence is simply a matrix-sequence $\{Z_{\boldsymbol{n}}\}_n$ such that $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$. Theorem 2.5 provides a characterization of zero-distributed sequences, which will allow us to show in Section 4 that any zero-distributed sequence is a sLT sequence. Theorem 2.6 gives a sufficient condition, formulated in terms of Schatten p-norms, that ensure a matrix-sequence $\{Z_{\boldsymbol{n}}\}_n$ to be zero-distributed.

Theorem 2.5. *Let $\{Z_{\boldsymbol{n}}\}_n$ be a matrix-sequence. Then, the following conditions are equivalent.*

1. $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$.
2. *For every $\epsilon > 0$,* $\displaystyle\lim_{n\to\infty} \frac{\#\{j \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_j(Z_{\boldsymbol{n}}) > \epsilon\}}{N(\boldsymbol{n})} = 0.$
3. *For every n we have $Z_{\boldsymbol{n}} = R_{\boldsymbol{n}} + N_{\boldsymbol{n}}$, where* $\displaystyle\lim_{n\to\infty} \frac{\mathrm{rank}(R_{\boldsymbol{n}})}{N(\boldsymbol{n})} = \lim_{n\to\infty} \|N_{\boldsymbol{n}}\| = 0.$

Proof. $(1 \Rightarrow 2)$ For every $\epsilon > 0$, take $F_\epsilon \in C_c(\mathbb{R})$ such that $F_\epsilon = 1$ over $[0, \epsilon/2]$, $F_\epsilon = 0$ over $[\epsilon, \infty)$ and $0 \le F_\epsilon \le 1$ over $[0, \infty)$. Note that $F_\epsilon \le \chi_{[0,\epsilon]}$ over $[0, \infty)$. Since $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$ by assumption, we have

$$\begin{aligned}\frac{\#\{j \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_j(Z_{\boldsymbol{n}}) > \epsilon\}}{N(\boldsymbol{n})} &= 1 - \frac{\#\{j \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_j(Z_{\boldsymbol{n}}) \le \epsilon\}}{N(\boldsymbol{n})} = 1 - \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} \chi_{[0,\epsilon]}(\sigma_j(Z_{\boldsymbol{n}}))\\ &\le 1 - \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(Z_{\boldsymbol{n}})) \overset{n\to\infty}{\longrightarrow} 1 - F_\epsilon(0) = 0.\end{aligned}$$

$(2 \Rightarrow 3)$ By assumption, for every $\epsilon > 0$ the quantity

$$q_{\boldsymbol{n}}(\epsilon) = \frac{\#\{j \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_j(Z_{\boldsymbol{n}}) > \epsilon\}}{N(\boldsymbol{n})}$$

tends to 0 as $n \to \infty$. Hence, there exists a sequence $\{\epsilon_{\boldsymbol{n}}\}_n$ of positive numbers such that, when $n \to \infty$,

$$\epsilon_{\boldsymbol{n}} \to 0, \qquad q_{\boldsymbol{n}}(\epsilon_{\boldsymbol{n}}) = \frac{\#\{j \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_j(Z_{\boldsymbol{n}}) > \epsilon_{\boldsymbol{n}}\}}{N(\boldsymbol{n})} \to 0.$$

Let $Z_{\boldsymbol{n}} = U_{\boldsymbol{n}}\Sigma_{\boldsymbol{n}}V_{\boldsymbol{n}}^*$ be an SVD of $Z_{\boldsymbol{n}}$. Let $\hat{\Sigma}_{\boldsymbol{n}}$ be the matrix obtained from $\Sigma_{\boldsymbol{n}}$ by setting to 0 all the singular values of $Z_{\boldsymbol{n}}$ that are less than or equal to $\epsilon_{\boldsymbol{n}}$, and let $\tilde{\Sigma}_{\boldsymbol{n}} = \Sigma_{\boldsymbol{n}} - \hat{\Sigma}_{\boldsymbol{n}}$ be the matrix obtained from $\Sigma_{\boldsymbol{n}}$ by setting to 0 all the singular values of $Z_{\boldsymbol{n}}$ that exceed $\epsilon_{\boldsymbol{n}}$. Then,

$$Z_{\boldsymbol{n}} = U_{\boldsymbol{n}}\Sigma_{\boldsymbol{n}}V_{\boldsymbol{n}}^* = U_{\boldsymbol{n}}\hat{\Sigma}_{\boldsymbol{n}}V_{\boldsymbol{n}}^* + U_{\boldsymbol{n}}\tilde{\Sigma}_{\boldsymbol{n}}V_{\boldsymbol{n}}^* = R_{\boldsymbol{n}} + N_{\boldsymbol{n}},$$

where $R_{\boldsymbol{n}} = U_{\boldsymbol{n}}\hat{\Sigma}_{\boldsymbol{n}}V_{\boldsymbol{n}}^*$ and $N_{\boldsymbol{n}} = U_{\boldsymbol{n}}\tilde{\Sigma}_{\boldsymbol{n}}V_{\boldsymbol{n}}^*$ satisfy

$$\frac{\mathrm{rank}(R_{\boldsymbol{n}})}{N(\boldsymbol{n})} = \frac{\#\{j \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_j(Z_{\boldsymbol{n}}) > \epsilon_{\boldsymbol{n}}\}}{N(\boldsymbol{n})} = q_{\boldsymbol{n}}(\epsilon_{\boldsymbol{n}}) \overset{n\to\infty}{\longrightarrow} 0, \qquad \|N_{\boldsymbol{n}}\| \le \epsilon_{\boldsymbol{n}} \overset{n\to\infty}{\longrightarrow} 0.$$

$(3 \Rightarrow 1)$ In order to give an elegant and extremely short proof of this implication, we use some results from Section 3. The assumption in item 3 ensures that $\{\{O_{N(\boldsymbol{n})}\}_n\}_m$ is an approximating class of sequences for $\{Z_{\boldsymbol{n}}\}_n$ according to Definition 3.1. Moreover, it is clear that $\{O_{N(\boldsymbol{n})}\}_n \sim_\sigma 0$. Hence, $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$ by Corollary 3.1. □

With the terminology of clustering (see, e.g., [30, p. 86] or [24, Section 1.3]), condition 2 in Theorem 2.5 is expressed by saying that $\{Z_{\boldsymbol{n}}\}_n$ is weakly clustered at 0 in the sense of the singular values.

In the statement of the next theorem, we use the natural convention $1/\infty = 0$.

Theorem 2.6. *Let $\{Z_{\boldsymbol{n}}\}_n$ be a matrix-sequence and suppose that, for some $p \in [1, \infty]$,*

$$\lim_{n\to\infty} \frac{\|Z_{\boldsymbol{n}}\|_p}{N(\boldsymbol{n})^{1/p}} = 0.$$

Then $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$. In particular, if $\|Z_{\boldsymbol{n}}\| \to 0$ then $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$.

Proof. In view of (2.12)–(2.13), for all $p \in [1, \infty]$ we have

$$\frac{\|Z_{\boldsymbol{n}}\|_1}{N(\boldsymbol{n})} \le \frac{\|Z_{\boldsymbol{n}}\|_p}{N(\boldsymbol{n})^{1/p}}.$$

Hence, it suffices to prove the theorem under the assumption that

$$\lim_{n\to\infty} \frac{\|Z_{\boldsymbol{n}}\|_1}{N(\boldsymbol{n})} = 0. \tag{2.29}$$

What we have to show is that

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(Z_{\boldsymbol{n}})) = F(0) \tag{2.30}$$

for all $F \in C_c(\mathbb{R})$. The proof of (2.30) is easy if F is a real-valued function in $C_c^1(\mathbb{R})$, because in this case we have

$$\left|\frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(Z_{\boldsymbol{n}})) - F(0)\right| \le \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} |F(\sigma_j(Z_{\boldsymbol{n}})) - F(0)| \le \frac{\|F'\|_\infty}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} \sigma_j(Z_{\boldsymbol{n}}) = \frac{\|F'\|_\infty}{N(\boldsymbol{n})} \|Z_{\boldsymbol{n}}\|_1,$$

which tends to 0 by (2.29). In the case of a general $F \in C_c(\mathbb{R})$, we approximate (in ∞-norm) the real and the imaginary part of F by means of real-valued functions in $C_c^1(\mathbb{R})$, and we see that (2.30) continues to hold. Let us work out the details. For every $\epsilon > 0$, let $\Re_\epsilon$ and $\Im_\epsilon$ be real-valued functions in $C_c^1(\mathbb{R})$ such that $\|\Re(F) - \Re_\epsilon\|_\infty \le \epsilon$ and $\|\Im(F) - \Im_\epsilon\|_\infty \le \epsilon$. Set $F_\epsilon = \Re_\epsilon + \mathrm{i}\,\Im_\epsilon$. Then, $\|F - F_\epsilon\|_\infty \le 2\epsilon$ and

$$\left|\left|\frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(Z_{\boldsymbol{n}})) - F(0)\right| - \left|\frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(Z_{\boldsymbol{n}})) - F_\epsilon(0)\right|\right| \le 4\epsilon.$$

Since (2.30) holds for $\Re_\epsilon$, $\Im_\epsilon$ (and hence also for F_ϵ), the previous inequality implies that

$$\limsup_{n\to\infty} \left|\frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(Z_{\boldsymbol{n}})) - F(0)\right| \le 4\epsilon,$$

for every $\epsilon > 0$. Thus, (2.30) holds for F. □

2.6 Multilevel Toeplitz matrices

In this section we recall the definition and some properties of multilevel Toeplitz matrices. Of course, we do not pretend to cover here, in a few pages, all the details of this extensive topic. We just provide the results that will be used in this work.

Let $\boldsymbol{n} \in \mathbb{N}^d$. A matrix of the form

$$[a_{\boldsymbol{i}-\boldsymbol{j}}]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}} \in \mathbb{C}^{N(\boldsymbol{n})\times N(\boldsymbol{n})}, \tag{2.31}$$

whose $(\boldsymbol{i}, \boldsymbol{j})$-th entry depends only on the difference between the d-indices $\boldsymbol{i}$ and $\boldsymbol{j}$, is called a multilevel Toeplitz matrix, or, more precisely, a d-level Toeplitz matrix. If $f : [-\pi, \pi]^d \to \mathbb{C}$ is a function in $L^1([-\pi, \pi]^d)$, we denote its Fourier coefficients by

$$f_{\boldsymbol{k}} = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} f(\boldsymbol{\theta}) \mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}} d\boldsymbol{\theta}, \qquad \boldsymbol{k} \in \mathbb{Z}^d. \tag{2.32}$$

The $\boldsymbol{n}$-th (multilevel) Toeplitz matrix associated with f is defined as

$$T_{\boldsymbol{n}}(f) = [f_{\boldsymbol{i}-\boldsymbol{j}}]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}}. \tag{2.33}$$

We call $\{T_{\boldsymbol{n}}(f)\}_{\boldsymbol{n}\in\mathbb{N}^d}$ the family of Toeplitz matrices associated with f, which, in turn, is called the generating function of $\{T_{\boldsymbol{n}}(f)\}_{\boldsymbol{n}\in\mathbb{N}^d}$.

For each fixed $\boldsymbol{n} \in \mathbb{N}^d$, the application $T_{\boldsymbol{n}}(\cdot) : L^1([-\pi, \pi]^d) \to \mathbb{C}^{N(\boldsymbol{n})\times N(\boldsymbol{n})}$ is linear:

$$T_{\boldsymbol{n}}(\alpha f + \beta g) = \alpha T_{\boldsymbol{n}}(f) + \beta T_{\boldsymbol{n}}(g), \qquad \alpha, \beta \in \mathbb{C}, \qquad f, g \in L^1([-\pi, \pi]^d).$$

This follows from the relation $(\alpha f + \beta g)_{\boldsymbol{k}} = \alpha f_{\boldsymbol{k}} + \beta g_{\boldsymbol{k}}$, $\boldsymbol{k} \in \mathbb{Z}^d$, which is a consequence of the linearity of the integral in (2.32). Another nice property of the Toeplitz operator $T_{\boldsymbol{n}}(\cdot)$ is that $T_{\boldsymbol{n}}(1) = I_{N(\boldsymbol{n})}$. For every $f \in L^1([-\pi, \pi]^d)$, the Fourier coefficients of f are related to the Fourier coefficients of its conjugate $\overline{f}$ by

$$\overline{f_{\boldsymbol{j}}} = \frac{1}{(2\pi)^d} \overline{\int_{[-\pi,\pi]^d} f(\boldsymbol{\theta}) \mathrm{e}^{-\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}} d\boldsymbol{\theta}} = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} \overline{f(\boldsymbol{\theta})} \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}} d\boldsymbol{\theta} = (\overline{f})_{-\boldsymbol{j}}, \qquad \boldsymbol{j} \in \mathbb{Z}^d.$$

Therefore, for all $\boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{n}$,

$$[T_{\boldsymbol{n}}(\overline{f})]_{\boldsymbol{i}\boldsymbol{j}} = (\overline{f})_{\boldsymbol{i}-\boldsymbol{j}} = \overline{f_{\boldsymbol{j}-\boldsymbol{i}}} = [T_{\boldsymbol{n}}(f)^*]_{\boldsymbol{i}\boldsymbol{j}},$$

i.e.,

$$T_{\boldsymbol{n}}(f)^* = T_{\boldsymbol{n}}(\overline{f}).$$

From this identity, which holds for all $\boldsymbol{n} \in \mathbb{N}^d$ and $f \in L^1([-\pi,\pi]^d)$, we infer that, if f is real-valued, or if f is real a.e.,[3] then all the matrices $T_{\boldsymbol{n}}(f)$ are Hermitian.

Theorem 2.7 is a fundamental result concerning multilevel Toeplitz matrices. It is known in the literature as the Szegö–Tilli theorem. We refer the reader to [14] for a rich account concerning the history of the Szegö theorem, which originally appeared in [31] and has undergone several extensions [14, 51, 54, 55]. In particular, Tilli's proof of Theorem 2.7 can be found in [51]. We also refer the reader to [29] for a proof of Theorem 2.7 based on the notion of approximating class of sequences (see Section 3); the proof in [29] is made only in the case of eigenvalues for $d = 1$, but the argument is general and can be extended to singular values and to higher dimensionalities d.

Theorem 2.7. *Let $f \in L^1([-\pi,\pi]^d)$ and consider the sequence of Toeplitz matrices $\{T_{\boldsymbol{n}}(f)\}_n$, where $\boldsymbol{n} \to \infty$ as $n \to \infty$. Then $\{T_{\boldsymbol{n}}(f)\}_n \sim_\sigma f$. Moreover, if f is real a.e., then $\{T_{\boldsymbol{n}}(f)\}_n \sim_\lambda f$.*

We stress that in Theorem 2.7, $\{T_{\boldsymbol{n}}(f)\}_n$ is any sequence of Toeplitz matrices extracted from the family $\{T_{\boldsymbol{n}}(f)\}_{\boldsymbol{n}\in\mathbb{N}^d}$ and such that $\boldsymbol{n} = \boldsymbol{n}(n) \to \infty$ when $n \to \infty$. Although we added the specification '$\boldsymbol{n} \to \infty$ as $n \to \infty$', we recall that this condition is verified, *by definition,* for every matrix-sequence $\{A_{\boldsymbol{n}}\}_n$; see Section 2.1.2.

Important inequalities involving Toeplitz matrices and Schatten p-norms were proved in [48, Corollary 4.2] and generalized in [43, Corollary 3.5]. We report them in the next theorem for future use.

Theorem 2.8. *Let $f \in L^p([-\pi,\pi]^d)$ and $\boldsymbol{n} \in \mathbb{N}^d$. Then,*

$$\|T_{\boldsymbol{n}}(f)\| = \|T_{\boldsymbol{n}}(f)\|_\infty \le \|f\|_{L^\infty}, \tag{2.34}$$

$$\|T_{\boldsymbol{n}}(f)\|_p \le \frac{1}{(2\pi)^{d/p}} \|f\|_{L^p}\, N(\boldsymbol{n})^{1/p}, \qquad 1 \le p < \infty. \tag{2.35}$$

In particular, using the natural convention $1/\infty = 0$, the inequality

$$\|T_{\boldsymbol{n}}(f)\|_p \le N(\boldsymbol{n})^{1/p} \|f\|_{L^p} \tag{2.36}$$

holds for all $p \in [1,\infty]$.

Lemma 2.10 provides a relation between tensor products and Toeplitz matrices.

Lemma 2.10. *Let $f_1, \dots, f_d \in L^1([-\pi,\pi])$ and $\boldsymbol{n} \in \mathbb{N}^d$. Then,*

$$T_{n_1}(f_1) \otimes \cdots \otimes T_{n_d}(f_d) = T_{\boldsymbol{n}}(f_1 \otimes \cdots \otimes f_d). \tag{2.37}$$

Note that the tensor-product function $f_1 \otimes \cdots \otimes f_d : [-\pi,\pi]^d \to \mathbb{C}$ belongs to $L^1([-\pi,\pi]^d)$ by Fubini's theorem.

Proof. The proof is simple if we use the fundamental property (2.19). The Fourier coefficients of $f_1 \otimes \cdots \otimes f_d$ are given by

$$(f_1 \otimes \cdots \otimes f_d)_{\boldsymbol{k}} = (f_1)_{k_1} \cdots (f_d)_{k_d}, \qquad \boldsymbol{k} \in \mathbb{Z}^d.$$

Hence, for all $\boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{n}$,

$$\begin{aligned}[T_{n_1}(f_1) \otimes \cdots \otimes T_{n_d}(f_d)]_{\boldsymbol{i}\boldsymbol{j}} &= [T_{n_1}(f_1)]_{i_1 j_1} \cdots [T_{n_d}(f_d)]_{i_d j_d} = (f_1)_{i_1-j_1} \cdots (f_d)_{i_d-j_d} = (f_1 \otimes \cdots \otimes f_d)_{\boldsymbol{i}-\boldsymbol{j}} \\ &= [T_{\boldsymbol{n}}(f_1 \otimes \cdots \otimes f_d)]_{\boldsymbol{i}\boldsymbol{j}},\end{aligned}$$

and (2.37) follows. □

Lemma 2.11 generalizes [20, Proposition 2] and will be used in Section 4.1.1 to study the so-called Locally Toeplitz operator. By Hölder's inequality [39], if $f \in L^p([-\pi,\pi]^d)$, $g \in L^q([-\pi,\pi]^d)$, and $p, q \in [1,\infty]$ are conjugate exponents, then $fg \in L^1([-\pi,\pi]^d)$. In this case, we can consider the three matrices $T_{\boldsymbol{n}}(f), T_{\boldsymbol{n}}(g), T_{\boldsymbol{n}}(fg)$.

Lemma 2.11. *Let $f \in L^p([-\pi,\pi]^d)$ and $g \in L^q([-\pi,\pi]^d)$, where $1 \le p, q \le \infty$ are conjugate exponents. Then,*

$$\lim_{\boldsymbol{n}\to\infty} \frac{\|T_{\boldsymbol{n}}(f)T_{\boldsymbol{n}}(g) - T_{\boldsymbol{n}}(fg)\|_1}{N(\boldsymbol{n})} = 0. \tag{2.38}$$

[3] Note that two functions $f, g \in L^1([-\pi,\pi]^d)$ which coincide a.e. give rise to the same multilevel Toeplitz matrices $T_{\boldsymbol{n}}(f) = T_{\boldsymbol{n}}(g)$, $\boldsymbol{n} \in \mathbb{N}^d$, because the Fourier coefficients of f and g coincide.

Proof. If f, g are in $L^\infty([-\pi,\pi]^d)$, eq. (2.38) holds by [20, Proposition 2]. In the general case where $f \in L^p([-\pi,\pi]^d)$ and $g \in L^q([-\pi,\pi]^d)$, the proof requires some work.

Take two sequences $\{f_m\}_m$ and $\{g_m\}_m$ such that $f_m,\, g_m \in L^\infty([-\pi,\pi]^d)$ for all m, $f_m \to f$ in $L^p([-\pi,\pi]^d)$ and $g_m \to g$ in $L^q([-\pi,\pi]^d)$; for example, one can choose $f_m = f\,\chi_{[-m,m]}$ and $g_m = g\,\chi_{[-m,m]}$. By the linearity of $T_{\boldsymbol{n}}(\cdot)$ and the inequalities (2.15), (2.36), for every m and every $\boldsymbol{n} \in \mathbb{N}^d$ we have

$$\begin{aligned}
&\|T_{\boldsymbol{n}}(f)T_{\boldsymbol{n}}(g) - T_{\boldsymbol{n}}(fg)\|_1 \\
&\le \|T_{\boldsymbol{n}}(f-f_m)T_{\boldsymbol{n}}(g)\|_1 + \|T_{\boldsymbol{n}}(f_m)T_{\boldsymbol{n}}(g-g_m)\|_1 + \|T_{\boldsymbol{n}}(f_m)T_{\boldsymbol{n}}(g_m) - T_{\boldsymbol{n}}(f_m g_m)\|_1 + \|T_{\boldsymbol{n}}(f_m g_m - fg)\|_1 \\
&\le N(\boldsymbol{n})^{1/p}\|f-f_m\|_{L^p} N(\boldsymbol{n})^{1/q}\|g\|_{L^q} + N(\boldsymbol{n})^{1/p}\|f_m\|_{L^p} N(\boldsymbol{n})^{1/q}\|g-g_m\|_{L^q} \\
&\qquad + \|T_{\boldsymbol{n}}(f_m)T_{\boldsymbol{n}}(g_m) - T_{\boldsymbol{n}}(f_m g_m)\|_1 + N(\boldsymbol{n})\|f_m g_m - fg\|_{L^1} \\
&\le N(\boldsymbol{n}) \left[\|f-f_m\|_{L^p}\|g\|_{L^q} + \sup_i \|f_i\|_{L^p}\|g-g_m\|_{L^q} + \frac{\|T_{\boldsymbol{n}}(f_m)T_{\boldsymbol{n}}(g_m) - T_{\boldsymbol{n}}(f_m g_m)\|_1}{N(\boldsymbol{n})} + \|f_m g_m - fg\|_{L^1} \right].
\end{aligned} \tag{2.39}$$

Note that $\sup_i \|f_i\|_{L^p} < \infty$, because $f_i \to f$ in $L^p([-\pi,\pi]^d)$ and $\|f_i\|_{L^p} \to \|f\|_{L^p}$. Since $f_m,\, g_m \in L^\infty([-\pi,\pi]^d)$, by [20, Proposition 2] we have

$$\lim_{\boldsymbol{n}\to\infty} \frac{\|T_{\boldsymbol{n}}(f_m)T_{\boldsymbol{n}}(g_m) - T_{\boldsymbol{n}}(f_m g_m)\|_1}{N(\boldsymbol{n})} = 0.$$

Dividing (2.39) by $N(\boldsymbol{n})$ and passing to the limit as $\boldsymbol{n} \to \infty$, we obtain

$$\limsup_{\boldsymbol{n}\to\infty} \frac{\|T_{\boldsymbol{n}}(f)T_{\boldsymbol{n}}(g) - T_{\boldsymbol{n}}(fg)\|_1}{N(\boldsymbol{n})} \le \|f-f_m\|_{L^p}\|g\|_{L^q} + \sup_i \|f_i\|_{L^p}\|g-g_m\|_{L^q} + \|f_m g_m - fg\|_{L^1}. \tag{2.40}$$

This relation holds for every m. When $m \to \infty$, $f_m \to f$ in L^p and $g_m \to g$ in L^q by construction. Moreover, $f_m g_m \to fg$ in $L^1([-\pi,\pi]^d)$ by Hölder's inequality; indeed,

$$\begin{aligned}
\|fg - f_m g_m\|_{L^1} &\le \|(f-f_m)g\|_{L^1} + \|f_m(g-g_m)\|_{L^1} \le \|f-f_m\|_{L^p}\|g\|_{L^q} + \|f_m\|_{L^p}\|g-g_m\|_{L^q} \\
&\le \|f-f_m\|_{L^p}\|g\|_{L^q} + \sup_i \|f_i\|_{L^p}\|g-g_m\|_{L^q}.
\end{aligned}$$

Passing to the limit as $m \to \infty$ in (2.40), we get the thesis. □

3 Approximating classes of sequences (a.c.s.)

In this section, we introduce the fundamental definition on which the theory of GLT sequences is based: the notion of approximating class of sequences, first introduced in [42]. This notion lays the foundations for a (spectral) approximation theory for matrix-sequences and provides general tools (Theorems 3.1 and 3.3) for computing the asymptotic spectral or singular value distribution of a 'difficult' matrix-sequence $\{A_{\boldsymbol{n}}\}_n$ from the one of 'simpler' matrix-sequences $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ that approximate $\{A_{\boldsymbol{n}}\}_n$ in a suitable sense when $m \to \infty$. We refer the reader to the introduction of [29] for a more detailed discussion on this subject.

Definition 3.1 (approximating class of sequences). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. An approximating class of sequences (a.c.s.) for $\{A_{\boldsymbol{n}}\}_n$ is a sequence of matrix-sequences $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ with the following property: for every m there exists n_m such that, for $n \ge n_m$,

$$\begin{gathered}
A_{\boldsymbol{n}} = B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m}, \\
\operatorname{rank}(R_{\boldsymbol{n},m}) \le c(m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \le \omega(m),
\end{gathered} \tag{3.1}$$

where the quantities n_m, $c(m)$, $\omega(m)$ depend only on m, and

$$\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0.$$

Roughly speaking, $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ if $A_{\boldsymbol{n}}$ is equal to $B_{\boldsymbol{n},m}$ plus a small-rank matrix (with respect to the matrix size $N(\boldsymbol{n})$), plus a small-norm matrix.

Remark 3.1. An equivalent definition of a.c.s. is obtained by replacing, in Definition 3.1, 'for every m' with 'for every sufficiently large m' (i.e., 'for every m greater than or equal to some number M'). Indeed, suppose that the splitting (3.1) holds for $m \ge M$. For $m < M$, define $n_m = 1,\ c(m) = 1,\ \omega(m) = 0$ and $R_{\boldsymbol{n},m} = A_{\boldsymbol{n},m} - B_{\boldsymbol{n},m},\ N_{\boldsymbol{n},m} = O_{N(\boldsymbol{n})}$. Then, we see that (3.1) actually holds for every m.

3.1 The a.c.s. machinery as a tool for computing singular value and eigenvalue distributions

The importance of the a.c.s. notion resides in Theorems 3.1 and 3.3. Theorem 3.1 provides a general tool for determining the singular value distribution of a 'difficult' matrix-sequence $\{A_{\boldsymbol{n}}\}_n$, starting from the knowledge of the singular value distribution of 'simpler' matrix-sequences $\{B_{\boldsymbol{n},m}\}_n$. For the proof of this theorem, we need the following fundamental result about a.c.s.

Lemma 3.1. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence and let $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ be an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Then, for every $F \in C_c(\mathbb{R})$,*

$$\lim_{m\to\infty} \limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) \right| = 0. \tag{3.2}$$

Proof. We first observe that it suffices to prove (3.2) for all real-valued functions $F \in C_c^1(\mathbb{R})$. Indeed, suppose that (3.2) holds for this kind of functions and fix any $F \in C_c(\mathbb{R})$. For every $\epsilon > 0$, choose two real-valued functions $\Re_\epsilon, \Im_\epsilon \in C_c^1(\mathbb{R})$ such that $\|\Re(F) - \Re_\epsilon\|_\infty \le \epsilon$ and $\|\Im(F) - \Im_\epsilon\|_\infty \le \epsilon$, and set $F_\epsilon = \Re_\epsilon + \mathrm{i}\,\Im_\epsilon$. Then, we have $\|F - F_\epsilon\|_\infty \le 2\epsilon$ and, for all m, n,

$$\left| \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) \right| - \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(B_{\boldsymbol{n},m})) \right| \right| \le 4\epsilon.$$

It follows that, for every m,

$$\limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) \right| \le \limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(B_{\boldsymbol{n},m})) \right| + 4\epsilon.$$

Passing to the limit as $m \to \infty$, and taking into account that (3.2) holds for $\Re_\epsilon$, $\Im_\epsilon$ (and hence also for F_ϵ), we obtain

$$\lim_{m\to\infty} \limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) \right| \le 4\epsilon,$$

which is true for every $\epsilon > 0$. Thus, (3.2) holds for F.

Now, fix a real-valued function $F \in C_c^1(\mathbb{R})$. By hypothesis, $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Hence, for every m there exists n_m such that, for $n \ge n_m$,

$$\begin{gathered} A_{\boldsymbol{n}} = B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m}, \\ \mathrm{rank}(R_{\boldsymbol{n},m}) \le c(m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \le \omega(m), \end{gathered} \tag{3.3}$$

where $\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0$. For every m and every $n \ge n_m$,

$$\begin{aligned} &\left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) \right| \\ &\le \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})) \right| \\ &\quad + \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) \right|. \end{aligned} \tag{3.4}$$

We will consider separately the two terms in the right-hand side of (3.4), and we will show that each of them is bounded from above by a quantity depending only on m and tending to 0 as $m \to \infty$. After this, the thesis is proved.

In order to estimate the first term in the right-hand side of (3.4), we use the perturbation theorem for singular values (Theorem 2.3). We have

$$\begin{aligned} &\left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})) \right| \le \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} |F(\sigma_j(A_{\boldsymbol{n}})) - F(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m}))| \\ &\le \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} \|F'\|_\infty |\sigma_j(A_{\boldsymbol{n}}) - \sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})| \le \|F'\|_\infty \|A_{\boldsymbol{n}} - B_{\boldsymbol{n},m} - R_{\boldsymbol{n},m}\| = \|F'\|_\infty \|N_{\boldsymbol{n},m}\| \le \|F'\|_\infty \omega(m), \end{aligned}$$

which tends to 0 as $m \to \infty$.

In order to estimate the second term in the right-hand side of (3.4), we use the interlacing theorem for singular values (Theorem 2.1). We first observe that F can be expressed as the difference between two non-negative, non-decreasing, bounded functions:

$$F = H - K, \qquad H(x) = \int_{-\infty}^{x} (F')_+(t)dt, \qquad K(x) = \int_{-\infty}^{x} (F')_-(t)dt,$$

where $(F')_+ = \max(F', 0)$ and $(F')_- = \max(-F', 0)$. Hence, for the second term in the right-hand side of (3.4) we have

$$\begin{aligned}
&\left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) \right| \\
&\le \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} H(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} H(\sigma_j(B_{\boldsymbol{n},m})) \right| \\
&\quad + \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} K(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} K(\sigma_j(B_{\boldsymbol{n},m})) \right|. \qquad (3.5)
\end{aligned}$$

Let $r_{\boldsymbol{n},m} = \text{rank}(R_{\boldsymbol{n},m}) \le c(m)N(\boldsymbol{n})$. By Theorem 2.1,

$$\sigma_{j-r_{\boldsymbol{n},m}}(B_{\boldsymbol{n},m}) \ge \sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m}) \ge \sigma_{j+r_{\boldsymbol{n},m}}(B_{\boldsymbol{n},m}), \qquad j = 1, \ldots, N(\boldsymbol{n}).$$

Moreover, it is clear from our notation that

$$\sigma_{j-r_{\boldsymbol{n},m}}(B_{\boldsymbol{n},m}) \ge \sigma_j(B_{\boldsymbol{n},m}) \ge \sigma_{j+r_{\boldsymbol{n},m}}(B_{\boldsymbol{n},m}), \qquad j = 1, \ldots, N(\boldsymbol{n}).$$

Recalling the monotonicity and non-negativity of H, we get

$$\begin{aligned}
&\left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} H(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} H(\sigma_j(B_{\boldsymbol{n},m})) \right| \le \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} |H(\sigma_j(B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m})) - H(\sigma_j(B_{\boldsymbol{n},m}))| \\
&\le \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} \left| H(\sigma_{j-r_{\boldsymbol{n},m}}(B_{\boldsymbol{n},m})) - H(\sigma_{j+r_{\boldsymbol{n},m}}(B_{\boldsymbol{n},m})) \right| = \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} H(\sigma_{j-r_{\boldsymbol{n},m}}(B_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} H(\sigma_{j+r_{\boldsymbol{n},m}}(B_{\boldsymbol{n},m})) \\
&= \frac{1}{N(\boldsymbol{n})} \sum_{j=1-r_{\boldsymbol{n},m}}^{N(\boldsymbol{n})-r_{\boldsymbol{n},m}} H(\sigma_j(B_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1+r_{\boldsymbol{n},m}}^{N(\boldsymbol{n})+r_{\boldsymbol{n},m}} H(\sigma_j(B_{\boldsymbol{n},m})) \\
&= \frac{1}{N(\boldsymbol{n})} \sum_{j=1-r_{\boldsymbol{n},m}}^{\min(r_{\boldsymbol{n},m},\, N(\boldsymbol{n})-r_{\boldsymbol{n},m})} H(\sigma_j(B_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1+\max(r_{\boldsymbol{n},m},\, N(\boldsymbol{n})-r_{\boldsymbol{n},m})}^{N(\boldsymbol{n})+r_{\boldsymbol{n},m}} H(\sigma_j(B_{\boldsymbol{n},m})) \\
&\le \frac{1}{N(\boldsymbol{n})} \sum_{j=1-r_{\boldsymbol{n},m}}^{r_{\boldsymbol{n},m}} H(\sigma_j(B_{\boldsymbol{n},m})) \le \frac{2r_{\boldsymbol{n},m} H(\infty)}{N(\boldsymbol{n})} \le 2c(m)\|H\|_\infty.
\end{aligned}$$

Similarly, one can show that the second term in the right-hand side of (3.5) is bounded from above by $2c(m)\|K\|_\infty$, implying that the quantity in the left-hand side of (3.5), namely the second term in the right-hand side of (3.4), is less than or equal to $2(\|H\|_\infty + \|K\|_\infty)c(m)$. Since the latter tends to 0 as $m \to \infty$, the thesis is proved. □

Theorem 3.1. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence and let ϕ be a functional on $C_c(\mathbb{R})$. Assume that:*

1. *$\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$;*
2. *$\{B_{\boldsymbol{n},m}\}_n \sim_\sigma \phi_m$ for every m;*
3. *$\phi_m \to \phi$ pointwise over $C_c(\mathbb{R})$.*

Then $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \phi$.

Proof. Let $F \in C_c(\mathbb{R})$. For all n, m we have

$$\left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \phi(F)\right| \le \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m}))\right| + \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) - \phi_m(F)\right| + |\phi_m(F) - \phi(F)|. \tag{3.6}$$

By hypothesis, the second term in the right-hand side tends to 0 for $n \to \infty$, while the third one tends to 0 for $m \to \infty$. Therefore, passing first to the $\limsup_{n\to\infty}$ and then to the $\lim_{m\to\infty}$ in (3.6), and using Lemma 3.1, we get the thesis. □

Theorem 3.1 admits the following interesting converse, which we report for future use.

Theorem 3.2. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. Assume that:*

1. $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \phi$;
2. $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$;
3. $\{B_{\boldsymbol{n},m}\}_n \sim_\sigma \phi_m$ *for every* m;

Then $\phi_m \to \phi$ pointwise over $C_c(\mathbb{R})$.

Proof. Let $F \in C_c(\mathbb{R})$. For all n, m we have

$$|\phi_m(F) - \phi(F)| \le \left|\phi_m(F) - \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m}))\right| + \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) - \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}}))\right| + \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \phi(F)\right|. \tag{3.7}$$

By hypothesis, the first term in the right-hand side tends to 0 for $m \to \infty$, while the third one tends to 0 for $n \to \infty$. Therefore, passing first to the $\limsup_{n\to\infty}$ and then to the $\lim_{m\to\infty}$ in (3.7), and using Lemma 3.1, we get the thesis. □

Remark 3.2. Let $\{A_{\boldsymbol{n}}\}_n$, $\{B_{\boldsymbol{n},m}\}_n$ be matrix-sequences and let ϕ, $\phi_m : C_c(\mathbb{R}) \to \mathbb{C}$ be functionals. Consider the following four conditions:

(1) $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \phi$;

(2) $\{B_{\boldsymbol{n},m}\}_n \sim_\sigma \phi_m$ for every m;

(3) $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$;

(4) $\phi_m \to \phi$ pointwise over $C_c(\mathbb{R})$.

Theorems 3.1–3.2 show that '(1) ∧ (2) ∧ (3) ⇒ (4)' and '(2) ∧ (3) ∧ (4) ⇒ (1)'.

The implication '(1) ∧ (2) ∧ (4) ⇒ (3)' is false in general. As a counterexample, take $A_n = I_n$ and $B_{n,m} = \mathrm{diag}_{i=1,\ldots,n}(-1)^i$. Then, $\{A_n\}_n \sim_\sigma 1$ and $\{B_{n,m}\}_n \sim_\sigma 1$, so that $\phi = \phi_m = \phi_{[1]}$ is the evaluation functional at 1, $\phi_{[1]}(F) = F(1)$. Moreover, it can be directly verified that $\{A_n - B_{n,m}\}_n \sim_\sigma \varphi(F) = \frac{1}{2}\big[F(0) + F(2)\big]$. Therefore, $\{\{B_{n,m}\}_n\}_m$ cannot be an a.c.s. for $\{A_n\}_n$, because otherwise $\{\{A_n - B_{n,m}\}_n\}_m$ would be an a.c.s. of $\{O_n\}_n$ and so, considering that $\{O_n\}_n \sim_\sigma \phi_{[0]}$, by Theorem 3.2 we would have $\{A_n - B_{n,m}\}_n \sim_\sigma \phi_{[0]}$.

The implication '(1) ∧ (3) ∧ (4) ⇒ (2)', written in this way, is meaningless. However, a natural modification reads as follows: '(1) ∧ (3) ⇒ there exists a functional ϕ_m such that $\phi_m \to \phi$ pointwise over $C_c(\mathbb{R})$ and $\{B_{\boldsymbol{n},m}\}_n \sim_\sigma \phi_m$ for all sufficiently large m'. This statement is false in general. As a counterexample, take $A_n = O_n$ and $B_{n,m} = (1 + (-1)^n)\frac{1}{m}I_n$. Then, $\|B_{n,m}\| \le \frac{2}{m}$, so $\{\{B_{n,m}\}_n\}_m$ is an a.c.s. of $\{O_n\}_n$. Nevertheless, the limit

$$\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n} F(\sigma_j(B_{n,m})) = \lim_{n\to\infty} F((1 + (-1)^n)\tfrac{1}{m})$$

does not exist for any function $F \in C_c(\mathbb{R})$ such that $F(0) \ne F(\frac{2}{m})$. Therefore, the relation $\{B_{n,m}\}_n \sim_\sigma \phi_m$ cannot hold for any functional ϕ_m.

In Theorem 3.3 we prove the analogue of Theorem 3.1 for the case of the eigenvalues, but we need to add the assumption that $A_{\boldsymbol{n}}$ and $B_{\boldsymbol{n},m}$ are Hermitian. Theorem 3.3 is then a general tool for determining the spectral distribution of a 'difficult' matrix-sequence $\{A_{\boldsymbol{n}}\}_n$ *formed by Hermitian matrices,* starting from the spectral distribution of simpler matrix-sequences $\{B_{\boldsymbol{n},m}\}_n$, *again formed by Hermitian matrices.*

The next lemma shows that, whenever $A_{\boldsymbol{n}}$ and $B_{\boldsymbol{n},m}$ are Hermitian, the small-rank matrix $R_{\boldsymbol{n},m}$ and the small-norm matrix $N_{\boldsymbol{n},m}$ in the splitting (3.1) may be supposed to be Hermitian. In the following, we say that $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$ *formed by Hermitian matrices* if $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ and every $B_{\boldsymbol{n},m}$ is Hermitian.

