Generalized Locally Toeplitz sequences: a review and an extension

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.

May 18, 2015

Abstract

We review the theory of Generalized Locally Toeplitz (GLT) sequences, hereinafter called 'the GLT theory', 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: every GLT sequence has a measurable symbol; the singular value distrbution of any GLT sequence is identified by the symbol (also the eigenvalue distribution if the sequence is made by Hermitian matrices); the GLT sequences form an algebra, closed under linear combinations, (pseudo)-inverse if the symbol vanishes in a set of zero measure, product and the symbol obeys to the same algebraic manipulations.

As already proved in several contexts, this theory is a powerful tool for computing/analyzing the asymptotic spectral distribution of the discretization matrices arising from the numerical approximation of continuous problems, such as Integral Equations and, especially, Partial Differential Equations, including variable coefficients, irregular domains, different approximation schemes such as Finite Differences, Finite Elements, Collocation/Galerkin Isogeometric Analysis etc. However, in this review we are not concerned with the applicative interest of the GLT theory, for which we limit to refer the reader to the numerous applications available in the literature. On the contrary, we focus on the theoretical foundations. We propose slight (but relevant) modifications of the original definitions, which allow us to enlarge the applicability of the GLT theory. In particular, we remove a certain '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. With the new definitions, we introduce new technical and useful results and we provide a formal proof of the fact that sequences formed by multilevel diagonal sampling matrices, as well as multilevel Toeplitz sequences, fall in the class of LT sequences; the latter results were mentioned in previous papers, but no direct proof was given especially regarding the case of multilevel diagonal sampling matrix-sequences. As a final step, we extend the GLT theory: 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 latter theoretical property has important implications, e.g. in proving that the geometric means of GLT sequences are still GLT, so obtaining for free that the spectral distribution of the mean is just the geometric mean of the symbols.

Contents

1 Introduction

We review, improve, and extend the theory of Generalized Locally Toeplitz (GLT) sequences, hereinafter called ‘the GLT theory’, which stems from Tilli’s work on Locally Toeplitz (LT) sequences [54] and from the theory of classical Toeplitz operators [8, 31, 53], and was developed by the second author in [45, 46].

As already proved in several contexts, this theory is a powerful tool 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, so that we are left with a sequence of discretization matrices $\{A_n\}_n$ with increasing size. What is often verified in practice is that $\{A_n\}_n$ 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 relations hold (see Subsection 2.4):

$$\lim_{n\to\infty} \frac{1}{d_n} \sum_{j=1}^{d_n} F(\lambda_j(A_n)) = \frac{1}{\mu_\delta(D)} \int_D F(\kappa(\mathbf{y}))d\mathbf{y}, \tag{1.1}$$

$$\lim_{n\to\infty} \frac{1}{d_n} \sum_{j=1}^{d_n} F(\sigma_j(A_n)) = \frac{1}{\mu_\delta(D)} \int_D F(|\kappa(\mathbf{y})|)d\mathbf{y}, \tag{1.2}$$

where $\lambda_j(A_n),\ j = 1, \ldots, d_n$, are the eigenvalues of A_n, $\sigma_j(A_n),\ j = 1, \ldots, d_n$, are the singular values of A_n, μ_δ is the Lebesgue measure in $\mathbb{R}^\delta$, $\mu_\delta(D) \in (0, \infty)$, and $\kappa : D \subset \mathbb{R}^\delta \to \mathbb{C}$ is a measurable function. If relations (1.1) hold, then κ is called the spectral symbol of the sequence of matrices $\{A_n\}_n$ and we write $\{A_n\}_n \sim_\lambda (\kappa, D)$; in such a case, κ provides a 'compact' description of the asymptotic spectral distribution of $\{A_n\}_n$. If relations (1.2) hold, then κ is called the singular value symbol of the sequence of matrices $\{A_n\}_n$; in this setting, the function $|\kappa|$ provides a 'compact' description of the asymptotic singular value distribution of $\{A_n\}_n$ and we write $\{A_n\}_n \sim_\sigma (\kappa, D)$.

Often, for the sake of of brevity and when the information is clear from the context, we will write $\{A_n\}_n \sim_\sigma \kappa$ in place of $\{A_n\}_n \sim_\sigma (\kappa, D)$ and $\{A_n\}_n \sim_\lambda \kappa$ in place of $\{A_n\}_n \sim_\lambda (\kappa, D)$, respectively.

The informal, but important, meaning of (1.1) can be given as follows: if κ is continuous, then a suitable ordering of the eigenvalues $\{\lambda_j(A_n)\}$, in correspondence with a equispaced gridding on D, reconstructs approximately the surface $t \to \kappa(t)$. On the other hand, the informal meaning of (1.2) can be summarized in perfect analogy: if κ is continuous, then a suitable ordering of the singular values $\{\sigma_j(A_n)\}$, in correspondence with a equispaced gridding on D, reconstructs approximately the surface $t \to |\kappa(t)|$.

The GLT theory, in combination with the result of [29, Theorem 3.4] concerning the spectral distribution of perturbed sequences of matrices, is one of the most powerful and successful tools for computing the spectral symbol κ for wide classes of approximated IEs and PDEs. Indeed, the sequence $\{A_n\}_n$ is often a GLT sequence. We refer the reader to [45, 46] for applications of the GLT theory in the context of Finite Difference discretizations of PDEs; to [6, 46, 49] for the Finite Element and Finite Difference collocation settings; to [1, 38] for the GLT aaproach to IE matrix-sequences; and to [15, 24, 25] for recent applications to the case of B-spline Isogeometric Analysis (IgA) approximations of PDEs, both in the collocation and Galerkin frameworks. We remind at this point that IgA is a modern paradigm introduced in [12, 33] for analyzing problems governed by PDEs, where the main idea is to combine the Computed Aided Geometric Design (CAGD), widely used by engineers for modelling, and the numerical methods for the solution of PDEs: the core of the success of the quoted approach is that the functions used for describing the numerical solution of the considered PDE are exactly the same as those used for modelling the geometric objects, so resulting in a saving of more than the 80% of the CPU time usually employed in the translation between two different languages (e.g. either between Finite Elements and CAGD or between Finite Differences and CAGD).

Now we summarize five main theoretical features of the GLT class of matrix-sequences and then, in Subsection 1.1, we show the reason why they have high relevance to the approximation of PDEs, by using some elementary examples.

GLT1 Each GLT sequence $\{A_n\}_n$ has a symbol $\kappa = \kappa(\mathbf{x}, \boldsymbol{\theta})$, $\mathbf{x} = (x_1, \ldots, x_d) \in [0, 1]^d$ physical variables, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d) \in (-\pi, \pi)^d$ Fourier variables, $\delta = 2d$, and relation (1.2) holds, that is κ is the singular value symbol of the sequence of matrices $\{A_n\}_n$. If, in addition, the sequence is Hermitian then (1.1) holds and κ is also the spectral symbol of the sequence of matrices $\{A_n\}_n$.

GLT2 The set of GLT sequences form a $*$-algebra that is close under linear combinations, conjugation, products, (pseudo-) inversion, whenever the sequence is invertible. In the case of a sequence having a singular value distribution κ, as in the the GLT setting thanks item **GLT1**, a sequence is invertible if κ vanishes, at most, in a set of zero Lebesgue measure. Hence, the sequence obtained via algebraic operations on a finite set of input GLT sequences is still a GLT sequence and its symbol is obtained by the same algebraic manipulations on the corresponding symbols of the input GLT sequences.

GLT3 Every multilevel Toeplitz sequence generated by a L^1 function f is a GLT sequence and its symbol is $\kappa(\mathbf{x}, \boldsymbol{\theta}) \equiv f(\boldsymbol{\theta})$, under the conditions given in item **GLT1**.

GLT4 Every sequence formed by multilevel diagonal sampling matrices related to the Riemann integrable function a is a GLT sequence and its symbol is $\kappa(\mathbf{x}, \boldsymbol{\theta}) \equiv a(\mathbf{x})$, with the eigenvalue distribution formula holding also in the non Hermitian case.

GLT5 $\{X_n\}$ which is distributed as the zero function in the singular value sense, i.e. $\{X_n\} \sim_\sigma (0, D)$ according to (1.2), is a GLT sequence with symbol $\kappa \equiv 0$.

1.1 GLT and approximated PDEs: basic examples

We first consider a scalar 1D boundary value problem and then a simple vector problem, that is, the linear elasticity in saddle-point form, both in 1D and 2D. The idea is to show that items **GLT1-GLT5** are a tool for giving an automatic way of deducing the asymptotic spectral features of approximated PDEs, just using (multilevel) diagonal sampling matrix-sequences, (multilevel) Toeplitz matrix-sequences, and zero distributed sequences as building blocks.

1.1.1 The model of a rod with variable section

Let us consider the simple differential operator $\mathcal{L}(u) = -u^{('')}$ and the corresponding boundary value problem $\mathcal{L}(u) = g$ on $(0,1)$ with Dirichlet boundary conditions. If we use centered equispaced standard Finite Differences for approximating the previous equation, then we obtain a linear system whose coefficient matrix has the form

$$\Delta_n = \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix}.$$

Such a matrix is a Toeplitz matrix according to the notation in (2.20)–(2.21), is Hermitian (in fact real symmetric), and its generating function, directly reconstructed from the entries of the matrix is $f(\theta) = 2 - 2\cos(\theta)$; see Section 2.5 for the general definitions. Therefore

$$\Delta_n = T_n(f)$$

and according to **GLT3** the sequence $\{\Delta_n\}$ is a GLT sequence with symbol $\kappa(x,\theta) = f(\theta)$ and finally by item **GLT1**, the eigenvalues of Δ_n behave as a uniform sampling of $f(\theta)$. However this is known also algebraically and indeed the eigenvalues of Δ_n are exactly given by $f\left(\theta_i^{(n)}\right) = 2 - 2\cos\left(\theta_i^{(n)}\right)$, $\theta_i^{(n)} = \dfrac{i\pi}{n+1}$.