Lemma 3.2. *Let $\{A_{\boldsymbol{n}}\}_n$ be a sequence of Hermitian matrices, and let $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ be an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ formed by Hermitian matrices. Then, for every m there exists n_m such that, for $n \geq n_m$,*

$$A_{\boldsymbol{n}} = B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$
$$\text{rank}(R_{\boldsymbol{n},m}) \leq c(m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \leq \omega(m),$$

where $R_{\boldsymbol{n},m}, N_{\boldsymbol{n},m}$ are Hermitian, the quantities n_m, $c(m)$, $\omega(m)$ depend only on m, and

$$\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0.$$

Proof. Take the real part in (3.1) and use the inequalities $\text{rank}(\Re(X)) \leq 2\,\text{rank}(X)$ and $\|\Re(X)\| \leq \|X\|$ to conclude that, by replacing $R_{\boldsymbol{n},m}$, $N_{\boldsymbol{n},m}$ with $\Re(R_{\boldsymbol{n},m})$, $\Re(N_{\boldsymbol{n},m})$ (if necessary), we can assume $R_{\boldsymbol{n},m}$, $N_{\boldsymbol{n},m}$ to be Hermitian. □

Lemma 3.3 is the 'eigenvalue version' of Lemma 3.1.

Lemma 3.3. *Let $\{A_{\boldsymbol{n}}\}_n$ be a sequence of Hermitian matrices and let $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ be an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ formed by Hermitian matrices. Then, for every $F \in C_c(\mathbb{C})$,*

$$\lim_{m\to\infty} \limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m})) \right| = 0. \tag{3.8}$$

Proof. The proof is essentially the same as the proof of Lemma 3.1; we just outline the main steps, emphasizing the analogies with Lemma 3.1 and pointing out where we need the assumption that $A_{\boldsymbol{n}}$ and $B_{\boldsymbol{n},m}$ are Hermitian.

Noting that all the eigenvalues $\lambda_j(A_{\boldsymbol{n}})$, $\lambda_j(B_{\boldsymbol{n},m})$, $j = 1, \ldots, N(\boldsymbol{n})$, are real, it suffices to prove (3.8) for all real-valued functions $F \in C_c^1(\mathbb{R})$. Indeed, suppose that (3.8) holds for this kind of functions and fix any $F \in C_c(\mathbb{C})$. For every $\epsilon > 0$, choose two real-valued functions $\Re_\epsilon$, $\Im_\epsilon \in C_c^1(\mathbb{R})$ such that $\|\Re(F) - \Re_\epsilon\|_{\infty,\mathbb{R}} \leq \epsilon$ and $\|\Im(F) - \Im_\epsilon\|_{\infty,\mathbb{R}} \leq \epsilon$, and set $F_\epsilon = \Re_\epsilon + \mathrm{i}\,\Im_\epsilon$. Then, we have $\|F - F_\epsilon\|_{\infty,\mathbb{R}} \leq 2\epsilon$ and, for all m, n,

$$\left| \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m})) \right| - \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\lambda_j(B_{\boldsymbol{n},m})) \right| \right| \leq 4\epsilon.$$

It follows that, for every m,

$$\limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m})) \right| \leq \limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\lambda_j(B_{\boldsymbol{n},m})) \right| + 4\epsilon.$$

Passing to the limit as $m \to \infty$, and taking into account that (3.8) holds for $\Re_\epsilon$, $\Im_\epsilon$ (and hence also for F_ϵ), we obtain

$$\lim_{m\to\infty} \limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m})) \right| \leq 4\epsilon,$$

which is true for every $\epsilon > 0$. Thus, (3.8) holds for F.

Now we fix a real-valued function $F \in C_c^1(\mathbb{R})$. Since $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ and $A_{\boldsymbol{n}}$, $B_{\boldsymbol{n},m}$ are Hermitian, we can use Lemma 3.2 to see that for every m there exists n_m such that, for $n \geq n_m$, the splitting (3.3) holds with Hermitian $R_{\boldsymbol{n},m}$, $N_{\boldsymbol{n},m}$. Following the proof of Lemma 3.1, we arrive at the inequality (3.4), with 'σ_j' replaced by 'λ_j', and the thesis is proved if we are able to bound the two terms in the right-hand side by a quantity depending only on m and tending to 0 as $m \to \infty$.

The first term is bounded exactly as in Lemma 3.1, using the perturbation theorem for eigenvalues (Theorem 2.4) instead of the perturbation theorem for singular values (Theorem 2.3). Note that the perturbation theorem for eigenvalues, contrary to the perturbation theorem for singular values, applies only to Hermitian matrices.

Also the second term is bounded exactly as in Lemma 3.1, using the interlacing theorem for eigenvalues (Theorem 2.2) instead of the interlacing theorem for singular values (Theorem 2.1). Even in this case, the interlacing theorem for eigenvalues, contrary to the interlacing theorem for singular values, applies only to Hermitian matrices. □

Theorem 3.3. *Let $\{A_{\boldsymbol{n}}\}_n$ be a sequence of Hermitian matrices and let ϕ be a functional on $C_c(\mathbb{C})$. Assume that:*

1. *$\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ formed by Hermitian matrices;*
2. *$\{B_{n,m}\}_n \sim_\lambda \phi_m$ for every m;*
3. *$\phi_m \to \phi$ pointwise over $C_c(\mathbb{C})$.*

Then $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \phi$.

Proof. Let $F \in C_c(\mathbb{C})$. For all n, m we have

$$\left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \phi(F)\right| \le \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m}))\right| + \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m})) - \phi_m(F)\right| + |\phi_m(F) - \phi(F)|. \tag{3.9}$$

By hypothesis, the second term in the right-hand side tends to 0 for $n \to \infty$, while the third one tends to 0 for $m \to \infty$. Therefore, passing first to the $\limsup_{n\to\infty}$ and then to the $\lim_{m\to\infty}$ in (3.9), and using Lemma 3.3, we get the thesis. □

Two important corollaries of Theorems 3.1 and 3.3 are given in the following. They will be used in Section 5.2 to prove the asymptotic spectral and singular value distribution results for GLT sequences.

Corollary 3.1. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. Assume that:*

1. *$\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$;*
2. *for every m, $\{B_{\boldsymbol{n},m}\}_n \sim_\sigma f_m$ for some measurable function $f_m : D \subset \mathbb{R}^k \to \mathbb{C}$;*
3. *$|f_m| \to |f|$ in measure over D, being $f : D \to \mathbb{C}$ another measurable function.*

Then $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$.

Proof. Apply Theorem 3.1 with $\phi_m = \phi_{[|f_m|]}$ and $\phi = \phi_{[|f|]}$. Note that $\phi_m \to \phi$ pointwise over $C_c(\mathbb{R})$ by Lemma 2.4. □

Corollary 3.2. *Let $\{A_{\boldsymbol{n}}\}_n$ be a sequence of Hermitian matrices. Assume that:*

1. *$\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ formed by Hermitian matrices;*
2. *for every m, $\{B_{\boldsymbol{n},m}\}_n \sim_\lambda f_m$ for some measurable function $f_m : D \subset \mathbb{R}^k \to \mathbb{C}$;*
3. *$f_m \to f$ in measure over D, being $f : D \to \mathbb{C}$ another measurable function.*

Then $\{A_{\boldsymbol{n}}\}_n \sim_\lambda f$.

Proof. Apply Theorem 3.3 with $\phi_m = \phi_{[f_m]}$ and $\phi = \phi_{[f]}$. Note that $\phi_m \to \phi$ pointwise over $C_c(\mathbb{C})$ by Lemma 2.4. □

3.2 The a.c.s. algebra

In this section, we investigate the algebraic properties possessed by the approximating classes of sequences. These properties form the basis of the so-called GLT algebra, which will be studied later on, in Section 5.5. We begin with the following observation, whose proof is very simple and is left to the reader.

Remark 3.3. Let $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ be an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Then $\{\{B^*_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A^*_{\boldsymbol{n}}\}_n$.

Proposition 3.1. *Let $\{A_{\boldsymbol{n}}\}_n$, $\{A'_{\boldsymbol{n}}\}_n$ be matrix-sequences, and let*

- *$\{\{B_{\boldsymbol{n},m}\}_n\}_m$ an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$,*
- *$\{\{B'_{\boldsymbol{n},m}\}_n\}_m$ an a.c.s. for $\{A'_{\boldsymbol{n}}\}_n$.*

Then $\{\{\alpha B_{\boldsymbol{n},m} + \beta B'_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{\alpha A_{\boldsymbol{n}} + \beta A'_{\boldsymbol{n}}\}_n$, for all $\alpha, \beta \in \mathbb{C}$.

Proof. By hypothesis,

- for every m there exists n_m such that, for $n \geq n_m$,

$$A_{\boldsymbol{n}} = B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$
$$\text{rank}(R_{\boldsymbol{n},m}) \leq c(m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \leq \omega(m),$$

  where $\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0$;

- for every m there exists n'_m such that, for $n \geq n'_m$,

$$A'_{\boldsymbol{n}} = B'_{\boldsymbol{n},m} + R'_{\boldsymbol{n},m} + N'_{\boldsymbol{n},m},$$
$$\text{rank}(R'_{\boldsymbol{n},m}) \leq c'(m)N(\boldsymbol{n}), \qquad \|N'_{\boldsymbol{n},m}\| \leq \omega'(m),$$

  where $\lim_{m\to\infty} c'(m) = \lim_{m\to\infty} \omega'(m) = 0$.

Hence, for every m and every $n \geq \max(n_m, n'_m)$,

$$\alpha A_{\boldsymbol{n}} + \beta A'_{\boldsymbol{n}} = (\alpha B_{\boldsymbol{n},m} + \beta B'_{\boldsymbol{n},m}) + (\alpha R_{\boldsymbol{n},m} + \beta R'_{\boldsymbol{n},m}) + (\alpha N_{\boldsymbol{n},m} + \beta N'_{\boldsymbol{n},m}),$$
$$\text{rank}(\alpha R_{\boldsymbol{n},m} + \beta R'_{\boldsymbol{n},m}) \leq \big(c(m) + c'(m)\big)N(\boldsymbol{n}), \qquad \|\alpha N_{\boldsymbol{n},m} + \beta N'_{\boldsymbol{n},m}\| \leq |\alpha|\omega(m) + |\beta|\omega'(m),$$

where

$$\lim_{m\to\infty} \big(c(m) + c'(m)\big) = \lim_{m\to\infty} \big(|\alpha|\omega(m) + |\beta|\omega'(m)\big) = 0.$$

Hence, by Definition 3.1, $\{\{\alpha B_{\boldsymbol{n},m} + \beta B'_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{\alpha A_{\boldsymbol{n}} + \beta A'_{\boldsymbol{n}}\}_n$. □

Proposition 3.1 addresses the case of a linear combination $\{\alpha A_{\boldsymbol{n}} + \beta A'_{\boldsymbol{n}}\}_n$ of two matrix-sequences $\{A_{\boldsymbol{n}}\}_n$ and $\{A'_{\boldsymbol{n}}\}_n$. We would like to prove an analogous result for the product $\{A_{\boldsymbol{n}}A'_{\boldsymbol{n}}\}_n$. To this end, an additional (mild) assumption on $\{A_{\boldsymbol{n}}\}_n$ and $\{A'_{\boldsymbol{n}}\}_n$ is required, namely that $\{A_{\boldsymbol{n}}\}_n$ and $\{A'_{\boldsymbol{n}}\}_n$ are sparsely unbounded.

Definition 3.2 (sparsely unbounded matrix-sequence). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. We say that $\{A_{\boldsymbol{n}}\}_n$ is sparsely unbounded (s.u.) if for every $M > 0$ there exists n_M such that, for $n \geq n_M$,

$$\frac{\#\{i \in \{1, \ldots, N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \leq r(M),$$

where $\lim_{M\to\infty} r(M) = 0$.

The following proposition provides equivalent characterizations of sparsely unbounded matrix-sequences.

Proposition 3.2. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. The following conditions are equivalent.*

1. $\{A_{\boldsymbol{n}}\}_n$ *is s.u.*
2. $\displaystyle\lim_{M\to\infty} \limsup_{n\to\infty} \frac{\#\{i \in \{1, \ldots, N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} = 0.$
3. *For every $M > 0$ there exists n_M such that, for $n \geq n_M$,*

$$A_{\boldsymbol{n}} = \hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M},$$
$$\text{rank}(\hat{A}_{\boldsymbol{n},M}) \leq r(M)N(\boldsymbol{n}), \qquad \|\tilde{A}_{\boldsymbol{n},M}\| \leq M,$$

   where $\lim_{M\to\infty} r(M) = 0$.

Note that condition 2 can be rewritten as

$$\lim_{M\to\infty} \limsup_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{i=1}^{N(\boldsymbol{n})} \chi_{(M,\infty)}(\sigma_i(A_{\boldsymbol{n}})) = 0.$$

Proof. $(1 \Rightarrow 2)$ Suppose that $\{A_{\boldsymbol{n}}\}_n$ is s.u. Then, for every $M > 0$ there exists n_M such that, for $n \geq n_M$,

$$\frac{\#\{i \in \{1, \ldots, N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \leq r(M),$$

where $\lim_{M\to\infty} r(M) = 0$. Therefore,

$$\limsup_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \le r(M)$$

and, consequently,

$$\lim_{M\to\infty} \limsup_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} = 0.$$

(2 ⇒ 1) Suppose that condition 2 is met. For every $M > 0$, define

$$\delta(M) = \limsup_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \in [0,1]$$

and note that (obviously)

$$\limsup_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} < \delta(M) + \frac{1}{M}.$$

Hence, by definition of $\limsup$, for every $M > 0$ the sequence $\frac{\#\{i\in\{1,\ldots,N(\boldsymbol{n})\}:\sigma_i(A_{\boldsymbol{n}})>M\}}{N(\boldsymbol{n})}$ is eventually less than $r(M) = \delta(M) + \frac{1}{M}$, i.e., there exists n_M such that, for $n \ge n_M$,

$$\frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \le r(M).$$

Since $r(M) \to 0$ as $M \to \infty$, item 1 is proved.

(1 ⇒ 3) Suppose that $\{A_{\boldsymbol{n}}\}_n$ is s.u.: for every $M > 0$ there exists n_M such that, for $n \ge n_M$,

$$\frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \le r(M),$$

where $\lim_{M\to\infty} r(M) = 0$. Let $A_{\boldsymbol{n}} = U_{\boldsymbol{n}}\Sigma_{\boldsymbol{n}}V_{\boldsymbol{n}}^*$ be an SVD of $A_{\boldsymbol{n}}$. Let $\hat{\Sigma}_{\boldsymbol{n},M}$ be the matrix obtained from $\Sigma_{\boldsymbol{n}}$ by setting to 0 all the singular values of $A_{\boldsymbol{n}}$ that are less than or equal to M, and let $\tilde{\Sigma}_{\boldsymbol{n},M} = \Sigma_{\boldsymbol{n}} - \hat{\Sigma}_{\boldsymbol{n},M}$ be the matrix obtained from $\Sigma_{\boldsymbol{n}}$ by setting to 0 all the singular values of $A_{\boldsymbol{n}}$ that exceed M. Then,

$$A_{\boldsymbol{n}} = U_{\boldsymbol{n}}\Sigma_{\boldsymbol{n}}V_{\boldsymbol{n}}^* = U_{\boldsymbol{n}}\hat{\Sigma}_{\boldsymbol{n},M}V_{\boldsymbol{n}}^* + U_{\boldsymbol{n}}\tilde{\Sigma}_{\boldsymbol{n},M}V_{\boldsymbol{n}}^* = \hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M},$$

where $\hat{A}_{\boldsymbol{n},M} = U_{\boldsymbol{n}}\hat{\Sigma}_{\boldsymbol{n},M}V_{\boldsymbol{n}}^*$ and $\tilde{A}_{\boldsymbol{n},M} = U_{\boldsymbol{n}}\tilde{\Sigma}_{\boldsymbol{n},M}V_{\boldsymbol{n}}^*$ satisfy, for $n \ge n_M$,

$$\text{rank}(\hat{A}_{\boldsymbol{n},M}) = \#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\} \le r(M)N(\boldsymbol{n}), \qquad \|\tilde{A}_{\boldsymbol{n},M}\| = \sigma_{\max}(\tilde{A}_{\boldsymbol{n},M}) \le M.$$

(3 ⇒ 1) Suppose that condition 3 holds. Then, for every $M > 0$ there exists n_M such that, for $n \ge n_M$,

$$A_{\boldsymbol{n}} = \hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M},$$
$$\text{rank}(\hat{A}_{\boldsymbol{n},M}) \le r(M)N(\boldsymbol{n}), \qquad \|\tilde{A}_{\boldsymbol{n},M}\| \le M,$$

where $\lim_{M\to\infty} r(M) = 0$. By the minimax principle for singular values [8, Problem III.6.1], for all $i = 1,\ldots,N(\boldsymbol{n})$ we have

$$\begin{aligned}\sigma_i(A_{\boldsymbol{n}}) &= \max_{\dim V = i}\ \min_{\mathbf{x}\in V,\ \|\mathbf{x}\|=1} \|A_{\boldsymbol{n}}\mathbf{x}\| \le \max_{\dim V = i}\ \min_{\mathbf{x}\in V,\ \|\mathbf{x}\|=1} \Big(\|\hat{A}_{\boldsymbol{n},M}\mathbf{x}\| + \|\tilde{A}_{\boldsymbol{n},M}\mathbf{x}\|\Big) \\ &\le \max_{\dim V = i}\ \min_{\mathbf{x}\in V,\ \|\mathbf{x}\|=1} \Big(\|\hat{A}_{\boldsymbol{n},M}\mathbf{x}\| + \|\tilde{A}_{\boldsymbol{n},M}\|\Big) = \sigma_i(\hat{A}_{\boldsymbol{n},M}) + \|\tilde{A}_{\boldsymbol{n},M}\| \le \sigma_i(\hat{A}_{\boldsymbol{n},M}) + M. \end{aligned} \tag{3.10}$$

Since $\text{rank}(\hat{A}_{\boldsymbol{n},M}) \le r(M)N(\boldsymbol{n})$, $\hat{A}_{\boldsymbol{n},M}$ has at most $r(M)N(\boldsymbol{n})$ nonzero singular values. Therefore, by (3.10), $A_{\boldsymbol{n}}$ has at most $r(M)N(\boldsymbol{n})$ singular values greater than M, i.e., $\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\} \le r(M)N(\boldsymbol{n})$, or, equivalently,

$$\frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \le r(M).$$

This implies that $\{A_{\boldsymbol{n}}\}_n$ is s.u. □

We now show that any matrix-sequence enjoying an asymptotic singular value distribution is s.u. A proof of this useful result, analogous to the one that we are going to see, appeared in [49, Proposition 2.7].

Proposition 3.3. *If $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$ then $\{A_{\boldsymbol{n}}\}_n$ is s.u.*

Proof. Let $D \subset \mathbb{R}^k$ be the domain of the function f. If we could choose $F = \chi_{(M,\infty)}$ as a test function in (2.28), then we would obtain

$$\begin{aligned}\lim_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\}:\ \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} &= \lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{i=1}^{N(\boldsymbol{n})} \chi_{(M,\infty)}(\sigma_i(A_{\boldsymbol{n}})) = \frac{1}{\mu_k(D)} \int_D \chi_{(M,\infty)}(|f(\mathbf{x})|)d\mathbf{x} \\ &= \frac{\mu_k\{|f| > M\}}{\mu_k(D)}. \end{aligned} \tag{3.11}$$

Since $\mu_k\{|f| > M\} \to 0$ as $M \to \infty$ (by the dominated convergence theorem), eq. (3.11) would imply that condition 2 in Proposition 3.2 is met and the proof would be over. However, $\chi_{(M,\infty)}$ cannot be chosen as a test function in (2.28), because it does not belong to $C_c(\mathbb{R})$. Hence, to obtain the thesis we need some work.

Fix $M > 0$ and take $F_M \in C_c(\mathbb{R})$ such that $F_M = 1$ over $[0, M/2]$, $F_M = 0$ over $[M,\infty)$ and $0 \le F_M \le 1$ over $\mathbb{R}$. Note that $F_M \le \chi_{[0,M]}$ over $[0,\infty)$. Then,

$$\begin{aligned}\frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} &= 1 - \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) \le M\}}{N(\boldsymbol{n})} = 1 - \frac{1}{N(\boldsymbol{n})} \sum_{i=1}^{N(\boldsymbol{n})} \chi_{[0,M]}(\sigma_i(A_{\boldsymbol{n}})) \\ &\le 1 - \frac{1}{N(\boldsymbol{n})} \sum_{i=1}^{N(\boldsymbol{n})} F_M(\sigma_i(A_{\boldsymbol{n}})) \overset{n\to\infty}{\longrightarrow} 1 - \frac{1}{\mu_k(D)} \int_D F_M(|f(\mathbf{x})|)d\mathbf{x}\end{aligned}$$

and

$$\limsup_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \le 1 - \frac{1}{\mu_k(D)} \int_D F_M(|f(\mathbf{x})|)d\mathbf{x}.$$

Since $F_M(|f(\mathbf{x})|) \to 1$ a.e. and $|F_M(|f(\mathbf{x})|)| \le 1$, by the dominated convergence theorem we get

$$\lim_{M\to\infty} \int_D F_M(|f(\mathbf{x})|)d\mathbf{x} = \mu_k(D),$$

and so

$$\lim_{M\to\infty} \limsup_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} = 0.$$

This means that condition 2 in Proposition 3.2 is satisfied, i.e., $\{A_{\boldsymbol{n}}\}_n$ is s.u. □

We now prove the analogue of Proposition 3.1 for the case of the product of two a.c.s. This important result appeared for the first time in [42].

Proposition 3.4. *Let $\{A_{\boldsymbol{n}}\}_n$, $\{A'_{\boldsymbol{n}}\}_n$ be s.u. matrix-sequences, and let*

- *$\{\{B_{\boldsymbol{n},m}\}_n\}_m$ an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$,*
- *$\{\{B'_{\boldsymbol{n},m}\}_n\}_m$ an a.c.s. for $\{A'_{\boldsymbol{n}}\}_n$.*

Then, $\{\{B_{\boldsymbol{n},m}B'_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}A'_{\boldsymbol{n}}\}_n$.

Proof. By hypothesis, for every m there exists n_m such that, for $n \ge n_m$,

$$A_{\boldsymbol{n}} = B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m}, \qquad A'_{\boldsymbol{n}} = B'_{\boldsymbol{n},m} + R'_{\boldsymbol{n},m} + N'_{\boldsymbol{n},m},$$

$$\text{rank}(R_{\boldsymbol{n},m}),\ \text{rank}(R'_{\boldsymbol{n},m}) \le c(m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\|,\ \|N'_{\boldsymbol{n},m}\| \le \omega(m),$$

where $\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0$. Hence,

$$A_{\boldsymbol{n}}A'_{\boldsymbol{n}} = B_{\boldsymbol{n},m}B'_{\boldsymbol{n},m} + B_{\boldsymbol{n},m}R'_{\boldsymbol{n},m} + \boxed{B_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}} + R_{\boldsymbol{n},m}A'_{\boldsymbol{n}} + \boxed{N_{\boldsymbol{n},m}A'_{\boldsymbol{n}}}.$$

Since $\{A_{\boldsymbol{n}}\}_n$ and $\{A'_{\boldsymbol{n}}\}_n$ are s.u., for every $M > 0$ there exists $n(M)$ such that, for $n \ge n(M)$,

$$A_{\boldsymbol{n}} = \hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M}, \qquad A'_{\boldsymbol{n}} = \hat{A}'_{\boldsymbol{n},M} + \tilde{A}'_{\boldsymbol{n},M},$$

$$\text{rank}(\hat{A}_{\boldsymbol{n},M}),\ \text{rank}(\hat{A}'_{\boldsymbol{n},M}) \le r(M)N(\boldsymbol{n}), \qquad \|\tilde{A}_{\boldsymbol{n},M}\|,\ \|\tilde{A}'_{\boldsymbol{n},M}\| \le M,$$

where $\lim_{M\to\infty} r(M)=0$. Setting $M_m=[\omega(m)]^{-1/2}$, for every m and every $n\ge\max(n_m,n(M_m))$ we have

$$\begin{aligned}
B_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}+N_{\boldsymbol{n},m}A'_{\boldsymbol{n}} &= (A_{\boldsymbol{n}}-R_{\boldsymbol{n},m}-N_{\boldsymbol{n},m})N'_{\boldsymbol{n},m}+N_{\boldsymbol{n},m}(\hat{A}'_{\boldsymbol{n},M_m}+\tilde{A}'_{\boldsymbol{n},M_m})\\
&= (\hat{A}_{\boldsymbol{n},M_m}+\tilde{A}_{\boldsymbol{n},M_m}-R_{\boldsymbol{n},m}-N_{\boldsymbol{n},m})N'_{\boldsymbol{n},m}+N_{\boldsymbol{n},m}\hat{A}'_{\boldsymbol{n},M_m}+N_{\boldsymbol{n},m}\tilde{A}'_{\boldsymbol{n},M_m}\\
&= \hat{A}_{\boldsymbol{n},M_m}N'_{\boldsymbol{n},m}+\tilde{A}_{\boldsymbol{n},M_m}N'_{\boldsymbol{n},m}-R_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}-N_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}+N_{\boldsymbol{n},m}\hat{A}'_{\boldsymbol{n},M_m}+N_{\boldsymbol{n},m}\tilde{A}'_{\boldsymbol{n},M_m},
\end{aligned}$$

and so

$$\begin{aligned}
A_{\boldsymbol{n}}A'_{\boldsymbol{n}} &= B_{\boldsymbol{n},m}B'_{\boldsymbol{n},m}+B_{\boldsymbol{n},m}R'_{\boldsymbol{n},m}+\boxed{B_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}}+R_{\boldsymbol{n},m}A'_{\boldsymbol{n}}+\boxed{N_{\boldsymbol{n},m}A'_{\boldsymbol{n}}}\\
&= B_{\boldsymbol{n},m}B'_{\boldsymbol{n},m}+B_{\boldsymbol{n},m}R'_{\boldsymbol{n},m}+R_{\boldsymbol{n},m}A'_{\boldsymbol{n}}+\hat{A}_{\boldsymbol{n},M_m}N'_{\boldsymbol{n},m}+\tilde{A}_{\boldsymbol{n},M_m}N'_{\boldsymbol{n},m}-R_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}-N_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}\\
&\quad +N_{\boldsymbol{n},m}\hat{A}'_{\boldsymbol{n},M_m}+N_{\boldsymbol{n},m}\tilde{A}'_{\boldsymbol{n},M_m},
\end{aligned}$$

where

$$\mathrm{rank}(B_{\boldsymbol{n},m}R'_{\boldsymbol{n},m}+R_{\boldsymbol{n},m}A'_{\boldsymbol{n}}+\hat{A}_{\boldsymbol{n},M_m}N'_{\boldsymbol{n},m}-R_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}+N_{\boldsymbol{n},m}\hat{A}'_{\boldsymbol{n},M_m})\le[3c(m)+2r(M_m)]N(\boldsymbol{n}),$$

$$\|\tilde{A}_{\boldsymbol{n},M_m}N'_{\boldsymbol{n},m}-N_{\boldsymbol{n},m}N'_{\boldsymbol{n},m}+N_{\boldsymbol{n},m}\tilde{A}'_{\boldsymbol{n},M_m}\|\le 2[\omega(m)]^{1/2}+[\omega(m)]^2.$$

Thus, $\{\{B_{\boldsymbol{n},m}B'_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}A'_{\boldsymbol{n}}\}_n$. □

3.3 Some criterions to identify a.c.s.

In practical applications, it happens that a matrix-sequence $\{A_{\boldsymbol{n}}\}_n$ is given together with a sequence of matrix-sequences $\{\{B_{\boldsymbol{n},m}\}_n\}_m$, and one would like to show that $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$, without constructing the splitting (3.1). In this section, we provide two useful criterions to solve this problem.

The first criterion, expressed in Theorem 3.4 and Corollary 3.3, is formulated in terms of Schatten p-norms. The second criterion, expressed in Theorem 3.5 and Corollary 3.4, is formulated in terms of the singular value distribution.

Lemma 3.4. *Let C be a square matrix of size s. Suppose that*

$$\|C\|_p^p\le\epsilon s$$

for some $p\in[1,\infty)$. Then

$$C=R+N,$$

with

$$\mathrm{rank}(R)\le\epsilon^{\frac{1}{p+1}}s,\qquad \|N\|\le\epsilon^{\frac{1}{p+1}}.$$

Proof. Since $\|C\|_p^p=\sum_{i=1}^s[\sigma_i(C)]^p\le\epsilon s$, the number of singular values of C that exceed $\epsilon^{\frac{1}{p+1}}$ cannot be larger than $\epsilon^{\frac{1}{p+1}}s$. Let $C=U\Sigma V^*$ be an SVD of C and write

$$C=U\Sigma V^*=U\hat{\Sigma}V^*+U\tilde{\Sigma}V^*,$$

where $\hat{\Sigma}$ is obtained from Σ by setting to 0 all the singular values that are less than or equal to $\epsilon^{\frac{1}{p+1}}$, while $\tilde{\Sigma}=\Sigma-\hat{\Sigma}$ is obtained from Σ by setting to 0 all the singular values that exceed $\epsilon^{\frac{1}{p+1}}$. Then

$$C=R+N,$$

where $R=U\hat{\Sigma}V^*$ and $N=U\tilde{\Sigma}V^*$ satisfy $\mathrm{rank}(R)\le\epsilon^{\frac{1}{p+1}}s$ and $\|N\|\le\epsilon^{\frac{1}{p+1}}$. □

By Definition 3.1, a sequence of matrix-sequences $\{\{C_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$ if and only if the following condition is met: for every m there exists n_m such that, for $n\ge n_m$,

$$\begin{gathered}
C_{\boldsymbol{n},m}=R_{\boldsymbol{n},m}+N_{\boldsymbol{n},m},\\
\mathrm{rank}(R_{\boldsymbol{n},m})\le c(m)N(\boldsymbol{n}),\qquad \|N_{\boldsymbol{n},m}\|\le\omega(m),
\end{gathered}\tag{3.12}$$

where $\lim_{m\to\infty}c(m)=\lim_{m\to\infty}\omega(m)=0$. Moreover, it is clear that $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ if and only if $\{\{A_{\boldsymbol{n}}-B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$.

Theorem 3.4. *Let $\{\{C_{\boldsymbol{n},m}\}_n\}_m$ be a sequence of matrix-sequences and let $1\le p<\infty$. Suppose that for every m there exists n_m such that, for $n\ge n_m$,*

$$\|C_{\boldsymbol{n},m}\|_p^p\le\epsilon(m,\boldsymbol{n})N(\boldsymbol{n}),$$

where $\lim_{m\to\infty}\limsup_{n\to\infty}\epsilon(m,\boldsymbol{n})=0$. Then $\{\{C_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$.

Proof. By Lemma 3.4, for every m and every $n \geq n_m$ we have

$$C_{\boldsymbol{n},m} = R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$

$$\mathrm{rank}(R_{\boldsymbol{n},m}) \leq [\epsilon(m,\boldsymbol{n})]^{\frac{1}{p+1}} N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \leq [\epsilon(m,\boldsymbol{n})]^{\frac{1}{p+1}}.$$

Let

$$\epsilon(m) = \limsup_{n\to\infty} \epsilon(m,\boldsymbol{n}).$$

By definition of lim sup, for every m there exists n'_m such that, for $n \geq n'_m$,

$$\epsilon(m,\boldsymbol{n}) \leq \epsilon(m) + \frac{1}{m}.$$

Setting $\hat{n}_m = \max(n_m, n'_m)$, for every m and every $n \geq \hat{n}_m$ we have

$$C_{\boldsymbol{n},m} = R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$

$$\mathrm{rank}(R_{\boldsymbol{n},m}) \leq \Big(\epsilon(m) + \frac{1}{m}\Big)^{\frac{1}{p+1}} N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \leq \Big(\epsilon(m) + \frac{1}{m}\Big)^{\frac{1}{p+1}}.$$

Since $\epsilon(m) \to 0$ by assumption, $\{\{C_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$. □

Corollary 3.3. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, let $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ be a sequence of matrix-sequences, and let $1 \leq p < \infty$. Suppose that for every m there exists n_m such that, for $n \geq n_m$,*

$$\|A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\|_p^p \leq \epsilon(m,\boldsymbol{n}) N(\boldsymbol{n}),$$

where $\lim_{m\to\infty} \limsup_{n\to\infty} \epsilon(m,\boldsymbol{n}) = 0$. Then $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.

Theorem 3.5. *Let $\{\{C_{\boldsymbol{n},m}\}_n\}_m$ be a sequence of matrix-sequences. Suppose that $\{C_{\boldsymbol{n},m}\}_n \sim_\sigma g_m$ for some measurable function $g_m : D \subset \mathbb{R}^k \to \mathbb{C}$ such that $g_m \to 0$ in measure as $m \to \infty$. Then $\{\{C_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$.*

Proof. For any $\ell \in \mathbb{N}$, let $F_\ell \in C_c(\mathbb{R})$ such that $F_\ell = 1$ over $[0, 1/2\ell]$, $F_\ell = 0$ outside $[-1/\ell, 1/\ell]$, and $0 \leq F_\ell \leq 1$ over $\mathbb{R}$. Note that $F_\ell \leq \chi_{[0,1/\ell]}$ over $[0,\infty)$. For every m, ℓ, we have

$$\begin{aligned}\frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} :\ \sigma_i(C_{\boldsymbol{n},m}) > 1/\ell\}}{N(\boldsymbol{n})} &= 1 - \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} :\ \sigma_i(C_{\boldsymbol{n},m}) \leq 1/\ell\}}{N(\boldsymbol{n})} = 1 - \frac{1}{N(\boldsymbol{n})}\sum_{i=1}^{N(\boldsymbol{n})} \chi_{[0,1/\ell]}(\sigma_i(C_{\boldsymbol{n},m}))\\ &\leq 1 - \frac{1}{N(\boldsymbol{n})}\sum_{i=1}^{N(\boldsymbol{n})} F_\ell(\sigma_i(C_{\boldsymbol{n},m})) \overset{n\to\infty}{\longrightarrow} c(m,\ell),\end{aligned} \tag{3.13}$$

where

$$c(m,\ell) = 1 - \frac{1}{\mu_k(D)} \int_D F_\ell(|g_m(\mathbf{x})|) d\mathbf{x}.$$

By Lemma 2.4, $c(m,\ell) \to 0$ as $m \to \infty$, for every fixed ℓ. Hence, there exists a sequence $\{\ell_m\}_m$ of natural numbers such that

$$\lim_{m\to\infty} \ell_m = \infty, \qquad \lim_{m\to\infty} c(m,\ell_m) = 0.$$

By (3.13), for each m we have

$$\limsup_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} :\ \sigma_i(C_{\boldsymbol{n},m}) > 1/\ell_m\}}{N(\boldsymbol{n})} \leq c(m,\ell_m). \tag{3.14}$$

Let $C_{\boldsymbol{n},m} = U_{\boldsymbol{n},m}\Sigma_{\boldsymbol{n},m}V^*_{\boldsymbol{n},m}$ be an SVD of $C_{\boldsymbol{n},m}$. Let $\hat{\Sigma}_{\boldsymbol{n},m}$ be the matrix obtained from $\Sigma_{\boldsymbol{n},m}$ by setting to 0 all the singular vaules that are less than or equal to $1/\ell_m$, and let $\tilde{\Sigma}_{\boldsymbol{n},m} = \Sigma_{\boldsymbol{n},m} - \hat{\Sigma}_{\boldsymbol{n},m}$ be the matrix obtained from $\Sigma_{\boldsymbol{n},m}$ by setting to 0 all the singular values that exceed $1/\ell_m$. Then we can write

$$C_{\boldsymbol{n},m} = R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$

where $R_{\boldsymbol{n},m} = U_{\boldsymbol{n},m}\hat{\Sigma}_{\boldsymbol{n},m}V^*_{\boldsymbol{n},m}$ and $N_{\boldsymbol{n},m} = U_{\boldsymbol{n},m}\tilde{\Sigma}_{\boldsymbol{n},m}V^*_{\boldsymbol{n},m}$. By definition, $\|N_{\boldsymbol{n},m}\| \leq 1/\ell_m$. Moreover, (3.14) says that

$$\limsup_{n\to\infty} \frac{\mathrm{rank}(R_{\boldsymbol{n},m})}{N(\boldsymbol{n})} \leq c(m,\ell_m),$$

implying the existence of a n_m such that, for $n \geq n_m$,

$$\mathrm{rank}(R_{\boldsymbol{n},m}) \leq \big(c(m,\ell_m) + 1/m\big)N(\boldsymbol{n}).$$

This shows that $\{\{C_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$. □

Corollary 3.4. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence and let $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ be a sequence of matrix-sequences. Suppose that $\{A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\}_n \sim_\sigma g_m$ for some measurable function $g_m : D \subset \mathbb{R}^k \to \mathbb{C}$ such that $g_m \to 0$ in measure as $m \to \infty$. Then $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.*

We conclude this section with the following converse of Theorem 3.5.

Theorem 3.6. *Let $\{\{C_{\boldsymbol{n},m}\}_n\}_m$ be an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$ and suppose that $\{C_{\boldsymbol{n},m}\}_n \sim_\sigma g_m$ for every m. Then $g_m \to 0$ in measure.*

Proof. Since $\{O_{N(\boldsymbol{n})}\}_n \sim_\sigma 0$, by Theorem 3.2 we have $\phi_{[g_m]} \to \phi_{[0]}$ pointwise over $C_c(\mathbb{R})$. By Lemma 2.5, this implies that $g_m \to 0$ in measure. □

3.4 An extension of the concept of a.c.s.

We provide in this section an extension of the definition of a.c.s. that will be used to define LT and GLT sequences in Sections 4–5. The extension is plain. The underlying idea is that, in Definition 3.1, one could choose to approximate $\{A_{\boldsymbol{n}}\}_n$ by a class of sequences $\{\{B_{\boldsymbol{n},\alpha}\}_n\}_{\alpha\in\mathcal{A}}$ parameterized by a not necessarily integer parameter α. For example, one may want to use a parameter $\epsilon > 0$ and to claim that a given class of sequences $\{\{B_{\boldsymbol{n},\epsilon}\}_n\}_{\epsilon>0}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ when $\epsilon \to 0$. Intuitively, this assertion should have the following meaning: for every $\epsilon > 0$ there exists n_ϵ such that, for $n \ge n_\epsilon$,

$$\begin{gathered} A_{\boldsymbol{n}} = B_{\boldsymbol{n},\epsilon} + R_{\boldsymbol{n},\epsilon} + N_{\boldsymbol{n},\epsilon}, \\ \text{rank}(R_{\boldsymbol{n},\epsilon}) \le c(\epsilon)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},\epsilon}\| \le \omega(\epsilon), \end{gathered}$$

where the quantities n_ϵ, $c(\epsilon)$, $\omega(\epsilon)$ depend only on ϵ and $\lim_{\epsilon\to 0} c(\epsilon) = \lim_{\epsilon\to 0}\omega(\epsilon) = 0$. Actually, this is the correct meaning!

Definition 3.3 (approximating class of sequences for $\epsilon \to 0$). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. We say that $\{\{B_{\boldsymbol{n},\epsilon}\}_n\}_{\epsilon>0}$ is an a.c.s. of $\{A_{\boldsymbol{n}}\}_n$ for $\epsilon \to 0$ if the following property holds: for every $\epsilon > 0$ there exists n_ϵ such that, for $n \ge n_\epsilon$,

$$\begin{gathered} A_{\boldsymbol{n}} = B_{\boldsymbol{n},\epsilon} + R_{\boldsymbol{n},\epsilon} + N_{\boldsymbol{n},\epsilon}, \\ \text{rank}(R_{\boldsymbol{n},\epsilon}) \le c(\epsilon)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},\epsilon}\| \le \omega(\epsilon), \end{gathered}$$

where the quantities n_ϵ, $c(\epsilon)$, $\omega(\epsilon)$ depend only on ϵ and

$$\lim_{\epsilon\to 0} c(\epsilon) = \lim_{\epsilon\to 0}\omega(\epsilon) = 0.$$

Definition 3.3 will be used to define GLT sequences in Section 5. Note that, if $\{\{B_{\boldsymbol{n},\epsilon}\}_n\}_{\epsilon>0}$ is an a.c.s. of $\{A_{\boldsymbol{n}}\}_n$ for $\epsilon \to 0$, then $\{\{B_{\boldsymbol{n},\epsilon(m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ (in the sense of the classical Definition 3.1) for all sequences of positive numbers $\{\epsilon(m)\}_m$ such that $\epsilon(m) \to 0$ as $m \to \infty$.

For the definition of LT sequences, we do not need the concept of a.c.s. parameterized by a positive $\epsilon \to 0$. On the contrary, we need the concept of a.c.s. parameterized by a multi-index $\boldsymbol{m} \to \infty$. Since any $m \in \mathbb{N}$ is a special multi-index, the definition of a.c.s. parameterized by a multi-index $\boldsymbol{m} \to \infty$ (Definition 3.4) is a true extension of Definition 3.1. In the following, a multi-index sequence of matrix-sequences is any class of sequences of the form $\{\{B_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$, where:

1. $\mathcal{M} \subseteq \mathbb{N}^q$ for some $q \ge 1$, and $\mathcal{M} \cap \{\boldsymbol{i} \in \mathbb{N}^q : \ \boldsymbol{i} \ge \boldsymbol{h}\} \ne \emptyset$ for every $\boldsymbol{h} \in \mathbb{N}^q$. We express the latter condition by saying that ∞ is an *accumulation point* for $\mathcal{M}$. This is required to ensure that $\boldsymbol{m}$ can tend to ∞ inside $\mathcal{M}$;
2. for every $\boldsymbol{m} \in \mathcal{M}$, $\{B_{\boldsymbol{n},\boldsymbol{m}}\}_n$ is a matrix-sequence.

Definition 3.4 (approximating class of sequences). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. An approximating class of sequences (a.c.s.) for $\{A_{\boldsymbol{n}}\}_n$ is a multi-index sequence of matrix-sequences $\{\{B_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ with the following property: for every $\boldsymbol{m} \in \mathcal{M}$ there exists $n_{\boldsymbol{m}}$ such that, for $n \ge n_{\boldsymbol{m}}$,

$$\begin{gathered} A_{\boldsymbol{n}} = B_{\boldsymbol{n},\boldsymbol{m}} + R_{\boldsymbol{n},\boldsymbol{m}} + N_{\boldsymbol{n},\boldsymbol{m}}, \\ \text{rank}(R_{\boldsymbol{n},\boldsymbol{m}}) \le c(\boldsymbol{m})N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},\boldsymbol{m}}\| \le \omega(\boldsymbol{m}), \end{gathered} \tag{3.15}$$

where the quantities $n_{\boldsymbol{m}}$, $c(\boldsymbol{m})$, $\omega(\boldsymbol{m})$ depend only on $\boldsymbol{m}$, and

$$\lim_{\boldsymbol{m}\to\infty} c(\boldsymbol{m}) = \lim_{\boldsymbol{m}\to\infty}\omega(\boldsymbol{m}) = 0.$$

Definition 3.4 extends the classical definition of a.c.s. (Definition 3.1). Indeed, a classical a.c.s. $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ for $\{A_{\boldsymbol{n}}\}_n$ is an a.c.s. also in the sense of Definition 3.4 (take $\mathcal{M}$ as the infinite subset of $\mathbb{N}$ where m varies).

Remark 3.4. If $\{\{B_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ in the sense of Definition 3.4, then $\{\{B_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ (in the sense of the classical Definition 3.1) for all sequences $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m \subseteq \mathcal{M}$ such that $\boldsymbol{m} \to \infty$ when $m \to \infty$.

Remark 3.5. An equivalent definition of a.c.s. is obtained by replacing, in Definition 3.4, 'for all $\boldsymbol{m} \in \mathcal{M}$' with 'for all sufficiently large $\boldsymbol{m} \in \mathcal{M}$' (i.e., 'for every $\boldsymbol{m} \in \mathcal{M}$ that is greater than or equal to some $\hat{\boldsymbol{m}} \in \mathbb{N}^q$, being $\mathcal{M} \subseteq \mathbb{N}^q$'). Indeed, suppose that the splitting (3.15) and the related conditions on $R_{\boldsymbol{n},\boldsymbol{m}}$ and $N_{\boldsymbol{n},\boldsymbol{m}}$ hold for $\boldsymbol{m} \geq \hat{\boldsymbol{m}}$; then, defining $n_{\boldsymbol{m}} = 1,\ c(\boldsymbol{m}) = 1,\ \omega(\boldsymbol{m}) = 0$ and $R_{\boldsymbol{n},\boldsymbol{m}} = A_{\boldsymbol{n}} - B_{\boldsymbol{n},\boldsymbol{m}},\ N_{\boldsymbol{n},\boldsymbol{m}} = O$ for all the other values of $\boldsymbol{m}$, we see that they actually hold for every $\boldsymbol{m} \in \mathcal{M}$.

Remark 3.6. Suppose that $\{\{B_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ and $\{\{B'_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ is an a.c.s. for $\{A'_{\boldsymbol{n}}\}_n$. Then:

1. $\{\{B^*_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ is an a.c.s. for $\{A^*_{\boldsymbol{n}}\}_n$;
2. $\{\{\alpha B_{\boldsymbol{n},\boldsymbol{m}} + \beta B'_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ is an a.c.s. for $\{\alpha A_{\boldsymbol{n}} + \beta A'_{\boldsymbol{n}}\}_n$, for all $\alpha, \beta \in \mathbb{C}$;
3. if $\{A_{\boldsymbol{n}}\}_n$ and $\{A'_{\boldsymbol{n}}\}_n$ are s.u., then $\{\{B_{\boldsymbol{n},\boldsymbol{m}} B'_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ is an a.c.s. for $\{A_{\boldsymbol{n}} A'_{\boldsymbol{n}}\}_n$.

The proof of these results is omitted, because it is essentially the same as the proof of the analogous results for standard a.c.s.; see Remark 3.3 and Propositions 3.1, 3.4.

4 LT and sLT sequences

In this section, we develop the theory of Locally Toeplitz sequences. We first introduce and analyze the Locally Toeplitz operator in Section 4.1. Then, in Section 4.2, we define the Locally Toeplitz and separable Locally Toeplitz sequences, and we study their properties. The results contained in this section are of fundamental importance for the theory of Generalized Locally Toeplitz sequences, which will be the subject of Section 5.

4.1 The Locally Toeplitz operator $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)$

Definition 4.1.

- Let $m, n \in \mathbb{N}$, let $a : [0, 1] \to \mathbb{C}$, and let $f : [-\pi, \pi] \to \mathbb{C}$ in $L^1([-\pi, \pi])$. Then, we define the $n \times n$ matrix
$$LT_n^m(a, f) = D_m(a) \otimes T_{\lfloor n/m \rfloor}(f) \ \oplus\ O_{n \bmod m} = \operatorname*{diag}_{j=1,\ldots,m} a\Big(\frac{j}{m}\Big) T_{\lfloor n/m \rfloor}(f) \ \oplus\ O_{n \bmod m}.$$
It is understood that $LT_n^m(a, f) = O_n$ when $n < m$ and that the term $O_{n \bmod m}$ is not present when n is a multiple of m. Moreover, here and in the following, the tensor product operation $\otimes$ is always applied before the direct sum $\oplus$, exactly as in the case of numbers, where multiplication is always applied before addition.
- Let $\boldsymbol{m}, \boldsymbol{n} \in \mathbb{N}^d$, let $a : [0, 1]^d \to \mathbb{C}$, and let $f_1, \ldots, f_d : [-\pi, \pi] \to \mathbb{C}$ in $L^1([-\pi, \pi])$. Then, we define the $N(\boldsymbol{n}) \times N(\boldsymbol{n})$ matrix
$$\begin{aligned} LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f_1 \otimes \cdots \otimes f_d) &= LT_{n_1,\ldots,n_d}^{m_1,\ldots,m_d}(a(x_1, \ldots, x_d), f_1 \otimes \cdots \otimes f_d) \\ &= \operatorname*{diag}_{j_1=1,\ldots,m_1} T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes LT_{n_2,\ldots,n_d}^{m_2,\ldots,m_d}\Big(a\Big(\frac{j_1}{m_1}, x_2, \ldots, x_d\Big), f_2 \otimes \cdots \otimes f_d\Big) \ \oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d}. \end{aligned}$$
This is a recursive definition, whose base case has been considered in the previous item. For example, in the case $d = 2$ we have
$$LT_{n_1,n_2}^{m_1,m_2}(a, f_1 \otimes f_2) = \operatorname*{diag}_{j_1=1,\ldots,m_1} T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes \left[\operatorname*{diag}_{j_2=1,\ldots,m_2} a\Big(\frac{j_1}{m_1}, \frac{j_2}{m_2}\Big) T_{\lfloor n_2/m_2 \rfloor}(f_2) \ \oplus\ O_{n_2 \bmod m_2}\right] \ \oplus\ O_{(n_1 \bmod m_1)n_2}.$$

In this section, especially in Subsection 4.1.1, we investigate the properties of $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)$ that will be of interest later on. We write $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)$ instead of $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f_1 \otimes \cdots \otimes f_d)$ because we are going to see that $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)$ is well-defined (in a unique way) for any function $f \in L^1([-\pi, \pi]^d)$; see Definition 4.2. In particular, if f is separable, the definition is independent of the factorization of f as a tensor product of the form $f_1 \otimes \cdots \otimes f_d,\ f_1, \ldots, f_d \in L^1([-\pi, \pi])$.

The main result about the Locally Toeplitz operator is stated in Theorem 4.1. It shows that $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f_1 \otimes \cdots \otimes f_d)$ coincides with $D_{\boldsymbol{m}}(a) \otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f_1 \otimes \cdots \otimes f_d) \ \oplus\ O$ up to a permutation transformation $\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}$ which only depends on $\boldsymbol{m}, \boldsymbol{n}$ and not on the specific functions $a, f_1, \ldots, f_d$. This result allows us to extend the definition of the Locally Toeplitz operator as in Definition 4.2. With this extension, we will be able to define in Section 4.2 the notion of Locally Toeplitz sequences in the multilevel setting. It is worth noting that such a notion is introduced here for the first time, because, so far, Locally Toeplitz sequences were considered in [52] only in the unilevel case, whereas the multilevel setting addressed in [44, 45] only deals with separable Locally Toeplitz sequences.

Theorem 4.1. *For any $\boldsymbol{m}, \boldsymbol{n} \in \mathbb{N}^d$ there exists a permutation matrix $\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}$ of size $N(\boldsymbol{n})$ such that*
$$LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f_1 \otimes \cdots \otimes f_d) = \Pi_{\boldsymbol{n}}^{\boldsymbol{m}} \Big[D_{\boldsymbol{m}}(a) \otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f_1 \otimes \cdots \otimes f_d) \ \oplus\ O_{N(\boldsymbol{n})-N(\boldsymbol{m})N(\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor)}\Big] (\Pi_{\boldsymbol{n}}^{\boldsymbol{m}})^T$$
for every $a : [0, 1]^d \to \mathbb{C}$ and every $f_1, \ldots, f_d \in L^1([-\pi, \pi])$.

Proof. The proof is done by induction on d. For $d=1$ the result holds with $\Pi_n^m = I_n$. For $d \geq 2$, set $\boldsymbol{\nu} = (n_2, \ldots, n_d)$ and $\boldsymbol{\mu} = (m_2, \ldots, m_d)$. By definition,

$$LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f_1 \otimes \cdots \otimes f_d) = \operatorname*{diag}_{j_1=1,\ldots,m_1} T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes LT_{\boldsymbol{\nu}}^{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big), f_2 \otimes \cdots \otimes f_d\Big) \ \oplus\ O_{(n_1 \bmod m_1) n_2 \cdots n_d}, \tag{4.1}$$

where $a(j_1/m_1, \cdot) : [0,1]^{d-1} \to \mathbb{C}$ is the function $(x_2, \ldots, x_d) \mapsto a(j_1/m_1, x_2, \ldots, x_d)$. By induction hypothesis, setting $N(\boldsymbol{\nu}, \boldsymbol{\mu}) = N(\boldsymbol{\nu}) - N(\boldsymbol{\mu}) N(\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor)$, we have

$$LT_{\boldsymbol{\nu}}^{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big), f_2 \otimes \cdots \otimes f_d\Big) = \Pi_{\boldsymbol{\nu}}^{\boldsymbol{\mu}} \Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor}(f_2 \otimes \cdots \otimes f_d) \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})}\Big] (\Pi_{\boldsymbol{\nu}}^{\boldsymbol{\mu}})^T. \tag{4.2}$$

Let us work on the argument of the 'diag operator' in (4.1). From Lemma 2.7, eq. (4.2) and the properties of tensor products (see Section 2.4.1), we get

$$\begin{aligned}
&T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes LT_{\boldsymbol{\nu}}^{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big), f_2 \otimes \cdots \otimes f_d\Big) \\
&= \Pi_{(\lfloor n_1/m_1 \rfloor, N(\boldsymbol{\nu}));[2,1]} \Big\{ LT_{\boldsymbol{\nu}}^{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big), f_2 \otimes \cdots \otimes f_d\Big) \otimes T_{\lfloor n_1/m_1 \rfloor}(f_1) \Big\} (\Pi_{(\lfloor n_1/m_1 \rfloor, N(\boldsymbol{\nu}));[2,1]})^T \\
&= \Pi_{(\lfloor n_1/m_1 \rfloor, N(\boldsymbol{\nu}));[2,1]} \Big\{ \Pi_{\boldsymbol{\nu}}^{\boldsymbol{\mu}} \Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor}(f_2 \otimes \cdots \otimes f_d) \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})}\Big] (\Pi_{\boldsymbol{\nu}}^{\boldsymbol{\mu}})^T \\
&\qquad\qquad \otimes\ T_{\lfloor n_1/m_1 \rfloor}(f_1) \Big\} (\Pi_{(\lfloor n_1/m_1 \rfloor, N(\boldsymbol{\nu}));[2,1]})^T \\
&= \Pi_{(\lfloor n_1/m_1 \rfloor, N(\boldsymbol{\nu}));[2,1]} (\Pi_{\boldsymbol{\nu}}^{\boldsymbol{\mu}} \otimes I_{\lfloor n_1/m_1 \rfloor}) \Big\{ \Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor}(f_2 \otimes \cdots \otimes f_d) \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})}\Big] \\
&\qquad\qquad \otimes\ T_{\lfloor n_1/m_1 \rfloor}(f_1) \Big\} (\Pi_{\boldsymbol{\nu}}^{\boldsymbol{\mu}} \otimes I_{\lfloor n_1/m_1 \rfloor})^T (\Pi_{(\lfloor n_1/m_1 \rfloor, N(\boldsymbol{\nu}));[2,1]})^T.
\end{aligned} \tag{4.3}$$

Using eq. (2.23), Lemma 2.7, Lemma 2.10 and the properties of tensor products and direct sums (see Section 2.4.1), we obtain

$$\begin{aligned}
&\Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor}(f_2 \otimes \cdots \otimes f_d) \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})}\Big] \otimes T_{\lfloor n_1/m_1 \rfloor}(f_1) \\
&= D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor}(f_2 \otimes \cdots \otimes f_d) \otimes T_{\lfloor n_1/m_1 \rfloor}(f_1) \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor} \\
&= \Pi_{(N(\boldsymbol{\mu}), \lfloor n_1/m_1 \rfloor, N(\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor));[1,3,2]} \Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes T_{\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor}(f_2 \otimes \cdots \otimes f_d)\Big] \\
&\qquad \cdot\ (\Pi_{(N(\boldsymbol{\mu}), \lfloor n_1/m_1 \rfloor, N(\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor));[1,3,2]})^T \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor} \\
&= \Pi_{(N(\boldsymbol{\mu}), \lfloor n_1/m_1 \rfloor, N(\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor));[1,3,2]} \Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f_1 \otimes \cdots \otimes f_d)\Big] \\
&\qquad \cdot\ (\Pi_{(N(\boldsymbol{\mu}), \lfloor n_1/m_1 \rfloor, N(\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor));[1,3,2]})^T \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor} \\
&= (\Pi_{(N(\boldsymbol{\mu}), \lfloor n_1/m_1 \rfloor, N(\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor));[1,3,2]} \oplus I_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor}) \Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f_1 \otimes \cdots \otimes f_d) \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor}\Big] \\
&\qquad \cdot\ (\Pi_{(N(\boldsymbol{\mu}), \lfloor n_1/m_1 \rfloor, N(\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor));[1,3,2]} \oplus I_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor})^T.
\end{aligned} \tag{4.4}$$

Substituting (4.4) into (4.3), we arrive at

$$\begin{aligned}
&T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes LT_{\boldsymbol{\nu}}^{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big), f_2 \otimes \cdots \otimes f_d\Big) \\
&= P_{\boldsymbol{n}}^{\boldsymbol{m}} \Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f_1 \otimes \cdots \otimes f_d) \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor}\Big] (P_{\boldsymbol{n}}^{\boldsymbol{m}})^T,
\end{aligned} \tag{4.5}$$

where $P_{\boldsymbol{n}}^{\boldsymbol{m}} = \Pi_{(\lfloor n_1/m_1 \rfloor, N(\boldsymbol{\nu}));[2,1]} (\Pi_{\boldsymbol{\nu}}^{\boldsymbol{\mu}} \otimes I_{\lfloor n_1/m_1 \rfloor}) (\Pi_{(N(\boldsymbol{\mu}), \lfloor n_1/m_1 \rfloor, N(\lfloor \boldsymbol{\nu}/\boldsymbol{\mu} \rfloor));[1,3,2]} \oplus I_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor})$. Combining (4.5) and (4.1), we obtain

$$\begin{aligned}
&LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f_1 \otimes \cdots \otimes f_d) \\
&= \Big(\bigoplus_{j_1=1}^{m_1} P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big) \operatorname*{diag}_{j_1=1,\ldots,m_1} \Big[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1}, \cdot\Big)\Big) \otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f_1 \otimes \cdots \otimes f_d) \ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu}) \lfloor n_1/m_1 \rfloor}\Big] \Big(\bigoplus_{j_1=1}^{m_1} P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)^T \ \oplus\ O_{(n_1 \bmod m_1) n_2 \cdots n_d}.
\end{aligned}$$

From Lemma 2.9,

$$
\begin{aligned}
&\operatorname*{diag}_{j_1=1,\ldots,m_1}\left[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor}\right]\\
&=\bigoplus_{j_1=1}^{m_1}\left[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor}\right]\\
&=V_{\boldsymbol{n}}^{\boldsymbol{m}}\left[\bigoplus_{j_1=1}^{m_1}\left[D_{\boldsymbol{\mu}}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\right]\ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor m_1}\right](V_{\boldsymbol{n}}^{\boldsymbol{m}})^T\\
&=V_{\boldsymbol{n}}^{\boldsymbol{m}}\left[D_{\boldsymbol{m}}(a)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor m_1}\right](V_{\boldsymbol{n}}^{\boldsymbol{m}})^T,
\end{aligned}
$$

where

$$
\begin{aligned}
V_{\boldsymbol{n}}^{\boldsymbol{m}}&=V_{\boldsymbol{h}(\boldsymbol{m},\boldsymbol{n});\sigma},\\
\sigma&=[1,m_1+1,2,m_1+2,\ldots,m_1,2m_1],\\
\boldsymbol{h}(\boldsymbol{m},\boldsymbol{n})&=\big(\underbrace{N(\boldsymbol{\mu})N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor),N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor}_{1},\ldots,\underbrace{N(\boldsymbol{\mu})N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor),N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor}_{m_1}\big).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f_1\otimes\cdots\otimes f_d)\\
&=\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)V_{\boldsymbol{n}}^{\boldsymbol{m}}\left[D_{\boldsymbol{m}}(a)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor m_1}\right](V_{\boldsymbol{n}}^{\boldsymbol{m}})^T\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)^T\ \oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d}\\
&=\left[\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)V_{\boldsymbol{n}}^{\boldsymbol{m}}\ \oplus\ I_{(n_1 \bmod m_1)n_2\cdots n_d}\right]\left[D_{\boldsymbol{m}}(a)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor m_1+(n_1 \bmod m_1)n_2\cdots n_d}\right]\\
&\qquad\cdot\left[(V_{\boldsymbol{n}}^{\boldsymbol{m}})^T\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)^T\ \oplus\ I_{(n_1 \bmod m_1)n_2\cdots n_d}\right].
\end{aligned}
$$

This concludes the proof; note that the permutation matrix $\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}$ is given by

$$
\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}=\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)V_{\boldsymbol{n}}^{\boldsymbol{m}}\ \oplus\ I_{(n_1 \bmod m_1)n_2\cdots n_d}
$$

and, moreover, $N(\boldsymbol{\nu},\boldsymbol{\mu})\lfloor n_1/m_1\rfloor m_1+(n_1 \bmod m_1)n_2\cdots n_d=N(\boldsymbol{n})-N(\boldsymbol{m})N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor)$. □

As a consequence of Theorem 4.1, we can extend Definition 4.1 in the following way.

Definition 4.2. Let $\boldsymbol{m},\boldsymbol{n}\in\mathbb{N}^d$, let $a:[0,1]^d\to\mathbb{C}$ and let $f\in L^1([-\pi,\pi]^d)$. Then, we define

$$
LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)=\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}\left[D_{\boldsymbol{m}}(a)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f)\ \oplus\ O_{N(\boldsymbol{n})-N(\boldsymbol{m})N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor)}\right](\Pi_{\boldsymbol{n}}^{\boldsymbol{m}})^T,
$$

where $\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}$ is the permutation matrix appearing in Theorem 4.1.

Remark 4.1. Note that $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)=LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,g)$ whenever $f=g$ a.e. Moreover, suppose that $f=f_1\otimes\cdots\otimes f_d$ a.e., with $f_1,\ldots,f_d\in L^1([-\pi,\pi])$; then $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)$ is equal to $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f_1\otimes\cdots\otimes f_d)$, as defined by Definition 4.1.

4.1.1 Properties of $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)$

We now derive a lot of interesting properties of $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)$ that we shall use in the study of LT, sLT and GLT sequences. We first not that, for any $\boldsymbol{n},\boldsymbol{m}\in\mathbb{N}^d$ and any pair of functions $a:[0,1]^d\to\mathbb{C}$ and $f\in L^1([-\pi,\pi]^d)$,

$$
[LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)]^*=LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\overline{a},\overline{f}). \tag{4.6}
$$

This follows from Definition 4.2, from the relations $(X\otimes Y)^*=X^*\otimes Y^*$, $(X\oplus Y)^*=X^*\oplus Y^*$, and from the equality $[T_{\boldsymbol{k}}(f)]^*=T_{\boldsymbol{k}}(\overline{f})$ (see Section 2.6).

Proposition 4.1. *Let $\boldsymbol{m}, \boldsymbol{n} \in \mathbb{N}^d$, let $a : [0,1]^d \to \mathbb{C}$ and let $f \in L^1([-\pi,\pi]^d)$. Then,*

$$\|LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)\| = \|D_{\boldsymbol{m}}(a)\| \, \|T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f)\| = \max_{\boldsymbol{j}=\boldsymbol{1},\ldots,\boldsymbol{m}} \left| a\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)\right| \|T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f)\|, \tag{4.7}$$

$$\|LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)\|_p = \|D_{\boldsymbol{m}}(a)\|_p \, \|T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f)\|_p = \left(\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}} \left| a\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)\right|^p \right)^{1/p} \|T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f)\|_p, \qquad 1 \le p < \infty. \tag{4.8}$$

Proof. Use Definition 4.2, the invariance of $\|\cdot\|$ and $\|\cdot\|_p$ by unitary transformations (such as permutations), and eqs. (2.20)–(2.21). □

We denote by $\mathbb{C}^{[0,1]^d}$ the vector space of all functions $a : [0,1]^d \to \mathbb{C}$.

Proposition 4.2. *Let $\boldsymbol{m}, \boldsymbol{n} \in \mathbb{N}^d$. Then, the operator*

$$LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,\cdot) : L^1([-\pi,\pi]^d) \to \mathbb{C}^{N(\boldsymbol{n})\times N(\boldsymbol{n})}$$

is linear for any $a : [0,1]^d \to \mathbb{C}$, and the operator

$$LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\cdot,f) : \mathbb{C}^{[0,1]^d} \to \mathbb{C}^{N(\boldsymbol{n})\times N(\boldsymbol{n})}$$

is linear for any $f \in L^1([-\pi,\pi]^d)$.

Proof. Use Definition 4.2, the linearity of the operators $D_{\boldsymbol{m}}(\cdot)$ and $T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(\cdot)$, and the bilinearity of tensor produtcs. □

By Hölder's inequality [39], if $f \in L^p([-\pi,\pi]^d)$ and $\tilde f \in L^q([-\pi,\pi]^d)$, where $1 \le p,q \le \infty$ are conjugate exponents, then $f\tilde f \in L^1([-\pi,\pi]^d)$. In this case, for any $a, \tilde a : [0,1]^d \to \mathbb{C}$, we can consider the three matrices $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)$, $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde a,\tilde f)$ and $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a\tilde a, f\tilde f)$. In Proposition 4.3 we show that $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde a,\tilde f)$ is 'close' to $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a\tilde a, f\tilde f)$.

Proposition 4.3. *Let $a, \tilde a : [0,1]^d \to \mathbb{C}$ be bounded, and let $f \in L^p([-\pi,\pi]^d)$ and $\tilde f \in L^q([-\pi,\pi]^d)$, where $1 \le p,q \le \infty$ are conjugate exponents. Then, for every $\boldsymbol{n}, \boldsymbol{m} \in \mathbb{N}^d$,*

$$\|LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde a,\tilde f) - LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a\tilde a, f\tilde f)\|_1 \le \epsilon(\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor) N(\boldsymbol{n}), \tag{4.9}$$

where

$$\epsilon(\boldsymbol{k}) = \|a\tilde a\|_\infty \frac{\|T_{\boldsymbol{k}}(f)T_{\boldsymbol{k}}(\tilde f) - T_{\boldsymbol{k}}(f\tilde f)\|_1}{N(\boldsymbol{k})}$$

and $\lim_{\boldsymbol{k}\to\infty} \epsilon(\boldsymbol{k}) = 0$ by Lemma 2.11. In particular, for every $\boldsymbol{m} \in \mathbb{N}^d$ there exists $\boldsymbol{n}_{\boldsymbol{m}} \in \mathbb{N}^d$ such that, for $\boldsymbol{n} \ge \boldsymbol{n}_{\boldsymbol{m}}$,

$$\|LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde a,\tilde f) - LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a\tilde a, f\tilde f)\|_1 \le \frac{N(\boldsymbol{n})}{N(\boldsymbol{m})}, \tag{4.10}$$

$$LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde a,\tilde f) = LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a\tilde a, f\tilde f) + R_{\boldsymbol{n},\boldsymbol{m}} + N_{\boldsymbol{n},\boldsymbol{m}}, \qquad \mathrm{rank}(R_{\boldsymbol{n},\boldsymbol{m}}) \le \frac{N(\boldsymbol{n})}{\sqrt{N(\boldsymbol{m})}}, \qquad \|N_{\boldsymbol{n},\boldsymbol{m}}\| \le \frac{1}{\sqrt{N(\boldsymbol{m})}}. \tag{4.11}$$

Proof. By Definition 4.2 and the properties of tensor products and direct sums,

$$LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde a,\tilde f) - LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a\tilde a, f\tilde f) = \Pi_{\boldsymbol{n}}^{\boldsymbol{m}} \Big[D_{\boldsymbol{m}}(a\tilde a) \otimes \Big(T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f)T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(\tilde f) - T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f\tilde f) \Big) \oplus O \Big] (\Pi_{\boldsymbol{n}}^{\boldsymbol{m}})^T.$$

Hence,

$$\begin{aligned}\|LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde a,\tilde f) - LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a\tilde a, f\tilde f)\|_1 &= \|D_{\boldsymbol{m}}(a\tilde a)\|_1 \, \|T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f)T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(\tilde f) - T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f\tilde f)\|_1 \\ &\le N(\boldsymbol{n})\|a\tilde a\|_\infty \frac{\|T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f)T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(\tilde f) - T_{\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor}(f\tilde f)\|_1}{N(\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor)},\end{aligned}$$

and (4.9) is proved. Since $\epsilon(\boldsymbol{k}) \to 0$ when $\boldsymbol{k} \to \infty$, for every $\boldsymbol{m} \in \mathbb{N}^d$ there exists $\boldsymbol{n}_{\boldsymbol{m}} \in \mathbb{N}^d$ such that, for $\boldsymbol{n} \ge \boldsymbol{n}_{\boldsymbol{m}}$, (4.10) holds. (4.11) follows directly from (4.10) and Lemma 3.4. □

Theorems 4.2–4.3 provide information about the asymptotic singular value and eigenvalue distribution of a finite sum of the form $\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)$. Together with Theorems 3.1, 3.3 and Corollaries 3.1–3.2, they play a central role in the computation of the singular value and eigenvalue distribution of GLT sequences.

Theorem 4.2. *Let $a_1,\dots,a_p:[0,1]^d\to\mathbb{C}$ and let $f_1,\dots,f_p\in L^1([-\pi,\pi]^d)$. Then, for every $\boldsymbol{m}\in\mathbb{N}^d$ and every $F\in C_c(\mathbb{R})$,*

$$\lim_{\boldsymbol{n}\to\infty}\frac{1}{N(\boldsymbol{n})}\sum_{r=1}^{N(\boldsymbol{n})}F\Big(\sigma_r\Big(\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)\Big)\Big)=\phi_{\boldsymbol{m}}(F)=\frac{1}{N(\boldsymbol{m})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}\frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d}F\Big(\Big|\sum_{i=1}^p a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i(\boldsymbol{\theta})\Big|\Big)d\boldsymbol{\theta}. \tag{4.12}$$

Moreover, if $a_1,\dots,a_p$ are Riemann-integrable, then, for every $F\in C_c(\mathbb{R})$,

$$\lim_{\boldsymbol{m}\to\infty}\phi_{\boldsymbol{m}}(F)=\phi(F)=\frac{1}{(2\pi)^d}\int_{[0,1]^d\times[-\pi,\pi]^d}F\Big(\Big|\sum_{i=1}^p a_i(\mathbf{x})f_i(\boldsymbol{\theta})\Big|\Big)d\mathbf{x}d\boldsymbol{\theta}. \tag{4.13}$$

Proof. By Definition 4.2,

$$(\Pi_{\boldsymbol{n}}^{\boldsymbol{m}})^T\left(\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)\right)\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}=\left(\sum_{i=1}^p D_{\boldsymbol{m}}(a_i)\otimes T_{\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor}(f_i)\right)\,\oplus\,O_{N(\boldsymbol{n})-N(\boldsymbol{m})N(\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor)}.$$

For $\boldsymbol{j}=\boldsymbol{1},\dots,\boldsymbol{m}$, the $\boldsymbol{j}$-th block of this matrix is given by

$$\sum_{i=1}^p a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)T_{\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor}(f_i)=T_{\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor}\Big(\sum_{i=1}^p a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i\Big).$$

It follows that the singular values of $\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)$ are

$$\sigma_k\Big(T_{\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor}\Big(\sum_{i=1}^p a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i\Big)\Big),\quad k=1,\dots,N(\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor),\quad \boldsymbol{j}=\boldsymbol{1},\dots,\boldsymbol{m},$$

plus further $N(\boldsymbol{n})-N(\boldsymbol{m})N(\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor)$ singular values equal to 0. Therefore, by Theorem 2.7, for any $F\in C_c(\mathbb{R})$ we have

$$\begin{aligned}
&\lim_{\boldsymbol{n}\to\infty}\frac{1}{N(\boldsymbol{n})}\sum_{r=1}^{N(\boldsymbol{n})}F\Big(\sigma_r\Big(\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)\Big)\Big)\\
&=\lim_{\boldsymbol{n}\to\infty}\frac{N(\boldsymbol{m})N(\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor)}{N(\boldsymbol{n})}\frac{1}{N(\boldsymbol{m})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}\frac{1}{N(\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor)}\sum_{k=1}^{N(\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor)}F\Big(\sigma_k\Big(T_{\lfloor\boldsymbol{n}/\boldsymbol{m}\rfloor}\Big(\sum_{i=1}^p a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i\Big)\Big)\Big)\\
&=\frac{1}{N(\boldsymbol{m})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}\frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d}F\Big(\Big|\sum_{i=1}^p a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i(\boldsymbol{\theta})\Big|\Big)d\boldsymbol{\theta}.
\end{aligned}\tag{4.14}$$

This proves (4.12).

If $a_1,\dots,a_p$ are Riemann-integrable, then the function $\mathbf{x}\mapsto F\left(\left|\sum_{i=1}^p a_i(\mathbf{x})\,f_i(\boldsymbol{\theta})\right|\right)$ is Riemann-integrable for each fixed $\boldsymbol{\theta}\in[-\pi,\pi]^d$, being the composition of a continuous function with a Riemann-integrable function. Hence,

$$\lim_{\boldsymbol{m}\to\infty}\frac{1}{N(\boldsymbol{m})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}F\Big(\Big|\sum_{i=1}^p a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i(\boldsymbol{\theta})\Big|\Big)=\int_{[0,1]^d}F\Big(\Big|\sum_{i=1}^p a_i(\mathbf{x})f_i(\boldsymbol{\theta})\Big|\Big)d\mathbf{x}.$$

Passing to the limit as $\boldsymbol{m}\to\infty$ in (4.14), and using the dominated convergence theorem, we get (4.13). □

Theorem 4.3. *Let $a_1,\dots,a_p:[0,1]^d\to\mathbb{C}$ and let $f_1,\dots,f_p\in L^1([-\pi,\pi]^d)$. Then, for every $\boldsymbol{m}\in\mathbb{N}^d$ and every $F\in C_c(\mathbb{C})$,*

$$\lim_{\boldsymbol{n}\to\infty}\frac{1}{N(\boldsymbol{n})}\sum_{r=1}^{N(\boldsymbol{n})}F\Big(\lambda_r\Big(\Re\Big(\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)\Big)\Big)\Big)=\phi_{\boldsymbol{m}}(F)=\frac{1}{N(\boldsymbol{m})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}\frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d}F\Big(\Re\Big(\sum_{i=1}^p a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i(\boldsymbol{\theta})\Big)\Big)d\boldsymbol{\theta}. \tag{4.15}$$

Moreover, if $a_1,\dots,a_p$ are Riemann-integrable, then, for every $F\in C_c(\mathbb{C})$,

$$\lim_{\boldsymbol{m}\to\infty}\phi_{\boldsymbol{m}}(F)=\phi(F)=\frac{1}{(2\pi)^d}\int_{[0,1]^d\times[-\pi,\pi]^d}F\Big(\Re\Big(\sum_{i=1}^p a_i(\mathbf{x})f_i(\boldsymbol{\theta})\Big)\Big)d\mathbf{x}d\boldsymbol{\theta}. \tag{4.16}$$

Proof. The proof follows the same pattern as the proof of Theorem 4.2. By (4.6) and Definition 4.2,

$$(\Pi_{\boldsymbol{n}}^{\boldsymbol{m}})^T\left(\Re\Big(\sum_{i=1}^{p} LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)\Big)\right)\Pi_{\boldsymbol{n}}^{\boldsymbol{m}} = (\Pi_{\boldsymbol{n}}^{\boldsymbol{m}})^T\left(\frac{1}{2}\left(\sum_{i=1}^{p} LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)+\sum_{i=1}^{p} LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\overline{a_i},\overline{f_i})\right)\right)\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}$$
$$=\frac{1}{2}\left(\sum_{i=1}^{p} D_{\boldsymbol{m}}(a_i)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_i)+\sum_{i=1}^{p} D_{\boldsymbol{m}}(\overline{a_i})\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(\overline{f_i})\right)\ \oplus\ O_{N(\boldsymbol{n})-N(\boldsymbol{m})N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor)}.$$

The $\boldsymbol{j}$-th block of this matrix, $\boldsymbol{1}\le \boldsymbol{j}\le \boldsymbol{m}$, is given by

$$\frac{1}{2}\left(\sum_{i=1}^{p} a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_i)+\sum_{i=1}^{p}\overline{a_i}\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(\overline{f_i})\right)=T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}\Big(\Re\Big(\sum_{i=1}^{p} a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i\Big)\Big).$$

It follows that the eigenvalues of $\Re\left(\sum_{i=1}^{p} LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)\right)$ are

$$\lambda_k\Big(T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}\Big(\Re\Big(\sum_{i=1}^{p} a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i\Big)\Big)\Big),\quad k=1,\ldots,N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor),\quad \boldsymbol{j}=\boldsymbol{1},\ldots,\boldsymbol{m},$$

plus further $N(\boldsymbol{n})-N(\boldsymbol{m})N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor)$ eigenvalues equal to 0. Therefore, by Theorem 2.7, for any $F\in C_c(\mathbb{C})$ we have

$$\lim_{\boldsymbol{n}\to\infty}\frac{1}{N(\boldsymbol{n})}\sum_{r=1}^{N(\boldsymbol{n})}F\Big(\lambda_r\Big(\Re\Big(\sum_{i=1}^{p} LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i,f_i)\Big)\Big)\Big)$$
$$=\lim_{\boldsymbol{n}\to\infty}\frac{N(\boldsymbol{m})N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor)}{N(\boldsymbol{n})}\,\frac{1}{N(\boldsymbol{m})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}\frac{1}{N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor)}\sum_{k=1}^{N(\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor)}F\Big(\lambda_k\Big(T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}\Big(\Re\Big(\sum_{i=1}^{p} a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i\Big)\Big)\Big)\Big)$$
$$=\frac{1}{N(\boldsymbol{m})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}\frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d}F\Big(\Re\Big(\sum_{i=1}^{p} a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i(\boldsymbol{\theta})\Big)\Big)d\boldsymbol{\theta}. \tag{4.17}$$

This proves (4.15).

If $a_1,\ldots,a_p$ are Riemann-integrable, then the function $\mathbf{x}\mapsto F\left(\Re\left(\sum_{i=1}^{p} a_i(\mathbf{x})\,f_i(\boldsymbol{\theta})\right)\right)$ is Riemann-integrable for each fixed $\boldsymbol{\theta}\in[-\pi,\pi]^d$, and so

$$\lim_{\boldsymbol{m}\to\infty}\frac{1}{N(\boldsymbol{m})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}F\Big(\Re\Big(\sum_{i=1}^{p} a_i\Big(\frac{\boldsymbol{j}}{\boldsymbol{m}}\Big)f_i(\boldsymbol{\theta})\Big)\Big)=\int_{[0,1]^d}F\Big(\Re\Big(\sum_{i=1}^{p} a_i(\mathbf{x})f_i(\boldsymbol{\theta})\Big)\Big)d\mathbf{x}.$$

Passing to the limit as $\boldsymbol{m}\to\infty$ in (4.17), and using the dominated convergence theorem, we get (4.16). □

4.2 LT and sLT sequences

We recall that, unless otherwise specified, the multi-index that parameterizes a matrix-sequence $\{A_{\boldsymbol{n}}\}_n$ is always assumed to be a d-index, $\boldsymbol{n}=(n_1,\ldots,n_d)$. This convention, which was adopted from the beginning of this work (see Section 2.1.2), must be kept in mind especially in the present section and in Section 5. We will not repeat this anymore (although, sometimes, we will add the specification that $\boldsymbol{n}\in\mathbb{N}^d$, as a reminder for the reader).

4.2.1 Definitions and basic examples: zero-distributed sequences, sequences of diagonal sampling matrices and sequences of Toeplitz matrices

Definition 4.3 (LT sequence). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, $\boldsymbol{n}\in\mathbb{N}^d$. We say that $\{A_{\boldsymbol{n}}\}_n$ is a Locally Toeplitz (LT) sequence if there exist a Riemann-integrable function $a:[0,1]^d\to\mathbb{C}$ and a function $f\in L^1([-\pi,\pi]^d)$ such that $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)\}_n\}_{\boldsymbol{m}\in\mathbb{N}^d}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ (in the sense of Definition 3.4). This means that, for all $\boldsymbol{m}\in\mathbb{N}^d$ there is $n_{\boldsymbol{m}}$ such that, for $n\ge n_{\boldsymbol{m}}$,

$$A_{\boldsymbol{n}}=LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)+R_{\boldsymbol{n},\boldsymbol{m}}+N_{\boldsymbol{n},\boldsymbol{m}},$$
$$\mathrm{rank}(R_{\boldsymbol{n},\boldsymbol{m}})\le c(\boldsymbol{m})N(\boldsymbol{n}),\qquad \|N_{\boldsymbol{n},\boldsymbol{m}}\|\le\omega(\boldsymbol{m}), \tag{4.18}$$

where the quantities $n_{\boldsymbol{m}}$, $c(\boldsymbol{m})$, $\omega(\boldsymbol{m})$ are independent of n, and $\lim_{\boldsymbol{m}\to\infty}c(\boldsymbol{m})=\lim_{\boldsymbol{m}\to\infty}\omega(\boldsymbol{m})=0$. In this case, we write $\{A_{\boldsymbol{n}}\}_n\sim_{\rm LT}a\otimes f$. The function $a\otimes f$ is referred to as the *symbol* of the LT sequence $\{A_{\boldsymbol{n}}\}_n$, a is the *weight function* and f is the *generating function*.[4]

[4]We refer the reder to the introduction of Tilli's paper [52] for the origin and the meaning of this terminology.