Now let us consider the variable coefficient version of $\mathcal{L}(u) = -u^{('')}$ in divergence form that is $\mathcal{L}_a(u) = -\left(a(x)u^{'}\right)^{'}$, $a(x)$ bounded and positive on $[0,1]$, which models the behavior of a rod with a variable section. By employing the same equispaced Finite Differences, setting $a_j = a((j-1/2)h), d_j = a_{j-1}+a_j$, we find that the corresponding coefficient matrix is given by

$$\Delta_n^{(2)}(a) = \begin{pmatrix} d_1 & -a_2 & & \\ -a_2 & \ddots & \ddots & \\ & \ddots & \ddots & -a_n \\ & & -a_n & d_n \end{pmatrix}.$$

Even if $\Delta_n^{(2)}(1) = \Delta_n = T_n(f)$, for nonconstant a the Toeplitz character seems to be completely lost. In fact, we find it again 'in a approximated sense' and at 'local scale' as explained below. First of all we have a nice 'dyadic representation' i.e.

$$\Delta_n^{(2)}(a) = \sum_{j=1}^{n+1} a_j \Psi_j = \sum_{j=1}^{n+1} a_j \mathrm{diag}(0,\ldots,0,\Theta,0,\ldots,0),$$

$$\Psi_i = \mathbb{O} \oplus \Theta \oplus \mathbb{O}, \quad \Theta = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

From the latter it is easy to deduce the inequalities

$$[\min a]\, T_n(f) \le \Delta_n^{(2)}(a) \le [\max a]\, T_n(f),$$

which implies that the minimal eigenvalue of $\Delta_n^{(2)}(a)$ is positive and is larger than $[\min a]\lambda_{\text{min}}(T_n(f))$ and the maximal eigenvalue is bounded from above by $[\max a]\lambda_{\text{max}}(T_n(f)) \le \max a(x)f(\theta)$. These bounds show that the new function $a(x)f(\theta)$ has an important role, as deduced in the next derivations using the GLT theory. In fact, setting $D_n(a)$ the diagonal sampling matrix related to the function a that is

$$D_n(a) = \begin{pmatrix} a(h) & & & \\ & a(2h) & & \\ & & \ddots & \\ & & & a(nh) \end{pmatrix}, \quad h = \frac{1}{n},$$

and defining $E_n = \Delta_n^{(2)}(a) - D_n(a)T_n(f)$, $F_n = \Delta_n^{(2)}(a) - D_n^{1/2}(a)T_n(f)D_n^{1/2}(a)$, we can plainly see that the spectral norms $\|E_n\|, \|F_n\|$ are infinitesimal as n tends to infinity, under the assumption the a is continuous. Therefore

1. $\{E_n\} \sim_\sigma (0,D)$, $\{F_n\} \sim_\sigma (0,D)$ for every D and hence by item **GLT5** both $\{E_n\}$, $\{F_n\}$ are GLT sequences with symbol $\kappa \equiv 0$;

2. since $D_n^{1/2}(a) = D_n(\sqrt{a})$, by item **GLT4** , $\{D_n^{1/2}(a)\}$, $\{D_n(a)\}$ are GLT sequences with symbols $\kappa(x,\theta) = \sqrt{a(x)}$, $\kappa(x,\theta) = a(x)$, respectively.

In conclusion, since F_n is Hermitian and

$$\Delta_n(a) = D_n^{1/2}(a)T_n(f)D_n^{1/2}(a) + F_n,$$

also $\{\Delta_n(a)\}$ is a Hermitian GLT sequence with symbol $\kappa(x,\theta) = a(x)f(\theta)$ by item **GLT2**, which states that GLT sequences form a $*$-algebra. Thus, as a consequence of item **GLT1**, we deduce that $\kappa(x,\theta) = a(x)f(\theta)$ is both the spectral and the singular value symbol according to (1.1) and (1.2). In other words, for $n = p^2$ large enough, the eigenvalues of $\Delta_n(a)$ are given approximately by the values $a(x_i)(2-2\cos(\theta_j))$, $i,j=1,\ldots,p$, $x_i=i/p, \theta_j=j\pi/p$.

Along the same lines, we may consider the same problem with non-equaspaced grid points and the same Finite Difference approximation. In fact, if the internal nodes are given as $t_j=g(jh)$, $j=1,\ldots,n$, $h=\frac{1}{n+1}$, and g such that $g([0,1])=[0,1]$, diffeomorfism, then setting

$$\phi(a,g)=\frac{a(g)}{(g')^2},$$

we find that the coefficient matrix $\Delta_n(a,g)$ has the form

$$\Delta_n(a,g)=\Delta_n(\phi(a,g))+\hat{E}_n,$$
$$\Delta_n(\phi(a,g))=D_n(\phi(a,g))T_n(f)+\hat{F}_n,$$

with $\{\hat{E}_n\}\sim_\sigma(0,D)$, $\{\hat{F}_n\}\sim_\sigma(0,D)$ for every D. Therefore, by following the same reasoning as before and using items **GLT1-GLT5**, we obtain that $\{\Delta_n(a,g)\}$ is a GLT sequence with symbol $\kappa(x,\theta)=\phi(a,g)(x)f(\theta)$: thus $\kappa(x,\theta)=\phi(a,g)(x)f(\theta)$ is both the spectral and the singular value symbol of $\{\Delta_n(a,g)\}$, according to (1.1) and (1.2). We observe that the map g is an example of the Geometric Maps used in the IgA, Finite Elements with 'graded' grids, etc, and that the above machinery has a natural extension in the multidimensional.

In this direction, as an example, we just mention that the same Finte Difference approximation of the operator

$$\mathcal{L}_a(u)=-\sum_{i,j=1}^{d}\frac{\partial}{\partial x_i}\left(a_{i,j}(\mathbf{x})\frac{\partial u}{\partial x_j}\right),\qquad a(\mathbf{x})=\left(a_{i,j}(\mathbf{x})\right)_{i,j=1}^{d},$$

leads to a matrix $\Delta_{\mathbf{n}}(a)$ that can be seen, up to squences distributed as the zero function, as $\sum_{i,j=1}^{d}D_{\mathbf{n}}(a_{i,j})\Delta_{\mathbf{n},i,j}$, where, $\mathbf{n}=(n_1,\ldots,n_d)$ is a multi-index and each $\Delta_{\mathbf{n},i,j}$ is a proper multilevel Toeplitz matrix generated by multi-variate trigonometric polynomial $p_{i,j}$ (see Subsection 2.5): again. by items **GLT1-GLT5**, this is enough to conclude that $\{\Delta_{\mathbf{n}}(a)\}$ is a GLT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta})=\sum_{i,j=1}^{d}.a_{i,j}(\mathbf{x})p_{i,j}(\boldsymbol{\theta})$.

1.1.2 The linear elasticity in saddle-point form

In this section we introduce the with the help of

We present two simplified examples, both in connection with the linear elasticity in saddle-point form: in the first case we consider one dimension and linear Finite Elements, while in the the second we consider two dimensions and standard Finite Differences. We use again the GLT technique for deriving spectral symbol of the related matrices and of the corresponding Schur complement: we recall that Schur complement, key tool for the numerical treatment of the underlying linear systems, implies the inversion of a block, but this is no problem since the GLT sequences are stable also under (pseudo)-inversion, as long as the symbol of the inverted block vanishes in a set of zero Lebesgue measure.

The analysis of the elasticity problem, using the GLT tools and taking into account stable approximation techniques, is given in [21].

Consider the coupled system of scalar equations

$$\begin{cases} -(\psi(x)u')'+v' &= g_1(x),\\ u'-\rho v &= g_2(x),\end{cases}\tag{1.3}$$

with Dirichlet boundary conditions. Here $\rho>0$ and the function $\psi(x)$ is positive and continuous on the domain $\Omega=[0,1]$.

Assume first that $\psi(x) = \kappa_0$ is constant in Ω. The use of linear Finite Element basis functions on a uniform mesh with a step size h and a proper scaling leads to a linear system of equations with a coefficient matrix that admits the following structure

$$\mathcal{A} = \begin{bmatrix} K & B^T \\ B & -\rho M \end{bmatrix}.$$

Here, the blocks K, B, M are square of size n and are given below

$$K = \kappa_0 \cdot \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix}, \quad M = \frac{h^2}{6} \begin{bmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & 1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & 1 & 4 \end{bmatrix}, \tag{1.4}$$

$$B = h \begin{bmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}. \tag{1.5}$$

Following the definitions in (2.20)-(2.21) as in Section 2.5, we find

$$\begin{aligned} K &= \kappa_0 T_n(2 - 2\cos(\theta)), & B &= h\, T_n(1 - \mathrm{e}^{\mathrm{i}\theta}), \\ B^T &= h\, T_n(1 - \mathrm{e}^{-\mathrm{i}\theta}), & M &= \tfrac{h^2}{3} T_n(2 + \cos(\theta)), \end{aligned} \tag{1.6}$$

where the negative Schur complement of $\mathcal{A}$ is expressed as

$$\mathcal{S}_n = \rho M + B^T K^{-1} B. \tag{1.7}$$

Since $\kappa_0 > 0$, the matrix K is invertible. Considering (1.7), we deduce that $\mathcal{S}_n$ is given in terms of Toeplitz structures

$$\mathcal{S} = \frac{\rho}{3} T_n(2 + \cos(\theta)) + \frac{1}{\kappa_0} T_n(1 - \mathrm{e}^{-\mathrm{i}\theta}) T_n^{-1}(2 - 2\cos(\theta)) T_n(1 - \mathrm{e}^{\mathrm{i}\theta}). \tag{1.8}$$

According to **GLT3**, the sequences $\{T_n(2 + \cos(\theta))\}$, $\{T_n(1 - \mathrm{e}^{\mathrm{i}\theta})\}$, $\{T_n(1 - \mathrm{e}^{-\mathrm{i}\theta})\}$ and $\{T_n(2 - 2\cos(\theta))\}$ are GLT sequences with symbols $2 + \cos(\theta)$, $1 - \mathrm{e}^{\mathrm{i}\theta}$, $1 - \mathrm{e}^{-\mathrm{i}\theta}$, $2 - 2\cos(\theta)$, respectively. As a consequence, by **GLT2**, taking into account that $2 - 2\cos(\theta)$ vanishes in a set of zero Lebesgue measure, we obtain that the inverse of $T_n(2 - 2\cos(\theta))$ is still in the GLT algebra and therefore also $\{\mathcal{S}_n\}_n$ is a GLT sequence generated by the symbol

$$\kappa^{\mathcal{S}}(\theta) = \frac{\rho}{3}(2 + \cos(\theta)) + \frac{1}{\kappa_0}(1 - \mathrm{e}^{-\mathrm{i}\theta}) \frac{1}{2 - 2\cos(\theta)} (1 - \mathrm{e}^{\mathrm{i}\theta}) = \frac{\rho}{3}(2 + \cos(\theta)) + \frac{1}{\kappa_0}. \tag{1.9}$$

Since $\mathcal{S}_n$ is Hermitian independently of its size, according to **GLT1**, we deduce that (1.1) holds for $\{\mathcal{S}_n\}_n$, i.e., $\{\mathcal{S}_n\}_n \sim_\lambda (\kappa^{\mathcal{S}}, (-\pi, \pi)$, where $\kappa^{\mathcal{S}}$ is an even trigonometric polynomial. Figure 1 shows the agreement between the asymptotic forecast and the eigenvalues of $\mathcal{S}_n$ for a couple of values of the parameters κ_0 and ρ, where both the evaluations of $\kappa^{\mathcal{S}}$ over $\theta_j^{(n)} = \frac{\pi j}{n+1}$, $j = 1, \ldots, n$. In the plots, the eigenvalues of $\mathcal{S}_n$ are sorted in an increasing order.