Definition 4.4 (sLT sequence). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, $\boldsymbol{n} \in \mathbb{N}^d$. We say that $\{A_{\boldsymbol{n}}\}_n$ is a separable Locally Toeplitz (sLT) sequence if there exist a Riemann-integrable function $a : [0,1]^d \to \mathbb{C}$ and a *separable* function $f \in L^1([-\pi,\pi]^d)$ such that $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$. In this case, we write $\{A_{\boldsymbol{n}}\}_n \sim_{\rm sLT} a \otimes f$.

It is clear from the definition that a sLT sequence is just a LT sequence with separable generating function. From now on, if we write $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$ (resp. $\{A_{\boldsymbol{n}}\}_n \sim_{\rm sLT} a \otimes f$), it is understood that $a : [0,1]^d \to \mathbb{C}$ is Riemann-integrable and $f \in L^1([-\pi,\pi]^d)$ (resp. $f \in L^1([-\pi,\pi]^d)$ is separable).

Remark 4.2. If $\{A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} a \otimes f_i$ for $i = 1,\ldots,r$, then $\{\sum_{i=1}^r A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} a \otimes (\sum_{i=1}^r f_i)$. Similarly, if $\{A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} a_i \otimes f$ for $i = 1,\ldots,r$, then $\{\sum_{i=1}^r A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} (\sum_{i=1}^r a_i) \otimes f$. The proof of these results relies on the linearity of the Locally Toeplitz operator with respect to both its arguments (see Proposition 4.2); we leave it as an exercise for the reader.

Let us now provide basic examples of LT sequences: zero-distributed sequences, sequences formed by multilevel diagonal sampling matrices, and sequences of multilevel Toeplitz matrices. These may be regarded as the building blocks of the theory of GLT sequences, because, starting from them, we can construct through algebraic operations a lot of other matrix-sequences which will be seen in Section 5 to be GLT sequences. We recall that, according to our terminology, ∞ is an accumulation point for a subset $\mathcal{M} \subseteq \mathbb{N}^q$ if $\mathcal{M} \cap \{\boldsymbol{i} \in \mathbb{N}^q : \boldsymbol{i} \geq \boldsymbol{h}\} \neq \emptyset$ for all $\boldsymbol{h} \in \mathbb{N}^q$.

Theorem 4.4. *Let $\{Z_{\boldsymbol{n}}\}_n$ be a matrix-sequence. Let $\mathcal{M}$ be a subset of some $\mathbb{N}^q$ such that ∞ is an accumulation point for $\mathcal{M}$. Then, the following conditions are equivalent.*

1. $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$.
2. $\{\{O_{N(\boldsymbol{n})}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ *is an a.c.s. for* $\{Z_{\boldsymbol{n}}\}_n$.
3. $\{\{Z_{\boldsymbol{n}}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ *is an a.c.s. for* $\{O_{N(\boldsymbol{n})}\}_n$.

In particular, a matrix-sequence $\{Z_{\boldsymbol{n}}\}_n$ is zero-distributed if and only if $\{Z_{\boldsymbol{n}}\}_n \sim_{\rm sLT} 0$.

Proof. The equivalence $2 \Leftrightarrow 3$ is obvious from Definition 3.4.

Let us prove that $1 \Rightarrow 2$. By Theorem 2.5, we can write $Z_{\boldsymbol{n}} = R_{\boldsymbol{n}} + N_{\boldsymbol{n}}$ for all n, where

$$\lim_{n\to\infty} \frac{\text{rank}(R_{\boldsymbol{n}})}{N(\boldsymbol{n})} = \lim_{n\to\infty} \|N_{\boldsymbol{n}}\| = 0.$$

It follows that $\{\{O_{N(\boldsymbol{n})}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ is an a.c.s. for $\{Z_{\boldsymbol{n}}\}_n$. To see this, take, in (3.15), $R_{\boldsymbol{n},\boldsymbol{m}} = R_{\boldsymbol{n}},\ N_{\boldsymbol{n},\boldsymbol{m}} = N_{\boldsymbol{n}},\ c(\boldsymbol{m})$ and $\omega(\boldsymbol{m})$ any two positive functions of $\boldsymbol{m}$ that converge to 0 as $\boldsymbol{m} \to \infty$ (for instance, $c(\boldsymbol{m}) = \omega(\boldsymbol{m}) = 1/\min(\boldsymbol{m})$), and $n_{\boldsymbol{m}}$ any integer such that $\text{rank}(R_{\boldsymbol{n}})/N(\boldsymbol{n}) \leq c(\boldsymbol{m})$ and $\|N_{\boldsymbol{n}}\| \leq \omega(\boldsymbol{m})$ for $n \geq n_{\boldsymbol{m}}$.

Let us now prove that $2 \Rightarrow 1$. By assumption, $\{\{O_{N(\boldsymbol{n})}\}_n\}_{\boldsymbol{m}\in\mathcal{M}}$ is an a.c.s. for $\{Z_{\boldsymbol{n}}\}_n$. Hence, if we take any sequence $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m \subseteq \mathcal{M}$ such that $\boldsymbol{m} \to \infty$ as $m \to \infty$, $\{\{O_{N(\boldsymbol{n})}\}_n\}_m$ is a (classical) a.c.s. for $\{Z_{\boldsymbol{n}}\}_n$. Moreover, it is clear that $\{O_{N(\boldsymbol{n})}\}_n \sim_\sigma 0$. Hence, $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$ by Corollary 3.1.

The fact that $\{Z_{\boldsymbol{n}}\}_n$ is zero-distributed if and only if $\{Z_{\boldsymbol{n}}\}_n \sim_{\rm sLT} 0$ follows from the equivalence $1 \Leftrightarrow 2$ (applied with $\mathcal{M} = \mathbb{N}^d$) and from the observation that $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(0,0)\}_n\}_{\boldsymbol{m}\in\mathbb{N}^d} = \{\{O_{N(\boldsymbol{n})}\}_n\}_{\boldsymbol{m}\in\mathbb{N}^d}$ and $0 \otimes 0 = 0$. ∎

For the proof of Theorems 4.5–4.6, we need the following technical lemma.

Lemma 4.1. *Let $\hat{\mathbb{N}}$ be an infinite subset of $\mathbb{N}$. Let $x(\cdot,\cdot) : \hat{\mathbb{N}} \times \mathbb{N}^d \to \mathbb{R}$ be any function satisfying*

$$\lim_{m\to\infty} \lim_{\boldsymbol{h}\to\infty} x(m,\boldsymbol{h}) = \xi \in \mathbb{R}.$$

Then, there exists a function $m(\cdot) : \mathbb{N}^d \to \hat{\mathbb{N}}$ such that, when $\boldsymbol{h} \to \infty$, $m(\boldsymbol{h}) \to \infty$ and $x(m(\boldsymbol{h}),\boldsymbol{h}) \to \xi$.

Proof. Let $\hat{\mathbb{N}} = \{m_1, m_2, m_3, \ldots\}$. We denote by m_+ (m_-) the number that follows (preceeds) m in $\hat{\mathbb{N}}$. Set

$$x(m) = \lim_{\boldsymbol{h}\to\infty} x(m,\boldsymbol{h}), \qquad m \in \hat{\mathbb{N}}.$$

Since $x(m) \to \xi$ by assumption, $x(m)$ is eventually a real number (different from $-\infty$ or $+\infty$). Suppose, for instance, that $x(m) \in \mathbb{R}$ for all $m \geq m_r$. We construct an injective function $\boldsymbol{h}(\cdot) : \{m_r, m_{r+1}, m_{r+2}, \ldots\} \to \mathbb{N}^d$ as follows: we set $\boldsymbol{h}(m_r) = \boldsymbol{1}$ and, for every $m \in \hat{\mathbb{N}} \backslash \{m_r\}$, we choose $\boldsymbol{h}(m) > \boldsymbol{h}(m_-)$ such that, for $\boldsymbol{h} \geq \boldsymbol{h}(m)$,

$$|x(m,\boldsymbol{h}) - x(m)| \leq \frac{1}{m} \qquad \Rightarrow \qquad |x(m,\boldsymbol{h}) - \xi| \leq |x(m) - \xi| + \frac{1}{m}.$$

Hence, we have constructed a sequence

$$\mathbf{1} = \boldsymbol{h}(m_r) < \boldsymbol{h}(m_{r+1}) < \boldsymbol{h}(m_{r+2}) < \boldsymbol{h}(m_{r+3}) < \dots$$

and this sequence has the following property: if $m \in \{m_{r+1}, m_{r+2}, \dots\}$ and $\boldsymbol{h} \geq \boldsymbol{h}(m)$, then

$$|x(m, \boldsymbol{h}) - \xi| \leq |x(m) - \xi| + \frac{1}{m}. \tag{4.19}$$

In view of this, we define $m(\cdot) : \mathbb{N}^d \to \hat{\mathbb{N}}$ as follows:

$$m(\boldsymbol{h}) = \begin{cases} m_r & \text{if } \boldsymbol{h} \in \{\boldsymbol{j} \in \mathbb{N}^d : \boldsymbol{j} \geq \boldsymbol{h}(m_r) = \mathbf{1}\} \backslash \{\boldsymbol{j} \in \mathbb{N}^d : \boldsymbol{j} \geq \boldsymbol{h}(m_{r+1})\}, \\ m_{r+1} & \text{if } \boldsymbol{h} \in \{\boldsymbol{j} \in \mathbb{N}^d : \boldsymbol{j} \geq \boldsymbol{h}(m_{r+1})\} \backslash \{\boldsymbol{j} \in \mathbb{N}^d : \boldsymbol{j} \geq \boldsymbol{h}(m_{r+2})\}, \\ m_{r+2} & \text{if } \boldsymbol{h} \in \{\boldsymbol{j} \in \mathbb{N}^d : \boldsymbol{j} \geq \boldsymbol{h}(m_{r+2})\} \backslash \{\boldsymbol{j} \in \mathbb{N}^d : \boldsymbol{j} \geq \boldsymbol{h}(m_{r+3})\}, \\ \vdots & \quad \vdots \end{cases}$$

Clearly, $m(\boldsymbol{h}) \to \infty$ when $\boldsymbol{h} \to \infty$. Moreover, for all $\boldsymbol{h} \geq \boldsymbol{h}(m_{r+1})$ we have $m(\boldsymbol{h}) \in \{m_{r+1}, m_{r+2}, \dots\}$ and $\boldsymbol{h} \geq \boldsymbol{h}(m(\boldsymbol{h}))$. Hence, by (4.19), for all $\boldsymbol{h} \geq \boldsymbol{h}(m_{r+1})$ we have $|x(m(\boldsymbol{h}), \boldsymbol{h}) - \xi| \leq |x(m(\boldsymbol{h})) - \xi| + 1/m(\boldsymbol{h})$, which tends to 0 as $\boldsymbol{h} \to \infty$. □

Theorem 4.5. *Let $a : [0,1]^d \to \mathbb{C}$ be Riemann-integrable and consider the sequence of multilevel diagonal sampling matrices $\{D_{\boldsymbol{n}}(a)\}_n$, where, of course, $\boldsymbol{n} \to \infty$ as $n \to \infty$. Then $\{D_{\boldsymbol{n}}(a)\}_n \sim_{\rm sLT} a \otimes 1$.*

Proof. The proof is organized in two steps: we first show by induction on d that the thesis holds if a is continuous; then, by using an approximation argument, we show that it holds even in the case where a is any Riemann-integrable function. As we shall see, the approximation argument heavily relies on the Riemann-integrability of a.

1. We prove by induction on d that, if $a \in C([0,1]^d)$, then

$$D_{\boldsymbol{n}}(a) = LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,1) + R_{\boldsymbol{n},\boldsymbol{m}} + N_{\boldsymbol{n},\boldsymbol{m}}, \qquad \text{rank}(R_{\boldsymbol{n},\boldsymbol{m}}) \leq N(\boldsymbol{n}) \sum_{i=1}^d \frac{m_i}{n_i}, \qquad \|N_{\boldsymbol{n},\boldsymbol{m}}\| \leq \sum_{i=1}^d \omega_a\Big(\frac{1}{m_i} + \frac{m_i}{n_i}\Big), \tag{4.20}$$

where ω_a is the modulus of continuity of a:

$$\omega_a : (0,\infty) \to \mathbb{R}, \qquad \omega_a(\delta) = \max_{\substack{\mathbf{x},\mathbf{y} \in [0,1]^d \\ \|\mathbf{x}-\mathbf{y}\| \leq \delta}} |a(\mathbf{x}) - a(\mathbf{y})|.$$

Since $\omega_a(\delta) \to 0$ as $\delta \to 0$, eq. (4.20) implies that the thesis holds for any continuous function $a \in C([0,1]^d)$; it suffices to choose, in Definition 4.3, an index $n_{\boldsymbol{m}}$ such that $\boldsymbol{n} \geq \boldsymbol{m}^2$ for $n \geq n_{\boldsymbol{m}}$, and to take $c(\boldsymbol{m}) = \sum_{i=1}^d (1/m_i),\ \omega(\boldsymbol{m}) = \sum_{i=1}^d \omega_a(2/m_i)$.

In the case $d = 1$, $LT_n^m(a,1)$ is the $n \times n$ diagonal matrix given by

$$LT_n^m(a,1) = D_m(a) \otimes I_{\lfloor n/m \rfloor} \ \oplus\ O_{n \bmod m} = a(1/m) I_{\lfloor n/m \rfloor} \oplus a(2/m) I_{\lfloor n/m \rfloor} \oplus \dots \oplus a(1) I_{\lfloor n/m \rfloor} \oplus O_{n \bmod m}.$$

For every $i = 1, \dots, m\lfloor n/m \rfloor$, let $j = j(i)$ be the index in $\{1, \dots, m\}$ such that $(j-1)\lfloor n/m \rfloor + 1 \leq i \leq j\lfloor n/m \rfloor$. Then,

$$|[LT_n^m(a,1)]_{ii} - [D_n(a)]_{ii}| = |a(j/m) - a(i/n)| \leq \omega_a(1/m + m/n),$$

because

$$\left|\frac{j}{m} - \frac{i}{n}\right| \leq \frac{j}{m} - \frac{(j-1)\lfloor n/m \rfloor}{n} \leq \frac{j}{m} - \frac{(j-1)(n/m-1)}{n} = \frac{1}{m} + \frac{j-1}{n} \leq \frac{1}{m} + \frac{m}{n}. \tag{4.21}$$

Define the following $n \times n$ diagonal matrices:

$$\tilde{D}_{n,m}(a) = \operatorname*{diag}_{i=1,\dots,m\lfloor n/m \rfloor} a(i/n) \ \oplus\ O_{n \bmod m}, \qquad \hat{D}_{n,m}(a) = O_{m\lfloor n/m \rfloor} \ \oplus \operatorname*{diag}_{i=m\lfloor n/m \rfloor+1,\dots,n} a(i/n).$$

Then,

$$D_n(a) - LT_n^m(a,1) = \hat{D}_{n,m}(a) + \tilde{D}_{n,m}(a) - LT_n^m(a,1) = R_{n,m} + N_{n,m},$$

where $R_{n,m} = \hat{D}_{n,m}(a)$ and $N_{n,m} = \tilde{D}_{n,m}(a) - LT_n^m(a,1)$ satisfy

$$\text{rank}(R_{n,m}) \leq n \bmod m \leq m, \qquad \|N_{n,m}\| = \max_{i=1,\dots,m\lfloor n/m \rfloor} |[LT_n^m(a,1)]_{ii} - [D_n(a)]_{ii}| \leq \omega_a(1/m + m/n).$$

This shows that (4.20) holds for $d = 1$.

In the case $d>1$, $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,1)$ is the $N(\boldsymbol{n})\times N(\boldsymbol{n})$ diagonal matrix given by

$$LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,1)=\operatorname*{diag}_{j_1=1,\ldots,m_1} I_{\lfloor n_1/m_1\rfloor}\otimes LT_{n_2,\ldots,n_d}^{m_2,\ldots,m_d}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big),1\Big)\ \oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d}, \tag{4.22}$$

where, for any $\hat{x}_1\in[0,1]$, $a(\hat{x}_1,\cdot):[0,1]^{d-1}\to\mathbb{C}$ is the function $(x_2,\ldots,x_d)\mapsto a(\hat{x}_1,x_2,\ldots,x_d)$. For every $j_1=1,\ldots,m_1$ and every $i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor$, by induction hypothesis we have

$$\begin{aligned}&LT_{n_2,\ldots,n_d}^{m_2,\ldots,m_d}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big),1\Big)-D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big)\\&=\Big[D_{n_2,\ldots,n_d}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)-D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big)\Big]+R_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]}+N_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]},\end{aligned}$$

where

$$\operatorname{rank}(R_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]})\le n_2\cdots n_d\sum_{k=2}^{d}\frac{m_k}{n_k},\qquad \|N_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]}\|\le\sum_{k=2}^{d}\omega_{a(j_1/m_1,\cdot)}\Big(\frac{1}{m_k}+\frac{m_k}{n_k}\Big)\le\sum_{k=2}^{d}\omega_a\Big(\frac{1}{m_k}+\frac{m_k}{n_k}\Big).$$

Moreover,

$$\Big\|D_{n_2,\ldots,n_d}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)-D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big)\Big\|\le\omega_a\Big(\frac{1}{m_1}+\frac{m_1}{n_1}\Big),$$

because one can show as in (4.21) that

$$\Big|\frac{j_1}{m_1}-\frac{i_1}{n_1}\Big|\le\frac{1}{m_1}+\frac{m_1}{n_1}.$$

Thus,

$$\begin{aligned}&LT_{n_2,\ldots,n_d}^{m_2,\ldots,m_d}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big),1\Big)-D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big)=R_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]}+N_{\boldsymbol{n},\boldsymbol{m}}^{[j_1/m_1,\,i_1/n_1]}\\&\operatorname{rank}(R_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]})\le n_2\cdots n_d\sum_{k=2}^{d}\frac{m_k}{n_k},\qquad \|N_{\boldsymbol{n},\boldsymbol{m}}^{[j_1/m_1,\,i_1/n_1]}\|\le\sum_{k=1}^{d}\omega_a\Big(\frac{1}{m_k}+\frac{m_k}{n_k}\Big).\end{aligned} \tag{4.23}$$

Now we observe that the diagonal matrices $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,1)$ and $D_{\boldsymbol{n}}(a)$ can be written as

$$\begin{aligned}LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,1)&=\operatorname*{diag}_{j_1=1,\ldots,m_1}\Big[\operatorname*{diag}_{i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor}LT_{n_2,\ldots,n_d}^{m_2,\ldots,m_d}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big),1\Big)\Big]\ \oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d},\\D_{\boldsymbol{n}}(a)&=\operatorname*{diag}_{j_1=1,\ldots,m_1}\Big[\operatorname*{diag}_{i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor}D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big)\Big]\ \oplus\ \operatorname*{diag}_{i_1=m_1\lfloor n_1/m_1\rfloor+1,\ldots,n_1}D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big);\end{aligned}$$

see (4.22) and (2.3). Hence,

$$\begin{aligned}D_{\boldsymbol{n}}(a)-LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,1)&=\operatorname*{diag}_{j_1=1,\ldots,m_1}\Big[\operatorname*{diag}_{i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor}\Big[D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big)-LT_{n_2,\ldots,n_d}^{m_2,\ldots,m_d}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big),1\Big)\Big]\Big]\\&\quad\oplus\operatorname*{diag}_{i_1=m_1\lfloor n_1/m_1\rfloor+1,\ldots,n_1}D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big)\\&=\operatorname*{diag}_{j_1=1,\ldots,m_1}\Big[\operatorname*{diag}_{i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor}\Big[-R_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]}-N_{\boldsymbol{n},\boldsymbol{m}}^{[j_1/m_1,\,i_1/n_1]}\Big]\Big]\\&\quad\oplus\operatorname*{diag}_{i_1=m_1\lfloor n_1/m_1\rfloor+1,\ldots,n_1}D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big)\\&=R_{\boldsymbol{n},\boldsymbol{m}}+N_{\boldsymbol{n},\boldsymbol{m}},\end{aligned}$$

where

$$\begin{aligned}R_{\boldsymbol{n},\boldsymbol{m}}&=\operatorname*{diag}_{j_1=1,\ldots,m_1}\Big[\operatorname*{diag}_{i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor}-R_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]}\Big]\ \oplus\ \operatorname*{diag}_{i_1=m_1\lfloor n_1/m_1\rfloor+1,\ldots,n_1}D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big),\\N_{\boldsymbol{n},\boldsymbol{m}}&=\operatorname*{diag}_{j_1=1,\ldots,m_1}\Big[\operatorname*{diag}_{i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor}-N_{\boldsymbol{n},\boldsymbol{m}}^{[j_1/m_1,\,i_1/n_1]}\Big]\ \oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d}.\end{aligned}$$

By (4.23) and (2.20), (2.22), we have

$$\operatorname{rank}(R_{\boldsymbol{n},\boldsymbol{m}}) \le m_1 \left\lfloor \frac{n_1}{m_1} \right\rfloor n_2 \cdots n_d \sum_{k=2}^{d} \frac{m_k}{n_k} + (n_1 \bmod m_1) n_2 \cdots n_d \le n_1 n_2 \cdots n_d \sum_{k=2}^{d} \frac{m_k}{n_k} + m_1 n_2 \cdots n_d = N(\boldsymbol{n}) \sum_{k=1}^{d} \frac{m_k}{n_k},$$

$$\|N_{\boldsymbol{n},\boldsymbol{m}}\| \le \sum_{k=1}^{d} \omega_a\Big(\frac{1}{m_k} + \frac{m_k}{n_k}\Big),$$

and (4.20) is proved.

2. Let $a : [0,1]^d \to \mathbb{C}$ be any Riemann-integrable function. Take any sequence of continuous functions $a_m : [0,1]^d \to \mathbb{C}$ such that $a_m \to a$ in $L^1([0,1]^d)$. Note that such a sequence exists because $C([0,1]^d)$ is dense in $L^1([0,1]^d)$; see [39]. By the first part of the proof, $\{D_{\boldsymbol{n}}(a_m)\}_n \sim_{\rm sLT} a_m \otimes 1$. Hence, for each m and each $\boldsymbol{h} \in \mathbb{N}^d$ there is $n_{m,\boldsymbol{h}}$ such that, for $n \ge n_{m,\boldsymbol{h}}$,

$$D_{\boldsymbol{n}}(a_m) = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m, 1) + R_{\boldsymbol{n},m,\boldsymbol{h}} + N_{\boldsymbol{n},m,\boldsymbol{h}},$$
$$\operatorname{rank}(R_{\boldsymbol{n},m,\boldsymbol{h}}) \le c(m,\boldsymbol{h}) N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m,\boldsymbol{h}}\| \le \omega(m,\boldsymbol{h}),$$

where

$$\lim_{\boldsymbol{h}\to\infty} c(m,\boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} \omega(m,\boldsymbol{h}) = 0.$$

Moreover, $\{\{D_{\boldsymbol{n}}(a_m)\}_n\}_m$ is an a.c.s. for $\{D_{\boldsymbol{n}}(a)\}_n$. Indeed,

$$\|D_{\boldsymbol{n}}(a) - D_{\boldsymbol{n}}(a_m)\|_1 = \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}} \left| a\Big(\frac{\boldsymbol{j}}{\boldsymbol{n}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{n}}\Big) \right| = N(\boldsymbol{n}) \, \frac{1}{N(\boldsymbol{n})} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}} \left| a\Big(\frac{\boldsymbol{j}}{\boldsymbol{n}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{n}}\Big) \right|.$$

By the Riemann-integrability of $|a - a_m|$, which follows from the Riemann-integrability of a and a_m, and by the fact that $a_m \to a$ in $L^1([0,1]^d)$, the quantity

$$\epsilon(m,\boldsymbol{n}) = \frac{1}{N(\boldsymbol{n})} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}} \left| a\Big(\frac{\boldsymbol{j}}{\boldsymbol{n}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{n}}\Big) \right|$$

satisfies

$$\lim_{m\to\infty} \lim_{\boldsymbol{n}\to\infty} \epsilon(m,\boldsymbol{n}) = \lim_{m\to\infty} \int_{[0,1]^d} |a(\mathbf{x}) - a_m(\mathbf{x})| d\mathbf{x} = \lim_{m\to\infty} \|a - a_m\|_{L^1} = 0.$$

By Corollary 3.3, this implies that $\{\{D_{\boldsymbol{n}}(a_m)\}_n\}_m$ is an a.c.s. for $\{D_{\boldsymbol{n}}(a)\}_n$. Thus, for every m there exists n_m such that, for $n \ge n_m$,

$$D_{\boldsymbol{n}}(a) = D_{\boldsymbol{n}}(a_m) + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$
$$\operatorname{rank}(R_{\boldsymbol{n},m}) \le c(m) N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \le \omega(m),$$

where

$$\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0.$$

It follows that, for every m, every $\boldsymbol{h} \in \mathbb{N}^d$ and every $n \ge \max(n_m, n_{m,\boldsymbol{h}})$,

$$D_{\boldsymbol{n}}(a) = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,1) + \big[LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m,1) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,1)\big] + (R_{\boldsymbol{n},m} + R_{\boldsymbol{n},m,\boldsymbol{h}}) + (N_{\boldsymbol{n},m} + N_{\boldsymbol{n},m,\boldsymbol{h}}),$$
$$\operatorname{rank}(R_{\boldsymbol{n},m} + R_{\boldsymbol{n},m,\boldsymbol{h}}) \le (c(m) + c(m,\boldsymbol{h})) N(\boldsymbol{n}),$$
$$\|N_{\boldsymbol{n},m} + N_{\boldsymbol{n},m,\boldsymbol{h}}\| \le \omega(m) + \omega(m,\boldsymbol{h}),$$
$$\big\|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m,1) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,1)\big\|_1 \le \frac{N(\boldsymbol{n})}{N(\boldsymbol{h})} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}} \left| a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) \right| = \epsilon(m,\boldsymbol{h}) N(\boldsymbol{n}),$$

where in the last inequality we used (4.8) and the identity $T_{\lfloor \boldsymbol{n}/\boldsymbol{h} \rfloor}(1) = I_{N(\lfloor \boldsymbol{n}/\boldsymbol{h} \rfloor)}$. For every $\boldsymbol{h} \in \mathbb{N}^d$, choose $m(\boldsymbol{h})$ such that $m(\boldsymbol{h}) \to \infty$ when $\boldsymbol{h} \to \infty$ and

$$\lim_{\boldsymbol{h}\to\infty} \epsilon(m(\boldsymbol{h}),\boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} c(m(\boldsymbol{h}),\boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} \omega(m(\boldsymbol{h}),\boldsymbol{h}) = 0.$$

An explicit construction of such a function $m(\boldsymbol{h})$ has been given in Lemma 4.1 (apply the lemma with $x(m,\boldsymbol{h}) = \epsilon(m,\boldsymbol{h}) + c(m,\boldsymbol{h}) + \omega(m,\boldsymbol{h})$ and $\xi = 0$). Then, for every $\boldsymbol{h} \in \mathbb{N}^d$ and every $n \geq \max(n_{m(\boldsymbol{h})}, n_{m(\boldsymbol{h}),\boldsymbol{h}})$,

$$D_{\boldsymbol{n}}(a) = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,1) + \left[LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_{m(\boldsymbol{h})},1) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,1)\right] + (R_{\boldsymbol{n},m(\boldsymbol{h})} + R_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}) + (N_{\boldsymbol{n},m(\boldsymbol{h})} + N_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}),$$
$$\text{rank}(R_{\boldsymbol{n},m(\boldsymbol{h})} + R_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}) \leq (c(m(\boldsymbol{h})) + c(m(\boldsymbol{h}),\boldsymbol{h}))N(\boldsymbol{n}),$$
$$\|N_{\boldsymbol{n},m(\boldsymbol{h})} + N_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}\| \leq \omega(m(\boldsymbol{h})) + \omega(m(\boldsymbol{h}),\boldsymbol{h}),$$
$$\left\|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_{m(\boldsymbol{h})},1) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,1)\right\|_1 \leq \epsilon(m(\boldsymbol{h}),\boldsymbol{h})N(\boldsymbol{n}).$$

The application of Lemma 3.4 allows one to decompose $LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_{m(\boldsymbol{h})},1) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,1)$ as the sum of a small-rank term $\hat{R}_{\boldsymbol{n},\boldsymbol{h}}$, with rank bounded from above by $\sqrt{\epsilon(m(\boldsymbol{h}),\boldsymbol{h})}\,N(\boldsymbol{n})$, plus a small-norm term $\hat{N}_{\boldsymbol{n},\boldsymbol{h}}$, with norm bounded from above by $\sqrt{\epsilon(m(\boldsymbol{h}),\boldsymbol{h})}$. This concludes the proof. □

Theorem 4.6. *Let $f \in L^1([-\pi,\pi]^d)$ and consider the sequence of multilevel Toeplitz matrices $\{T_{\boldsymbol{n}}(f)\}_n$, where, of course, $\boldsymbol{n} \to \infty$ as $n \to \infty$. Then $\{T_{\boldsymbol{n}}(f)\}_n \sim_{\rm LT} 1 \otimes f$.*

Proof. The proof is organized in three steps: we first show by induction on d that the thesis holds if f is a separable d-variate trigonometric polynomial; then, by linearity, we show that it also holds if f is an arbitrary d-variate trigonometric polynomial; finally, using an approximation argument, we prove the theorem under the sole assumption that $f \in L^1([-\pi,\pi]^d)$.

1. We show by induction on d that, if f is a separable d-variate trigonometric polynomial, say $f = f_1 \otimes \cdots \otimes f_d$ with $f_1,\ldots,f_d$ univariate trigonometric polynomials, then

$$T_{\boldsymbol{n}}(f) = LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1,f) + R_{\boldsymbol{n},\boldsymbol{m}}, \qquad \text{rank}(R_{\boldsymbol{n},\boldsymbol{m}}) \leq N(\boldsymbol{n})\sum_{i=1}^d \frac{(2r_i+1)m_i}{n_i}, \tag{4.24}$$

where r_i is the degree of f_i. From (4.24), it follows that the theorem holds for any separable trigonometric polynomial f; it suffices to choose, in Definition 4.3, an index $n_{\boldsymbol{m}}$ such that $\boldsymbol{n} \geq \boldsymbol{m}^2$ for $n \geq n_{\boldsymbol{m}}$, and to take $c(\boldsymbol{m}) = \sum_{i=1}^d (2r_i+1)/m_i,\ \omega(\boldsymbol{m}) = 0$.

In the case $d = 1$, let $f(\theta) = \sum_{j=-r}^r f_j \mathrm{e}^{\mathrm{i}j\theta}$. Then,

$$LT_n^m(1,f) = I_m \otimes T_{\lfloor n/m\rfloor}(f) \ \oplus\ O_{n \bmod m} = T_{\lfloor n/m\rfloor}(f) \oplus \cdots \oplus T_{\lfloor n/m\rfloor}(f) \oplus O_{n \bmod m}.$$

Looking carefully at the structure of $T_n(f)$ and $LT_n^m(1,f)$, we see that the number of nonzero rows of the difference $T_n(f) - LT_n^m(1,f)$ is at most $2rm - r + (n \bmod m)$. Hence,

$$T_n(f) = LT_n^m(1,f) + R_{n,m}, \qquad \text{rank}(R_{n,m}) \leq 2rm - r + (n \bmod m) \leq (2r+1)m, \tag{4.25}$$

and so (4.24) holds for $d = 1$.

In the case $d > 1$, let $f = f_1 \otimes \cdots \otimes f_d$ with $f_1,\ldots,f_d$ univariate trigonometric polynomials of degrees $r_1,\ldots,r_d$, respectively. By induction hypothesis we have

$$LT_{n_2,\ldots,n_d}^{m_2,\ldots,m_d}(1, f_2 \otimes \cdots \otimes f_d) = T_{n_2,\ldots,n_d}(f_2 \otimes \cdots \otimes f_d) + R_{n_2,\ldots,n_d,m_2,\ldots,m_d},$$
$$\text{rank}(R_{n_2,\ldots,n_d,m_2,\ldots,m_d}) \leq n_2 \cdots n_d \sum_{i=2}^d \frac{(2r_i+1)m_i}{n_i}.$$

From the definition of $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1,f)$ and the properties of tensor products and direct sums (see Section 2.4.1), we obtain

$$\begin{aligned}
LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1,f) &= \underset{j_1=1,\ldots,m_1}{\text{diag}} T_{\lfloor n_1/m_1\rfloor}(f_1) \otimes LT_{n_2,\ldots,n_d}^{m_2,\ldots,m_d}(1, f_2 \otimes \cdots \otimes f_d) \ \oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d} \\
&= \Big[\underset{j_1=1,\ldots,m_1}{\text{diag}} T_{\lfloor n_1/m_1\rfloor}(f_1)\Big] \otimes \big[T_{n_2,\ldots,n_d}(f_2 \otimes \cdots \otimes f_d) + R_{n_2,\ldots,n_d,m_2,\ldots,m_d}\big] \ \oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d} \\
&= \Big[\underset{j_1=1,\ldots,m_1}{\text{diag}} T_{\lfloor n_1/m_1\rfloor}(f_1) \ \oplus\ O_{n_1 \bmod m_1}\Big] \otimes \big[T_{n_2,\ldots,n_d}(f_2 \otimes \cdots \otimes f_d) + R_{n_2,\ldots,n_d,m_2,\ldots,m_d}\big] \\
&= LT_{n_1}^{m_1}(1,f_1) \otimes \big[T_{n_2,\ldots,n_d}(f_2 \otimes \cdots \otimes f_d) + R_{n_2,\ldots,n_d,m_2,\ldots,m_d}\big] \\
&= LT_{n_1}^{m_1}(1,f_1) \otimes T_{n_2,\ldots,n_d}(f_2 \otimes \cdots \otimes f_d) + \tilde{R}_{n_1,\ldots,n_d,m_1,\ldots,m_d},
\end{aligned}$$

where $\tilde{R}_{n_1,\ldots,n_d,m_1,\ldots,m_d} = LT_{n_1}^{m_1}(1,f_1) \otimes R_{n_1,\ldots,n_d,m_1,\ldots,m_d}$ satisfies

$$\text{rank}(\tilde{R}_{n_1,\ldots,n_d,m_1,\ldots,m_d}) \leq N(\boldsymbol{n})\sum_{i=2}^d \frac{(2r_i+1)m_i}{n_i}.$$

Using (4.25), we can decompose $LT_{n_1}^{m_1}(1,f_1)$ as the sum of $T_{n_1}(f_1)$ plus a small-rank matrix R_{n_1,m_1}, whose rank is bounded by $(2r_1+1)m_1$. Recalling Lemma 2.10, we arrive at

$$LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1,f) = (T_{n_1}(f_1) + R_{n_1,m_1}) \otimes T_{n_2,\ldots,n_d}(f_2 \otimes \cdots \otimes f_d) + \tilde{R}_{n_1,\ldots,n_d,m_1,\ldots,m_d} = T_{\boldsymbol{n}}(f) + R_{\boldsymbol{n},\boldsymbol{m}},$$

where $R_{\boldsymbol{n},\boldsymbol{m}} = R_{n_1,m_1} \otimes T_{n_2,\ldots,n_d}(f_2 \otimes \cdots \otimes f_d) + \tilde{R}_{n_1,\ldots,n_d,m_1,\ldots,m_d}$ satisfies

$$\mathrm{rank}(R_{\boldsymbol{n},\boldsymbol{m}}) \le (2r_1+1)m_1 n_2 \cdots n_d + N(\boldsymbol{n}) \sum_{i=2}^{d} \frac{(2r_i+1)m_i}{n_i} = N(\boldsymbol{n}) \sum_{i=1}^{d} \frac{(2r_i+1)m_i}{n_i}.$$

This completes the proof of (4.24).

2. Let f be any d-variate trigonometric polynomial. By definition, f is a finite linear combination of the Fourier frequencies $\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}$, $\boldsymbol{j} \in \mathbb{Z}$, and so we can write $f(\boldsymbol{\theta}) = \sum_{\boldsymbol{j}=-\boldsymbol{r}}^{\boldsymbol{r}} f_{\boldsymbol{j}} \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}$ for some separable trigonometric polynomials $f_{\boldsymbol{j}} \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}$. By linearity,

$$T_{\boldsymbol{n}}(f) = \sum_{\boldsymbol{j}=-\boldsymbol{r}}^{\boldsymbol{r}} f_{\boldsymbol{j}} T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}), \qquad LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1,f) = \sum_{\boldsymbol{j}=-\boldsymbol{r}}^{\boldsymbol{r}} f_{\boldsymbol{j}} LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1,\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}).$$

By the first part of the proof, $\{T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})\}_n \sim_{\mathrm{LT}} 1 \otimes \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}$, hence $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1,\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})\}_n\}_{\boldsymbol{m}\in\mathbb{N}^d}$ is an a.c.s. for $\{T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})\}_n$. It follows that $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1,f)\}_n\}_{\boldsymbol{m}\in\mathbb{N}^d}$ is an a.c.s. for $\{T_{\boldsymbol{n}}(f)\}_n$; see Remark 3.6. Thus, $\{T_{\boldsymbol{n}}(f)\}_n \sim_{\mathrm{LT}} 1 \otimes f$ for every trigonometric polynomial f.

3. Let $f \in L^1([-\pi,\pi]^d)$. Since the set of d-variate trigonometric polynomials is dense in $L^1([-\pi,\pi]^d)$, there is a sequence $\{f_m\}$ of d-variate trigonometric polynomials such that $f_m \to f$ in $L^1([-\pi,\pi]^d)$. By the second part of the proof, $\{T_{\boldsymbol{n}}(f_m)\}_n \sim_{\mathrm{LT}} 1 \otimes f_m$. Hence, for each m and each $\boldsymbol{h} \in \mathbb{N}^d$ there is $n_{m,\boldsymbol{h}}$ such that, for $n \ge n_{m,\boldsymbol{h}}$,

$$T_{\boldsymbol{n}}(f_m) = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f_m) + R_{\boldsymbol{n},m,\boldsymbol{h}} + N_{\boldsymbol{n},m,\boldsymbol{h}},$$
$$\mathrm{rank}(R_{\boldsymbol{n},m,\boldsymbol{h}}) \le c(m,\boldsymbol{h}) N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m,\boldsymbol{h}}\| \le \omega(m,\boldsymbol{h}),$$

where

$$\lim_{\boldsymbol{h}\to\infty} c(m,\boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} \omega(m,\boldsymbol{h}) = 0.$$

Moreover, by Theorem 2.8,

$$\|T_{\boldsymbol{n}}(f) - T_{\boldsymbol{n}}(f_m)\|_1 = \|T_{\boldsymbol{n}}(f - f_m)\|_1 \le N(\boldsymbol{n}) \|f - f_m\|_{L^1}$$

and so $\{\{T_{\boldsymbol{n}}(f_m)\}_n\}_m$ is an a.c.s. for $\{T_{\boldsymbol{n}}(f)\}_n$ by Corollary 3.3: for every m there exists n_m such that, for $n \ge n_m$,

$$T_{\boldsymbol{n}}(f) = T_{\boldsymbol{n}}(f_m) + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$
$$\mathrm{rank}(R_{\boldsymbol{n},m}) \le c(m) N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \le \omega(m),$$

where

$$\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0.$$

It follows that, for every m, every $\boldsymbol{h} \in \mathbb{N}^d$ and every $n \ge \max(n_m, n_{m,\boldsymbol{h}})$,

$$T_{\boldsymbol{n}}(f) = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f) + \left[LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f_m) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f)\right] + (R_{\boldsymbol{n},m} + R_{\boldsymbol{n},m,\boldsymbol{h}}) + (N_{\boldsymbol{n},m} + N_{\boldsymbol{n},m,\boldsymbol{h}}),$$
$$\mathrm{rank}(R_{\boldsymbol{n},m} + R_{\boldsymbol{n},m,\boldsymbol{h}}) \le (c(m) + c(m,\boldsymbol{h})) N(\boldsymbol{n}),$$
$$\|N_{\boldsymbol{n},m} + N_{\boldsymbol{n},m,\boldsymbol{h}}\| \le \omega(m) + \omega(m,\boldsymbol{h}),$$
$$\left\|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f_m) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f)\right\|_1 = \left\|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f_m - f)\right\|_1 \le N(\boldsymbol{n}) \|f - f_m\|_{L^1}.$$

Choose, for every $\boldsymbol{h} \in \mathbb{N}^d$, a $m(\boldsymbol{h})$ such that $m(\boldsymbol{h}) \to \infty$ when $\boldsymbol{h} \to \infty$ and

$$\lim_{\boldsymbol{h}\to\infty} c(m(\boldsymbol{h}),\boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} \omega(m(\boldsymbol{h}),\boldsymbol{h}) = 0.$$

An explicit construction of such a function $m(\boldsymbol{h})$ is given in Lemma 4.1 (apply the lemma with $x(m,\boldsymbol{h}) = c(m,\boldsymbol{h}) + \omega(m,\boldsymbol{h})$ and $\xi = 0$). Then, for every $\boldsymbol{h} \in \mathbb{N}^d$ and every $n \ge \max(n_{m(\boldsymbol{h})}, n_{m(\boldsymbol{h}),\boldsymbol{h}})$,

$$T_{\boldsymbol{n}}(f) = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f) + \left[LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f_{m(\boldsymbol{h})}) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f)\right] + (R_{\boldsymbol{n},m(\boldsymbol{h})} + R_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}) + (N_{\boldsymbol{n},m(\boldsymbol{h})} + N_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}),$$
$$\mathrm{rank}(R_{\boldsymbol{n},m(\boldsymbol{h})} + R_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}) \le (c(m(\boldsymbol{h})) + c(m(\boldsymbol{h}),\boldsymbol{h})) N(\boldsymbol{n}),$$
$$\|N_{\boldsymbol{n},m(\boldsymbol{h})} + N_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}\| \le \omega(m(\boldsymbol{h})) + \omega(m(\boldsymbol{h}),\boldsymbol{h}),$$
$$\left\|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f_{m(\boldsymbol{h})}) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1,f)\right\|_1 \le \|f_{m(\boldsymbol{h})} - f\|_{L^1} N(\boldsymbol{n}).$$

The application of Lemma 3.4 allows one to decompose $LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1, f_{m(\boldsymbol{h})}) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(1, f)$ as the sum of a small-rank term $\hat{R}_{\boldsymbol{n},\boldsymbol{h}}$, with rank bounded from above by $\sqrt{\|f_{m(\boldsymbol{h})} - f\|_{L^1}}\, N(\boldsymbol{n})$, plus a small-norm term $\hat{N}_{\boldsymbol{n},\boldsymbol{h}}$, with norm bounded from above by $\sqrt{\|f_{m(\boldsymbol{h})} - f\|_{L^1}}$. This concludes the proof. □

It follows from Theorem 4.6 that $\{T_{\boldsymbol{n}}(f)\}_n \sim_{\rm sLT} 1 \otimes f$ whenever $f \in L^1([-\pi,\pi]^d)$ is separable.

4.2.2 Properties and characterizations of LT and sLT sequences

We begin with a basic spectral result for LT sequences, which will be used in the proof of Theorem 4.7. Since it will be generalized afterwards, in the more general context of GLT sequences, it is not necessary to remember it.

Lemma 4.2. *If $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$ then $\{A_{\boldsymbol{n}}\}_n \sim_\sigma a \otimes f$.*

Proof. Take any sequence of multi-indices $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m \subseteq \mathbb{N}^d$ such that $\boldsymbol{m} \to \infty$ as $m \to \infty$. By Definition 4.3, $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. By Theorem 4.2, we have $\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)\}_n \sim_\sigma \phi_{\boldsymbol{m}}$ and $\phi_{\boldsymbol{m}} \to \phi_{[|a\otimes f|]}$ pointwise over $C_c(\mathbb{R})$, where we recall that $\phi_{[g]}$ is defined in (2.6). Therefore, Theorem 3.1 gives $\{A_{\boldsymbol{n}}\}_n \sim_\sigma a \otimes f$. □

As a consequence of Lemma 4.2 and Proposition 3.3, every LT sequence is s.u. (see Definition 3.2). We now show, under mild assumptions, that the product of LT sequences is again a LT sequence with symbol given by the product of the symbols.

Theorem 4.7. *Suppose that*

$$\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f, \qquad \{\tilde{A}_{\boldsymbol{n}}\}_n \sim_{\rm LT} \tilde{a} \otimes \tilde{f},$$

where $f \in L^p([-\pi,\pi]^d)$, $\tilde{f} \in L^q([-\pi,\pi]^d)$, and p, q are conjugate exponents ($1 \le p, q \le \infty$). Then

$$\{A_{\boldsymbol{n}}\tilde{A}_{\boldsymbol{n}}\}_n \sim_{\rm LT} a\tilde{a} \otimes f\tilde{f}.$$

Proof. As noted before the statement of the theorem, Lemma 4.2 and Proposition 3.3 imply that every LT sequence is s.u., so in particular $\{A_{\boldsymbol{n}}\}_n$ and $\{\tilde{A}_{\boldsymbol{n}}\}_n$ are s.u. Since $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)\}_n\}_{\boldsymbol{m}\in\mathbb{N}^d}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ and $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde{a}, \tilde{f})\}_n\}_{\boldsymbol{m}\in\mathbb{N}^d}$ is an a.c.s. for $\{\tilde{A}_{\boldsymbol{n}}\}_n$, the product $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f) LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde{a}, \tilde{f})\}_n\}_{\boldsymbol{m}\in\mathbb{N}^d}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\tilde{A}_{\boldsymbol{n}}\}_n$; see Remark 3.6. The thesis now follows from Definition 4.3 and Proposition 4.3 (see in particular eq. (4.11) in Proposition 4.3). □

As a consequence of Theorem 4.7 and Theorems 4.5–4.6, we immediately obtain the following result.

Theorem 4.8. *Let $a : [0,1]^d \to \mathbb{C}$ be Riemann-integrable, let $f \in L^1([-\pi,\pi]^d)$, and consider the sequence of matrices $\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n$, where, of course, $\boldsymbol{n} \to \infty$ as $n \to \infty$. Then $\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n \sim_{\rm LT} a \otimes f$.*

Theorem 4.8 shows that, for any a, f as in Definition 4.3, there always exists a matrix-sequence $\{A_{\boldsymbol{n}}\}_n$ such that $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$. Indeed, it suffices to take $A_{\boldsymbol{n}} = D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)$. Theorem 4.9 shows that the sequences of the form $\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n$ play a central role in the world of LT sequences. Indeed, $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$ if and only if $A_{\boldsymbol{n}}$ equals $D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)$ up to a small-rank plus small-norm correction. In fact, any LT sequence $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$ admits the fixed matrix-sequence $\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n$ as an a.c.s., and, vice versa, any sequence $\{A_{\boldsymbol{n}}\}_n$ admitting $\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n$ as an a.c.s. is a LT sequence with symbol $a \otimes f$. Moreover, any LT sequence $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$ also admits an a.c.s. of the form $\{\{D_{\boldsymbol{n}}(a_m)T_{\boldsymbol{n}}(f_m)\}_n\}_m$, with a_m continuous and f_m trigonometric polynomial; as one could guess, a_m will be chosen as an approximation of a, converging to a for $m \to \infty$, while f_m will be chosen as an approximation of f, converging to f for $m \to \infty$.

Theorem 4.9 (characterizations of LT sequences). *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, let $a : [0,1]^d \to \mathbb{C}$ be a Riemann-integrable function and let $f \in L^1([-\pi,\pi]^d)$. Then, the following conditions are equivalent.*

1. *$\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$.*
2. *For all sequences $\{a_m\}_m$, $\{f_m\}_m$, $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ with the following properties:*
   - *$a_m : [0,1]^d \to \mathbb{C}$ is Riemann-integrable and $a_m \to a$ in $L^1([0,1]^d)$;*
   - *$f_m \in L^1([-\pi,\pi]^d)$ and $f_m \to f$ in $L^1([-\pi,\pi]^d)$;*
   - *$\{A_{\boldsymbol{n}}^{(m)}\}_n \sim_{\rm LT} a_m \otimes f_m$;*

   it holds that $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.
3. *There exist sequences $\{a_m\}_m$, $\{f_m\}_m$ such that:*
   - *$a_m : [0,1]^d \to \mathbb{C}$ is continuous, $\|a_m\|_\infty \le \|a\|_{L^\infty}$ for all m and $a_m \to a$ a.e.;*
   - *$f_m : [-\pi,\pi]^d \to \mathbb{C}$ is a trigonometric polynomial and $f_m \to f$ a.e. and in $L^1([-\pi,\pi]^d)$;*

* $\{\{D_{\boldsymbol{n}}(a_m)T_{\boldsymbol{n}}(f_m)\}_n\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$.

4. *There exist sequences* $\{a_m\}_m$, $\{f_m\}_m$, $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ *such that:*
   * $a_m : [0,1]^d \to \mathbb{C}$ *is Riemann-integrable and* $a_m \to a$ *in* $L^1([0,1]^d)$*;*
   * $f_m \in L^1([-\pi,\pi]^d)$ *and* $f_m \to f$ *in* $L^1([-\pi,\pi]^d)$*;*
   * $\{A_{\boldsymbol{n}}^{(m)}\}_n \sim_{\rm LT} a_m \otimes f_m$ *and* $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$.
5. $\{\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$.
6. *For every* n *we have* $A_{\boldsymbol{n}} = D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f) + R_{\boldsymbol{n}} + N_{\boldsymbol{n}}$*, where* $\lim_{n\to\infty} \frac{\mathrm{rank}(R_{\boldsymbol{n}})}{N(\boldsymbol{n})} = \lim_{n\to\infty} \|N_{\boldsymbol{n}}\| = 0$.

Proof. (1 ⇒ 2) Let $\{a_m\}$, $\{f_m\}$, $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ be sequences with the properties specified in item 2. Since $\{A_{\boldsymbol{n}}^{(m)}\}_n \sim_{\rm LT} a_m \otimes f_m$, for each m and each $\boldsymbol{h} \in \mathbb{N}^d$ there is $n_{m,\boldsymbol{h}}$ such that, for $n \ge n_{m,\boldsymbol{h}}$,

$$A_{\boldsymbol{n}}^{(m)} = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m, f_m) + R_{\boldsymbol{n},m,\boldsymbol{h}} + N_{\boldsymbol{n},m,\boldsymbol{h}},$$
$$\mathrm{rank}(R_{\boldsymbol{n},m,\boldsymbol{h}}) \le c(m,\boldsymbol{h})N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m,\boldsymbol{h}}\| \le \omega(m,\boldsymbol{h}),$$

where

$$\lim_{\boldsymbol{h}\to\infty} c(m,\boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} \omega(m,\boldsymbol{h}) = 0.$$

Moreover, since $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$, for every $\boldsymbol{h} \in \mathbb{N}^d$ there is $n_{\boldsymbol{h}}$ such that, for $n \ge n_{\boldsymbol{h}}$,

$$A_{\boldsymbol{n}} = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a, f) + R_{\boldsymbol{n},\boldsymbol{h}} + N_{\boldsymbol{n},\boldsymbol{h}},$$
$$\mathrm{rank}(R_{\boldsymbol{n},\boldsymbol{h}}) \le c(\boldsymbol{h})N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},\boldsymbol{h}}\| \le \omega(\boldsymbol{h}),$$

where

$$\lim_{\boldsymbol{h}\to\infty} c(\boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} \omega(\boldsymbol{h}) = 0.$$

Hence, for every m, every $\boldsymbol{h} \in \mathbb{N}^d$ and every $n \ge \max(n_{m,\boldsymbol{h}}, n_{\boldsymbol{h}})$,

$$\begin{aligned} A_{\boldsymbol{n}} &= A_{\boldsymbol{n}}^{(m)} + \big[LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m,f_m)\big] + (R_{\boldsymbol{n},\boldsymbol{h}} - R_{\boldsymbol{n},m,\boldsymbol{h}}) + (N_{\boldsymbol{n},\boldsymbol{h}} - N_{\boldsymbol{n},m,\boldsymbol{h}}), \\ \mathrm{rank}(R_{\boldsymbol{n},\boldsymbol{h}} - R_{\boldsymbol{n},m,\boldsymbol{h}}) &\le (c(\boldsymbol{h}) + c(m,\boldsymbol{h}))N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},\boldsymbol{h}} - N_{\boldsymbol{n},m,\boldsymbol{h}}\| \le \omega(\boldsymbol{h}) + \omega(m,\boldsymbol{h}). \end{aligned} \tag{4.26}$$

Thanks to Propositions 4.1–4.2 and to Theorem 2.8, we have

$$\begin{aligned} \|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m,f_m)\|_1 &\le \|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f-f_m)\|_1 + \|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a-a_m,f_m)\|_1 \\ &= \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}} \Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big| \|T_{\lfloor \boldsymbol{n}/\boldsymbol{h}\rfloor}(f-f_m)\|_1 + \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}} \Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big| \|T_{\lfloor \boldsymbol{n}/\boldsymbol{h}\rfloor}(f_m)\|_1 \\ &\le N(\boldsymbol{n})\|a\|_\infty \|f-f_m\|_{L^1} + \|f_m\|_{L^1} \frac{N(\boldsymbol{n})}{N(\boldsymbol{h})} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}} \Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big| \\ &\le \left[\|a\|_\infty \|f-f_m\|_{L^1} + \sup_k \|f_k\|_{L^1} \frac{1}{N(\boldsymbol{h})} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}} \Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big|\right] N(\boldsymbol{n}); \end{aligned} \tag{4.27}$$

note that $\|f_k\|_{L^1}$ is uniformly bounded with respect to k, because $f_k \to f$ in $L^1([-\pi,\pi]^d)$. By the Riemann-integrability of $|a - a_m|$, which follows from the Riemann-integrability of a and a_m, and by the fact that $a_m \to a$ in $L^1([0,1]^d)$ and $f_m \to f$ in $L^1([-\pi,\pi]^d)$, the quantity

$$\varepsilon(m,\boldsymbol{h}) = \|a\|_\infty \|f-f_m\|_{L^1} + \sup_k \|f_k\|_{L^1} \frac{1}{N(\boldsymbol{h})} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}} \Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big| \tag{4.28}$$

satisfies

$$\lim_{m\to\infty} \lim_{\boldsymbol{h}\to\infty} \varepsilon(m,\boldsymbol{h}) = \lim_{m\to\infty} \left(\|a\|_\infty \|f-f_m\|_{L^1} + \sup_k \|f_k\|_{L^1} \int_{[0,1]^d} |a(\mathbf{x}) - a_m(\mathbf{x})| d\mathbf{x}\right) = 0. \tag{4.29}$$

Choose any sequence of multi-indices $\{\boldsymbol{h}(m)\}_m$ such that $\boldsymbol{h}(m) \to \infty$ for $m \to \infty$ and

$$\lim_{m\to\infty} c(m, \boldsymbol{h}(m)) = \lim_{m\to\infty} \omega(m, \boldsymbol{h}(m)) = \lim_{m\to\infty} \varepsilon(m, \boldsymbol{h}(m)) = 0.$$

Then, by (4.26)–(4.27), for every m and every $n \geq \max(n_{m,\boldsymbol{h}(m)}, n_{\boldsymbol{h}(m)})$,

$$A_{\boldsymbol{n}} = A_{\boldsymbol{n}}^{(m)} + \left[LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a, f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a_m, f_m)\right] + (R_{\boldsymbol{n},\boldsymbol{h}(m)} - R_{\boldsymbol{n},m,\boldsymbol{h}(m)}) + (N_{\boldsymbol{n},\boldsymbol{h}(m)} - N_{\boldsymbol{n},m,\boldsymbol{h}(m)}),$$
$$\text{rank}(R_{\boldsymbol{n},\boldsymbol{h}(m)} - R_{\boldsymbol{n},m,\boldsymbol{h}(m)}) \leq [c(\boldsymbol{h}(m)) + c(m, \boldsymbol{h}(m))]\, N(\boldsymbol{n}),$$
$$\|N_{\boldsymbol{n},\boldsymbol{h}(m)} - N_{\boldsymbol{n},m,\boldsymbol{h}(m)}\| \leq \omega(\boldsymbol{h}(m)) + \omega(m, \boldsymbol{h}(m)),$$
$$\left\|LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a, f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a_m, f_m)\right\|_1 \leq \varepsilon(m, \boldsymbol{h}(m))N(\boldsymbol{n}).$$

Using Lemma 3.4, we can decompose $LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a, f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a_m, f_m)$ as the sum of a small-rank term $\hat{R}_{\boldsymbol{n},m}$, with rank bounded from above by $\sqrt{\varepsilon(m, \boldsymbol{h}(m))}\, N(\boldsymbol{n})$, plus a small-norm term $\hat{N}_{\boldsymbol{n},m}$, with norm bounded from above by $\sqrt{\varepsilon(m, \boldsymbol{h}(m))}$. This concludes the proof of the implication 1 ⇒ 2.

(2 ⇒ 3) Since any Riemann-integrable function is bounded, we have $a \in L^\infty([0,1]^d)$. Hence, by the Lusin theorem [39], there exists a sequence of continuous functions $\hat{a}_m : [0,1]^d \to \mathbb{C}$ such that $\|\hat{a}_m\|_\infty \leq \|a\|_{L^\infty}$ for all m and $\hat{a}_m \to a$ in measure. This implies that $\hat{a}_m \to a$ also in $L^1([0,1]^d)$, due to the uniform boundedness of $\|\hat{a}_m\|_\infty$. Thus, there exists a subsequence of $\{\hat{a}_m\}$, say $\{a_m\}$, which converges to a a.e. in $[0,1]^d$. The sequence $\{a_m\}$ satisfies all the properties required in item 3.

Since $f \in L^1([-\pi,\pi]^d)$ and the set of d-variate trigonometric polynomials is dense in $L^1([-\pi,\pi]^d)$, there exists a sequence $\{\hat{f}_m\}$ of d-variate trigonometric polynomials such that $\hat{f}_m \to f$ in $L^1([-\pi,\pi]^d)$. Choosing a subsequence $\{f_m\}$ of $\{\hat{f}_m\}$ which converges to f a.e., $\{f_m\}$ satisfies all the properties required in item 3.

By item 2 and Theorem 4.8, $\{\{D_{\boldsymbol{n}}(a_m)T_{\boldsymbol{n}}(f_m)\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$, and the proof is finished.

(3 ⇒ 4) Obvious. We just recall that, under the assumptions in item 3, $a_m \to a$ in $L^1([0,1]^d)$ by the dominated convergence theorem. Moreover, $\{D_{\boldsymbol{n}}(a_m)T_{\boldsymbol{n}}(f_m)\}_n \sim_{\rm LT} a_m \otimes f_m$ by Theorem 4.8.

(4 ⇒ 1) Since $\{A_{\boldsymbol{n}}^{(m)}\}_n \sim_{\rm LT} a_m \otimes f_m$, for each m and each $\boldsymbol{h} \in \mathbb{N}^d$ there is $n_{m,\boldsymbol{h}}$ such that, for $n \geq n_{m,\boldsymbol{h}}$,

$$A_{\boldsymbol{n}}^{(m)} = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m, f_m) + R_{\boldsymbol{n},m,\boldsymbol{h}} + N_{\boldsymbol{n},m,\boldsymbol{h}},$$
$$\text{rank}(R_{\boldsymbol{n},m,\boldsymbol{h}}) \leq c(m, \boldsymbol{h})N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m,\boldsymbol{h}}\| \leq \omega(m, \boldsymbol{h}),$$

where

$$\lim_{\boldsymbol{h}\to\infty} c(m, \boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} \omega(m, \boldsymbol{h}) = 0.$$

Moreover, since $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$, for every m there exists n_m such that, for $n \geq n_m$,

$$A_{\boldsymbol{n}} = A_{\boldsymbol{n}}^{(m)} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$
$$\text{rank}(R_{\boldsymbol{n},m}) \leq c(m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \leq \omega(m),$$

where

$$\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0.$$

Thus, for every m, every $\boldsymbol{h} \in \mathbb{N}^d$ and every $n \geq \max(n_m, n_{m,\boldsymbol{h}})$,

$$A_{\boldsymbol{n}} = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a, f) + \left[LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m, f_m) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a, f)\right] + (R_{\boldsymbol{n},m} + R_{\boldsymbol{n},m,\boldsymbol{h}}) + (N_{\boldsymbol{n},m} + N_{\boldsymbol{n},m,\boldsymbol{h}}),$$
$$\text{rank}(R_{\boldsymbol{n},m} + R_{\boldsymbol{n},m,\boldsymbol{h}}) \leq (c(m) + c(m, \boldsymbol{h}))N(\boldsymbol{n}),$$
$$\|N_{\boldsymbol{n},m} + N_{\boldsymbol{n},m,\boldsymbol{h}}\| \leq \omega(m) + \omega(m, \boldsymbol{h}),$$
$$\left\|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m, f_m) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a, f)\right\|_1 \leq \varepsilon(m, \boldsymbol{h})N(\boldsymbol{n}),$$

where in the last inequalities we used (4.27); the quantity $\varepsilon(m, \boldsymbol{h})$ is defined in (4.28) and satisfies (4.29). Choose, for every $\boldsymbol{h} \in \mathbb{N}^d$, a $m(\boldsymbol{h})$ such that $m(\boldsymbol{h}) \to \infty$ when $\boldsymbol{h} \to \infty$ and

$$\lim_{\boldsymbol{h}\to\infty} \varepsilon(m(\boldsymbol{h}), \boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} c(m(\boldsymbol{h}), \boldsymbol{h}) = \lim_{\boldsymbol{h}\to\infty} \omega(m(\boldsymbol{h}), \boldsymbol{h}) = 0.$$

A construction of such a function $m(\boldsymbol{h})$ is provided in Lemma 4.1 (apply the lemma with $x(m,\boldsymbol{h}) = \varepsilon(m,\boldsymbol{h}) + c(m,\boldsymbol{h}) + \omega(m,\boldsymbol{h})$ and $\xi = 0$). Then, for every $\boldsymbol{h} \in \mathbb{N}^d$ and every $n \geq \max(n_{m(\boldsymbol{h})}, n_{m(\boldsymbol{h}),\boldsymbol{h}})$,

$$A_{\boldsymbol{n}} = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f) + \left[LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_{m(\boldsymbol{h})}, f_{m(\boldsymbol{h})}) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f)\right] + (R_{\boldsymbol{n},m(\boldsymbol{h})} + R_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}) + (N_{\boldsymbol{n},m(\boldsymbol{h})} + N_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}),$$
$$\mathrm{rank}(R_{\boldsymbol{n},m(\boldsymbol{h})} + R_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}) \leq (c(m(\boldsymbol{h})) + c(m(\boldsymbol{h}),\boldsymbol{h}))N(\boldsymbol{n}),$$
$$\|N_{\boldsymbol{n},m(\boldsymbol{h})} + N_{\boldsymbol{n},m(\boldsymbol{h}),\boldsymbol{h}}\| \leq \omega(m(\boldsymbol{h})) + \omega(m(\boldsymbol{h}),\boldsymbol{h}),$$
$$\left\|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_{m(\boldsymbol{h})}, f_{m(\boldsymbol{h})}) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f)\right\|_1 \leq \varepsilon(m(\boldsymbol{h}),\boldsymbol{h})N(\boldsymbol{n}).$$

The application of Lemma 3.4 allows one to decompose $LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_{m(\boldsymbol{h})}, f_{m(\boldsymbol{h})}) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f)$ as the sum of a small-rank term $\hat{R}_{\boldsymbol{n},\boldsymbol{h}}$, with rank bounded from above by $\sqrt{\varepsilon(m(\boldsymbol{h}),\boldsymbol{h})}\,N(\boldsymbol{n})$, plus a small-norm term $\hat{N}_{\boldsymbol{n},\boldsymbol{h}}$, with norm bounded from above by $\sqrt{\varepsilon(m(\boldsymbol{h}),\boldsymbol{h})}$. This concludes the proof of the implication $4 \Rightarrow 1$.

($5 \Leftrightarrow 6$) We note that $\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n$ is a fixed matrix-sequence, independent of m. Item 5 is equivalent to saying that $\{\{A_{\boldsymbol{n}} - D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n\}_m$ is an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$, which, by Theorem 4.4, is equivalent to saying that $\{A_{\boldsymbol{n}} - D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n$ is zero-distributed. The latter assertion is equivalent to item 6 (by Theorem 2.5).

($2 \Rightarrow 5$) Obvious (take $a_m = a$, $f_m = f$ and $A_{\boldsymbol{n}}^{(m)} = D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)$).

($5 \Rightarrow 4$) Obvious (take $a_m = a$, $f_m = f$ and $A_{\boldsymbol{n}}^{(m)} = D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)$). □

Remark 4.3. Theorem 4.9 continues to hold if f is assumed to be separable, item 1 is replaced by '$\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{sLT}} a \otimes f$', and we add in item 3 the requirement that each f_m is separable. The proof is left as an exercise for the reader.

We end with a result that provides a relation between LT and sLT sequences. This result will be used in the next section to show that any LT sequence is a GLT sequence, and, implicitly, that the definition of GLT sequences, originally formulated in [44, 45] in terms of sLT sequences, can be equivalently formulated in terms of LT sequences.

Proposition 4.4. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{LT}} a \otimes f$. Then, for any $m \in \mathbb{N}$ there exist matrix-sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\mathrm{sLT}} a \otimes f_{i,m},\ i = 1, \ldots, N_m$, such that $\sum_{i=1}^{N_m} f_{i,m} \to f$ in $L^1([-\pi,\pi]^d)$ when $m \to \infty$ and $\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.*

Proof. Take any sequence of d-variate trigonometric polynomials f_m such that $f_m \to f$ in $L^1([-\pi,\pi]^d)$. We recall that such a sequence exists because the set of d-variate trigonometric polynomials is dense in $L^1([-\pi,\pi]^d)$. By definition, any d-variate trigonometric polynomial is a finite sum of separable d-variate trigonometric polynomials. Hence, we can write

$$f_m = \sum_{i=1}^{N_m} f_{i,m},$$

for some separable d-variate trigonometric polynomials $f_{i,m},\ i = 1, \ldots, N_m$. Take arbitrary matrix-sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\mathrm{sLT}} a \otimes f_{i,m},\ i = 1, \ldots, N_m$. In view of Theorem 4.8, one can choose, for example, $A_{\boldsymbol{n}}^{(i,m)} = D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f_{i,m})$. By Remark 4.2, $\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\mathrm{LT}} a \otimes (\sum_{i=1}^{N_m} f_{i,m}) = a \otimes f_m$. Hence, $\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ by Theorem 4.9. □

5 GLT sequences

In this section, we develop the theory of Generalized Locally Toeplitz sequences. In particular, we prove all the statements contained in items **GLT 1 – GLT 9**.

5.1 Definition

We first report a 'corrected' version of the original definition of GLT sequences; cf. [44, Definition 2.3] and [45, Definition 1.5]. This definition is formulated in terms of a.c.s. parameterized by a positive $\epsilon \to 0$ (see Definition 3.3).

Definition 5.1 (GLT sequence). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, $\boldsymbol{n} \in \mathbb{N}^d$, and let $\kappa : [0,1]^d \times [-\pi,\pi]^d \to \mathbb{C}$ be a measurable function. We say that $\{A_{\boldsymbol{n}}\}_n$ is a Generalized Locally Toeplitz (GLT) sequence with *symbol* κ, and we write $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$, if the following condition is met.

For every $\epsilon > 0$ there exists a finite number of sLT sequences $\{A_{\boldsymbol{n}}^{(i,\epsilon)}\}_n \sim_{\mathrm{sLT}} a_{i,\epsilon} \otimes f_{i,\epsilon},\ i = 1, \ldots, N_\epsilon$, such that:

- $\sum_{i=1}^{N_\epsilon} a_{i,\epsilon} \otimes f_{i,\epsilon} \to \kappa$ in measure over $[0,1]^d \times [-\pi,\pi]^d$ when $\epsilon \to 0$;
- $\big\{\{\sum_{i=1}^{N_\epsilon} A_{\boldsymbol{n}}^{(i,\epsilon)}\}_n\big\}_{\epsilon>0}$ is an a.c.s. of $\{A_{\boldsymbol{n}}\}_n$ for $\epsilon \to 0$.

From now on, if we write $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$, it is understood that $\kappa : [0,1]^d \times [-\pi,\pi]^d \to \mathbb{C}$ is measurable.

It is clear that any sLT sequence is a GLT sequence. Indeed, if $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{sLT}} a \otimes f$ then $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} a \otimes f$. To see this, it suffices to take, in Definition 5.1, $N_\epsilon = 1$, $\{A_{\boldsymbol{n}}^{(1,\epsilon)}\}_n = \{A_{\boldsymbol{n}}\}_n$, $a_{1,\epsilon} = a$ and $f_{i,\epsilon} = f$, for all $\epsilon > 0$. Proposition 4.4 and the first characterization of GLT sequences (Proposition 5.1) imply that any LT sequence is a GLT sequence. More precisely,

$$\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{LT}} a \otimes f \quad \Rightarrow \quad \{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} a \otimes f. \tag{5.1}$$

Proposition 5.1. *We have $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$ if and only if the following condition is met.*

For every m varying in some infinite subset of $\mathbb{N}$ there exists a finite number of sLT sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\mathrm{sLT}} a_{i,m} \otimes f_{i,m},\ i = 1,\ldots,N_m$, such that:

- $\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ *in measure over* $[0,1]^d \times [-\pi,\pi]^d$ *when* $m \to \infty$*;*
- $\big\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\big\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$.

Proof. If $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$, then the condition of the proposition hold with

$$a_{i,m} = a_{i,\epsilon(m)}, \qquad f_{i,m} = f_{i,\epsilon(m)}, \qquad \{A_{\boldsymbol{n}}^{(i,m)}\}_n = \{A_{\boldsymbol{n}}^{(i,\epsilon(m))}\}_n, \qquad N_m = N_{\epsilon(m)},$$

where $a_{i,\epsilon},\ f_{i,\epsilon},\ \{A_{\boldsymbol{n}}^{(i,\epsilon)}\}_n$ are as in Definition 5.1 and $\{\epsilon(m)\}_m$ is any sequence of positive numbers such that $\epsilon(m) \to 0$ as $m \to \infty$.

Conversely, suppose that the condition of the proposition hold. Let $\mathcal{M} \subseteq \mathbb{N}$ be the infinite subset of $\mathbb{N}$ where m varies. Then, the condition of Definition 5.1 hold with

$$a_{i,\epsilon} = a_{i,m(\epsilon)}, \qquad f_{i,\epsilon} = f_{i,m(\epsilon)}, \qquad \{A_{\boldsymbol{n}}^{(i,\epsilon)}\}_n = \{A_{\boldsymbol{n}}^{(i,m(\epsilon))}\}_n, \qquad N_\epsilon = N_{m(\epsilon)},$$

where $\{m(\epsilon)\}_{\epsilon>0} \subseteq \mathcal{M}$ is any family of indices such that $m(\epsilon) \to \infty$ when $\epsilon \to 0$, and $a_{i,m},\ f_{i,m},\ \{A_{\boldsymbol{n}}^{(i,m)}\}_n$ are as in the statement of the proposition. Thus, $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$. □

Proposition 5.1 is essentially the same as Definition 5.1, but it is easier to handle, because it is based on the standard notion of a.c.s. (Definition 3.1).

Corollary 5.1. *If $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{LT}} a \otimes f$ then $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} a \otimes f$.*

Proof. It follows directly from Proposition 5.1 and Proposition 4.4. □

5.2 Singular value and eigenvalue distribution of GLT sequences

We begin with a lemma concerning the singular value distribution of a finite sum of LT sequences. The lemma will be used in the proof of the singular value distribution result for GLT sequences (Theorem 5.1).

Lemma 5.1. *Let $\{A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\mathrm{LT}} a_i \otimes f_i,\ i = 1,\ldots,p$. Then $\{\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)}\}_n \sim_\sigma \sum_{i=1}^p a_i \otimes f_i$.*

Proof. Choose a sequence $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m \subseteq \mathbb{N}^d$ such that $\boldsymbol{m} \to \infty$ when $m \to \infty$. From the properties of a.c.s., see Proposition 3.1, and from the definition of LT sequences, we know that $\{\{\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i, f_i)\}_n\}_m$ is an a.c.s. for $\{\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)}\}_n$. By Theorem 4.2, $\{\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i, f_i)\}_n \sim_\sigma \phi_{\boldsymbol{m}}$ and $\phi_{\boldsymbol{m}} \to \phi_{[\,|\sum_{i=1}^p a_i \otimes f_i|\,]}$ pointwise over $C_c(\mathbb{R})$, where we recall that the functional $\phi_{[g]}$ is defined in (2.6). Hence, by Theorem 3.1, $\{\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)}\}_n \sim_\sigma \sum_{i=1}^p a_i \otimes f_i$. □

Theorem 5.1. *If $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$ then $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \kappa$.*

Proof. By Proposition 5.1, there exist matrix-sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\mathrm{LT}} a_{i,m} \otimes f_{i,m},\ i = 1,\ldots,N_m$, such that $\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ in measure and $\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. By Lemma 5.1, we have $\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_\sigma \sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m}$. Since $\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ in measure, all the assumptions of Corollary 3.1 are satisfied and so $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \kappa$. □

As a consequence of Theorem 5.1, every GLT sequence is s.u. in the sense of Definition 3.2 (see Proposition 3.3). Using Theorem 5.1, we show in Proposition 5.3 that the symbol of a GLT sequence is unique. For the proof of Proposition 5.3 we point out that any linear combination of GLT sequences is again a GLT sequence with symbol given by the same linear combination of the symbols. This is one of the most elementary results in the world of the algebraic properties possessed by GLT sequences. These properties will be investigated in Section 5.5 and give rise to the so-called GLT algebra.

Proposition 5.2. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$ and $\{B_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \xi$. Then, $\{A_{\boldsymbol{n}}^*\}_n \sim_{\mathrm{GLT}} \overline{\kappa}$ and $\{\alpha A_{\boldsymbol{n}} + \beta B_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \alpha\kappa + \beta\xi$ for all $\alpha, \beta \in \mathbb{C}$.*

The proof of Proposition 5.2 is easy: it suffices to write the meaning of $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and $\{B_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \xi$ (using the characterization of Proposition 5.1), and to apply Proposition 3.1; the details are left to the reader.

Proposition 5.3. *Assume that $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \xi$. Then $\kappa = \xi$ a.e. in $[0,1]^d \times [-\pi,\pi]^d$.*

Proof. By Proposition 5.2, $\{O_{N(\boldsymbol{n})}\}_n \sim_{\rm GLT} \kappa - \xi$. Therefore, by Theorem 5.1, for all test functions $F \in C_c(\mathbb{R})$ we have

$$F(0) = \frac{1}{(2\pi)^d} \int_{[0,1]^d \times [-\pi,\pi]^d} F(|\kappa(\mathbf{x},\boldsymbol{\theta}) - \xi(\mathbf{x},\boldsymbol{\theta})|) d\mathbf{x} d\boldsymbol{\theta}. \tag{5.2}$$

This means that $\phi_{[|\kappa-\xi|]} = \phi_{[0]}$ and so, by Remark 2.1, $|\kappa - \xi| = 0$ a.e. □

Proposition 5.4. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and assume that the matrices $A_{\boldsymbol{n}}$ are Hermitian. Then $\kappa \in \mathbb{R}$ a.e.*

Proof. Since the matrices $A_{\boldsymbol{n}}$ are Hermitian, by Proposition 5.2 we have $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \overline{\kappa}$. Thus, by Proposition 5.3, $\kappa = \overline{\kappa}$ a.e., i.e., $\kappa \in \mathbb{R}$ a.e. □

The next lemma deals with the spectral distribution of the real part of a finite sum of LT sequences. The lemma will be used in the proof of the eigenvalue distribution result for (Hermitian) GLT sequences (Theorem 5.2).

Lemma 5.2. *Let $\{A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} a_i \otimes f_i,\ i = 1,\ldots,p$. Then $\{\Re(\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)})\}_n \sim_\lambda \Re(\sum_{i=1}^p a_i \otimes f_i)$.*

Proof. Choose a sequence $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m \subseteq \mathbb{N}^d$ such that $\boldsymbol{m} \to \infty$ when $m \to \infty$. From the properties of a.c.s., see Remark 3.3 and Proposition 3.1, and from the definition of LT sequences, $\{\{\Re(\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i, f_i))\}_n\}_m$ is an a.c.s. for $\{\Re(\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)})\}_n$. By Theorem 4.3, $\{\Re(\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i, f_i))\}_n \sim_\lambda \phi_{\boldsymbol{m}}$ and $\phi_{\boldsymbol{m}} \to \phi_{[\Re(\sum_{i=1}^p a_i \otimes f_i)]}$. Hence, by Theorem 3.3, $\{\Re(\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)})\}_n \sim_\lambda \Re(\sum_{i=1}^p a_i \otimes f_i)$. □

Theorem 5.2. *If $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and the matrices $A_{\boldsymbol{n}}$ are Hermitian, then $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \kappa$.*

Proof. By Proposition 5.1, there exist matrix-sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm LT} a_{i,m} \otimes f_{i,m},\ i = 1,\ldots,N_m$, such that $\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ in measure and $\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Since the matrices $A_{\boldsymbol{n}}$ are Hermitian, $\{\{\Re(\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)})\}_n\}_m$ is another a.c.s. for $\{A_{\boldsymbol{n}} = \Re(A_{\boldsymbol{n}})\}_n$, and it is formed by Hermitian matrices. By Lemma 5.2, $\{\Re(\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)})\}_n \sim_\lambda \Re(\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m})$. The function κ is real a.e. by Proposition 5.4, and so from $\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ (in measure) we get $\Re(\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m}) \to \kappa$ (in measure). All the assumptions of Corollary 3.2 are then satisfied, and it follows that $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \kappa$. □

Remark 5.1. By Proposition 5.2, Lemmas 4.2 and 5.1 are particular cases of Theorem 5.1, and Lemma 5.2 is a particular case of Theorem 5.2.

5.3 Approximation results for GLT sequences

Theorem 5.3 is the main approximation result for GLT sequences. It is the same as Corollaries 3.1–3.2 with '$\sim_\sigma$' and '$\sim_\lambda$' replaced by '$\sim_{\rm GLT}$', and it is particularly useful to show that a given matrix-sequence $\{A_{\boldsymbol{n}}\}_n$ is a GLT sequence. Applications of Theorem 5.3 will be seen in Section 5.4 and, especially, in Section 5.5.

Theorem 5.3. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence and let $\kappa : [0,1]^d \times [-\pi,\pi]^d \to \mathbb{C}$ be a measurable function. Suppose that:*

1. *$\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$;*
2. *$\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$ for every m;*
3. *$\kappa_m \to \kappa$ in measure.*

Then $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$.

Proof. Since $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$, for every m and every h (varying in some infinite subset $\mathcal{H} \subseteq \mathbb{N}$) there exists a finite number of matrix-sequences $\{A_{\boldsymbol{n},m}^{(i,h)}\}_n \sim_{\rm sLT} a_{i,h,m} \otimes f_{i,h,m},\ i = 1,\ldots,N_{h,m}$, such that:

- $\sum_{i=1}^{N_{h,m}} a_{i,h,m} \otimes f_{i,h,m} \to \kappa_m$ in measure over $[0,1]^d \times [-\pi,\pi]^d$ when $h \to \infty$;
- $\big\{\{\sum_{i=1}^{N_{h,m}} A_{\boldsymbol{n},m}^{(i,h)}\}_n\big\}_h$ is an a.c.s. for $\{B_{\boldsymbol{n},m}\}_n$.