Next we consider equation (1.3) with a non-constant diffusivity coefficient $\psi(x)$. The resulting matrix K is no longer Toeplitz but, as shown in [8] and in analogy with the Finite Difference case reported in Subsection 1.1.1 , the related sequence $\{K_n\}_n$ belongs to the GLT class with symbol $\psi(x)(2 - 2\cos(\theta))$. Therefore, following

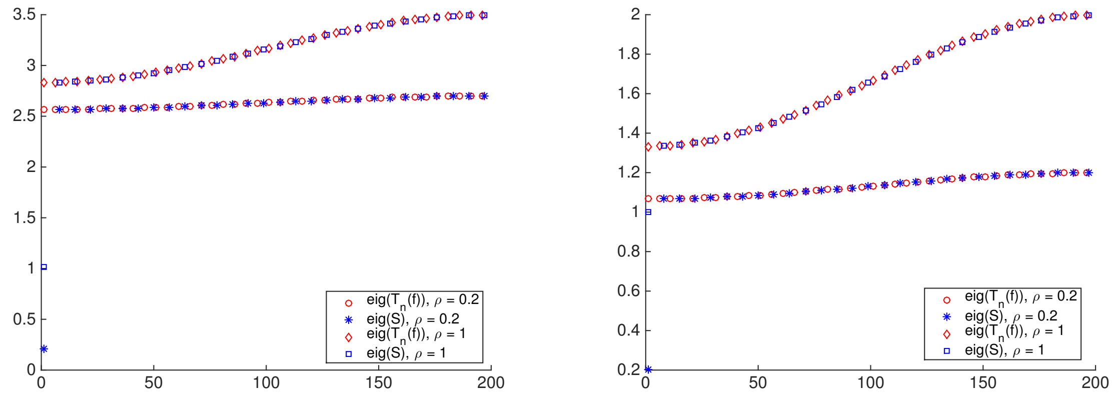

Figure 1: 1D Elasticity Problem: Spectrum of $\mathcal{S}_n$ vs sampling of its symbol $\kappa^{\mathcal{S}}$, const. coeff. $\kappa_0 = 0.4$ and $\kappa_0 = 1$

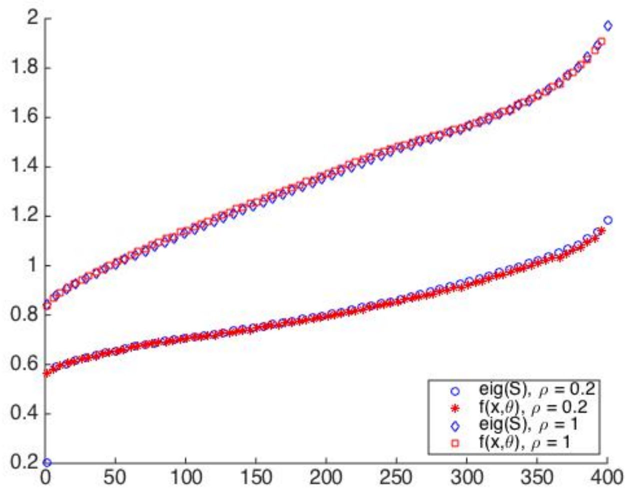

Figure 2: 1D Elasticity Problem: Spectrum of $\mathcal{S}$ vs sampling of its symbol $\kappa^{\mathcal{S}}$, variable coefficient $\psi(x) = 1 + x$.

verbatim the reasoning for the case of $\psi(x) = \kappa_0$, we deduce that the sequence of Schur complements forms a GLT sequence with the symbol

$$\begin{aligned} \kappa^{\mathcal{S}}(x,\theta) &= \frac{\rho}{3}(2+\cos(\theta)) + \frac{1}{\psi(x)}(1-\mathrm{e}^{-\mathrm{i}\theta})\frac{1}{\psi(x)(2-2\cos(\theta))}(1-\mathrm{e}^{\mathrm{i}\theta}) \\ &= \frac{\rho}{3}(2+\cos(\theta)) + \frac{1}{\psi(x)}. \end{aligned} \tag{1.10}$$

The latter is illustrated in Figure 2, showing the agreement between the asymptotic forecast and the eigenvalues of $\mathcal{S}_n$ for a pair of choices of ρ and $\psi(x)$, where both the evaluations of $\kappa^{\mathcal{S}}$ over a n-sized uniform gridding over $[0,1] \times [0,\pi]$ of size n and the eigenvalues of $\mathcal{S}$ have been sorted in an increasing order.

We formulate next a two-dimensional coupled test system of two PDEs and two unknowns, the first of which, referred to as the $displacements$ is a vector with two components.

Consider an elasticity-like problem in saddle point form, defined in $\Omega = [0,1]^2$,

$$\mathcal{A} = \begin{bmatrix} K & B^T \\ B & -\rho M \end{bmatrix} \begin{array}{l} \}\text{displacements} \\ \}\text{pressure,} \end{array}$$

where K and M are symmetric and positive definite.

Under the so-called separate displacement ordering (SDO) of the components of the displacements, K itself attains a two-by-two block structure,

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{matrix} \}\text{displacements in } x \\ \}\text{displacements in } y. \end{matrix}$$

The SDO ordering of the displacements induces a corresponding block structure in the block B, which we denote as $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$. Thus, we have

$$\mathcal{A} = \begin{bmatrix} K_{11} & K_{12} & B_1^T \\ K_{21} & K_{22} & B_2^T \\ B_1 & B_2 & -\rho M \end{bmatrix}. \tag{1.11}$$

Further, we assume that, respectively, K_{11}, K_{22} approximate two anisotropic Laplacians of the form $-\left(2\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)$, $-\left(\frac{\partial^2}{\partial x^2} + 2\frac{\partial^2}{\partial y^2}\right)$, K_{12}, K_{21} approximate the operators $-\frac{\partial^2}{\partial y \partial x}$, $-\frac{\partial^2}{\partial x \partial y}$ and B_1^T, B_2^T approximate the operators $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$. M is the mass matrix, approximating the identity operator.

We use standard Finite Differences on a square mesh, even though this approximation does not possess the stability properties, required for mixed problems. The blocks of $\mathcal{A}$ are described as follows:

$$\begin{aligned}
K_{11} &= 2\,T_n(2-2\cos(\theta_1)) \otimes I_n + I_n \otimes T_n(2-2\cos(\theta_2)), \\
K_{12} &= T_n(1-\mathrm{e}^{-\mathrm{i}\theta_1}) \otimes T_n(1-\mathrm{e}^{-\mathrm{i}\theta_2}), \\
K_{21} &= K_{12}^T, \\
K_{22} &= T_n(2-2\cos(\theta_1)) \otimes I_n + 2I_n \otimes T_n(2-2\cos(\theta_2)), \\
B_1 &= h\,T_n(1-\mathrm{e}^{-\mathrm{i}\theta_1}) \otimes I_n, \\
B_2 &= h\,I_n \otimes T_n(1-\mathrm{e}^{-\mathrm{i}\theta_2}), \\
M &= h^2\,I_n.
\end{aligned}$$

Here $\otimes$ denotes the Kronecker product (cf. [7]). We notice that the Kronecker product induces a two-level Toeplitz structure and in fact every matrix is Toeplitz, where each''entry' along the diagonal is a standard unilevel Toeplitz matrix, see Section 2.5. Hence, using the two- level notation introduced in [26, 56], we construct the two- variate generating functions as in Section 2.5, see (2.20)-(2.21), associated with each block. More precisely,

$$\begin{aligned}
K_{11} &= T_n((6-4\cos(\theta_1)-2\cos(\theta_2))), && (1.12)\\
K_{12} &= T_n((1-\mathrm{e}^{-\mathrm{i}\theta_1})(1-\mathrm{e}^{-\mathrm{i}\theta_2})), && (1.13)\\
K_{21} &= T_n((1-\mathrm{e}^{\mathrm{i}\theta_1})(1-\mathrm{e}^{\mathrm{i}\theta_2})), && (1.14)\\
K_{22} &= T_n((6-2\cos(\theta_1)-4\cos(\theta_2))), && (1.15)\\
B_1 &= h\,T_n(1-\mathrm{e}^{-\mathrm{i}\theta_1}), && (1.16)\\
B_2 &= h\,T_n(1-\mathrm{e}^{-\mathrm{i}\theta_2}), && (1.17)\\
M &= h^2\,T_n(1). && (1.18)
\end{aligned}$$

Consider next the negative Schur complement of $\mathcal{A}$, $S = \rho M + BK^{-1}B^T$. Given the rich block structure of the matrix $\mathcal{A}$, the formal expression of the Schur complement involves inversion of the spd block K and multiplication by rectangular blocks. Since we want to use the symbols (1.12)–(1.18) of the related sub-blocks, we utilize the following exact block-factorization of K and of its inverse:

$$K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} = \begin{bmatrix} I & \\ K_{21}K_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} K_{11} & \\ & S_K \end{bmatrix} \begin{bmatrix} I & K_{11}^{-1}K_{12} \\ & I \end{bmatrix}, \tag{1.19}$$

where $S_K = K_{22} - K_{21}K_{11}^{-1}K_{12}$. Since the positive definite character of S_K is guaranteed by the positive definite character of K, we find

$$K^{-1} = \begin{bmatrix} I & -K_{11}^{-1}K_{12} \\ & I \end{bmatrix} \begin{bmatrix} K_{11}^{-1} & \\ & S_K^{-1} \end{bmatrix} \begin{bmatrix} I & \\ -K_{21}K_{11}^{-1} & I \end{bmatrix}. \tag{1.20}$$

CLearly, the latter factorization holds for any nonsingular matrix K. Therefore, as it is well known, the explicit formal expression of $\mathcal{S}$ is given by

$$\begin{aligned} \mathcal{S} &= \rho M + \begin{bmatrix} B_1 & B_2 \end{bmatrix} K^{-1} \begin{bmatrix} B_1^T \\ B_2^T \end{bmatrix} \\ &= \rho M + B_1 K_{11}^{-1} B_1^T + (B_2 - B_1 K_{11}^{-1} K_{12}) S_K^{-1} (B_2 - B_1 K_{11}^{-1} K_{12})^T. \end{aligned} \tag{1.21}$$

As expected, constructing the symbol in 2D is by far more complicated than that in 1D. Nevertheless, the expression of $\mathcal{S}$ in (1.21) can be seen as a'rational noncommutative' formula involving the blocks in (1.12)–(1.18), which are two-level Toeplitz structures with trigonometric polynomial symbols. As a consequence, in view of **GLT1, GLT2, GLT3**, exactly as done in the one-dimensional example (1.3), we deduce that $\{\mathcal{S}\}$ is a GLT sequence with $\{\mathcal{S}\} \sim_\lambda \kappa^{\mathcal{S}}$, since $\mathcal{S}$ is Hermitian independently of its size, and $\kappa^{\mathcal{S}}$ is a rational function of the symbols given in (1.12)–(1.18). Interestingly enough, setting $\delta(\theta) = 2 - 2\cos(\theta)$ the symbol of the unilevel Laplacian and making tedious algebraic manipulations, we find

$$\kappa^{\mathcal{S}} = \rho + \mu \frac{q_1(\delta(\theta_1), \delta(\theta_2))}{q_2(\delta(\theta_1), \delta(\theta_2))},$$

where q_1, q_2 are two nonnegative homogeneous polynomials of degree two. Therefore, $\rho \le f(\theta_1, \theta_2) \le \rho + \phi$ with ϕ being the maximum of q_1/q_2.

Following the same procedure as in the one-dimensional setting we can treat the case of variable coefficients.

1.1.3 Some intermediate remarks

There are two further features of the GLT class of matrices that are a consequence of items **GLT1-GLT5**, as briefly sketched in Subsection 1.1 and that represent a powerful and general tool for identifying the spectral behavior of matrix-sequences, coming from the approximation by local methods of PDEs and IEs. We summarize them in the following.

GLT6 The approximation of PDEs with non-constant coefficients, general domains, nonuniform gridding by local methods (Finite Difference, Finite Elements, IgA etc), under very mild assumptions leads also to GLT sequences (see [54, 45, 46] for the case of Finite Differences, [8, 26] for the Finite Elements setting, and [15, 23] for the case of IgA approximations).

GLT7 We encounter GLT structures when dealing with preconditioning of iterative and semi-iterative solvers (see e.g. [4]), when the dealing with the stabilty analysis of implicit numerical methods for PDEs (see e.g. [34, 35]), when considering multigrid methods applied to approximations of PDEs and IEs by local methods (see [1, 2, 3, 20, 18] and references therein). Moreover, the symbol includes information about the coefficients and the domain of the PDE, as well as information on the discretization schemes for the derivatives including the used meshes, which have to be described, at least asymptotically, as a map of a reference equispaced mesh (see [36, 54, 48, 39] for the one-dimensional setting and [42, 40, 49, 45, 46, 8, 13, 14, 26] for the two-dimensional and multi-dimensional settings. Furthermore, also in presence of non-dominating advection terms the distribution result for the eigenvalues can be recovered, thanks to ad hoc results in [29, 27], heavily based on the majorization theory well explained in the remarkable book [7].

1.2 Technical introduction: new vs previous approach

The focus of this paper is on the theoretical foundations of the GLT theory and therefore we are not concerned with its applicative interest, already briefly emphasized in the former subsections.

Here we first propose slight (but relevant) modifications of the original definitions of separable Locally Toeplitz (sLT) sequences and GLT sequences appeared in [45, 46]. With the new definitions, which are based on the notion of approximating class of sequences [28, 41], we are able to enlarge the applicability of the GLT theory. In particular, we remove a certain '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 [45, Theorems 4.5 and 4.8] and [46, Theorem 2.2]. We also show that the product of two LT sequences with symbols $a\otimes f$, $\tilde{a}\otimes\tilde{f}$ is a LT sequence with symbol $a\tilde{a}\otimes f\tilde{f}$, 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 [45, Theorem 5.3] both the assumption that $f, \tilde{f}$ are in L^∞ and the technical condition in [45, Eq. (41)]. In addition, we provide a formal proof of the fact that sequences formed by multilevel diagonal sampling matrix-sequences, as well as multilevel Toeplitz sequences, fall in the class of LT sequences; this result was often used in previous papers, but no formal proof was given. The latter two results allows 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 generated by f, is a LT sequence with symbol $a\otimes f$, under the only assumptions that a is Riemann-integrable and f is in L^1.

As a final step, we also extend the GLT theory. This is the completely new part of the the review. We first provide an approximation result in Section 5.2, which is particularly useful to show that a given sequence of matrices is a GLT sequence. Then, by using this result, we provide in Section 5.4 a new and easier proof of the fact that $\{A_n^\dagger\}_n$ is a GLT sequence with symbol κ^{-1} whenever $\{A_n\}_n$ is a GLT sequence with symbol κ and $\kappa\neq 0$ almost everywhere; here, $A_n^\dagger$ denotes the (Moore–Penrose) pseudoinverse of A_n. Finally, using again the approximation result of Section 5.2, we prove in Section 5.4 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.

The paper is concluded with a final section where we summarize the results, we put in evidence the applicative interest, especially regarding the approaximation of PDEs, and we state open problems, both concrete and theoretical, to be considered in a future reserches.

2 Mathematical background

The section contains the mathematical background, including notations, terminology, preliminaries in analysis, linear algebra, and matrix theory. Of special importance is the definition of singular value and spectral symbol and the definition of multilevel Toeplitz matrices.

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.
- χ_E is the characteristic (or indicator) function of the set E, so $\chi_E(x)=1$ if $x\in E$ and $\chi_E(x)=0$ otherwise.
- μ_d denotes the Lebesgue measure in $\mathbb{R}^d$. Throughout this work, all the terminology of measure theory (such as 'measure', 'measurable', 'a.e.', 'in L^p', etc.) is always referred to the Lebesgue measure.
- If $f:D\subseteq\mathbb{R}^d\to\mathbb{C}$ is in $L^p(D)$ and if 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, which is defined by $\|f\|_{L^p}=(\int_D|f|^p)^{1/p}$ for $1\le p<\infty$, and by $\|f\|_{L^\infty}=\operatorname{ess\,sup}_D|f|$ for $p=\infty$.

- We use a notation that is common in probability theory to indicate the sets defined in terms of functions. For example, if $f : D \subseteq \mathbb{R}^d \to \mathbb{C}$, then $\{f \neq 1\} = \{\mathbf{x} \in D : \ f(\mathbf{x}) \neq 1\}$, $\{0 \le f \le 1\} = \{\mathbf{x} \in D : \ 1 \le f(\mathbf{x}) \le 1\}$, $\mu_d\{f > 0\}$ is the measure of the set $\{\mathbf{x} \in D : \ f(\mathbf{x}) > 0\}$, $\chi_{\{f=0\}}$ is the characteristic function of the set where f vanishes, and so on.

- 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 [7]. 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.

- $\Re(X)$ 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 $w_i : D_i \to \mathbb{C}, \ i = 1, \ldots, d$, 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

$$(w_1 \otimes \cdots \otimes w_d)(\xi_1, \ldots, \xi_d) = w_1(\xi_1) \cdots w_d(\xi_d), \qquad \xi_i \in D_i, \qquad i = 1, \ldots, d.$$

- $C_c(\mathbb{C})$ (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.

- If $H : \mathbb{R} \to \mathbb{R}$ and the limit $\lim_{x\to\infty} H(x)$ exists, we denote it by $H(\infty)$. Similarly, $H(-\infty) = \lim_{x\to-\infty} H(x)$.

- $a_n \overset{n\to\infty}{\longrightarrow} a$ means that $a_n \to a$ as $n \to \infty$.

- A function $a : [0, 1]^d \to \mathbb{C}$ is said to be Riemann-integrable if $\Re(a), \Im(a) : [0, 1]^d \to \mathbb{R}$ 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 $\|g\|_{\infty,E} = \sup_{x\in E} |g(x)|$. 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.

2.1.1 Multi-index notation

Throughout this work, we will systematically use the multi-index notation, expounded by Tyrtyshnikov in [56, Section 6]. A multi-index $\boldsymbol{i}$ is simply a vector in $\mathbb{Z}^d$; its components are denoted by $i_1, \ldots, i_d$. A multi-index $\boldsymbol{i} \in \mathbb{Z}^d$ is also called a d-index.

- $\mathbf{0}$, $\mathbf{1}$, $\mathbf{2}$, ... 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 $\min_{j=1,\ldots,d} m_j \to \infty$.
- If $\boldsymbol{i}, \boldsymbol{j}$ are d-indices, $\boldsymbol{i} \leq \boldsymbol{j}$ means that $i_r \leq j_r$ for all $r = 1, \ldots, d$.
- If $\boldsymbol{h}, \boldsymbol{k}$ are d-indices such that $\boldsymbol{h} \leq \boldsymbol{k}$, the multi-index range $\boldsymbol{h}, \ldots, \boldsymbol{k}$ is the set $\{\boldsymbol{j} \in \mathbb{Z}^d : \boldsymbol{h} \leq \boldsymbol{j} \leq \boldsymbol{k}\}$. We assume for the multi-index range $\boldsymbol{h}, \ldots, \boldsymbol{k}$ the standard lexicographic ordering:

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

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

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

- 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 always 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 $X = [x_{\boldsymbol{ij}}]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{m}}$, 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 over the multi-index range $\mathbf{1}, \ldots, \boldsymbol{m}$ according to (2.1). Similarly, if $\mathbf{x} = [x_{\boldsymbol{i}}]_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{m}}$ 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 $x_{\boldsymbol{m}} = x_{(m_1,\ldots,m_d)}$.
- 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 $\boldsymbol{h} \leq \boldsymbol{k}$, 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, $\boldsymbol{np} = (n_1 p_1, \ldots, n_d p_d)$, $\alpha \boldsymbol{i}/\boldsymbol{j} = (\alpha i_1/j_1, \ldots, \alpha i_d/j_d)$ 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:

- n varies in some infinite subset of $\mathbb{N}$;
- $\boldsymbol{n} = (n_1, \ldots, n_d) \in \mathbb{N}^d$ is a multi-index which depends on n, i.e., $\boldsymbol{n} = \boldsymbol{n}(n)$;
- $A_{\boldsymbol{n}}$ is a square matrix with $\mathrm{size}(A_{\boldsymbol{n}}) = N(\boldsymbol{n})$;

- $\boldsymbol{n} \to \infty$ when $n \to \infty$.