Hence, for every m and every h there exists $n_{h,m}$ such that, for $n \geq n_{h,m}$,

$$B_{\boldsymbol{n},m} = \sum_{i=1}^{N_{h,m}} A_{\boldsymbol{n},m}^{(i,h)} + R_{\boldsymbol{n},h,m} + N_{\boldsymbol{n},h,m},$$

$$\text{rank}(R_{\boldsymbol{n},h,m}) \leq c(h,m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},h,m}\| \leq \omega(h,m),$$

where $\lim_{h\to\infty} c(h,m) = \lim_{h\to\infty} \omega(h,m) = 0$. Let $\{\delta_m\}_m$ be a sequence such that $\delta_m \searrow 0$. Since $\sum_{i=1}^{N_{h,m}} a_{i,h,m} \otimes f_{i,h,m} \to \kappa_m$ in measure when $h \to \infty$, for every m we have

$$\mu(m,h,\delta_m) = \mu_{2d}\left\{\left|\sum_{i=1}^{N_{h,m}} a_{i,h,m} \otimes f_{i,h,m} - \kappa_m\right| \geq \delta_m\right\} \to 0 \qquad \text{as } h \to \infty.$$

Now we recall that $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$: for every m there exists n_m such that, for $n \geq n_m$,

$$A_{\boldsymbol{n}} = B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m},$$

$$\text{rank}(R_{\boldsymbol{n},m}) \leq c(m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \leq \omega(m),$$

where $\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0$. It follows that, for every m, every h, and every $n \geq \max(n_m, n_{h,m})$,

$$A_{\boldsymbol{n}} = \sum_{i=1}^{N_{h,m}} A_{\boldsymbol{n},m}^{(i,h)} + (R_{\boldsymbol{n},h,m} + R_{\boldsymbol{n},m}) + (N_{\boldsymbol{n},h,m} + N_{\boldsymbol{n},m}),$$

$$\text{rank}(R_{\boldsymbol{n},h,m} + R_{\boldsymbol{n},m}) \leq (c(h,m) + c(m))\,N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},h,m} + N_{\boldsymbol{n},m}\| \leq \omega(h,m) + \omega(m).$$

Choose a sequence $\{h_m\}_m \subseteq \mathcal{H}$ such that $h_m \nearrow \infty$ and

$$\lim_{m\to\infty} c(h_m,m) = \lim_{m\to\infty} \omega(h_m,m) = \lim_{m\to\infty} \mu(m,h_m,\delta_m) = 0.$$

Then, for every m and every $n \geq \max(n_m, n_{h_m,m})$,

$$A_{\boldsymbol{n}} = \sum_{i=1}^{N_{h_m,m}} A_{\boldsymbol{n},m}^{(i,h_m)} + (R_{\boldsymbol{n},h_m,m} + R_{\boldsymbol{n},m}) + (N_{\boldsymbol{n},h_m,m} + N_{\boldsymbol{n},m}),$$

$$\text{rank}(R_{\boldsymbol{n},h_m,m} + R_{\boldsymbol{n},m}) \leq c(h_m,m) + c(m), \qquad \|N_{\boldsymbol{n},h_m,m} + N_{\boldsymbol{n},m}\| \leq \omega(h_m,m) + \omega(m).$$

It follows that $\{\{\sum_{i=1}^{N_{h_m,m}} A_{\boldsymbol{n},m}^{(i,h_m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Moreover, $\{A_{\boldsymbol{n},m}^{(i,h_m)}\}_n \sim_{\text{sLT}} a_{i,h_m,m} \otimes f_{i,h_m,m}$ for all m and all $i = 1,\ldots,N_{h_m,m}$, and $\sum_{i=1}^{N_{h_m,m}} a_{i,h_m,m} \otimes f_{i,h_m,m} \to \kappa$ in measure over $[0,1]^d \times [-\pi,\pi]^d$ when $m \to \infty$. Indeed, for any $\delta > 0$,

$$\mu_{2d}\left\{\left|\sum_{i=1}^{N_{h_m,m}} a_{i,h_m,m} \otimes f_{i,h_m,m} - \kappa\right| \geq \delta\right\} \leq \mu_{2d}\left\{\left|\sum_{i=1}^{N_{h_m,m}} a_{i,h_m,m} \otimes f_{i,h_m,m} - \kappa_m\right| \geq \delta/2\right\} + \mu_{2d}\left\{|\kappa_m - \kappa| \geq \delta/2\right\},$$

$\mu_{2d}\{|\kappa_m - \kappa| \geq \delta/2\} \to 0$ by assumption (since $\kappa_m \to \kappa$ in measure), and

$$\mu_{2d}\left\{\left|\sum_{i=1}^{N_{h_m,m}} a_{i,h_m,m} \otimes f_{i,h_m,m} - \kappa_m\right| \geq \delta/2\right\} = \mu(m,h_m,\delta/2)$$

tends to 0, because it is eventually less than $\mu(m,h_m,\delta_m)$. Thus, $\{A_{\boldsymbol{n}}\}_n \sim_{\text{GLT}} \kappa$ by Proposition 5.1. □

The approximation result for GLT sequences stated in Theorem 5.3 admits the following converse, which can be seen as another approximation result for GLT sequences. In fact, Theorems 5.3–5.4 look like a characterization of GLT sequences in terms of a.c.s.

Theorem 5.4. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence and let $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ be a sequence of matrix-sequences. Suppose that:*

1. *$\{A_{\boldsymbol{n}}\}_n \sim_{\text{GLT}} \kappa$;*
2. *$\{B_{\boldsymbol{n},m}\}_n \sim_{\text{GLT}} \kappa_m$ for every m.*

Then, $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ if and only if $\kappa_m \to \kappa$ in measure.

Proof. ($\Leftarrow$) Assume that 1–2 hold and $\kappa_m \to \kappa$ in measure. By Proposition 5.2, $\{A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa - \kappa_m$ for each m. Hence, by Theorem 5.1, $\{A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\}_n \sim_\sigma \kappa - \kappa_m$, with $\kappa - \kappa_m$ tending to 0 in measure by hypothesis. Hence, by Corollary 3.4, $\{B_{\boldsymbol{n},m}\}_n$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.

($\Rightarrow$) Assume that 1–2 hold and $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Then, $\{\{A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. of $\{O_{N(\boldsymbol{n})}\}_n$. Moreover, $\{A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa - \kappa_m$ by Proposition 5.2, and so $\{A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\}_n \sim_\sigma \kappa - \kappa_m$ by Theorem 5.1. Since $\{O_{N(\boldsymbol{n})}\}_n \sim_\sigma 0$, it follows from Theorem 3.2 that $\phi_{[|\kappa-\kappa_m|]} \to \phi_{[0]}$ pointwise over $C_c(\mathbb{R})$. Hence, by Lemma 2.5, $|\kappa - \kappa_m| \to 0$ in measure. □

Corollary 5.2. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$. Then, for all functions $a_{i,m},\ f_{i,m},\ i = 1, \ldots, N_m$, with the following properties:*

- $*$ *$a_{i,m} : [0,1]^d \to \mathbb{C}$ is Riemann-integrable and $f_{i,m} \in L^1([-\pi,\pi]^d)$;*
- $*$ *$\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ in measure when $m \to \infty$;*

it holds that $\{\{\sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m}) T_{\boldsymbol{n}}(f_{i,m})\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. In particular, $\{A_{\boldsymbol{n}}\}_n$ admits an a.c.s. of the form

$$\left\{\left\{\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m} D_{\boldsymbol{n}}(a_{\boldsymbol{j}}^{(m)}) T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})\right\}_n\right\}_m, \qquad a_{\boldsymbol{j}}^{(m)} \in C^\infty([0,1]^d), \qquad \boldsymbol{N}_m \in \mathbb{N}^d, \tag{5.3}$$

where $\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m} a_{\boldsymbol{j}}^{(m)}(\mathbf{x}) \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}} \to \kappa(\mathbf{x}, \boldsymbol{\theta})$ a.e.

Proof. Since $\{D_{\boldsymbol{n}}(a_{i,m}) T_{\boldsymbol{n}}(f_{i,m})\}_n \sim_{\rm GLT} a_{i,m} \otimes f_{i,m}$, we have $\{\sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m}) T_{\boldsymbol{n}}(f_{i,m})\}_n \sim_{\rm GLT} \sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m}$ by Proposition 5.2. Therefore, the thesis follows from Theorem 5.4 applied with $B_{\boldsymbol{n},m} = \sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m}) T_{\boldsymbol{n}}(f_{i,m})$ and $\kappa_m = \sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m}$. To obtain an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ of the form (5.3), it suffices to use the result of the corollary in combination with Lemma 2.6. □

Remark 5.2. Let $\{A_{\boldsymbol{n}}\}_n,\ \{B_{\boldsymbol{n},m}\}_n$ be matrix-sequences and let $\kappa,\ \kappa_m : [0,1]^d \times [-\pi,\pi]^d \to \mathbb{C}$ be measurable functions. Consider the following four conditions:

(1) $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$;

(2) $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$ for every m;

(3) $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$;

(4) $\kappa_m \to \kappa$ in measure.

Theorems 5.3–5.4 show that '(1) ∧ (2) ∧ (3) ⇒ (4)', '(1) ∧ (2) ∧ (4) ⇒ (3)' and '(2) ∧ (3) ∧ (4) ⇒ (1)'.

The implication '(1) ∧ (3) ∧ (4) ⇒ (2)', written in this way, is meaningless. However, a natural modification reads as follows: '(1) ∧ (3) ⇒ there exists a measurable function κ_m, tending to κ in measure, such that $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$ for all sufficiently large m'. This statement is false in general. As a counterexample, take $A_n = O_n$ and $B_{n,m} = (1 + (-1)^n)\frac{1}{m} I_n$, as in Remark 3.2. Since we have seen in Remark 3.2 that the relation $\{B_{n,m}\}_n \sim_\sigma \phi_m$ cannot hold for any functional ϕ_m defined over $C_c(\mathbb{R})$, in particular there is no κ_m such that $\{B_{n,m}\}_n \sim_{\rm GLT} \kappa_m$.

5.4 Characterizations of GLT sequences

As a first application of Theorem 5.3, we show in Proposition 5.5 that GLT sequences could be defined in terms of LT sequences instead of sLT sequences. Proposition 5.5 is then a characterization of GLT sequences in terms of LT sequences and, specifically, it is the same as Proposition 5.1 with 'sLT' replaced by 'LT'.

Proposition 5.5. *We have $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ if and only if the following condition is met.*

> *For every m varying in some infinite subset of $\mathbb{N}$ there exists a finite number of LT sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm LT} a_{i,m} \otimes f_{i,m},\ i = 1, \ldots, N_m$, such that:*
>
> - *$\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ in measure over $[0,1]^d \times [-\pi,\pi]^d$ when $m \to \infty$;*
> - *$\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.*

Proof. It is clear that, if $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$, then the condition holds by Proposition 5.1. Conversely, suppose that the condition holds. Then, $\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ by hypothesis, $\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm GLT} \sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m}$ by Corollary 5.1 and Proposition 5.2, and $\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ in measure, by hypothesis. Hence, the thesis follows from Theorem 5.3. □

It is clear that Proposition 5.5, as well as Proposition 5.1, may be taken as the definition of GLT sequences.

The next result is a characterization theorem for GLT sequences. All the characterizations provided in the theorem have already been proved in the previous section, but it is anyway useful to collect them in a single statement.

Theorem 5.5 (characterizations of GLT sequences). *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence and let $\kappa : [0,1]^d \times [-\pi,\pi]^d \to \mathbb{C}$ be a measurable function. Then, the following conditions are equivalent.*

1. $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$.
2. *For all sequences $\{\kappa_m\}_m,\ \{\{B_{\boldsymbol{n},m}\}_m\}_m$ with the following properties:*
   - $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$ *for every m;*
   - $\kappa_m \to \kappa$ *in measure;*

   it holds that $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.
3. *There exist functions $a_{i,m},\ f_{i,m},\ i=1,\ldots,N_m,\ m\in\mathbb{N}$, such that:*
   - $a_{i,m} : [0,1]^d \to \mathbb{C}$ *belongs to $C^\infty([0,1]^d)$ and $f_{i,m}$ is a trigonometric monomial belonging to $\{{\rm e}^{{\rm i}\boldsymbol{j}\cdot\boldsymbol{\theta}} : \boldsymbol{j}\in\mathbb{Z}^d\}$;*
   - $\sum_{i=1}^{N_m} a_{i,m}\otimes f_{i,m} \to \kappa$ *a.e.;*
   - $\{\{\sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m})T_{\boldsymbol{n}}(f_{i,m})\}_n\}_m$ *is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.*
4. *There exist sequences $\{\kappa_m\}_m,\ \{\{B_{\boldsymbol{n},m}\}_n\}_m$ such that:*
   - $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$ *for every m;*
   - $\kappa_m \to \kappa$ *in measure;*
   - $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ *is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.*

Proof. The implication ($1 \Rightarrow 2$) follows from Theorem 5.4. The implication ($2 \Rightarrow 3$) follows from the observation that, by Lemma 2.6, we can find functions $a_{i,m},\ f_{i,m},\ i=1,\ldots,N_m,\ m\in\mathbb{N}$, with the first two properties specified in item 3. The implication ($3 \Rightarrow 4$) is obvious (take $B_{\boldsymbol{n},m} = \sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m})T_{\boldsymbol{n}}(f_{i,m})$ and $\kappa_m = \sum_{i=1}^{N_m} a_{i,m}\otimes f_{i,m}$). Finally, the implication ($4 \Rightarrow 1$) is Theorem 5.3. □

5.5 The GLT algebra

We investigate in this section the important algebraic properties possessed by GLT sequences, which give rise to the so-called GLT algebra. In short, these properties establish that, if $\{A_{\boldsymbol{n}}^{(1)}\}_n,\ldots,\{A_{\boldsymbol{n}}^{(r)}\}_n$ are given GLT sequences with symbols $\kappa_1,\ldots,\kappa_r$, respectively, and if $A_{\boldsymbol{n}} = {\rm ops}(A_{\boldsymbol{n}}^{(1)},\ldots,A_{\boldsymbol{n}}^{(r)})$ is obtained from $A_{\boldsymbol{n}}^{(1)},\ldots,A_{\boldsymbol{n}}^{(r)}$ by means of certain operations 'ops', then $\{A_{\boldsymbol{n}}\}_n$ is a GLT sequence with symbol $\kappa = {\rm ops}(\kappa_1,\ldots,\kappa_r)$ obtained by performing the same operations on the symbols $\kappa_1,\ldots,\kappa_r$.

Theorem 5.6. *Suppose that $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and $\{B_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \xi$. Then:*

1. $\{A_{\boldsymbol{n}}^*\}_n \sim_{\rm GLT} \overline{\kappa}$*;*
2. $\{\alpha A_{\boldsymbol{n}} + \beta B_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \alpha\kappa + \beta\xi$*, for all $\alpha,\beta\in\mathbb{C}$;*
3. $\{A_{\boldsymbol{n}}B_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa\xi$.

Proof. The first two statements have already been settled before (see Proposition 5.2). We prove the third statement. By assumption and Proposition 5.5, there exist LT sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm LT} a_{i,m}\otimes f_{i,m},\ i=1,\ldots,N_m$, and $\{B_{\boldsymbol{n}}^{(j,m)}\}_n \sim_{\rm sLT} b_{j,m}\otimes g_{j,m},\ j=1,\ldots,M_m$, such that:

- $\sum_{i=1}^{N_m} a_{i,m}\otimes f_{i,m} \to \kappa$ in measure and $\sum_{j=1}^{M_m} b_{j,m}\otimes g_{j,m} \to \xi$ in measure;
- $\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ and $\{\{\sum_{j=1}^{M_m} B_{\boldsymbol{n}}^{(j,m)}\}_n\}_m$ is an a.c.s. for $\{B_{\boldsymbol{n}}\}_n$.

Thanks to Theorem 5.5 (item 3), the functions $f_{i,m},\ g_{j,m}$ may be supposed to be in $L^\infty([-\pi,\pi]^d)$; actually, they might be supposed to be trigonometric monomials, the functions $a_{i,m},\ b_{j,m}$ might be supposed to belong to $C^\infty([0,1]^d)$, and $\{A_{\boldsymbol{n}}^{(i,m)}\}_n,\ \{B_{\boldsymbol{n}}^{(j,m)}\}_n$ might be chosen of the form $\{D_{\boldsymbol{n}}(a_{i,m})T_{\boldsymbol{n}}(f_{i,m})\}_n,\ \{D_{\boldsymbol{n}}(b_{j,m})T_{\boldsymbol{n}}(g_{j,m})\}_n$. By Theorem 5.1, $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \kappa$ and $\{B_{\boldsymbol{n}}\}_n \sim_\sigma \xi$, which implies, by Proposition 3.3, that $\{A_{\boldsymbol{n}}\}_n$ and $\{B_{\boldsymbol{n}}\}_n$ are s.u. Thus, by Proposition 3.4,

$$\left\{\left\{\left(\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\right)\left(\sum_{j=1}^{M_m} B_{\boldsymbol{n}}^{(j,m)}\right)\right\}_n\right\}_m = \left\{\left\{\sum_{i=1}^{N_m}\sum_{j=1}^{M_m} A_{\boldsymbol{n}}^{(i,m)}B_{\boldsymbol{n}}^{(j,m)}\right\}_n\right\}_m$$

is an a.c.s. for $\{A_{\boldsymbol{n}}B_{\boldsymbol{n}}\}_n$. Since $f_{i,m},\ g_{j,m} \in L^\infty([-\pi,\pi]^d)$, by Theorem 4.7 we have $\{A_{\boldsymbol{n}}^{(i,m)}B_{\boldsymbol{n}}^{(j,m)}\}_n \sim_{\rm LT} a_{i,m}b_{j,m} \otimes f_{i,m}g_{j,m},\ i = 1,\ldots,N_m,\ j = 1,\ldots,M_m$. Finally,

$$\sum_{i=1}^{N_m}\sum_{j=1}^{M_m} a_{i,m}b_{j,m} \otimes f_{i,m}g_{j,m} = \left(\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m}\right)\left(\sum_{j=1}^{M_m} b_{j,m} \otimes g_{j,m}\right) \to \kappa\xi$$

in measure by Lemma 2.3, and the proof is over. □

Corollary 5.3. *Let $r, q_1,\ldots,q_r \in \mathbb{N}$ and, for $i = 1,\ldots,r$ and $j = 1,\ldots,q_i$, let $\{A_{\boldsymbol{n}}^{(ij)}\}_n \sim_{\rm GLT} \kappa_{ij}$. Then,*

$$\left\{\sum_{i=1}^{r}\prod_{j=1}^{q_i} A_{\boldsymbol{n}}^{(ij)}\right\}_n \sim_{\rm GLT} \sum_{i=1}^{r}\prod_{j=1}^{q_i} \kappa_{ij}.$$

The results we have seen so far are enough to conclude that the set of GLT sequences is a *-algebra over the complex field $\mathbb{C}$. More precisely, fix any sequence of d-indices $\{\boldsymbol{n} = \boldsymbol{n}(n)\}_n \subseteq \mathbb{N}^d$ such that $\boldsymbol{n} \to \infty$ when $n \to \infty$; then,

$$\mathscr{A} = \big\{\{A_{\boldsymbol{n}}\}_n :\ \{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa \ \text{ for some measurable function } \kappa : [0,1]^d \times [-\pi,\pi]^d \to \mathbb{C}\big\} \tag{5.4}$$

is a *-algebra over $\mathbb{C}$, with respect to the natural operations of Hermitian transposition, addition, scalar-multiplication and product of matrix-sequences: $(\{A_{\boldsymbol{n}}\}_n)^* = \{A_{\boldsymbol{n}}^*\}_n,\ \{A_{\boldsymbol{n}}\}_n + \{B_{\boldsymbol{n}}\}_n = \{A_{\boldsymbol{n}} + B_{\boldsymbol{n}}\}_n,\ \alpha\{A_{\boldsymbol{n}}\}_n = \{\alpha A_{\boldsymbol{n}}\}_n,\ \{A_{\boldsymbol{n}}\}_n\{B_{\boldsymbol{n}}\}_n = \{A_{\boldsymbol{n}}B_{\boldsymbol{n}}\}_n$. We call $\mathscr{A}$ the GLT algebra. This algebra contains the algebra generated by zero-distributed sequences, Toeplitz sequences and sequences of diagonal sampling matrices, because we have seen that these matrix-sequences fall in the class of GLT sequences. To be precise, let

$$\mathscr{B} = \left\{\Big\{\sum_{i=1}^{r}\prod_{j=1}^{q_i} X_{\boldsymbol{n}}^{(i,j)}\Big\} :\ r, q_1,\ldots,q_r \in \mathbb{N},\ \{X_{\boldsymbol{n}}^{(i,j)}\}_n \in \mathcal{B} \text{ for all } i = 1,\ldots,r \text{ and } j = 1,\ldots,q_i\right\},$$

where

$$\mathcal{B} = \big\{\{T_{\boldsymbol{n}}(g)\}_n :\ g \in L^1([-\pi,\pi]^d)\big\} \cup \big\{\{D_{\boldsymbol{n}}(a)\}_n :\ a : [0,1]^d \to \mathbb{C} \text{ is Riemann-integrable}\big\} \cup \{\{Z_{\boldsymbol{n}}\}_n :\ \{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0\}.$$

Then, $\mathscr{B}$ is the algebra generated by $\mathcal{B}$ and $\mathscr{B} \subseteq \mathscr{A}$. We are going to see in Theorems 5.7–5.8 that the GLT algebra enjoys other nice properties, in addition to those of Theorem 5.6, which make it look like a 'big container', closed under any type of 'regular' operation.

Theorem 5.7 provides a positive answer to a question raised in [46]. Incidentally, we note that in [46] the authors proved that $\{\{f(B_{\boldsymbol{n},m})\}_n\}_m$ is an a.c.s. for $\{f(A_{\boldsymbol{n}})\}_n$ whenever $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ and $f : \mathbb{R} \to \mathbb{R}$ is continuous (and other mild assumptions are met); this result enlarges the algebraic properties of a.c.s. studied in Section 3.2.

Theorem 5.7. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and suppose that the matrices $A_{\boldsymbol{n}}$ are Hermitian. Then*

$$\{f(A_{\boldsymbol{n}})\}_n \sim_{\rm GLT} f(\kappa)$$

for any continuous function $f : \mathbb{R} \to \mathbb{R}$.[5]

Proof. For each $M > 0$, let $\{p_{m,M}\}_m$ be a sequence of polynomials that converges uniformly to f over the compact interval $[-M, M]$:

$$\lim_{m\to\infty} \|f - p_{m,M}\|_{\infty,[-M,M]} = 0.$$

For every $M > 0$ and every m, n, write

$$f(A_{\boldsymbol{n}}) = p_{m,M}(A_{\boldsymbol{n}}) + f(A_{\boldsymbol{n}}) - p_{m,M}(A_{\boldsymbol{n}}). \tag{5.5}$$

Since any GLT sequence is s.u. (by Theorem 5.1 and Proposition 3.3), the sequence $\{A_{\boldsymbol{n}}\}_n$ is s.u. Hence, by Proposition 3.2, for all $M > 0$ there exists n_M such that, for $n \geq n_M$,

$$A_{\boldsymbol{n}} = \hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M}, \qquad \text{rank}(\hat{A}_{\boldsymbol{n},M}) \leq r(M)N(\boldsymbol{n}), \qquad \|\tilde{A}_{\boldsymbol{n},M}\| \leq M, \tag{5.6}$$

where $\lim_{M\to\infty} r(M) = 0$. However, for the purpose of this proof we need a splitting of the form (5.6) such that $g(\hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M}) = g(\hat{A}_{\boldsymbol{n},M}) + g(\tilde{A}_{\boldsymbol{n},M})$ for all functions $g : \mathbb{R} \to \mathbb{R}$. Luckily, the matrices $A_{\boldsymbol{n}}$ are Hermitian and, consequently, such a splitting can be

[5] Recall from Proposition 5.4 that $\kappa \in \mathbb{R}$ a.e., because every $A_{\boldsymbol{n}}$ is Hermitian. Hence, $f(\kappa)$ is well-defined.

constructed by following the same argument used in the proof of Proposition 3.2. For the reader's convenience, we include the details of the construction. By definition, since $\{A_{\boldsymbol{n}}\}_n$ is s.u., for every $M > 0$ there exists n_M such that, for $n \geq n_M$,

$$\frac{\#\{i \in \{1, \ldots, N(\boldsymbol{n})\} :\ \sigma_i(A_{\boldsymbol{n}}) > M\}}{N(\boldsymbol{n})} \leq r(M),$$

where $\lim_{M\to\infty} r(M) = 0$. Let $A_{\boldsymbol{n}} = U_{\boldsymbol{n}}\Lambda_{\boldsymbol{n}}U_{\boldsymbol{n}}^*$ be a spectral decomposition of $A_{\boldsymbol{n}}$. Let $\hat{\Lambda}_{\boldsymbol{n},M}$ be the matrix obtained from $\Lambda_{\boldsymbol{n}}$ by setting to 0 all the eigenvalues of $A_{\boldsymbol{n}}$ whose absolute value is less than or equal to M, and let $\tilde{\Lambda}_{\boldsymbol{n},M} = \Lambda_{\boldsymbol{n}} - \hat{\Lambda}_{\boldsymbol{n},M}$ be the matrix obtained from $\Lambda_{\boldsymbol{n}}$ by setting to 0 all the eigenvalues of $A_{\boldsymbol{n}}$ whose absolute value is greater than M. Then, for $M > 0$ and $n \geq n_M$,

$$A_{\boldsymbol{n}} = U_{\boldsymbol{n}}\Lambda_{\boldsymbol{n}}U_{\boldsymbol{n}}^* = U_{\boldsymbol{n}}\hat{\Lambda}_{\boldsymbol{n},M}U_{\boldsymbol{n}}^* + U_{\boldsymbol{n}}\tilde{\Lambda}_{\boldsymbol{n},M}U_{\boldsymbol{n}}^* = \hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M},$$

where $\hat{A}_{\boldsymbol{n},M} = U_{\boldsymbol{n}}\hat{\Lambda}_{\boldsymbol{n},M}U_{\boldsymbol{n}}^*$ and $\tilde{A}_{\boldsymbol{n},M} = U_{\boldsymbol{n}}\tilde{\Lambda}_{\boldsymbol{n},M}U_{\boldsymbol{n}}^*$. The matrices $\hat{A}_{\boldsymbol{n},M}$, $\tilde{A}_{\boldsymbol{n},M}$ constructed in this way are Hermitian, satisfy the properties in (5.6) and, moreover,

$$g(\hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M}) = g(\hat{A}_{\boldsymbol{n},M}) + g(\tilde{A}_{\boldsymbol{n},M}) = U_{\boldsymbol{n}}\, g(\hat{\Lambda}_{\boldsymbol{n},M})U_{\boldsymbol{n}}^* + U_{\boldsymbol{n}}\, g(\tilde{\Lambda}_{\boldsymbol{n},M})U_{\boldsymbol{n}}^*$$

for all functions $g : \mathbb{R} \to \mathbb{R}$.

Going back to (5.5), for every $M > 0$, every m and every $n \geq n_M$ we can write

$$\begin{aligned} f(A_{\boldsymbol{n}}) &= p_{m,M}(A_{\boldsymbol{n}}) + f(\hat{A}_{\boldsymbol{n},M}) + f(\tilde{A}_{\boldsymbol{n},M}) - p_{m,M}(\hat{A}_{\boldsymbol{n},M}) - p_{m,M}(\tilde{A}_{\boldsymbol{n},M}) \\ &= p_{m,M}(A_{\boldsymbol{n}}) + (f - p_{m,M})(\hat{A}_{\boldsymbol{n},M}) + (f - p_{m,M})(\tilde{A}_{\boldsymbol{n},M}). \end{aligned} \tag{5.7}$$

The term $(f - p_{m,M})(\hat{A}_{\boldsymbol{n},M})$ can be split in the sum of two terms $R_{\boldsymbol{n},m,M} + N'_{\boldsymbol{n},m,M}$: $R_{\boldsymbol{n},m,M}$ is obtained from $(f - p_{m,M})(\hat{A}_{\boldsymbol{n},M})$ by setting to 0 all the eigenvalues that are equal to $(f - p_{m,M_m})(0)$, so that rank$(R_{\boldsymbol{n},m,M}) =$ rank$(\hat{A}_{\boldsymbol{n},M})$; while $N'_{\boldsymbol{n},m,M}$ is obtained from $(f - p_{m,M})(\hat{A}_{\boldsymbol{n},M})$ by setting to 0 all the eigenvalues that are different from $(f - p_{m,M_m})(0)$. Let $N''_{\boldsymbol{n},m,M} = (f - p_{m,M})(\tilde{A}_{\boldsymbol{n},M})$ and $N_{\boldsymbol{n},m,M} = N'_{\boldsymbol{n},m,M} + N''_{\boldsymbol{n},m,M}$. From (5.7), for every $M > 0$, every m and every $n \geq n_M$ we have

$$f(A_{\boldsymbol{n}}) = p_{m,M}(A_{\boldsymbol{n}}) + R_{\boldsymbol{n},m,M} + N_{\boldsymbol{n},m,M}, \tag{5.8}$$

and, by our construction,

$$\begin{gathered} \text{rank}(R_{\boldsymbol{n},m,M}) = \text{rank}(\hat{A}_{\boldsymbol{n},M}) \leq r(M)N(\boldsymbol{n}), \\ \|N_{\boldsymbol{n},m,M}\| \leq |f(0) - p_{m,M}(0)| + \|f - p_{m,M}\|_{\infty,[-M,M]} \leq 2\|f - p_{m,M}\|_{\infty,[-M,M]}. \end{gathered} \tag{5.9}$$

Choose a sequence $\{M_m\}_m$ such that, when $m \to \infty$,

$$M_m \to \infty, \qquad \|f - p_{m,M_m}\|_{\infty,[-M_m,M_m]} \to 0. \tag{5.10}$$

Then, for every m and every $n \geq n_{M_m}$,

$$f(A_{\boldsymbol{n}}) = p_{m,M_m}(A_{\boldsymbol{n}}) + R_{\boldsymbol{n},m,M_m} + N_{\boldsymbol{n},m,M_m},$$

$$\text{rank}(R_{\boldsymbol{n},m,M_m}) \leq r(M_m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m,M_m}\| \leq 2\|f - p_{m,M_m}\|_{\infty,[-M_m,M_m]},$$

which implies that $\{\{p_{m,M_m}(A_{\boldsymbol{n}})\}_n\}_m$ is an a.c.s. for $\{f(A_{\boldsymbol{n}})\}_n$. Moreover, $\{p_{m,M_m}(A_{\boldsymbol{n}})\}_n \sim_{\rm GLT} p_{m,M_m}(\kappa)$ by Theorem 5.6. Finally, $p_{m,M_m}(\kappa) \to f(\kappa)$ a.e. in $[0,1]^d \times [-\pi,\pi]^d$, due to (5.10). In conclusion, all the hypotheses of Theorem 5.3 are satisfied and so $\{f(A_{\boldsymbol{n}})\}_n \sim_{\rm GLT} f(\kappa)$. □

The last issue we are interested in is to know if $\{A_{\boldsymbol{n}}^{-1}\}_n \sim_{\rm GLT} \kappa^{-1}$ in the case where $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$, each $A_{\boldsymbol{n}}$ is invertible, and $\kappa \neq 0$ a.e. (so that κ^{-1} is a well-defined measurable function). More in general, we may ask if $\{A_{\boldsymbol{n}}^\dagger\}_n \sim_{\rm GLT} \kappa^{-1}$ when $\{A_{\boldsymbol{n}}\} \sim_{\rm GLT} \kappa$ and $\kappa \neq 0$ a.e., being $A_{\boldsymbol{n}}^\dagger$ the (Moore–Penrose) pseudoinverse of $\{A_{\boldsymbol{n}}\}_n$. The answer to both the previous questions is affirmative, but some work is needed to bring out the related proofs. Note that these results cannot be inferred from Theorem 5.7, because the matrices $A_{\boldsymbol{n}}$ may fail to be Hermitian and, moreover, $f(x) = x^{-1}$ is not continuous on $\mathbb{R}$. We begin by introducing the concept of sparsely vanishing matrix-sequences.

Definition 5.2 (sparsely vanishing matrix-sequence). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. We say that $\{A_{\boldsymbol{n}}\}_n$ is sparsely vanishing (s.v.) if for every $M > 0$ there exists n_M such that, for $n \geq n_M$,

$$\frac{\#\{i \in \{1, \ldots, N(\boldsymbol{n})\} :\ \sigma_i(A_{\boldsymbol{n}}) < 1/M\}}{N(\boldsymbol{n})} \leq r(M),$$

where $\lim_{M\to\infty} r(M) = 0$.

It is clear from Definition 5.2 that if $\{A_{\boldsymbol{n}}\}_n$ is s.v. then $\{A^\dagger_{\boldsymbol{n}}\}_n$ is s.u.; it suffices to recall that the singular values of $A^\dagger$ are $1/\sigma_1(A),\ldots,1/\sigma_r(A),0,\ldots,0$, where $\sigma_1(A)\ldots\sigma_r(A)$ are the nonzero singular values of A ($r=\mathrm{rank}(A)$).

Remark 5.3. Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. Then, $\{A_{\boldsymbol{n}}\}_n$ is s.v. if and only if

$$\lim_{M\to\infty}\limsup_{n\to\infty}\frac{\#\{i\in\{1,\ldots,N(\boldsymbol{n})\}:\,\sigma_i(A_{\boldsymbol{n}})<1/M\}}{N(\boldsymbol{n})}=0. \tag{5.11}$$

The proof of this equivalence is easy and follows the same line as the proof of the equivalence ($1\Leftrightarrow 2$) in Proposition 3.2; the details are left to the reader. Note that (5.11) can be rewritten as

$$\lim_{M\to\infty}\limsup_{n\to\infty}\frac{1}{N(\boldsymbol{n})}\sum_{i=1}^{N(\boldsymbol{n})}\chi_{[0,1/M)}(\sigma_i(A_{\boldsymbol{n}}))=0.$$

Proposition 5.6. *If $\{A_{\boldsymbol{n}}\}_n\sim_\sigma f$ then $\{A_{\boldsymbol{n}}\}_n$ is s.v. if and only if $f\neq 0$ a.e.*

Proof. Let $D\subset\mathbb{R}^k$ be the domain of the function f. Fix $M>0$ and take $F_M\in C_c(\mathbb{R})$ such that $F_M=1$ over $[0,1/M]$, $F_M=0$ over $[2/M,\infty]$ and $0\le F_M\le 1$ over $\mathbb{R}$. Note that $F_M\ge\chi_{[0,1/M)}$ over $[0,\infty)$. Then,

$$\begin{aligned}\frac{\#\{i\in\{1,\ldots,N(\boldsymbol{n})\}:\,\sigma_i(A_{\boldsymbol{n}})<1/M\}}{N(\boldsymbol{n})}&=\frac{1}{N(\boldsymbol{n})}\sum_{i=1}^{N(\boldsymbol{n})}\chi_{[0,1/M)}(\sigma_i(A_{\boldsymbol{n}}))\\&\le\frac{1}{N(\boldsymbol{n})}\sum_{i=1}^{N(\boldsymbol{n})}F_M(\sigma_i(A_{\boldsymbol{n}}))\overset{n\to\infty}{\longrightarrow}\frac{1}{\mu_k(D)}\int_D F_M(|f(\mathbf{x})|)d\mathbf{x}\end{aligned}$$

and

$$\limsup_{n\to\infty}\frac{\#\{i\in\{1,\ldots,N(\boldsymbol{n})\}:\,\sigma_i(A_{\boldsymbol{n}})<1/M\}}{N(\boldsymbol{n})}\le\frac{1}{\mu_k(D)}\int_D F_M(|f(\mathbf{x})|)d\mathbf{x}.$$

Since $F_M(|f(\mathbf{x})|)\to\chi_{\{f=0\}}(\mathbf{x})$ a.e. and $|F(|f(\mathbf{x})|)|\le 1$, by the dominated convergence theorem we get

$$\lim_{M\to\infty}\int_D F_M(|f(\mathbf{x})|)d\mathbf{x}=\frac{\mu_k\{f\neq 0\}}{\mu_k(D)}.$$

Thus,

$$\lim_{M\to\infty}\limsup_{n\to\infty}\frac{\#\{i\in\{1,\ldots,N(\boldsymbol{n})\}:\,\sigma_i(A_{\boldsymbol{n}})<1/M\}}{N(\boldsymbol{n})}=0$$

if and only if $f=0$ a.e. By Remark 5.3, this means that $\{A_{\boldsymbol{n}}\}_n$ is s.v. if and only if $f=0$ a.e. ☐

Theorem 5.8. *Let $\{A_{\boldsymbol{n}}\}_n\sim_{\mathrm{GLT}}\kappa$ with $\kappa\neq 0$ a.e., then $\{A^\dagger_{\boldsymbol{n}}\}_n\sim_{\mathrm{GLT}}\kappa^{-1}$.*

Proof. Take a sequence of matrix-sequences $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ such that $\{B_{\boldsymbol{n},m}\}_m\sim_{\mathrm{GLT}}\xi_m$ for every m and $\xi_m\to\kappa^{-1}$ in measure. Note that a sequence $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ with these properties exists. Indeed, by Lemma 2.6 there exists a sequence $\{\xi_m\}_m$, with ξ_m of the form

$$\xi_m(\mathbf{x},\boldsymbol{\theta})=\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m}a_{\boldsymbol{j}}^{(m)}(\mathbf{x})\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}},\qquad a_{\boldsymbol{j}}^{(m)}\in C^\infty([0,1]^d),\qquad \boldsymbol{N}_m\in\mathbb{N}^d,$$

such that $\xi_m\to\kappa^{-1}$ a.e. (and hence also in measure). Therefore, it suffices to take $B_{\boldsymbol{n},m}=\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m}D_{\boldsymbol{n}}(a_{\boldsymbol{j}}^{(m)})T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})$ and to observe that $\{B_{\boldsymbol{n},m}\}_n\sim_{\mathrm{GLT}}\xi_m$ (see Theorem 4.8, Corollary 5.1 and Theorem 5.6).

By Theorem 5.6, we have $\{B_{\boldsymbol{n},m}A_{\boldsymbol{n}}-I_{N(\boldsymbol{n})}\}_n\sim_{\mathrm{GLT}}\xi_m\kappa-1$ for every m, and $\xi_m\kappa-1\to 0$ a.e. (and hence also in measure). Therefore, by Theorem 3.5, for every m there exists n_m such that, for $n\ge n_m$,

$$\begin{gathered}B_{\boldsymbol{n},m}A_{\boldsymbol{n}}=I_{N(\boldsymbol{n})}+R_{\boldsymbol{n},m}+N_{\boldsymbol{n},m},\\ \mathrm{rank}(R_{\boldsymbol{n},m})\le c(m)N(\boldsymbol{n}),\qquad \|N_{\boldsymbol{n},m}\|\le\omega(m),\end{gathered}\tag{5.12}$$

where $\lim_{m\to\infty}c(m)\lim_{m\to\infty}\omega(m)=0$. Multiplying (5.12) by $A^\dagger_{\boldsymbol{n}}$, we obtain that, for every m and every $n\ge n_m$,

$$B_{\boldsymbol{n},m}\,A_{\boldsymbol{n}}A^\dagger_{\boldsymbol{n}}=A^\dagger_{\boldsymbol{n}}+(R_{\boldsymbol{n},m}+N_{\boldsymbol{n},m})A^\dagger_{\boldsymbol{n}}. \tag{5.13}$$

Since $\kappa \neq 0$ a.e. by hypothesis, $\{A_{\boldsymbol{n}}\}_n$ is s.v. (by Theorem 5.1 and Proposition 5.6). It follows that $\{A_{\boldsymbol{n}}^\dagger\}_n$ is s.u. and so, by Proposition 3.2, for all $M > 0$ there is $\overline{n}_M$ such that, for $n \ge \overline{n}_M$,

$$A_{\boldsymbol{n}}^\dagger = \hat{A}_{\boldsymbol{n},M}^\dagger + \tilde{A}_{\boldsymbol{n},M}^\dagger,$$
$$\text{rank}(\hat{A}_{\boldsymbol{n},M}^\dagger) \le r(M)N(\boldsymbol{n}), \qquad \|\tilde{A}_{\boldsymbol{n},M}^\dagger\| \le M,$$

where $\lim_{M\to\infty} r(M) = 0$. Choosing $M_m = [\omega(m)]^{-1/2}$, from (5.13) we see that, for every m and every $n \ge \max(n_m, \overline{n}_{M_m})$,

$$\begin{aligned} B_{\boldsymbol{n},m}\, A_{\boldsymbol{n}} A_{\boldsymbol{n}}^\dagger &= A_{\boldsymbol{n}}^\dagger + R'_{\boldsymbol{n},m} + N'_{\boldsymbol{n},m}, \\ \text{rank}(R'_{\boldsymbol{n},m}) &\le \big(c(m) + r(M_m)\big)N(\boldsymbol{n}), \qquad \|N'_{\boldsymbol{n},m}\| \le [\omega(m)]^{1/2}, \end{aligned} \tag{5.14}$$

where we have set $R'_{\boldsymbol{n},m} = R_{\boldsymbol{n},m}A_{\boldsymbol{n}}^\dagger + N_{\boldsymbol{n},m}\hat{A}_{\boldsymbol{n},M_m}^\dagger$ and $N'_{\boldsymbol{n},m} = N_{\boldsymbol{n},m}\tilde{A}_{\boldsymbol{n},M_m}^\dagger$.