Two 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: sequences of multilevel diagonal sampling matrices and sequences of multilevel Toeplitz matrices. Here, we introduce multilevel diagonal sampling matrices, while multilevel Toeplitz matrices will be considered in Section 2.5. 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 sort of 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

$$D_{\boldsymbol{n}}(a) = D_{n_1,\ldots,n_d}(a) = \operatorname*{diag}_{i_1=1,\ldots,n_1} D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big), \tag{2.2}$$

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

2.1.3 Separable functions and multivariate trigonometric polynomials

Let $I_1,\ldots,I_d \subseteq \mathbb{R}$ be measurable sets and let $f : I_1 \times \cdots \times I_d \to \mathbb{C}$ be measurable. 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 $f = c_1 f_1 \otimes \cdots \otimes c_d f_d$.

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 $f_i = 0$ a.e. for some i and $f = 0$ 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 $f_1 \in L^\infty(I_1)$ (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 $\alpha > 0$, implying that

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

for all $\alpha > 0$. This 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 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, for any separable function $f : I_1\times\cdots\times I_d \to \mathbb{C}$ in L^p there always exists a factorization $f = f_1 \otimes \cdots \otimes f_d$ such that $f_1, \ldots, f_d$ are in L^p.

A (d-variate) trigonometric polynomial is a finite linear combination of the Fourier frequencies $\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}$, $\boldsymbol{j} \in \mathbb{Z}^d$. Let $f : \mathbb{R}^d \to \mathbb{C}$ be a separable trigonometric polynomial, i.e., a 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 $f_2(\vartheta_2)\cdots f_d(\vartheta_d) \neq 0$. For obvious properties of trigonometric polynomials, $\theta_1 \mapsto f(\theta_1, \vartheta_2, \ldots, \vartheta_d) = f_1(\theta_1)f_2(\vartheta_2)\cdots f_d(\vartheta_d)$ is a (univariate) trigonometric polynomial, i.e., f_1 is a trigonometric polynomial. With the same argument, one can prove that $f_2, \ldots, f_d$ are trigonometric polynomials as well. This shows that, for a separable d-variate trigonometric polynomial f there always exists a factorization $f = f_1 \otimes \cdots \otimes f_d$ in which $f_1, \ldots, f_d$ are trigonometric polynomials.

2.2 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 that we shall use in the following.

Definition 2.1 (**convergence in measure**)**.** Let $D \subseteq \mathbb{R}^k$ be any measurable subset of $\mathbb{R}^k$, and let $f_m,\ f : D \to \mathbb{C}$ be measurable functions. We say that $f_m \to f$ in measure over D – or simply that $f_m \to f$ in measure, if the domain D can be inferred from the context – when the following condition is met: for every $\epsilon > 0$,

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

Using the definition, it can be shown that, if $f_m \to f$ in measure, then $|f_m| \to |f|$ in measure; and 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 $\alpha, \beta \in \mathbb{C}$. Moreover, 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.

The next lemma provide a relation between the convergence in measure and the L^1-convergence. Since the result is not so popular, for the reader's convenience we include the details of the proof.

Lemma 2.1. *Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$, let $F : \mathbb{K} \to \mathbb{C}$ be a uniformly continuous bounded function, and let $g_m,\ g : D \subseteq \mathbb{R}^d \to \mathbb{K}$ be measurable functions such that $g_m \to g$ in measure over D. Then, $F \circ g_m \to F \circ g$ in $L^1(D)$.*

Proof. For every m and every $\epsilon > 0$,

$$\begin{aligned}\int_D \bigl|F(g_m(\mathbf{x})) - F(g(\mathbf{x}))\bigr| d\mathbf{x} &= \int_{\{|g_m-g|\geq\epsilon\}} \bigl|F(g_m(\mathbf{x})) - F(g(\mathbf{x}))\bigr| d\mathbf{x} + \int_{\{|g_m-g|<\epsilon\}} \bigl|F(g_m(\mathbf{x})) - F(g(\mathbf{x}))\bigr| d\mathbf{x} \\ &\leq \frac{2\|F\|_\infty\, \mu_d\{|g_m - g| \geq \epsilon\}}{\mu_d(D)} + \omega_F(\epsilon), \end{aligned} \tag{2.3}$$

where ω_F is the modulus of continuity of F. Since

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

(because $g_m \to g$ in measure and F is uniformly continuous), passing first to the $\limsup_{m\to\infty}$ and then to the $\lim_{\epsilon\to 0}$ in (2.3), we conclude that $F \circ g_m \to F \circ g$ in $L^1(D)$. $\square$

Lemma 2.2 is the last result we need about the convergence in measure; it will play a crucial role in the proof of Theorem 5.7. For the proof of Lemma 2.2, we recall that the space generated by the monomials

$$\left\{\mathrm{e}^{\mathrm{i}(\frac{2\pi}{b_1-a_1}j_1\theta_1 + \ldots + \frac{2\pi}{b_k-a_k}j_k\theta_k)} \,:\, \boldsymbol{j}=(j_1,\ldots,j_k)\in\mathbb{Z}^k\right\},$$

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

Lemma 2.2. *Let $\kappa:[0,1]^d\times[-\pi,\pi]^d\to\mathbb{C}$ be a measurable function. Then, there exists a sequence $\{\kappa_m\}_m$, with $\kappa_m:[0,1]^d\times[-\pi,\pi]^d\to\mathbb{C}$ a function of the form*

$$\kappa_m(\mathbf{x},\boldsymbol{\theta})=\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m} a_{\boldsymbol{j}}^{(m)}(\mathbf{x})\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}},\qquad a_{\boldsymbol{j}}\in C^\infty([0,1]^d),\qquad \boldsymbol{N}_m\in\mathbb{N}^d, \tag{2.4}$$

such that $\kappa_m\to\kappa$ in measure.

Proof. The space generated by the monomials

$$\left\{\mathrm{e}^{\mathrm{i}2\pi\boldsymbol{\ell}\cdot\mathbf{x}}\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}\,:\,\boldsymbol{\ell},\boldsymbol{j}\in\mathbb{Z}^d\right\} \tag{2.5}$$

is dense in $L^1([0,1]^d\times[-\pi,\pi]^d)$. The function $\widetilde{\kappa}_m=\kappa\chi_{\{|\kappa|\le 1/m\}}$ belongs to $L^\infty([0,1]^d\times[-\pi,\pi]^d)\subset L^1([0,1]^d\times[-\pi,\pi]^d)$ and $\widetilde{\kappa}_m\to\kappa$ in measure. Choose κ_m in the space generated by the monomials (2.5) such that $\|\kappa_m-\widetilde{\kappa}_m\|_{L^1}\le 1/m$. Note that κ_m is of the form (2.4). Then, for all $\epsilon>0$,

$$\mu_{2d}\{|\kappa_m-\kappa|>\epsilon\}\le\mu_{2d}\{|\kappa_m-\widetilde{\kappa}_m|>\epsilon/2\}+\mu_{2d}\{|\widetilde{\kappa}_m-\kappa|>\epsilon/2\}\le\frac{\|\kappa_m-\widetilde{\kappa}_m\|_{L^1}}{(\epsilon/2)}+\mu_{2d}\{|\widetilde{\kappa}_m-\kappa|>\epsilon/2\},$$

which converges to 0 as $m\to\infty$. Hence, $\kappa_m\to\kappa$ in measure. □

2.3 Preliminaries on Linear Algebra and Matrix Analysis

Given $X\in\mathbb{C}^{m\times m}$, 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 $\|X\|=\sigma_{\max}(X)$, we obtain

$$\|X\|_1=\sum_{i=1}^m\sigma_i(X)\le\text{rank}(X)\|X\|\le m\|X\|,\qquad X\in\mathbb{C}^{m\times m}. \tag{2.6}$$

Let $1\le p,q\le\infty$ be conjugate exponents ($\frac{1}{p}+\frac{1}{q}=1$). As a consequence of the general Hölder-type inequalities for unitarily invariant norms, see [7, Corollary IV.2.6], the Schatten norms satisfy

$$\|XY\|_1\le\|X\|_p\|Y\|_q,\qquad X,Y\in\mathbb{C}^{m\times m}. \tag{2.7}$$

If $X\in\mathbb{C}^{m\times m}$ is a normal matrix, 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, it can be shown that the singular values of X coincide with the moduli of the eigenvalues, $|\lambda_j(X)|,\ j=1,\ldots,m$. Consequently, $\|X\|=\rho(X)$ and $\|X\|_1=\sum_{j=1}^m|\lambda_j(X)|$. Note that, if X is Hermitian ($X^*=X$) or skew-Hermitian ($X^*=-X$), then X is normal.

We now recall some important interlacing and perturbation theorems for both 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 of indices $1,\ldots,m$, 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* $\text{rank}(E) \le k$. *Then*

$$\sigma_{j-k}(X) \ge \sigma_j(Y) \ge \sigma_{j+k}(X), \qquad j = 1, \dots, m. \tag{2.8}$$

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, i.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 $\text{rank}(E) \le k$ *then*

$$\lambda_{j-k}(X) \ge \lambda_j(Y) \ge \lambda_{j+k}(X), \qquad j = 1, \dots, m. \tag{2.9}$$

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

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

are $\sigma_j(A), -\sigma_j(A),\ j = 1, \dots, m$; see [7, Exercise II.1.15]. Therefore, applying Theorem 2.2 with $\hat{Y}$, $\hat{X}$, $\hat{E}$ in place of Y, X, E, we obtain Theorem 2.1. The proof of Theorem 2.2 can be derived from the result in [7, 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 [7, Corollary III.2.6]. Theorem 2.3 can be obtained as a corollary of Theorem 2.4 by considering again the matrices $\hat{X}$ and $\hat{Y}$. Alternatively, Theorem 2.3 (resp. Theorem 2.4) can be proved by using the minimax principle for singular values [7, Problem III.6.1] (resp. the minimax principle for eigenvalues [7, Corollary III.1.2]). We also refer the reader to [7, Problem II.6.13] for a more general perturbation theorem for singular values, which extends Theorem 2.3.