If the matrices $A_{\boldsymbol{n}}$ were invertible, then $A_{\boldsymbol{n}}^\dagger = A_{\boldsymbol{n}}^{-1}$ and (5.14) would imply that $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}^{-1}\}_n$; this, in combination with the approximation result for GLT sequences (Theorem 5.3), would conclude the proof. In the general case where the matrices $A_{\boldsymbol{n}}$ are not invertible, the thesis will follow again from (5.14) and Theorem 5.3 as soon as we have proved the following: for every m there exists $\hat{n}_m$ such that, for $n \ge \hat{n}_m$,

$$A_{\boldsymbol{n}}A_{\boldsymbol{n}}^\dagger = I_{N(\boldsymbol{n})} + S_{\boldsymbol{n}}, \qquad \text{rank}(S_{\boldsymbol{n}}) \le \theta(m)N(\boldsymbol{n}),$$

where $\lim_{m\to\infty}\theta(m) = 0$. This is easy, because, by definition of $A_{\boldsymbol{n}}^\dagger$, the rank of the matrix $S_{\boldsymbol{n}} = A_{\boldsymbol{n}}A_{\boldsymbol{n}}^\dagger - I_{N(\boldsymbol{n})}$ is given by $\text{rank}(S_{\boldsymbol{n}}) = \#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) = 0\}$. Hence, the previous claim follows directly from the fact that $\{A_{\boldsymbol{n}}\}_n$ is s.v. (see Definition 5.2). □

5.5.1 The algebra generated by Toeplitz sequences

Fix a sequence of d-indices $\{\boldsymbol{n} = \boldsymbol{n}(n)\}_n$ such that $\boldsymbol{n} \to \infty$ as $n \to \infty$. In this section, we briefly talk about the algebra $\mathscr{T}$ over the complex field $\mathbb{C}$ generated by the Toeplitz sequences of the form $\{T_{\boldsymbol{n}}(g)\}_n,\ g \in L^1([-\pi,\pi]^d)$. It is not difficult to see that

$$\mathscr{T} = \left\{ \Big\{\sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij})\Big\}_n \,:\, r, q_1, \ldots, q_r \in \mathbb{N},\ g_{ij} \in L^1([-\pi,\pi]^d) \text{ for all } i = 1,\ldots,r \text{ and } j = 1,\ldots,q_i \right\}. \tag{5.15}$$

It is clear from Theorem 4.6 and Corollary 5.1 that $\mathscr{T}$ is a sub-algebra of the GLT algebra $\mathscr{A}$ defined in (5.4). Indeed, according to Corollary 5.3,

$$\Big\{\sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij})\Big\}_n \sim_{\text{GLT}} \sum_{i=1}^r \prod_{j=1}^{q_i} 1 \otimes g_{ij} = 1 \otimes \sum_{i=1}^r \prod_{j=1}^{q_i} g_{ij}.$$

Since $\int_{[0,1]^d\times[-\pi,\pi]^d}(1\otimes g) = \int_{[-\pi,\pi]^d} g$, Theorem 5.1 and Definition 2.2 immediately give

$$\Big\{\sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij})\Big\}_n \sim_\sigma \sum_{i=1}^r \prod_{j=1}^{q_i} g_{ij}.$$

This result was originally obtained in [42]. Similarly, if the matrices $\sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij})$ are Hermitian, Theorem 5.2 gives

$$\Big\{\sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij})\Big\}_n \sim_\lambda \sum_{i=1}^r \prod_{j=1}^{q_i} g_{ij}. \tag{5.16}$$

The extension of the spectral distribution relation (5.16) to the case where the matrices $\sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij})$ are not Hermitian has been the subject of a recent research [21]. Note that, if we remove the hypothesis of Hermitianess, then we necessarily have to add some additional assumption. Indeed, (5.16) does not hold in general; a counterexample is provided, e.g., by the sequence of (1-level) Toeplitz matrices $\{T_n(\mathrm{e}^{\mathrm{i}j\theta})\}_n$. The hypothesis added in [21] is a topological assumption on the range of the functions g_{ij}. A completely analogous hypothesis was already used in [20, 22] and, especially, in the pioneering work by Tilli [53], in order to extend the spectral distribution relation expressed in Theorem 2.7 to the case where the generating function f is not real (and hence the related Toeplitz matrices $T_{\boldsymbol{n}}(f)$ are not Hermitian).

5.6 Summary of the theory of GLT sequences

After developing the theory of GLT sequences, we note at this point that we have proved all the items **GLT 1 – GLT 9** of Section 1.2. In particular, item **GLT 1** was proved in Theorems 5.1–5.2, items **GLT 2 – GLT 4** were proved in Theorems 4.4–4.6, items **GLT 5 – GLT 8** were proved in Theorems 5.6–5.8, and item **GLT 9** is contained in Theorem 5.5.

We note however that in Section 1.2, we adopted for simplicity a not completely precise notation and we used 'n' instead of '$\boldsymbol{n}$'. For the reader's convenience, we report again items **GLT 1 – GLT 9** in the correct notation. We also simplify a little bit their statements.

GLT 1. If $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$, then $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \kappa$. If moreover the matrices $A_{\boldsymbol{n}}$ are Hermitian, then $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \kappa$.

GLT 2. $\{T_{\boldsymbol{n}}(f)\}_n \sim_{\rm GLT} 1 \otimes f$ for every $f \in L^1([-\pi,\pi]^d)$.

GLT 3. $\{D_{\boldsymbol{n}}(a)\}_n \sim_{\rm GLT} a \otimes 1$ for every Riemann-integrable function $a : [0,1]^d \to \mathbb{C}$.

GLT 4. If $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$ then $\{Z_{\boldsymbol{n}}\}_n \sim_{\rm GLT} 0$, and vice versa.

GLT 5. If $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ then $\{A_{\boldsymbol{n}}^*\}_n \sim_{\rm GLT} \overline{\kappa}$.

GLT 6. If $A_{\boldsymbol{n}} = \sum_{i=1}^r \alpha_i \prod_{j=1}^{q_i} A_{\boldsymbol{n}}^{(i,j)}$, where $r, q_1, \ldots, q_r \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_r \in \mathbb{C}$ and $\{A_{\boldsymbol{n}}^{(i,j)}\}_n \sim_{\rm GLT} \kappa_{ij}$, then $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa = \sum_{i=1}^r \alpha_i \prod_{j=1}^{q_i} \kappa_{ij}$.

GLT 7. If $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and $\kappa \neq 0$ a.e., then $\{A_{\boldsymbol{n}}^\dagger\}_n \sim_{\rm GLT} \kappa^{-1}$.

GLT 8. If $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and each $A_{\boldsymbol{n}}$ is Hermitian, then $\{f(A_{\boldsymbol{n}})\}_n \sim_{\rm GLT} f(\kappa)$ for all continuous functions $f : \mathbb{R} \to \mathbb{R}$.

GLT 9. $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ if and only if there exist GLT sequences $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$ such that $\kappa_m \to \kappa$ in measure and $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.

6 Applications

In this section, we outline a few applications of the theory of GLT sequences. This is done to give a flavour of the applicative interest of the theory. For more applications, we refer the reader to Section 1.1, where specific pointers to the available literature are provided.

6.1 Geometric means of matrices

Everyone knows that the geometric mean of two positive numbers a, b is $G(a,b) = (ab)^{1/2}$. But what is the geometric mean $G(A,B)$ of two Hermitian Positive Definite (HPD) matrices $A, B \in \mathbb{C}^{n\times n}$? An appropriate definition was proposed in a remarkable paper by Ando, Li and Mathias [2]. The approach of these authors was axiomatic: denoting by $\mathscr{P}_n$ the set of HPD matrices of size n, a function $G : \mathscr{P}_n \times \mathscr{P}_n \to \mathscr{P}_n$ is said to be a geometric mean if it verifies a suitable list of properties that any geometric mean worthy of the name should verify. Ando, Li and Mathias proposed a list of ten properties, which are referred to as the ALM axioms. Let us mention three of them.

1. *Permutation invariance:* $G(A,B) = G(B,A)$ for all $A, B \in \mathscr{P}_n$.
2. *Congruence invariance:* $G(M^*AM, M^*BM) = M^*G(A,B)M$ for all $A, B \in \mathscr{P}_n$ and all invertible matrices M.
3. *Consistency with scalars:* $G(A,B) = (AB)^{1/2}$ for all commuting matrices $A, B \in \mathscr{P}_n$.

It may be proved [9, Chapter 4] that the unique function $G : \mathscr{P}_n \times \mathscr{P}_n \to \mathscr{P}_n$ verifying both consistency with scalars and congruence invariance is $G(A,B) = A(A^{-1}B)^{1/2}$, which, moreover, verifies all the ALM axioms. Using some properties of matrix functions and taking into account that $G(A,B) = G(B,A)$, it can be shown that

$$G(A,B) = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} = B^{1/2}(B^{-1/2}AB^{-1/2})^{1/2}B^{1/2}, \qquad A, B \in \mathscr{P}_n. \tag{6.1}$$

Suppose now that $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and $\{B_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \xi$, where $A_{\boldsymbol{n}}, B_{\boldsymbol{n}} \in \mathscr{P}_{N(\boldsymbol{n})}$. Due to **GLT 1**, the positive definiteness of the matrices $A_{\boldsymbol{n}}, B_{\boldsymbol{n}}$, and Proposition 2.1, the essential ranges $\mathcal{ER}(\kappa)$, $\mathcal{ER}(\xi)$ are contained in $[0,\infty)$. Hence, $\kappa, \xi \geq 0$ a.e. We assume that at least one between κ and ξ is nonzero a.e. Under this assumption, we show that the sequence of geometric means $\{G(A_{\boldsymbol{n}}, B_{\boldsymbol{n}})\}_n$ is a GLT sequence whose symbol is given by the geometric mean of the symbols κ, ξ. In other words, we prove the following relation:

$$\{G(A_{\boldsymbol{n}}, B_{\boldsymbol{n}})\}_n \sim_{\rm GLT} (\kappa\xi)^{1/2}. \tag{6.2}$$

Let $f:\mathbb{R}\to\mathbb{R}$ be any continuous function such that $f(x)=x^{1/2}$ for all $x\ge 0$. If $\kappa\neq 0$ a.e., then $\kappa>0$ a.e., and so, using the first expression of $G(A_{\boldsymbol{n}},B_{\boldsymbol{n}})$ in (6.1), we see that

$$G(A_{\boldsymbol{n}},B_{\boldsymbol{n}})=A_{\boldsymbol{n}}^{1/2}(A_{\boldsymbol{n}}^{-1/2}B_{\boldsymbol{n}}A_{\boldsymbol{n}}^{-1/2})^{1/2}A_{\boldsymbol{n}}^{1/2}=f(A_{\boldsymbol{n}})f(f(A_{\boldsymbol{n}})^{-1}B_{\boldsymbol{n}}f(A_{\boldsymbol{n}})^{-1})f(A_{\boldsymbol{n}}).$$

By **GLT 6**, **GLT 7** and **GLT 8**, it follows that

$$\{G(A_{\boldsymbol{n}},B_{\boldsymbol{n}})\}_n\sim_{\rm GLT} f(\kappa)f(f(\kappa)^{-1}\xi f(\kappa)^{-1})f(\kappa).$$

Since $\kappa>0$ a.e., we have

$$f(\kappa)f(f(\kappa)^{-1}\xi f(\kappa)^{-1})f(\kappa)=\kappa^{1/2}(\kappa^{-1/2}\xi\kappa^{-1/2})^{1/2}\kappa^{1/2}=(\kappa\xi)^{1/2}\qquad\text{a.e.,}$$

and the relation (6.2) is proved. If $\xi\neq 0$ a.e., then the proof of (6.2) follows the same pattern as in the case $\kappa\neq 0$ a.e., with the only difference that now we use the second expression of $G(A_{\boldsymbol{n}},B_{\boldsymbol{n}})$ in (6.1) instead of the first one. Noting that $G(A_{\boldsymbol{n}},B_{\boldsymbol{n}})$ is HPD whenever $A_{\boldsymbol{n}},\ B_{\boldsymbol{n}}$ are HPD, from (6.2) and **GLT 1** we get

$$\{G(A_{\boldsymbol{n}},B_{\boldsymbol{n}})\}_n\sim_{\sigma,\,\lambda}(\kappa\xi)^{1/2}. \tag{6.3}$$

While it is easy to generalize the concept of geometric mean to the case where the numbers to be averaged are $k>2$, the same is not true for HPD matrices. In particular, the axiomatic approach by Ando, Li and Mathias is not satisfying for $k>2$, because the ten ALM axioms do not lead to a unique definition [12, 37]. The path to the right definition was different, involving a little bit of differential geometry. In fact, the geometric mean (or Karcher mean) of k matrices $A^{(1)},\ldots,A^{(k)}\in\mathscr{P}_n$ was defined as the barycenter of the matrices with respect to a certain Riemannian distance; see [10] and [9, Chapter 6]. More precisely, the Karcher mean $G(A^{(1)},\ldots,A^{(k)})$ is the unique minimizer over $\mathscr{P}_n$ of the functional

$$D(\cdot\,;A^{(1)},\ldots,A^{(k)}):\mathscr{P}_n\to\mathbb{R},\qquad D(X;A^{(1)},\ldots,A^{(k)})=\sum_{i=1}^{k}\big[\delta(X,A^{(i)})\big]^2, \tag{6.4}$$

where $\delta(A,B)$ is the distance given by the Riemannian structure,

$$\delta(A,B)=\|\log(A^{-1/2}BA^{-1/2})\|_2=\left(\sum_{\ell=1}^{n}\log^2(\lambda_\ell(A^{-1}B))\right)^{1/2}. \tag{6.5}$$

It was proved with some effort [11, 32, 35] that the Karcher mean verifies all the ALM axioms and some further properties, and thus now everyone agrees that the Karcher mean has the right to be called the geometric mean of matrices.

Suppose now that $\{A_{\boldsymbol{n}}^{(i)}\}_n\sim_{\rm GLT}\kappa_i$ for $i=1,\ldots,k$, where $A_{\boldsymbol{n}}^{(1)},\ldots,A_{\boldsymbol{n}}^{(k)}\in\mathscr{P}_{N(\boldsymbol{n})}$. By **GLT 1**, the positive definiteness of $A_{\boldsymbol{n}}^{(i)}$, and Proposition 2.1, each κ_i is non-negative a.e. In this situation, we have reason to believe that the sequence of Karcher means $\{G(A_{\boldsymbol{n}}^{(1)},\ldots,A_{\boldsymbol{n}}^{(k)})\}_n$ is a GLT sequence with symbol $(\kappa_1\cdots\kappa_k)^{1/k}$. The formal proof of this result, which might be achieved via **GLT 9**, is certainly an interesting subject for future research, also considering that geometric means of Toeplitz matrices are of interest in practical applications. For example, in a radar application one is interested in computing a geometric mean of HPD matrices which are Toeplitz or block Toeplitz with Toeplitz blocks (i.e., 2-level Toeplitz); see [3, 4, 34].

6.2 PDE discretizations

The main application of the theory of GLT sequences was already described in Section 1.1. It consists in the computation of the spectral distribution of the sequences of discretization matrices arising from the approximation of PDEs by numerical methods. In fact, these matrix-sequences are often GLT sequences, especially when the numerical method belongs to the class of the so-called ‘local methods’. Local methods are, for example, Finite Difference methods, collocation methods and Finite Element methods with ‘locally supported’ basis functions; in short, all standard numerical methods for the approximation of PDEs.

In Section 6.2.1, we present the GLT analysis of a simple 1-dimensional model problem, approximated by standard Finite Differences (FDs). We will discuss in Section 6.2.2 the extension of the analysis to the d-dimensional setting, but without entering into the (technical) details. Finally, in Section 6.2.3, we will consider the matrices arising from the Finite Element (FE) approximation of a system of PDEs somehow connected with the linear elasticity equations; these matrices naturally show up in saddle point form, and we will analyze their Schur complements through the theory of GLT sequences. The idea of all this section is to show that items **GLT 1 – GLT 9** are a powerful tool for computing the asymptotic spectral distribution of PDE discretization matrices.

6.2.1 Finite Difference discretization of 1-dimensional elliptic PDEs

Consider the following second order elliptic PDE with homogeneous Dirichlet boundary conditions:

$$\begin{cases} -(a(x)u'(x))' = f(x), & x \in (0,1), \\ u(0) = u(1) = 0, \end{cases} \tag{6.6}$$

where $a : [0,1] \to \mathbb{R}$ is a function in $C^1([0,1])$ such that $a(x) > 0$ for every $x \in [0,1]$. We consider the discretization of (6.6) by using the classical second order centered FD scheme. In the case where $a(x)$ is constant, this is also known as the $(-1, 2, -1)$ FD scheme. Let us describe it shortly; for more details on FD methods, we refer the reader to the available literature (see, e.g., [50] or any good book on FDs). We choose a discretization parameter $n \in \mathbb{N}$, we set $h = \frac{1}{n+1}$ and $x_j = jh$ for all $j \in [0, n+1]$, and we note that, for $j = 1, \ldots, n$, we can approximate $(a(x)u'(x))'|_{x=x_j}$ by the following FD formula:

$$\begin{aligned} -(a(x)u'(x))'|_{x=x_j} &\approx -\frac{a(x_{j+\frac{1}{2}})u'(x_{j+\frac{1}{2}}) - a(x_{j-\frac{1}{2}})u'(x_{j-\frac{1}{2}})}{h} \approx -\frac{a(x_{j+\frac{1}{2}})\dfrac{u(x_{j+1}) - u(x_j)}{h} - a(x_{j-\frac{1}{2}})\dfrac{u(x_j) - u(x_{j-1})}{h}}{h} \\ &= \frac{-a(x_{j+\frac{1}{2}})u(x_{j+1}) + \big(a(x_{j+\frac{1}{2}}) + a(x_{j-\frac{1}{2}})\big)u(x_j) - a(x_{j-\frac{1}{2}})u(x_{j-1})}{h^2}. \end{aligned} \tag{6.7}$$

Then, we approximate the solution of (6.6) by the piecewise linear function that takes the values u_j in x_j for $j = 0, \ldots, n+1$, where $u_0 = u_{n+1} = 0$ and $\mathbf{u} = (u_1, \ldots, u_n)^T$ is the (unique) solution of the linear system

$$-a(x_{j+\frac{1}{2}})u_{j+1} + \big(a(x_{j+\frac{1}{2}}) + a(x_{j-\frac{1}{2}})\big)u_j - a(x_{j-\frac{1}{2}})u_{j-1} = h^2 f(x_j), \qquad j = 1, \ldots, n. \tag{6.8}$$

The matrix corresponding to the linear system (6.8) is the tridiagonal symmetric matrix given by

$$A_n = \begin{bmatrix} a(x_{\frac{1}{2}}) + a(x_{\frac{3}{2}}) & -a(x_{\frac{3}{2}}) & & & & \\ -a(x_{\frac{3}{2}}) & a(x_{\frac{3}{2}}) + a(x_{\frac{5}{2}}) & -a(x_{\frac{5}{2}}) & & & \\ & -a(x_{\frac{5}{2}}) & a(x_{\frac{5}{2}}) + a(x_{\frac{7}{2}}) & -a(x_{\frac{7}{2}}) & & \\ & & -a(x_{\frac{7}{2}}) & \ddots & \ddots & \\ & & & \ddots & \ddots & -a(x_{n-\frac{1}{2}}) \\ & & & & -a(x_{n-\frac{1}{2}}) & a(x_{n-\frac{1}{2}}) + a(x_{n+\frac{1}{2}}) \end{bmatrix}. \tag{6.9}$$

In this example, we will see that the theory of GLT sequences allows us to compute the singular value and eigenvalue distribution of the sequence of discretization matrices $\{A_n\}_n$. Actually, this is the fundamental example that led to the birth of the theory of LT sequences [52] and, subsequently, of GLT sequences [44, 45]. Given the importance, we will compute the singular value and eigenvalue distribution of $\{A_n\}_n$ by two different methods, both of them instructive.

Method 1. Suppose first that $a(x)$ is constant, say $a(x) = 1$ identically. In this case, the matrix A_n becomes

$$A_n = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & -1 & \ddots & \ddots & \\ & & & \ddots & \ddots & -1 \\ & & & & -1 & 2 \end{bmatrix}.$$

Therefore, $A_n = T_n(2-2\cos\theta)$ is simply the Toeplitz matrix generated by the function $2-2\cos\theta$; see (2.32)–(2.33). Using Theorem 2.7, we have $\{A_n\}_n \sim_{\sigma,\lambda} 2 - 2\cos\theta$. Since the function $2 - 2\cos\theta$ is even, the relations $\{A_n\}_n \sim_{\sigma,\lambda} 2 - 2\cos\theta$ continue to hold even if we consider $[0,\pi]$ as the domain of $2 - 2\cos\theta$ instead of $[-\pi,\pi]$. According to the informal meaning behind the definition of spectral and singular value distribution, see Remark 2.2, we may conclude that the eigenvalues and the singular values of A_n are approximately a uniform sampling over $[0,\pi]$ of the non-negative function $2 - 2\cos\theta$. However, this result is known also analytically. Indeed, since $T_n(2 - 2\cos\theta)$ is symmetric positive definite, its eigenvalues coincide with the singular values and are given explicitly by $2 - 2\cos\frac{j\pi}{n+1}$, $j = 1, \ldots, n$; see [13, p. 35] or [50, p. 154].

Now let us turn to the case where $a(x)$ is not constant. In this case, the expression of A_n is given by (6.9) and the Toeplitzness seems to be completely lost. In reality, we find it again ‘in an approximated sense’ and ‘at a local scale’. Indeed, we note that $a(x)$ varies smoothly from $a(0)$ to $a(1)$, because it is uniformly continuous. Therefore, assuming that n is large with respect to k, any $k \times k$ leading principal submatrix of A_n shows an approximate Toeplitz structure. Let us be more quantitative. Fix a large $m \in \mathbb{N}$ and assume $n > m$. Then, n is large with respect to $\lfloor n/m \rfloor$, and so, according to the previous reasoning, any $\lfloor n/m \rfloor \times \lfloor n/m \rfloor$ leading principal submatrix of A_n shows an approximate Toeplitz structure. In fact, the evaluations of $a(x)$ appearing in the first $\lfloor n/m \rfloor \times \lfloor n/m \rfloor$ leading principal submatrix of A_n are approximately equal to $a(\frac{1}{m})$; the evaluations of $a(x)$ appearing in the second $\lfloor n/m \rfloor \times \lfloor n/m \rfloor$ leading principal submatrix of A_n are approximately equal to $a(\frac{2}{m})$; and so on until the evaluations of $a(x)$ appearing in the m-th $\lfloor n/m \rfloor \times \lfloor n/m \rfloor$ leading principal submatrix of A_n, which are approximately equal to $a(1)$. If, for all $j = 1, \ldots, m$, we replace by $a(\frac{j}{m})$ the evaluations of $a(x)$ in the j-th $\lfloor n/m \rfloor \times \lfloor n/m \rfloor$ leading principal submatrix of A_n, this submatrix becomes $a(\frac{j}{m})T_{\lfloor n/m \rfloor}(2 - 2\cos\theta)$. In conclusion, the matrix A_n is approximated by the Locally Toeplitz operator

$$LT_n^m(a(x), 2 - 2\cos\theta) = \begin{bmatrix} a(\frac{1}{m})T_{\lfloor n/m \rfloor}(2 - 2\cos\theta) & & & & \\ & a(\frac{2}{m})T_{\lfloor n/m \rfloor}(2 - 2\cos\theta) & & & \\ & & \ddots & & \\ & & & a(1)T_{\lfloor n/m \rfloor}(2 - 2\cos\theta) & \\ & & & & O_{n \bmod m} \end{bmatrix}.$$

In fact, $\{\{LT_n^m(a(x), 2 - 2\cos\theta)\}_n\}_m$ is an a.c.s. for $\{A_n\}_n$, because it can be shown that

$$A_n = LT_n^m(a(x), 2 - 2\cos\theta) + R_{n,m} + N_{n,m}$$

$$\mathrm{rank}(R_{n,m}) \le 3m, \qquad \|N_{n,m}\| \le \omega_a\Big(\frac{1}{m} + \frac{m+1}{n+1}\Big),$$

being $\omega_a(\cdot)$ the modulus of continuity of a. Thus, by definition, $\{A_n\}_n \sim_{\mathrm{LT}} a(x)(2 - 2\cos\theta)$, and so

$$\{A_n\}_n \sim_{\sigma,\,\lambda} a(x)(2 - 2\cos\theta). \tag{6.10}$$

Since $a(x)(2 - 2\cos\theta)$ is symmetric with respect to the Fourier variable θ, the relations (6.10) continue to hold even if we consider $[0, 1] \times [0, \pi]$ as the domain of $a(x)(2 - 2\cos\theta)$ instead of $[0, 1] \times [-\pi, \pi]$; this follows directly from the definition of spectral and singular value distribution. According to the informal meaning of (6.10), see Remark 2.2, if $n = \ell^2$ is large enough, the eigenvalues of A_n are approximately given by the uniform sampling $a(\frac{i}{\ell})(2 - 2\cos\frac{j\pi}{\ell+1})$, $i, j = 1, \ldots, \ell$.

Method 2. As already pointed out, the example we are dealing with led to the birth of the theory of GLT sequences. In particular, the procedure followed in *Method 1* to obtain (6.10) motivated the definition of Locally Toeplitz sequences, as well as the introduction of the Locally Toeplitz operator. However, now that we have fully developed the theory of GLT sequences, we should say that *Method 1* is probably not the most effective way to obtain (6.10). The method we are going to see now seems to be by far more effective.

Let $\hat{x}_j = \frac{j}{n}$, $j = 1, \ldots, n$, and note that $|x_j - \hat{x}_j| \le \frac{1}{n+1} = h$ for all $j = 1, \ldots, n$. Consider the matrix

$$D_n(a)T_n(2 - 2\cos\theta) = \begin{bmatrix} 2a(\hat{x}_1) & -a(\hat{x}_1) & & & & \\ -a(\hat{x}_2) & 2a(\hat{x}_2) & -a(\hat{x}_2) & & & \\ & -a(\hat{x}_3) & 2a(\hat{x}_3) & -a(\hat{x}_3) & & \\ & & -a(\hat{x}_4) & \ddots & \ddots & \\ & & & \ddots & \ddots & -a(\hat{x}_{n-1}) \\ & & & & -a(\hat{x}_n) & 2a(\hat{x}_n) \end{bmatrix}. \tag{6.11}$$

By a direct comparison between (6.11) and (6.9), we see that the modulus of each diagonal entry of the matrix $A_n - D_n(a)T_n(2-2\cos\theta)$ is bounded from above by $2\omega_a(h + h/2)$, and the modulus of each off-diagonal entry of $A_n - D_n(a)T_n(2 - 2\cos\theta)$ is bounded from above by $\omega_a(h+h/2)$. Therefore, the 1-norm and the infinity norm of $A_n - D_n(a)T_n(2-2\cos\theta)$ are bounded from above by $4\omega_a(3h/2)$, and so, using (2.16), we get

$$\|A_n - D_n(a)T_n(2 - 2\cos\theta)\| \le 4\omega_a(3h/2).$$

In particular, $\|A_n - D_n(a)T_n(2 - 2\cos\theta)\| \to 0$ as $n \to \infty$. Setting $Z_n = A_n - D_n(a)T_n(2 - 2\cos\theta)$, we have $\{Z_n\}_n \sim_\sigma 0$ by Theorem 2.6 and $\{Z_n\}_n \sim_{\mathrm{GLT}} 0$ by **GLT 4**. Moreover, by **GLT 2**, **GLT 3** and **GLT 6** we have $\{D_n(a)T_n(2 - 2\cos\theta)\}_n \sim_{\mathrm{GLT}} a(x)(2 - 2\cos\theta)$. Since

$$A_n = D_n(a)T_n(2 - 2\cos\theta) + Z_n, \tag{6.12}$$

again by **GLT 6** we have $\{A_n\}_n \sim_{\mathrm{GLT}} a(x)(2 - 2\cos\theta)$, and (6.10) follows from **GLT 1**.

Remark 6.1. Before passing to discuss the d-dimensional setting, some remarks are in order.

1. Problem (6.6) can be rewritten in the form

$$\begin{cases} -a(x)u''(x) - a'(x)u'(x) = f(x), \quad x \in (0,1), \\ u(0) = u(1) = 0. \end{cases} \tag{6.13}$$

From this reformulation, it appears more clearly that the symbol $a(x)(2-2\cos\theta)$ consists of two 'ingredients':

- the coefficient $a(x)$ of the higher-order differential operator in (6.13);
- the trigonometric polynomial $2-2\cos\theta = -\mathrm{e}^{\mathrm{i}\theta} + 2 - \mathrm{e}^{-\mathrm{i}\theta}$ associated with the FD formula $(-1, 2, -1)$ used to approximate the higher-order derivative $-u''(x)$.

In particular, the term $-a'(x)u'(x)$, which only depends on lower-order derivatives of $u(x)$, does not enter into the expression of the symbol.

2. Suppose that we add to the diffusion equation (6.6) a convenction and a reaction term. In this way, we obtain the following reaction-convection-diffusion PDE:

$$\begin{cases} -(a(x)u'(x))' + \beta(x)u'(x) + \gamma(x)u(x) = f(x), \quad x \in (0,1), \\ u(0) = u(1) = 0. \end{cases} \tag{6.14}$$

Based on the discussion in item 1, we expect that the term $\beta(x)u'(x) + \gamma(x)u(x)$, which only involves lower-order derivatives of $u(x)$, does not enter into the expression of the symbol. In other words, the symbol of the FD discretization matrices B_n associated with (6.14) should be again $a(x)(2-2\cos\theta)$. This is in fact the case. Let us sketch the proof. Consider the discretization of (6.14) by the FD scheme defined as follows:

- to approximate the higher-order term $-(a(x)u'(x))'$, use again the FD formula (6.7);
- to approximate the first order term $\beta(x)u'(x)$, use any (consistent) FD formula; a natural choice is the centered formula

$$\beta(x)u'(x)|_{x=x_j} \approx \beta(x_j)\frac{u(x_{j+1}) - u(x_{j-1})}{2h};$$

- to approximate the reaction term $\gamma(x)u(x)$, use the obvious equation

$$\gamma(x)u(x)|_{x=x_j} = \gamma(x_j)u(x_j).$$

Let B_n be the resulting discretization matrix. Then, we can write

$$B_n = A_n + Z_n', \tag{6.15}$$

where Z_n' is the matrix coming from the approximation of the term $\beta(x)u'(x) + \gamma(x)u(x)$. Since $\beta(x)u'(x) + \gamma(x)u(x)$ only involves lower-order derivatives of $u(x)$, it can be shown that $\|Z_n'\| \le C/n$ for some constant C. Hence, $\{Z_n'\}_n$ is zero-distributed, i.e., $\{Z_n'\}_n \sim_{\rm GLT} 0$ by **GLT 4**. In view of (6.12) and (6.15), we have

$$B_n = D_n(a)T_n(2-2\cos\theta) + Z_n + Z_n',$$

and **GLT 2**, **GLT 3**, **GLT 4**, **GLT 6** imply that $\{B_n\}_n \sim_{\rm GLT} a(x)(2-2\cos\theta)$. Alternatively, assuming to know in advance that $\{A_n\}_n \sim_{\rm GLT} a(x)(2-2\cos\theta)$, we could obtain the relation $\{B_n\}_n \sim_{\rm GLT} a(x)(2-2\cos\theta)$ directly from (6.15) and **GLT 6**. Now, if the convection term is not present, i.e. $\beta(x) = 0$ identically, then B_n is symmetric and so, by **GLT 1**, the relation $\{B_n\}_n \sim_{\rm GLT} a(x)(2-2\cos\theta)$ implies $\{B_n\}_n \sim_{\sigma,\lambda} a(x)(2-2\cos\theta)$. If $\beta(x)$ is not identically 0, B_n is not symmetric and the relation $\{B_n\}_n \sim_{\rm GLT} a(x)(2-2\cos\theta)$ only implies $\{B_n\}_n \sim_\sigma a(x)(2-2\cos\theta)$. However, since $\|Z_n'\| \le C/n$, combining **GLT 1** with the result of [30, Theorem 3.4] or [27, Theorem 3.3], one can show that $\{B_n\}_n \sim_\lambda a(x)(2-2\cos\theta)$ even if $\beta(x)$ is an arbitrary (bounded) function.

3. Based on the discussion in item 1, if we change the FD scheme to approximate (6.6) or (6.14), the symbol becomes $a(x)p(\theta)$, where $p(\theta)$ is the trigonometric polynomial associated with the new FD formula used to approximate the second derivative $-u''(x)$ (the higher-order differential operator).

6.2.2 Finite Difference discretization of d-dimensional elliptic PDEs

Let us consider the following second order d-dimensional elliptic problem:

$$\begin{cases} -\nabla \cdot A(\mathbf{x})\nabla u(\mathbf{x}) = f(\mathbf{x}), & \mathbf{x} \in (0,1)^d, \\ u(\mathbf{x}) = 0, & \mathbf{x} \in \partial((0,1)^d), \end{cases} \tag{6.16}$$

where $A : [0,1]^d \to \mathbb{R}^{d\times d}$ is a symmetric positive definite matrix of functions $a_{ij} \in C^1([0,1]^d)$. For simplicity, we only consider the case of a square domain, but the GLT analysis can be extended to the case of an arbitrary domain Ω, provided that Ω can be exactly described by a global geometry map $\mathbf{G} : [0,1]^d \to \overline{\Omega}$; note that this is precisely what happens in the IgA framework [15]. Problem (6.16) can be rewritten in the form

$$\begin{cases} -\displaystyle\sum_{i,j=1}^{d} a_{ij}(\mathbf{x})\frac{\partial^2 u}{\partial x_i \partial x_j}(\mathbf{x}) - \sum_{i,j=1}^{d} \frac{\partial a_{ij}}{\partial x_i}(\mathbf{x})\frac{\partial u}{\partial x_j}(\mathbf{x}) = f(\mathbf{x}), & \mathbf{x} \in (0,1)^d, \\ u(\mathbf{x}) = 0, & \mathbf{x} \in \partial((0,1)^d), \end{cases} \tag{6.17}$$

From (6.17) we see that the higher-order operator in (6.16)–(6.17), containing the second derivatives of $u(\mathbf{x})$, is

$$-\sum_{i,j=1}^{d} a_{ij}(\mathbf{x})\frac{\partial^2 u}{\partial x_i \partial x_j}(\mathbf{x}) = -\mathbf{1}\big(A(\mathbf{x}) \circ Hu(\mathbf{x})\big)\mathbf{1}^T, \tag{6.18}$$

where $\mathbf{1} = (1,\dots,1) \in \mathbb{N}^d$, $Hu(\mathbf{x})$ is the Hessian of $u(\mathbf{x})$, i.e.

$$(Hu(\mathbf{x}))_{ij} = \frac{\partial^2 u}{\partial x_i \partial x_j}(\mathbf{x}),$$

and $\circ$ denotes the componentwise Hadamard product of matrices.

Similarly to the 1-dimensional case, the second order centered FD scheme analogous to the one considered in Section 6.2.1 leads to a sequence of discretization matrices $\{A_{\boldsymbol{n}}\}_n$, where the parameters $\boldsymbol{n} = (n_1,\dots,n_d)$ denote the discretization steps in each direction and are chosen as a function of $n \in \mathbb{N}$. Usually, one gives to every direction the same attention and choose $n_j = n+1$ for all $j = 1,\dots,d$. The matrices $A_{\boldsymbol{n}}$ are symmetric and can be written in the form

$$A_{\boldsymbol{n}} = \sum_{i,j=1}^{d} D_{\boldsymbol{n}}(a_{ij})T_{\boldsymbol{n}}(p_{ij}) + Z_{\boldsymbol{n}},$$

where $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$ and p_{ij} is a (separable) d-variate trigonometric polynomial which only depends on the FD formula used to approximate the second order partial derivative $\frac{\partial^2}{\partial x_i \partial x_j}$. By **GLT 2**, **GLT 3**, **GLT 4** and **GLT 6**, it follows that

$$\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \sum_{i,j=1}^{d} a_{ij}(\mathbf{x})p_{ij}(\boldsymbol{\theta}) = \mathbf{1}\big(A(\mathbf{x}) \circ H(\boldsymbol{\theta})\big)\mathbf{1}^T, \tag{6.19}$$

where

$$H(\boldsymbol{\theta}) = \big[p_{ij}(\boldsymbol{\theta})\big]_{i,j=1}^{d}.$$

The matrix $H(\boldsymbol{\theta})$ collects the trigonometric polynomials associated with the FD formulas used to approximate the second order partial derivatives. Noting the formal analogy between the symbol in (6.19) and the higher-order differential operator (6.18), the matrix $H(\boldsymbol{\theta})$ is sometimes referred to as the 'symbol of the (negative) Hessian operator', although this terminology is not rigorous from the mathematical viewpoint. As a consequence of (6.19), **GLT 1** gives

$$\{A_{\boldsymbol{n}}\}_n \sim_{\sigma,\lambda} \mathbf{1}\big(A(\mathbf{x}) \circ H(\boldsymbol{\theta})\big)\mathbf{1}^T.$$

As in the 1-dimensional case, if we add to the diffusion problem (6.16) a convection term $\boldsymbol{\alpha}(\mathbf{x}) \cdot \nabla u(\mathbf{x})$ and a reaction term $\gamma(\mathbf{x})u(\mathbf{x})$, obtaining a reaction-convection-diffusion equation, the symbol remains the same. Indeed, the symbol is not affected by terms with lower-order derivatives (see item 2 in Remark 6.1).

6.2.3 Schur complements of matrices arising from the Finite Element approximation of a system of PDEs

We consider in this section the FE approximation of a system of PDEs somehow connected with the linear elasticity equations. The resulting discretization matrices show up in saddle point form [7], and we will describe the way to compute the asymptotic spectral distribution of their Schur complements using the theory of GLT sequences. We recall that the Schur complement is a key tool for the numerical treatment of the related linear systems; see [7, Section 5]. The analysis of this section is the same as the analysis in [23, Section 2], with the only difference that the discretization technique considered herein is a pure FE approximation, whereas in [23, Section 2] the authors adopted a mixed FD/FE technique. For an extension of the analysis of this section to the 2-dimensional setting, as well as for the GLT analysis of the linear elasticity equations, we refer the reader to [23].