2.3.1 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

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

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

$$X \oplus Y = \text{diag}(X, Y) = \left[\begin{array}{c|c} X & O \\ \hline O & Y \end{array}\right].$$

Tensor products and direct sums possess a lot of nice algebraic properties; see [22, Section 1.2.1] or [26, Section 2.4]. Here, we mention the bilinearity of tensor products; the associativity, which allows us to omit parentheses in expressions like $X_1\otimes X_2\otimes\cdots\otimes X_d$ or $X_1\oplus X_2\oplus\cdots\oplus X_d$; the relations $(X_1\otimes Y_1)(X_2\otimes Y_2)=(X_1X_2)\otimes(Y_1Y_2)$ and $(X_1\oplus Y_1)(X_2\oplus Y_2)=(X_1X_2)\oplus(Y_1Y_2)$, which hold whenever X_1,X_2 can be multiplied and Y_1,Y_2 can be multiplied; the identities $(X\otimes Y)^*=X^*\otimes Y^*$, $(X\oplus Y)^*=X^*\oplus Y^*$ and $(X\otimes Y)^T=X^T\otimes Y^T$, $(X\oplus Y)^T=X^T\oplus Y^T$, which hold for all matrices X,Y; and the (very important) multi-index formula for tensor products: if $X_k\in\mathbb{C}^{m_k\times m_k},\ k=1,\ldots,d$, then

$$(X_1\otimes\cdots\otimes X_d)_{\boldsymbol{ij}}=(X_1)_{i_1j_1}\cdots(X_d)_{i_dj_d},\qquad \boldsymbol{i},\boldsymbol{j}=\boldsymbol{1},\ldots,\boldsymbol{m}=(m_1,\ldots,m_d). \tag{2.10}$$

From these basic properties, a lot of other interesting results follow. For example, 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. $\lambda_i(X),\ \lambda_j(Y),\ i=1,\ldots,m,\ j=1,\ldots,\ell$ and $\sigma_i(X),\ \sigma_j(Y),\ i=1,\ldots,m,\ j=1,\ldots,\ell$). In particular, for all matrices $X\in\mathbb{C}^{m\times m}$ and $Y\in\mathbb{C}^{n\times n}$,

$$\|X\oplus Y\|=\max(\|X\|,\|Y\|),\qquad \|X\otimes Y\|=\|X\|\,\|Y\|, \tag{2.11}$$

$$\|X\oplus Y\|_p=\left(\|X\|_p^p+\|Y\|_p^p\right)^{1/p},\qquad \|X\otimes Y\|_p=\|X_1\|_p\,\|Y\|_p,\qquad 1\le p<\infty, \tag{2.12}$$

$$\text{rank}(X\oplus Y)=\text{rank}(X)+\text{rank}(Y)\qquad \text{rank}(X\otimes Y)=\text{rank}(X)\,\text{rank}(Y), \tag{2.13}$$

and if X, Y are Hermitian positive definite (HPD), then $X\otimes Y$ is HPD as well, with

$$\lambda_{\min}(X\otimes Y)=\lambda_{\min}(X)\lambda_{\min}(Y),\qquad \lambda_{\max}(X\otimes Y)=\lambda_{\max}(X)\lambda_{\max}(Y). \tag{2.14}$$

Moreover,

$$X\otimes Y\ge X'\otimes Y',\qquad \text{for all HPD matrices } X,Y,X',Y' \text{ such that } X\ge X' \text{ and } Y\ge Y', \tag{2.15}$$

because $X\otimes Y-X'\otimes Y'=(X-X')\otimes Y+X'\otimes(Y-Y')$ is a sum of two HPD matrices. We also highlight the following property [26, p. 7]: let $X_1,\ldots,X_d,Y_1,\ldots,Y_d$ be matrices with $X_i,Y_i\in\mathbb{C}^{m_i\times m_i}$ for all $i=1,\ldots,d$, then

$$\text{rank}(X_1\otimes\cdots\otimes X_d-Y_1\otimes\cdots\otimes Y_d)\le N(\boldsymbol{m})\sum_{i=1}^d\frac{\text{rank}(X_i-Y_i)}{m_i}. \tag{2.16}$$

A property of tensor products, which is not as popular as the previous ones, is given in Lemma 2.3. For the proof we refer the reader to [26, Lemma 2].

Lemma 2.3. *For all $\boldsymbol{m}\in\mathbb{N}^d$ and all permutations σ of the set $\{1,\ldots,d\}$, there exists a permutation matrix $\Pi_{\boldsymbol{m};\sigma}\in\mathbb{C}^{N(\boldsymbol{m})\times N(\boldsymbol{m})}$ such that*

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

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

Lemma 2.3 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 $X_1,\ldots,X_d$. Concerning the 'distributive properties' of tensor products with respect to direct sums, a result analogous to Lemma 2.3 holds: these properties hold modulo permutation transformations which depend only on the dimensions of the involved matrices. For the proof of Lemma 2.4, see [26, Lemma 4].

Lemma 2.4. *For all $\ell \in \mathbb{N}$, $\boldsymbol{m} \in \mathbb{N}^d$ there exists a permutation matrix $Q_{\ell,\boldsymbol{m}} \in \mathbb{C}^{\ell(m_1+\ldots+m_d)\times\ell(m_1+\ldots+m_d)}$ such that*

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

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

Lemma 2.4 gives the distributive law on the left; the distributive law on the right holds without permutation matrices, as stated in the following remark.

Remark 2.1. From the definition of tensor products and direct sums, for all matrices $X_1, \ldots, X_d, Y$ we have

$$(X_1 \oplus X_2 \oplus \cdots \oplus X_d) \otimes Y = (X_1 \otimes Y) \oplus (X_2 \otimes Y) \oplus \cdots \oplus (X_d \otimes Y).$$

We end this section with a result about direct sums, which is completely analogous to Lemma 2.3: it shows that the direct sum operation is 'almost' commutative. The proof is rather technical and may be skipped in a first reading; take also into account that Lemma 2.5 will be used only in Theorem 4.1, whose proof may be skipped as well.

Lemma 2.5. *For all $\boldsymbol{m} \in \mathbb{N}^d$ and all permutations σ of the set $\{1, \ldots, d\}$, there exists a permutation matrix $V_{\boldsymbol{m};\sigma} \in \mathbb{C}^{(m_1+\ldots+m_d)\times(m_1+\ldots+m_d)}$ such that*

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

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

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 such that $\sigma(i) = d$. Define the permutation τ of $\{1, \ldots, d-1\}$ by setting $\tau(j) = \sigma(j)$ 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,

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

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

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

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

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

and the thesis is proved with $V_{\boldsymbol{m};\sigma} = U_{\boldsymbol{m};\sigma}(V_{(m_1,\ldots,m_{d-1});\tau} \oplus I_{m_d})$. □

2.4 Singular value and eigenvalue distributions of matrix-sequences: the symbol

Definition 2.2 (spectral distribution of a matrix-sequence, spectral symbol). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, and let $f : D \to \mathbb{C}$ be a measurable function, defined on a measurable set $D \subset \mathbb{R}^d$ with $0 < \mu_d(D) < \infty$.

- We say that $\{A_{\boldsymbol{n}}\}_n$ has an asymptotic eigenvalue (or spectral) distribution described by f, in symbols $\{A_{\boldsymbol{n}}\}_n \sim_\lambda f$, if, for all $F \in C_c(\mathbb{C})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) = \frac{1}{\mu_d(D)} \int_D F(f(\mathbf{x}))d\mathbf{x}. \tag{2.17}$$

  In this case, f is referred to as the eigenvalue (or spectral) symbol of the matrix-sequence $\{A_{\boldsymbol{n}}\}_n$.

- We say that $\{A_{\boldsymbol{n}}\}_n$ has an asymptotic singular value distribution described by f, in symbols $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$, if, for all $F \in C_c(\mathbb{R})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) = \frac{1}{\mu_d(D)} \int_D F(|f(\mathbf{x})|)d\mathbf{x}. \tag{2.18}$$

  In this case, f is referred to as the singular value symbol of the matrix-sequence $\{A_{\boldsymbol{n}}\}_n$.

It is clear 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 with the test function $F(|\cdot|) \in C_c(\mathbb{C})$, we get

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

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

2.5 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 couple of pages, all the details of this extensive topic; we just report the results that will be used hereinafter.

Given $\boldsymbol{n} \in \mathbb{N}^d$, a matrix of the form

$$[a_{\boldsymbol{i}-\boldsymbol{j}}]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}} \in \mathbb{C}^{N(\boldsymbol{n})\times N(\boldsymbol{n})}, \tag{2.19}$$

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

$$f_{\boldsymbol{k}} = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} f(\boldsymbol{\theta})\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}} d\boldsymbol{\theta}, \qquad \boldsymbol{k} \in \mathbb{Z}^d, \tag{2.20}$$

where $\boldsymbol{k}\cdot\boldsymbol{\theta} = k_1\theta_1 + \ldots + k_d\theta_d$. For every $\boldsymbol{n} \in \mathbb{N}^d$, the $\boldsymbol{n}$-th Toeplitz matrix associated with f is defined as

$$T_{\boldsymbol{n}}(f) = [f_{\boldsymbol{i}-\boldsymbol{j}}]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}}. \tag{2.21}$$

We call $\{T_{\boldsymbol{n}}(f)\}_{\boldsymbol{n}\in\mathbb{N}^d}$ the family of (multilevel) 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: for all $\alpha,\beta\in\mathbb{C}$ and $f,g\in L^1([-\pi,\pi]^d)$,

$$T_{\boldsymbol{n}}(\alpha f+\beta g)=\alpha T_{\boldsymbol{n}}(f)+\beta T_{\boldsymbol{n}}(g).$$

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 (2.20). For every $f\in L^1([-\pi,\pi]^d)$, the Fourier coefficients of f are related to those of $\overline{f}$ 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 $\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1},\ldots,\boldsymbol{n}$,

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

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.,[1] then all the matrices $T_{\boldsymbol{n}}(f)$ are Hermitian. Another nice property of the Toeplitz operator $T_{\boldsymbol{n}}(\cdot)$ is that $T_{\boldsymbol{n}}(1)=I_{N(\boldsymbol{n})}$.

Theorem 2.5 is a fundamental result concerning multilevel Toeplitz matrices. It is known in the literature as the Szegö–Tilli theorem. We refer the reader to [8] for a rich account concerning the history of the Szegö theorem, originally appeared in [31]. Tilli's proof of Theorem 2.5 can be found in [53]. We also refer the reader to [28] for a proof of Theorem 2.5 based on the notion of a.c.s. (see Section 3); the proof in [28] 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 2.5. *Let $f\in L^1([-\pi,\pi]^d)$, then $\{T_{\boldsymbol{n}}(f)\}_n\sim_\sigma f$. If moreover f is real a.e., then $\{T_{\boldsymbol{n}}(f)\}_n\sim_\lambda f$.*

In Theorem 2.5, $\{T_{\boldsymbol{n}}(f)\}_n$ 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 $n\to\infty$. In this regard, we recall that, for all matrix-sequences $\{A_{\boldsymbol{n}}\}_n$, it is always understood that the multi-index $\boldsymbol{n}$ tends to ∞ when $n\to\infty$; see Section 2.1.2.

Important inequalities involving Toeplitz matrices and Schatten p-norms can be found in [44, Corollary 3.5]. We report them in the next theorem for future use.

Theorem 2.6. *Let $f\in L^p([-\pi,\pi]^d),\ 1\le p\le\infty$, and let $\boldsymbol{n}\in\mathbb{N}^d$. Then,*

$$\|T_{\boldsymbol{n}}(f)\|=\|T_{\boldsymbol{n}}(f)\|_\infty\le\|f\|_{L^\infty},\qquad \text{if } p=\infty, \tag{2.22}$$

$$\|T_{\boldsymbol{n}}(f)\|_p\le\frac{1}{(2\pi)^d}\|f\|_{L^p}\,N(\boldsymbol{n})^{1/p},\qquad \text{if } 1\le p<\infty. \tag{2.23}$$

In particular, using the convention that $N(\boldsymbol{n})^{1/\infty}=1$, the inequality

$$\|T_{\boldsymbol{n}}(f)\|_p\le N(\boldsymbol{n})^{1/p}\|f\|_{L^p} \tag{2.24}$$

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

Lemma 2.6 relates tensor products and Toeplitz matrices.

[1] 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.

Lemma 2.6. *Let $f_1, \ldots, f_d \in L^1([-\pi, \pi])$ and $\boldsymbol{n} \in \mathbb{N}^d$. Then,*

$$T_{n_1}(f_1) \otimes \cdots \otimes T_{n_d}(f_d) = T_{\boldsymbol{n}}(f_1 \otimes \cdots \otimes f_d) \tag{2.25}$$

(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 (2.10). 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 $\boldsymbol{i}, \boldsymbol{j} = \boldsymbol{1}, \ldots, \boldsymbol{n}$,

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

and (2.25) follows. □

Lemma 2.7 will be used in Section 4.1.1 to study the Locally Toeplitz operator. The result of this lemma generalizes [16, Proposition 2] in the case where the involved generating functions f, g are scalar-valued. By Hölder's inequality [37], if $f \in L^p([-\pi, \pi]^d)$ and $g \in L^q([-\pi, \pi]^d)$, where $1 \le p, q \le \infty$ are conjugate exponents ($\frac{1}{p} + \frac{1}{q} = 1$), then $fg \in L^1([-\pi, \pi]^d)$. In this case, we can consider the three matrices $T_{\boldsymbol{k}}(f)$, $T_{\boldsymbol{k}}(g)$ and $T_{\boldsymbol{k}}(fg)$.

Lemma 2.7. *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,*

$$\lim_{\boldsymbol{k} \to \infty} \frac{\|T_{\boldsymbol{k}}(f) T_{\boldsymbol{k}}(g) - T_{\boldsymbol{k}}(fg)\|_1}{N(\boldsymbol{k})} = 0. \tag{2.26}$$

Proof. If f, g were in $L^\infty([-\pi, \pi]^d)$, then (2.26) holds by [16, Proposition 2] and the proof is over. In the general case where $f \in L^p([-\pi, \pi]^d)$ and $g \in L^q([-\pi, \pi]^d)$, the proof requires a little more effort.

Take two sequences $\{f_m\}$ and $\{g_m\}$ 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 $T_{\boldsymbol{k}}(\cdot)$ and Eqs. (2.7), (2.24), for every m and every $\boldsymbol{k} \in \mathbb{N}^d$ we have

$$\begin{aligned}
&\|T_{\boldsymbol{k}}(f) T_{\boldsymbol{k}}(g) - T_{\boldsymbol{k}}(fg)\|_1 \\
&\le \|T_{\boldsymbol{k}}(f - f_m) T_{\boldsymbol{k}}(g)\|_1 + \|T_{\boldsymbol{k}}(f_m) T_{\boldsymbol{k}}(g - g_m)\|_1 + \|T_{\boldsymbol{k}}(f_m) T_{\boldsymbol{k}}(g_m) - T_{\boldsymbol{k}}(f_m g_m)\|_1 + \|T_{\boldsymbol{k}}(f_m g_m - fg)\|_1 \\
&\le N(\boldsymbol{k})^{1/p} \|f - f_m\|_{L^p} N(\boldsymbol{k})^{1/q} \|g\|_{L^q} + N(\boldsymbol{k})^{1/p} \|f_m\|_{L^p} N(\boldsymbol{k})^{1/q} \|g - g_m\|_{L^q} \\
&\quad + \|T_{\boldsymbol{k}}(f_m) T_{\boldsymbol{k}}(g_m) - T_{\boldsymbol{k}}(f_m g_m)\|_1 + N(\boldsymbol{k}) \|f_m g_m - fg\|_{L^1} \\
&\le N(\boldsymbol{k}) \left[\|f - f_m\|_{L^p} \|g\|_{L^q} + \sup_h \|f_h\|_{L^p} \|g - g_m\|_{L^q} + \frac{\|T_{\boldsymbol{k}}(f_m) T_{\boldsymbol{k}}(g_m) - T_{\boldsymbol{k}}(f_m g_m)\|_1}{N(\boldsymbol{k})} + \|f_m g_m - fg\|_{L^1} \right].
\end{aligned} \tag{2.27}$$

By [16, Proposition 2],

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

so, dividing (2.27) by $N(\boldsymbol{k})$ and passing to the limit as $\boldsymbol{k} \to \infty$, we get

$$\limsup_{\boldsymbol{k} \to \infty} \frac{\|T_{\boldsymbol{k}}(f) T_{\boldsymbol{k}}(g) - T_{\boldsymbol{k}}(fg)\|_1}{N(\boldsymbol{k})} \le \|f - f_m\|_{L^p} \|g\|_{L^q} + \sup_h \|f_h\|_{L^p} \|g - g_m\|_{L^q} + \|f_m g_m - fg\|_{L^1}.$$

This relation holds for every m. Passing to the limit as $m \to \infty$ and observing that $f_m g_m \to fg$ in $L^1([-\pi, \pi]^d)$ by Hölder's inequality, 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 [41]. This notion lays the foundations for a (spectral) approximation theory for matrix-sequences and provides general tools (Theorems 3.3 and 3.5) 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 [28] for a deeper insight 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 $n \geq n_m$,

$$A_{\boldsymbol{n}} = B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m}, \tag{3.1}$$

$$\operatorname{rank}(R_{\boldsymbol{n},m}) \leq c(m)N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m}\| \leq \omega(m),$$

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

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

In this definition, it is understood that the matrices $B_{\boldsymbol{n},m},\ R_{\boldsymbol{n},m},\ N_{\boldsymbol{n},m}$ have size $N(\boldsymbol{n})$, like the matrix $A_{\boldsymbol{n}}$, otherwise (3.1) would have no meaning. 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) and the related conditions on $R_{\boldsymbol{n},m}$ and $N_{\boldsymbol{n},m}$ hold for $m \geq M$; then, defining $n_m = 1,\ c(m) = 1,\ \omega(m) = 0$ and $R_{\boldsymbol{n},m} = A_{\boldsymbol{n},m} - B_{\boldsymbol{n},m},\ N_{\boldsymbol{n},m} = O$ for $m < M$, we see that they actually hold 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 lies in Theorems 3.3 and 3.5, whose proofs appeared in [41, 28]. Theorem 3.3 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 $\{B_{\boldsymbol{n},m}\}_n,\ m \in \mathbb{N}$. For its proof, some intermediate results are needed.

Theorem 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 and every $F \in C_c^1(\mathbb{R})$, there exists $\displaystyle\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) = \phi_m(F) \in \mathbb{C}$;*
3. *for every $F \in C_c^1(\mathbb{R})$, there exists $\displaystyle\lim_{m\to\infty} \phi_m(F) = \phi(F) \in \mathbb{C}$.*

Then, for all $F \in C_c^1(\mathbb{R})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) = \phi(F). \tag{3.2}$$

Proof. We first observe that it suffices to prove (3.2) for those test functions $F \in C_c^1(\mathbb{R})$ that are real-valued. Indeed, any (complex-valued) $F \in C_c^1(\mathbb{R})$ can be decomposed as $F = \Re(F) + \mathrm{i}\Im(F)$, where $\Re(F), \Im(F) \in C_c^1(\mathbb{R})$. Thus, once we have proved (3.2) for all real-valued functions in $C_c^1(\mathbb{R})$, we have

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) = \lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} [\Re(F(\sigma_j(A_{\boldsymbol{n}}))) + \mathrm{i}\Im(F(\sigma_j(A_{\boldsymbol{n}})))]$$
$$= \phi(\Re(F)) + \mathrm{i}\phi(\Im(F)) = \phi(F),$$

where the last equality holds by the linearity of the functional ϕ, which follows from its definition.

Let $F \in C_c^1(\mathbb{R})$ be real-valued. For all n, m we have

$$\left|\frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \phi(F)\right| \le \left|\frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m}))\right|$$
$$+ \left|\frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) - \phi_m(F)\right|$$
$$+ |\phi_m(F) - \phi(F)|. \tag{3.3}$$

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, if we prove that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \left|\frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m}))\right| = 0, \tag{3.4}$$

then, passing first to the $\limsup_{n\to\infty}$ and then to the $\lim_{m\to\infty}$ in (3.3), we get the thesis.

In conclusion, we only have to prove (3.4). To this end, we recall that $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Hence, for every m there exists n_m such that, for $n \ge n_m$,

$$A_{\boldsymbol{n}} = B_{\boldsymbol{n},m} + R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m}, \tag{3.5}$$

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

where $\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0$. We can then write, for every m and every $n \ge n_m$,

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

We will consider separately the two terms in the right-hand side of (3.6), 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, (3.4) is proved and the thesis follows.