Consider the system of PDEs

$$\begin{cases} -(a(x)u'(x))' + v'(x) = f(x), & x \in (0,1), \\ -u'(x) - \rho\, v(x) = g(x), & x \in (0,1), \end{cases} \tag{6.20}$$

with homogeneus Dirichlet boundary conditions. Here, $a : [0,1] \to \mathbb{R}$ is a positive continuous function over $[0,1]$ and ρ is a constant. We consider the approximation of (6.20) by linear FEs on the uniform mesh in $[0,1]$ with stepsize $h = \frac{1}{n+1}$. Let us briefly describe this approximation technique; for more details, we refer the reader to [38, Chapter 4]. The weak form of (6.20) reads as follows: find $u, v \in H_0^1([0,1])$ such that, for all $w \in H_0^1([0,1])$,

$$\begin{cases} \int_0^1 a(x)u'(x)w'(x)dx + \int_0^1 v'(x)w(x)dx = \int_0^1 f(x)w(x)dx, \\ -\int_0^1 u'(x)w(x)dx - \rho\int_0^1 v(x)w(x)dx = \int_0^1 g(x)w(x)dx. \end{cases} \tag{6.21}$$

Set $x_j = jh,\ j = 0, \ldots, n+1$, and fix the subspace $\mathscr{W}_n = \mathrm{span}(\varphi_1, \ldots, \varphi_n) \subset H_0^1([0,1])$, where $\varphi_1, \ldots, \varphi_n$ are the so-called 'hat functions'. For every $i = 1, \ldots, n$, the function φ_i is given explicitly by

$$\varphi_i(x) = \frac{x - x_{i-1}}{x_i - x_{i-1}}\chi_{[x_{i-1},x_i)}(x) + \frac{x_{i+1} - x}{x_{i+1} - x_i}\chi_{[x_i,x_{i+1})}(x) = \frac{1}{h}\left[(x - x_{i-1})\chi_{[x_{i-1},x_i)}(x) + (x_{i+1} - x)\chi_{[x_i,x_{i+1})}(x)\right]. \tag{6.22}$$

In the FE approach, we look for approximations $u_{\mathscr{W}_n}, v_{\mathscr{W}_n}$ of u, v by solving the following (Galerkin) problem: find $u_{\mathscr{W}_n}, v_{\mathscr{W}_n} \in \mathscr{W}_n$ such that, for all $w \in \mathscr{W}_n$,

$$\begin{cases} \int_0^1 a(x)u'_{\mathscr{W}_n}(x)w'(x)dx + \int_0^1 v'_{\mathscr{W}_n}(x)w(x)dx = \int_0^1 f(x)w(x)dx, \\ -\int_0^1 u'_{\mathscr{W}_n}(x)w(x)dx - \rho\int_0^1 v_{\mathscr{W}_n}(x)w(x)dx = \int_0^1 g(x)w(x)dx. \end{cases} \tag{6.23}$$

Since $\{\varphi_1, \ldots, \varphi_n\}$ is a basis of $\mathscr{W}_n$, we can write $u_{\mathscr{W}_n} = \sum_{j=1}^n u_j\varphi_j$ and $v_{\mathscr{W}_n} = \sum_{j=1}^n v_j\varphi_j$ for unique vectors $\mathbf{u} = (u_1, \ldots, u_n)^T$ and $\mathbf{v} = (v_1, \ldots, v_n)^T$. By linearity, the computation of $u_{\mathscr{W}_n}, v_{\mathscr{W}_n}$ (i.e., of $\mathbf{u}, \mathbf{v}$) reduces to solving the linear system

$$A_{2n}\begin{bmatrix}\mathbf{u}\\ \mathbf{v}\end{bmatrix} = \begin{bmatrix}\mathbf{f}\\ \mathbf{g}\end{bmatrix}. \tag{6.24}$$

Here, $\mathbf{f} = \left[\int_0^1 f(x)\varphi_i(x)dx\right]_{i=1}^n$, $\mathbf{g} = \left[\int_0^1 g(x)\varphi_i(x)dx\right]_{i=1}^n$, and A_{2n} is the stiffness matrix, which admits the following saddle point structure:

$$A_{2n} = \begin{bmatrix} K_n & H_n \\ H_n^T & -\rho M_n \end{bmatrix},$$

where the blocks K_n, H_n, M_n are square matrices of size n and are given by

$$K_n = \left[\int_0^1 a(x)\varphi_j'(x)\varphi_i'(x)dx\right]_{i,j=1}^n, \tag{6.25}$$

$$H_n = \left[\int_0^1 \varphi_j'(x)\varphi_i(x)dx\right]_{i,j=1}^n = \frac{1}{2}\begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -1 & 0 \end{bmatrix} = -\mathrm{i}\, T_n(\sin\theta), \tag{6.26}$$

$$M_n = \left[\int_0^1 \varphi_j(x)\varphi_i(x)dx\right]_{i,j=1}^n = \frac{h}{6}\begin{bmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 4 \end{bmatrix} = \frac{h}{3}\, T_n(2 + \cos\theta). \tag{6.27}$$

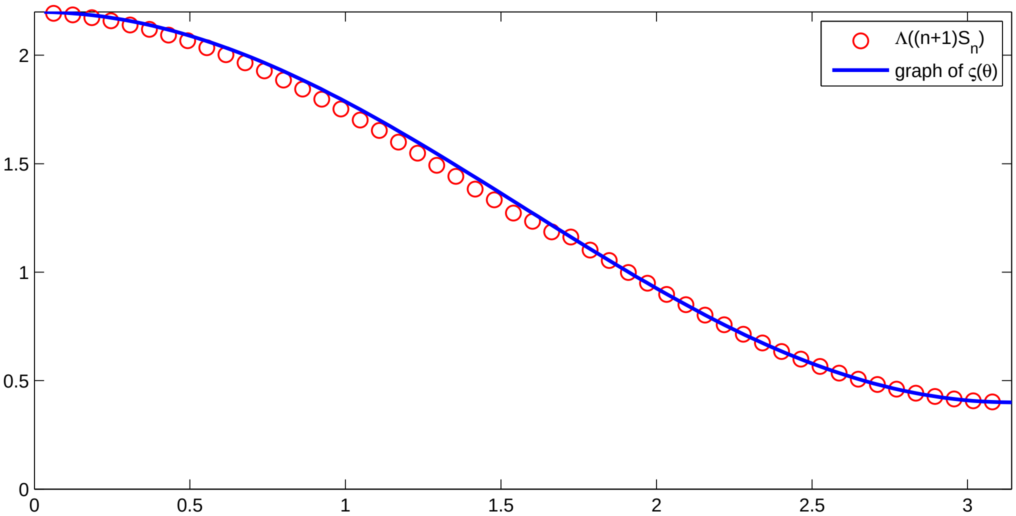

Figure 1: graph of $\varsigma(\theta)$ and spectrum of $(n+1)S_n$ for $n=50$, with $a_0=1,\ \rho=1.2$.

Note that $K_n,\ M_n$ are symmetric positive definite, while H_n is skew-symmetric. In particular, $H_n^T=-H_n=\mathrm{i}\,T_n(\sin\theta)$. In the case where $a(x)$ is constant, say $a(x)=a_0$ identically, the matrix K_n becomes

$$K_n=a_0\left[\int_0^1\varphi_j'(x)\varphi_i'(x)dx\right]_{i,j=1}^n=a_0\,\frac{1}{h}\begin{bmatrix}2&-1&&&\\-1&2&-1&&\\&\ddots&\ddots&\ddots&\\&&-1&2&-1\\&&&-1&2\end{bmatrix}=a_0\,\frac{1}{h}\,T_n(2-2\cos\theta). \tag{6.28}$$

The negative Schur complement of A_{2n} is the symmetric matrix given by

$$S_n=\rho M_n+H_nK_n^{-1}H_n^T. \tag{6.29}$$

In the following, we compute the asymptotic spectral and singular value distribution of the sequence of normalized Schur complements $\{(n+1)S_n\}_n$ using the theory of GLT sequences.

We first consider the case where $a(x)$ is constant, $a(x)=a_0$ identically. In this case, S_n can be expressed in terms of Toeplitz matrices as follows:

$$S_n=\frac{\rho h}{3}\,T_n(2+\cos\theta)+\frac{h}{a_0}\,T_n(\sin\theta)\big[T_n(2-2\cos\theta)\big]^{-1}T_n(\sin\theta). \tag{6.30}$$

According to **GLT 2**, we have $\{T_n(2+\cos\theta)\}_n\sim_{\rm GLT}2+\cos\theta,\ \{T_n(\sin\theta)\}_n\sim_{\rm GLT}\sin\theta$ and $\{T_n(2-2\cos\theta)\}_n\sim_{\rm GLT}2-2\cos\theta$. By **GLT 6 – GLT 7**, taking into account that $2-2\cos\theta\neq0$ a.e. and $(n+1)h=1$, we get

$$\{(n+1)S_n\}_n\sim_{\rm GLT}\varsigma(\theta)=\frac{\rho}{3}(2+\cos\theta)+\frac{1}{a_0}\,\frac{\sin^2\theta}{2-2\cos\theta}. \tag{6.31}$$

Since S_n is symmetric, (6.31) and **GLT 1** imply that $\{S_n\}_n\sim_{\sigma,\lambda}\varsigma(\theta)$. Considering that $\varsigma(\theta)$ is an even trigonometric polynomial, the relations $\{S_n\}_n\sim_{\sigma,\lambda}\varsigma(\theta)$ continue to hold if we consider $[0,\pi]$ as the domain of $\varsigma(\theta)$ instead of $[0,1]\times[-\pi,\pi]$. In particular, according to Remark 2.2, for large n the eigenvalues of $(n+1)S_n$ are approximately given by $\varsigma(\frac{j\pi}{n+1}),\ j=1,\ldots,n$. This theoretical forecast is confirmed by Figure 1, where we fixed $a_0=1$ and $\rho=1.2$, and we plotted the graph of $\varsigma(\theta)$ together with the spectrum of $(n+1)S_n$ for $n=50$. Note that, in the figure, the spectrum of $(n+1)S_n$ is represented by the pairs $(\frac{j\pi}{n+1},\lambda_j((n+1)S_n)),\ j=1,\ldots,n$, where the eigenvalues $\lambda_j((n+1)S_n)$ are labeled in decreasing order, according to our notational convention (see Section 2.1).

Now we consider the case where $a(x)$ is not constant. In this case, eq. (6.28) does not hold and K_n is no longer a Toeplitz matrix. However, as proved below, $\{\frac{1}{n+1}K_n\}_n$ is a GLT sequence with symbol $a(x)(2-2\cos\theta)$. Therefore, using the expression

$$(n+1)S_n=\frac{\rho}{3}\,T_n(2+\cos\theta)+T_n(\sin\theta)\,(n+1)K_n^{-1}\,T_n(\sin\theta),$$

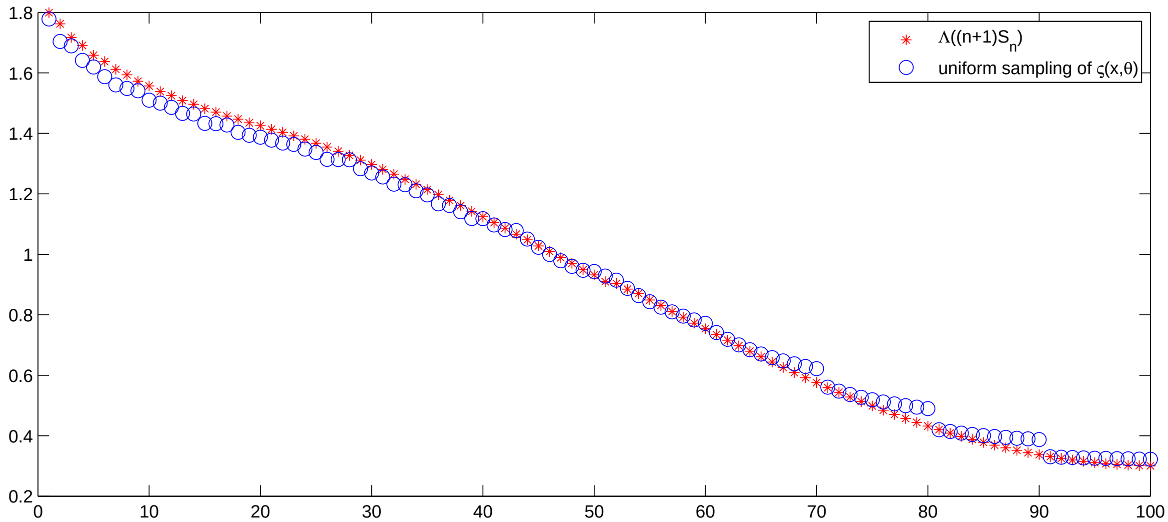

Figure 2: spectrum of $(n+1)S_n$ and samplings $\varsigma(\frac{i}{\ell}, \frac{j\pi}{\ell+1}),\ i,j=1,\ldots,\ell$, for $n=\ell^2=100,\ a(x)=1+x,\ \rho=0.9$.

and noting that $a(x)(2-2\cos\theta)\neq 0$ a.e., from **GLT 2** and **GLT 6 – GLT 7** we obtain

$$\{(n+1)S_n\}_n \sim_{\rm GLT} \varsigma(x,\theta)=\frac{\rho}{3}(2+\cos\theta)+\frac{\sin^2\theta}{a(x)(2-2\cos\theta)}. \tag{6.32}$$

As a consequence, by **GLT 1** we get

$$\{(n+1)S_n\}_n \sim_{\sigma,\lambda} \varsigma(x,\theta), \tag{6.33}$$

and these relations continue to hold even if we consider $[0,1]\times[0,\pi]$ as the domain of $\varsigma(x,\theta)$ instead of $[0,1]\times[-\pi,\pi]$, because $\varsigma(x,\theta)$ is symmetric with respect to the Fourier variable θ. According to Remark 2.2, if $n=\ell^2$ is large enough, the eigenvalues of $(n+1)S_n$ are approximately given by the uniform sampling $\varsigma(\frac{i}{\ell}, \frac{j\pi}{\ell+1}),\ i,j=1,\ldots,\ell$. This is confirmed by Figure 2, where we fixed $a(x)=1+x$ and $\rho=0.9$, and we plotted the spectrum of $(n+1)S_n$ together with the values $\varsigma(\frac{i}{\ell}, \frac{j\pi}{\ell+1}),\ i,j=1,\ldots,\ell$, for $n=\ell^2=100$. Note that, in the figure, the eigenvalues of $(n+1)S_n$, as well as the samplings of the symbol $\varsigma(x,\theta)$, are depicted in decreasing order.

Proof of the relation $\{\frac{1}{n+1}K_n\}_n \sim_{\rm GLT} a(x)(2-2\cos\theta)$. We first illustrate the idea of the proof, and then we go into the details. The proof is based on the fact that the hat functions (6.22) are 'locally supported'. Indeed, the support $[x_{i-1},x_{i+1}]$ of the i-th hat function $\varphi_i(x)$ is localized near the point $\hat{x}_i=\frac{i}{n}\in[x_i,x_{i+1}]$, and the amplitude of the support tends to 0 when $n\to\infty$. Since $a(x)$ varies continuously over $[0,1]$, the (i,j)-th entry of K_n can be approximated as follows, for all $i,j=1,\ldots,n$:

$$(K_n)_{ij}=\int_0^1 a(x)\varphi_j'(x)\varphi_i'(x)dx=\int_{x_{i-1}}^{x_{i+1}} a(x)\varphi_j'(x)\varphi_i'(x)dx\approx a(\hat{x}_i)\int_{x_{i-1}}^{x_{i+1}}\varphi_j'(x)\varphi_i'(x)dx=a(\hat{x}_i)\int_0^1\varphi_j'(x)\varphi_i'(x)dx. \tag{6.34}$$

After normalization, we can rewrite (6.34) in matrix form,

$$\frac{1}{n+1}K_n\approx D_n(a)\,\frac{1}{n+1}\left[\int_0^1\varphi_j'(x)\varphi_i'(x)dx\right]_{i,j=1}^n=D_n(a)T_n(2-2\cos\theta). \tag{6.35}$$

As we shall see, the approximation (6.35) implies that the matrix-sequence $\{\frac{1}{n+1}K_n-D_n(a)T_n(2-2\cos\theta)\}_n$ is zero-distributed, and so the relation $\{\frac{1}{n+1}K_n\}_n\sim_{\rm GLT} a(x)(2-2\cos\theta)$ follows from **GLT 2**, **GLT 3**, **GLT 4** and **GLT 6**.

Now let us go into the details. Since $|\varphi_i'(x)|\le n+1$ and $\left[\int_0^1\varphi_j'(x)\varphi_i'(x)dx\right]_{i,j=1}^n=(n+1)T_n(2-2\cos\theta)$, for all $i,j=1,\ldots,n$ we have

$$\begin{aligned}|(K_n)_{ij}-((n+1)D_n(a)T_n(2-2\cos\theta))_{ij}|&=\left|\int_0^1\big[a(x)-a(\hat{x}_i)\big]\varphi_j'(x)\varphi_i'(x)dx\right|\le(n+1)^2\int_{x_{i-1}}^{x_{i+1}}|a(x)-a(\hat{x}_i)|dx\\&\le(n+1)^2\,\omega_a\Big(\frac{2}{n+1}\Big)\int_{x_{i-1}}^{x_{i+1}}dx=2(n+1)\,\omega_a\Big(\frac{2}{n+1}\Big),\end{aligned}$$

where $\omega_a(\cdot)$ is the modulus of continuity of a. It follows that each component of the matrix $\frac{1}{n+1}K_n - D_n(a)T_n(2-2\cos\theta)$ is bounded from above in modulus by $2\,\omega_a(\frac{2}{n+1})$. Moreover, $\frac{1}{n+1}K_n - D_n(a)T_n(2-2\cos\theta)$ is banded (actually, tridiagonal), because $\mathrm{supp}(\varphi_i)\cap\mathrm{supp}(\varphi_j)=\emptyset$ whenever $|i-j|>1$. Thus,

$$\left|\frac{1}{n+1}K_n - D_n(a)T_n(2-2\cos\theta)\right|_1,\ \left|\frac{1}{n+1}K_n - D_n(a)T_n(2-2\cos\theta)\right|_\infty \le 6\,\omega_a\Big(\frac{2}{n+1}\Big),$$

and so, by (2.16), $\|\frac{1}{n+1}K_n - D_n(a)T_n(2-2\cos\theta)\| \le 6\,\omega_a(\frac{2}{n+1}) \to 0$ as $n\to\infty$. By Theorem 2.6, we conclude that the matrix-sequence $\{\frac{1}{n+1}K_n - D_n(a)T_n(2-2\cos\theta)\}_n$ is zero-distributed, and this implies $\{\frac{1}{n+1}K_n\}_n \sim_{\mathrm{GLT}} a(x)(2-2\cos\theta)$ by **GLT 2**, **GLT 3**, **GLT 4** and **GLT 6**. □

Remark 6.2. The relations (6.32)–(6.33) continue to hold even if $a(x)$ is only assumed to be in $L^\infty([0,1])$, because the relation $\{\frac{1}{n+1}K_n\}_n \sim_{\mathrm{GLT}} a(x)(2-2\cos\theta)$ remains true if $a(x)$ belongs to $L^\infty([0,1])$. To prove the latter statement, let us first introduce a more precise notation and denote by $K_n(a)$ the matrix K_n obtained from the FE approximation of (6.20); in this way, the dependence of K_n on the function $a(x)$ is highlighted. We want to show that $\{\frac{1}{n+1}K_n(a)\}_n \sim_{\mathrm{GLT}} a(x)(2-2\cos\theta)$ whenever $a(x)$ belongs to $L^\infty([0,1])$. By the Lusin theorem, there exists a sequence of continuous functions $a_m\in C([0,1])$ such that $\mu_1\{a_m\ne a\}\le\frac{1}{m}$ and $\|a_m\|_\infty \le \|a\|_{L^\infty}$. By the previous proof, $\{\frac{1}{n+1}K_n(a_m)\}_n \sim_{\mathrm{GLT}} a_m(x)(2-2\cos\theta)$ and, moreover, $a_m(x)(2-2\cos\theta)\to a(x)(2-2\cos\theta)$ in measure. In addition, $\{\{\frac{1}{n+1}K_n(a_m)\}_n\}_m$ is an a.c.s. for $\{\frac{1}{n+1}K_n(a)\}_n$. Indeed, using (2.14) and observing that $\sum_{i=1}^n|\varphi_i'(x)|\le 2(n+1)$ for all $x\in[0,1]$, we obtain

$$\begin{aligned}\|K_n(a)-K_n(a_m)\|_1 &\le \sum_{i,j=1}^n |(K_n(a))_{ij}-(K_n(a_m))_{ij}| = \sum_{i,j=1}^n \left|\int_0^1 \big[a(x)-a_m(x)\big]\varphi_j'(x)\varphi_i'(x)dx\right| \\ &\le 2\|a\|_{L^\infty}\int_{\{a\ne a_m\}}\sum_{i,j=1}^n |\varphi_j'(x)|\,|\varphi_i'(x)|dx \le 8(n+1)^2\|a\|_{L^\infty}\mu_1\{a\ne a_m\} \le \frac{8(n+1)^2\|a\|_{L^\infty}}{m},\end{aligned}$$

and so

$$\left\|\frac{1}{n+1}K_n(a)-\frac{1}{n+1}K_n(a_m)\right\|_1 \le C\,\frac{n}{m}$$

for some constant C independent of n and m. Thus, $\{\{\frac{1}{n+1}K_n(a_m)\}_n\}_m$ is an a.c.s. for $\{\frac{1}{n+1}K_n(a)\}_n$ by Corollary 3.3, and the relation $\{\frac{1}{n+1}K_n(a)\}_n \sim_{\mathrm{GLT}} a(x)(2-2\cos\theta)$ follows from **GLT 9**.

7 Conclusions and future works

In this work, we fully developed the theory of GLT sequences. We made a significant review of the original theory, by 'correcting' all the relevant definitions and by generalizing and/or simplifying a lot of key results. We also extended the theory itself: the main novelties of this work are the content of Sections 5.3–5.4, Theorem 5.7 and the new proof of Theorem 5.8. Finally, we provided a precise summary of the theory of GLT sequences in Section 5.6, with the purpose of giving to the reader an easy-to-use GLT manual. Some hints on how to use this manual in practical applications were given in Section 6.

We conclude this work with a list of possible future lines of research.

1. Prove (or disprove) the result mentioned at the end of Section 6.1, which we report here for the reader's convenience.

   Conjecture. Suppose that $\{A_{\boldsymbol n}^{(i)}\}_n \sim_{\mathrm{GLT}} \kappa_i$ for $i=1,\ldots,k$, where $A_{\boldsymbol n}^{(1)},\ldots,A_{\boldsymbol n}^{(k)}$ are Hermitian positive definite, and let $G(A_{\boldsymbol n}^{(1)},\ldots,A_{\boldsymbol n}^{(k)})$ be the geometric (Karcher) mean of $A_{\boldsymbol n}^{(1)},\ldots,A_{\boldsymbol n}^{(k)}$. Then $\{G(A_{\boldsymbol n}^{(1)},\ldots,A_{\boldsymbol n}^{(k)})\}_n \sim_{\mathrm{GLT}} (\kappa_1,\ldots,\kappa_k)^{1/k}$. In particular, $\{G(A_{\boldsymbol n}^{(1)},\ldots,A_{\boldsymbol n}^{(k)})\}_n \sim_{\sigma,\lambda} (\kappa_1,\ldots,\kappa_k)^{1/k}$.

2. Try to understand if the a.c.s. notion is related to some topology $\tau_{\text{a.c.s.}}$ defined on the algebra of matrix-sequences. If the answer is a affirmative, try to give a topological characterization of the sub-algebra formed by GLT sequences. Let us be more precise. Fix a sequence of d-indices $\{\boldsymbol n=\boldsymbol n(n)\}_n\subseteq\mathbb N^d$ such that $\boldsymbol n\to\infty$ as $n\to\infty$, and consider the complex *-algebra of all matrix-sequences

   $$\mathscr E = \big\{\{A_{\boldsymbol n}\}_n : \{A_{\boldsymbol n}\}_n \text{ is a matrix-sequence}\big\}. \tag{7.1}$$

   If $\{A_{\boldsymbol n}\}_n, \{B_{\boldsymbol n,m}\}_n\in\mathscr E$, we say that $\{B_{\boldsymbol n,m}\}_n$ converges to $\{A_{\boldsymbol n}\}_n$ when $m\to\infty$ if $\{\{B_{\boldsymbol n,m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol n}\}_n$. Note that this statement could be meaningless, because noone proved that the 'a.c.s. convergence' is a true notion of convergence. In other words, noone proved that there exists a topology on $\mathscr E$ inducing this kind of convergence. Trying to understand if such a topology exists could be an interesting theoretical subject for future research. Assuming for the moment that such a topology $\tau_{\text{a.c.s.}}$ exists, it is easy to see that $\tau_{\text{a.c.s.}}$ is not 'well-bahaved'. For instance, the limit of a converging sequence $\{\{B_{\boldsymbol n,m}\}_n\}_m$ is not unique in general. Indeed, any fixed sequence $\{B_{\boldsymbol{n}}\}_n$ converges to an infinite number of sequences, because $\{B_{\boldsymbol{n}}\}_n$ is an a.c.s. of $\{B_{\boldsymbol{n}}+Z_{\boldsymbol{n}}\}_n$ if (and only if) $\{Z_{\boldsymbol{n}}\}_n$ is zero-distributed (see Theorem 4.4). As a consequence, $\{A_{\boldsymbol{n}}\}_n \to \{A_{\boldsymbol{n}}+Z_{\boldsymbol{n}}\}_n$ for all zero-distributed sequences $\{Z_{\boldsymbol{n}}\}_n$. This implies that every open set in $\tau_{\text{a.c.s.}}$ which contains a sequence $\{A_{\boldsymbol{n}}\}_n$, actually contains all the sequences $\{A_{\boldsymbol{n}}+Z_{\boldsymbol{n}}\}_n$ such that $\{Z_{\boldsymbol{n}}\}_n$ is zero-distributed. In particular, $(\mathscr{E},\tau_{\text{a.c.s.}})$ is not an Housdorff space and so, a fortiori, it is not a topological vector space [40]. This means that the classical Functional Analysis [40], which starts from the notion of topological vector space, finds no applications in our context. Despite of this, if $\tau_{\text{a.c.s.}}$ exists, it would be interesting to provide a topological characterization of the GLT algebra

$$\mathscr{A}=\big\{\{A_{\boldsymbol{n}}\}_n:\ \{A_{\boldsymbol{n}}\}_n\sim_{\rm GLT}\kappa\ \text{ for some measurable function } \kappa:[0,1]^d\times[-\pi,\pi]^d\to\mathbb{C}\big\}\subseteq\mathscr{E}.$$

In particular, is the GLT algebra $\mathscr{A}$ a sequentially closed subset of $\mathscr{E}$? Is it closed? If it is not closed, what is its closure?

3. Develop the theory of block GLT sequences. Multilevel block Toeplitz matrices, defined, e.g., in [51] or [24, Section 1.4.1], naturally arise in the numerical approximation of constant coefficient systems of PDEs and, surprisingly enough, also in the Finite Element approximation of constant coefficient elliptic problems; see [28] or [24, Chapter 3]. In the case of non-constant coefficients, the counterpart of (multilevel) block Toeplitz sequences would be block GLT sequences, in the same way as GLT sequences are the counterpart of (multilevel) Toeplitz sequences; to understand better this point, examine carefully the procedure followed in Section 6.2.1 (*Method 1*). Hence, a theory of block GLT sequences, already mentioned in [45, Section 3.3], would be the ideal framework to deal with variable coefficient differential problems approximated by Finite Elements of any regularity, as well as to face the linear systems of non-constant coefficient PDEs. We note that the first step of such a theory has already been made. Indeed, thanks to the work of Böttcher, Silbermann, Miranda, and Tilli (see [14, 36, 51]), we know that any sequence of multilevel block Toeplitz matrices $\{T_{\boldsymbol{n}}(\mathbf{f})\}_n$, generated by a multivariate matrix-valued function $\mathbf{f}:[-\pi,\pi]^d\to\mathbb{C}^{s\times s}$ with components $f_{ij}\in L^1([-\pi,\pi]^d)$, has an asymptotic spectral and singular value distribution described by $\mathbf{f}$; we refer the reader to [51] for the meaning of this statement. In particular, in [14, p. 202] and [36] the monolevel block case was disposed of, and Tilli [51] finally proved the result in the multilevel block case under the sole assumption that the entries of $\mathbf{f}$ are in $L^1([-\pi,\pi]^d)$.

4. Revisit (and extend) the work of [45, Section 3.1.4] concerning *reduced* GLT sequences. A suitable theory of reduced GLT sequences would allow one to deal with sequences of discretization matrices associated with the Finite Element approximation of PDEs defined on non-rectangular domains Ω. Actually, also the theory of GLT sequences allows one to deal with such sequences, but under the additional assumption that: (a) the non-rectangular domain Ω is exactly decribed by a regular geometry map $\mathbf{G}:\widehat{\Omega}\to\overline{\Omega}$ defined on a rectangular domain $\widehat{\Omega}$; (b) the basis functions used in the Finite Element approximation are defined as the '$\mathbf{G}$-deformations' of basis functions defined over $\widehat{\Omega}$. The assumptions (a) and (b) are verified in the IgA approach [15], but not in the general Finite Element setting.

5. Try to design an automatic procedure for computing the symbol of a sequence of PDE discretization matrices, assuming to know that it is a GLT sequence. The idea would be to express the symbol as a function of the higher-order differential operator associated with the PDE, of the related coefficient, and of the used approximation technique. Some hints in this direction are given in [44, Section 2] and [45, Question 3.1].

References

[1] AL-FHAID A. S., SERRA-CAPIZZANO S., SESANA D., ULLAH M. Z. *Singular-value (and eigenvalue) distribution and Krylov preconditioning of sequences of sampling matrices approximating integral operators.* Numer. Linear Algebra Appl. **21** (2014) 722–743.

[2] ANDO T., LI C. K., MATHIAS R. *Geometric means.* Linear Algebra Appl. **385** (2004) 305–334.

[3] BARBARESCO F. *Robust statistical radar processing in Fréchet metric space: OS-HDR-CFAR and OS-STAP processing in Siegel homogeneous bounded domains.* Radar Symposium (IRS), 2011 Proceedings International, IEEE (2011) 639–644.

[4] BARBARESCO F. *Information geometry of covariance matrix: Cartan-Siegel homogeneous bounded domains, Mostow/Berger fibration and Fréchet median.* Chapter 9 of the book *Matrix Information Geometry*, edited by F. Nielsen and R. Bhatia, Springer-Verlag Berlin Heidelberg (2013).

[5] BECKERMANN B., KUIJLAARS A. B. J. *Superlinear convergence of Conjugate Gradients.* SIAM J. Numer. Anal. **39** (2001) 300–329.

[6] BECKERMANN B., SERRA-CAPIZZANO S. *On the asymptotic spectrum of Finite Element matrix sequences.* SIAM J. Numer. Anal. **45** (2007) 746–769.

[7] BENZI M., GOLUB G. H., LIESEN J. *Numerical solution of saddle point problems.* Acta Numerica **14** (2005) 1–137.

[8] BHATIA R. *Matrix Analysis.* Springer-Verlag, New York (1997).

[9] BHATIA R. *Positive definite matrices.* Princeton Series in Applied Mathematics, Princeton University Press, New Jersey (2007).

[10] BHATIA R., HOLBROOK J. *Riemannian geometry and matrix geometric means.* Linear Algebra Appl. **413** (2006) 594–618.

[11] BHATIA R., KARANDIKAR R. L. *Monotonicity of the matrix geometric mean.* Math. Ann. **353** (2012) 1453–1467.

[12] BINI D. A., MEINI B., POLONI F. *An effective matrix geometric mean satisfying the Ando-Li-Mathias properties.* Math. Comp. **79** (2010) 437–452.

[13] BÖTTCHER A., GRUDSKY S. M. *Spectral properties of banded Toeplitz matrices.* SIAM, Philadelphia (2005).

[14] BÖTTCHER A., SILBERMANN B. *Introduction to large truncated Toeplitz matrices.* Springer-Verlag, New York (1999).

[15] COTTRELL J. A., HUGHES T. J. R., BAZILEVS Y. *Isogeometric Analysis: toward integration of CAD and FEA.* John Wiley & Sons, Chichester (2009).

[16] DONATELLI M., GARONI C., MANNI C., SERRA-CAPIZZANO S., SPELEERS H. *Symbol-based multigrid methods for Galerkin B-spline Isogeometric Analysis.* Technical Report TW650, Department of Computer Science, KU Leuven (2014) full text available at:
http://www2.cs.kuleuven.be/publicaties/rapporten/tw/TW650.abs.html

[17] DONATELLI M., GARONI C., MANNI C., SERRA-CAPIZZANO S., SPELEERS H. *Robust and optimal multi-iterative techniques for IgA Galerkin linear systems.* Comput. Methods Appl. Mech. Engrg. **284** (2015) 230–264.

[18] DONATELLI M., GARONI C., MANNI C., SERRA-CAPIZZANO S., SPELEERS H. *Robust and optimal multi-iterative techniques for IgA collocation linear systems.* Comput. Methods Appl. Mech. Engrg. **284** (2015) 1120–1146.

[19] DONATELLI M., GARONI C., MANNI C., SERRA-CAPIZZANO S., SPELEERS H. *Spectral analysis and spectral symbol of matrices in isogeometric collocation methods.* Math. Comp. (in press).

[20] DONATELLI M., GARONI C., MAZZA M., SERRA-CAPIZZANO S., SESANA D. *Spectral behavior of preconditioned non-Hermitian multilevel block Toeplitz matrices with matrix-valued symbol.* Appl. Math. Comput. **245** (2014) 158–173.

[21] DONATELLI M., GARONI C., MAZZA M., SERRA-CAPIZZANO S., SESANA D. *Preconditioned HSS method for large multilevel block Toeplitz linear systems via the notion of matrix-valued symbol.* Numer. Linear Algebra Appl. (in press).

[22] DONATELLI M., NEYTCHEVA M., SERRA-CAPIZZANO S. *Canonical eigenvalue distribution of multilevel block Toeplitz sequences with non-Hermitian symbols.* Oper. Theory Adv. Appl. **221** (2012) 269–291.

[23] DOROSTKAR A., NEYTCHEVA M., SERRA-CAPIZZANO S. *Spectral analysis of coupled PDEs and of their Schur complements via the notion of Generalized Locally Toeplitz sequences.* Technical Report 2015-008, Department of Information Technology, Uppsala University, Sweden (2015) full text available at: http://www.it.uu.se/research/publications/reports/2015-008/

[24] GARONI C. *Structured matrices coming from PDE Approximation Theory: spectral analysis, spectral symbol and design of fast iterative solvers.* Ph.D. Thesis in Mathematics of Computation, University of Insubria, Como, Italy (2014) full text available at: http://hdl.handle.net/10277/568

[25] GARONI C., MANNI C., SERRA-CAPIZZANO S., SESANA D., SPELEERS H. *Spectral analysis and spectral symbol of matrices in isogeometric Galerkin methods.* Technical Report 2015-005, Department of Information Technology, Uppsala University, Sweden (2015) full text available at:
http://www.it.uu.se/research/publications/reports/2015-005/

[26] GARONI C., MANNI C., SERRA-CAPIZZANO S., SESANA D., SPELEERS H. *Lusin theorem, GLT sequences and matrix computations: an application to the spectral analysis of PDE discretization matrices.* Technical Report 2015-012, Department of Information Technology, Uppsala University (2015) full text available at: http://www.it.uu.se/research/publications/reports/2015-012/

[27] GARONI C., SERRA-CAPIZZANO S., SESANA D. *Tools for determining the asymptotic spectral distribution of non-Hermitian perturbations of Hermitian matrix-sequences and applications.* Integr. Equ. Oper. Theory **81** (2015) 213–225.

[28] GARONI C., SERRA-CAPIZZANO S., SESANA D. *Spectral analysis and spectral symbol of* d*-variate* $\mathbb{Q}_{\boldsymbol{p}}$ *Lagrangian FEM stiffness matrices.* SIAM J. Matrix Anal. Appl. **36** (2015) 1100–1128.

[29] Garoni C., Serra-Capizzano S., Vassalos P. *A general tool for determining the asymptotic spectral distribution of Hermitian matrix-sequences.* Oper. Matrices **9** (2015) 549–561.

[30] Golinskii L., Serra-Capizzano S. *The asymptotic properties of the spectrum of nonsymmetrically perturbed Jacobi matrix sequences.* J. Approx. Theory **144** (2007) 84–102.

[31] Grenander U., Szegö G. *Toeplitz forms and their applications.* Second Edition, AMS Chelsea Publishing, New York (1984).

[32] Holbrook J. *No dice: A deterministic approach to the Cartan centroid.* J. Ramanujan Math. Soc. **27** (2012) 509–521.

[33] Hughes T. J. R., Cottrell J. A., Bazilevs Y. *Isogeometric Analysis: CAD, Finite Elements, NURBS, exact geometry and mesh refinement.* Comput. Methods Appl. Mech. Engrg. **194** (2005) 4135–4195.

[34] Lapuyade-Lahorgue J., Barbaresco F. *Radar detection using Siegel distance between autoregressive processes, application to HF and X-band radar.* Radar Conference, 2008. RADAR '08. IEEE (2008) 1–6.

[35] Lawson J., Lim Y. *Monotonic properties of the least squares mean.* Math. Ann. **351** (2011) 267–279.

[36] Miranda M., Tilli P. *Asymptotic spectra of Hermitian block Toeplitz matrices and preconditioning results.* SIAM J. Matrix Anal. Appl. **21** (2000) 867–881.

[37] Nakamura N. *Geometric means of positive operators.* Kyungpook Math. J. **49** (2009) 167–181.

[38] Quarteroni A. *Numerical models for differential problems.* Second Edition, Springer-Verlag Italia, Milan (2014).

[39] Rudin W. *Real and complex analysis.* Third Edition, McGraw-Hill, Singapore (1987).

[40] Rudin W. *Functional Analysis.* Second Edition, McGraw-Hill, Singapore (1991).

[41] Salinelli E., Serra-Capizzano S., Sesana D. *Eigenvalue-eigenvector structure of Schoenmakers–Coffey matrices via Toeplitz technology and applications.* Linear Algebra Appl. http://dx.doi.org/10.1016/j.laa.2015.03.017

[42] Serra-Capizzano S. *Distribution results on the algebra generated by Toeplitz sequences: a finite dimensional approach.* Linear Algebra Appl. **328** (2001) 121–130.

[43] Serra-Capizzano S. *More inequalities and asymptotics for matrix valued linear positive operators: the noncommutative case.* Oper. Theory Adv. Appl. **135** (2002) 293–315.

[44] Serra-Capizzano S. *Generalized Locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations.* Linear Algebra Appl. **366** (2003) 371–402.

[45] Serra-Capizzano S. *The GLT class as a generalized Fourier Analysis and applications.* Linear Algebra Appl. **419** (2006) 180–233.

[46] Serra-Capizzano S., Sesana D. *Approximating classes of sequences: the Hermitian case.* Linear Algebra Appl. **434** (2011) 1163–1170.

[47] Serra-Capizzano S., Tablino-Possio C. *Analysis of preconditioning strategies for collocation linear systems.* Linear Algebra Appl. **369** (2003) 41–75.

[48] Serra-Capizzano S., Tilli P. *On unitarily invariant norms of matrix-valued linear positive operators.* J. Inequal. Appl. **7** (2002) 309–330.

[49] Sesana D. *Spectral distributions of structured matrix-sequences: tools and applications.* Ph.D. Thesis in Mathematics of Computation, University of Insubria, Como, Italy (2010) full text available at: http://hdl.handle.net/10277/318

[50] Smith G. D. *Numerical solution of partial differential equations: Finite Difference methods.* Third Edition, Oxford University Press, New York (1985).

[51] Tilli P. *A note on the spectral distribution of Toeplitz matrices.* Linear and Multilinear Algebra **45** (1998) 147–159.

[52] Tilli P. *Locally Toeplitz sequences: spectral properties and applications.* Linear Algebra Appl. **278** (1998) 91–120.

[53] Tilli P. *Some results on complex Toeplitz eigenvalues.* J. Math. Anal. Appl. **239** (1999) 390–401.

[54] Tyrtyshnikov E. E. *A unifying approach to some old and new theorems on distribution and clustering.* Linear Algebra Appl. **232** (1996) 1–43.

[55] Zamarashkin N. L., Tyrtyshnikov E. E. *Distribution of eigenvalues and singular values of Toeplitz matrices under weakened conditions on the generating function.* Sb. Math. **188** (1997) 1191–1201.