In order to estimate the first term in the right-hand side of (3.6), we use the the perturbation theorem for singular values (Theorem 2.3). We have

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

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

In order to estimate the second term in the right-hand side of (3.6), 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 (3.6) we have

$$\begin{aligned}
&\left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}F(\sigma_j(B_{\boldsymbol{n},m}+R_{n,m}))-\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}F(\sigma_j(B_{\boldsymbol{n},m}))\right|\\
&\leq\left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}H(\sigma_j(B_{\boldsymbol{n},m}+R_{n,m}))-\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}H(\sigma_j(B_{\boldsymbol{n},m}))\right|\\
&\quad+\left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}K(\sigma_j(B_{\boldsymbol{n},m}+R_{n,m}))-\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}K(\sigma_j(B_{\boldsymbol{n},m}))\right|.
\end{aligned}\tag{3.7}$$

Defining $r_{n,m}=\text{rank}(R_{n,m})\leq c(m)N(\boldsymbol{n})$, Theorem 2.1 gives

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

and, moreover, it is clear from our notation that

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

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

$$\begin{aligned}
&\left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}H(\sigma_j(B_{\boldsymbol{n},m}+R_{n,m}))-\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}H(\sigma_j(B_{\boldsymbol{n},m}))\right|\\
&\leq\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}|H(\sigma_j(B_{\boldsymbol{n},m}+R_{n,m}))-H(\sigma_j(B_{\boldsymbol{n},m}))|\\
&\leq\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})}\left|H(\sigma_{j-r_{n,m}}(B_{\boldsymbol{n},m}))-H(\sigma_{j+r_{n,m}}(B_{\boldsymbol{n},m}))\right|
\end{aligned}$$

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

Similarly, one can show that the second term in the right-hand side of (3.7) is bounded from above by $2c(m)\|K\|_\infty$, implying that the quantity in (3.7), namely the second term in the right-hand side of (3.6), 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. □

The only unpleasant point about Theorem 3.1 is that, in traditional formulations of asymptotic singular value distribution results, the usual set of test functions F is $C_c(\mathbb{R})$ and not $C_c^1(\mathbb{R})$; see Definition 2.2. However, this point is readily settled in Theorem 3.3. For the proof of Theorem 3.3, we shall use the following corollary of the Banach-Steinhaus theorem [37].

Theorem 3.2. *Let $\mathscr{E}, \mathscr{F}$ be normed vector spaces, with $\mathscr{E}$ a Banach space, and let $T_n : \mathscr{E} \to \mathscr{F}$ be a sequence of continuous linear operators. Assume that, for all $x \in \mathscr{E}$, there exists $\lim_{n\to\infty} T_n x = Tx \in \mathscr{F}$. Then,*

- $\sup \|T_n\| < \infty$;
- *$T : \mathscr{E} \to \mathscr{F}$ is a continuous linear operator with $\|T\| \le \liminf_{n\to\infty} \|T_n\|$.*

Theorem 3.3. *Let $\{A_{\boldsymbol{n}}\}_n$ be a sequence of matrices. Assume that:*

1. *$\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$;*
2. *for every m and every $F \in C_c(\mathbb{R})$, there exists $\displaystyle\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) = \phi_m(F) \in \mathbb{C}$;*
3. *for every $F \in C_c(\mathbb{R})$, there exists $\lim_{m\to\infty} \phi_m(F) = \phi(F) \in \mathbb{C}$.*

Then $\phi : (C_c(\mathbb{R}), \|\cdot\|_\infty) \to \mathbb{C}$ is a continuous linear functional with $\|\phi\| \le 1$, and, for all $F \in C_c(\mathbb{R})$,

$$
\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) = \phi(F). \tag{3.8}
$$

Proof. For fixed n, m, let

$$
\phi_{n,m}(F) = \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(B_{\boldsymbol{n},m})) : (C_c(\mathbb{R}), \|\cdot\|_\infty) \to \mathbb{C}.
$$

It is clear that each $\phi_{n,m}$ is a continuous linear functional on $(C_c(\mathbb{R}), \|\cdot\|_\infty)$ with $\|\phi_{n,m}\| \le 1$. Indeed, the linearity of $\phi_{n,m}$ is obvious and the inequality $|\phi_{n,m}(F)| \le \|F\|_\infty$, which is satisfied for all $F \in C_c(\mathbb{R})$, yields the continuity of $\phi_{n,m}$ as well as the bound $\|\phi_{n,m}\| \le 1$. The functional ϕ_m is the pointwise limit of $\phi_{n,m}$ as $n \to \infty$. Hence, by Theorem 3.2, ϕ_m is a continuous linear functional on $(C_c(\mathbb{R}), \|\cdot\|_\infty)$ with $\|\phi_m\| \le 1$. The functional ϕ is the pointwise limit of ϕ_m as $m \to \infty$. Hence, again by Theorem 3.2, ϕ is a continuous linear functional on $(C_c(\mathbb{R}), \|\cdot\|_\infty)$ with $\|\phi\| \le 1$.

Now, fix $F \in C_c(\mathbb{R})$. For all $\epsilon > 0$ we can find $F_\epsilon \in C_c^1(\mathbb{R})$ such that $\|F - F_\epsilon\|_\infty \le \epsilon$. As a consequence, for all $\epsilon > 0$ and for all n we have

$$\begin{aligned}
&\left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \phi(F)\right| \\
&\le \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\sigma_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(A_{\boldsymbol{n}}))\right| + \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(A_{\boldsymbol{n}})) - \phi(F_\epsilon)\right| + |\phi(F_\epsilon) - \phi(F)| \\
&\le \|F - F_\epsilon\|_\infty + \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\sigma_j(A_{\boldsymbol{n}})) - \phi(F_\epsilon)\right| + |\phi(F_\epsilon) - \phi(F)|.
\end{aligned}$$

Considering that (3.8) holds for F_ϵ by Theorem 3.1, we have

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

Passing to the limit as $\epsilon \to 0$ and taking into account the continuity of ϕ, we obtain

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

which means that (3.8) holds for every $F \in C_c(\mathbb{R})$. □

In Theorems 3.4 and 3.5 we prove analogous versions of Theorems 3.1 and 3.3 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 3.5 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 $\{B_{\boldsymbol{n},m}\}_n,\ m \in \mathbb{N}$, *again formed by Hermitian matrices.*

The next lemma shows that, whenever the matrices $A_{\boldsymbol{n}}$ 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.

Lemma 3.1. *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 (i.e. every $B_{\boldsymbol{n},m}$ is Hermitian). Then, for every m there exists n_m such that, for $n \ge n_m$,*

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

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

where the quantities $n_m,\ c(m),\ \omega(m)$ depend only on m, the matrices $R_{\boldsymbol{n},m}, N_{\boldsymbol{n},m}$ are Hermitian, 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 $\mathrm{rank}(\Re(X)) \le 2\,\mathrm{rank}(X)$ and $\|\Re(X)\| \le \|X\|$ 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. □

Theorem 3.4. *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 *and every* $F \in C_c^1(\mathbb{R})$*, there exists* $\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m})) = \phi_m(F) \in \mathbb{C}$*;*

3. *for every* $F \in C_c^1(\mathbb{R})$*, there exists* $\lim_{m\to\infty} \phi_m(F) = \phi(F) \in \mathbb{C}$.

Then, for all $F \in C_c^1(\mathbb{R})$*,*

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) = \phi(F). \tag{3.9}$$

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

As in the proof of Theorem 3.1, one can show that it suffices to prove (3.9) for all real-valued test functions $F \in C_c^1(\mathbb{R})$. We fix a real-valued function $F \in C_c^1(\mathbb{R})$ and we bound the quantity

$$\left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \phi(F)\right|$$

as in (3.3), with 'σ_j' replaced by 'λ_j'. 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, the thesis is proved if we show that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \left|\frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})}\sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m}))\right| = 0. \tag{3.10}$$

To prove (3.10), we recall that $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ and that $A_{\boldsymbol{n}}$, $B_{\boldsymbol{n},m}$ are Hermitian. Hence, by Lemma 3.1, for every m there exists n_m such that, for $n \geq n_m$, the splitting (3.5) holds with Hermitian $R_{\boldsymbol{n},m}$, $N_{\boldsymbol{n},m}$. Following the proof of Theorem 3.1, we arrive at the inequality (3.6), 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 Theorem 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 Theorem 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. □

The only unpleasant point about Theorem 3.4 is that, in traditional formulations of asymptotic spectral distribution results, the usual set of test functions F is $C_c(\mathbb{C})$ and not $C_c^1(\mathbb{R})$; see Definition 2.2. This point is readily settled in Theorem 3.5, which is analogous to Theorem 3.3.

Theorem 3.5. *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 *and every* $F \in C_c(\mathbb{C})$*, there exists* $\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m})) = \phi_m(F) \in \mathbb{C}$*;*

3. *for every* $F \in C_c(\mathbb{C})$*, there exists* $\lim_{m\to\infty} \phi_m(F) = \phi(F) \in \mathbb{C}$.

Then $\phi : (C_c(\mathbb{C}), \|\cdot\|_\infty) \to \mathbb{C}$ *is a continuous linear functional with* $\|\phi\| \le 1$*, and, for all* $F \in C_c(\mathbb{C})$*,*

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) = \phi(F). \tag{3.11}$$

Proof. The proof is similar to the one used for proving Theorem 3.3. For fixed n, m, let

$$\phi_{n,m}(F) = \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(B_{\boldsymbol{n},m})) : (C_c(\mathbb{C}), \|\cdot\|_\infty) \to \mathbb{C}.$$

It is clear that each $\phi_{n,m}$ is a continuous linear functional on $(C_c(\mathbb{C}), \|\cdot\|_\infty)$ with $\|\phi_{n,m}\| \le 1$. Indeed, the linearity of $\phi_{n,m}$ is obvious and the inequality

$$|\phi_{n,m}(F)| \le \|F\|_{\infty,\mathbb{R}}, \tag{3.12}$$

which is satisfied for all $F \in C_c(\mathbb{C})$, yields the continuity of $\phi_{n,m}$ as well as the bound $\|\phi_{n,m}\| \le 1$. The functional ϕ_m is the pointwise limit of $\phi_{n,m}$ as $n \to \infty$. Hence, by Theorem 3.2, ϕ_m is a continuous linear functional on $(C_c(\mathbb{C}), \|\cdot\|_\infty)$ with $\|\phi_m\| \le 1$; moreover, by (3.12), for all $F \in C_c(\mathbb{C})$ we have

$$|\phi_m(F)| \le \|F\|_{\infty,\mathbb{R}}. \tag{3.13}$$

The functional ϕ is the pointwise limit of ϕ_m as $m \to \infty$. Hence, again by Theorem 3.2, ϕ is a continuous linear functional on $(C_c(\mathbb{C}), \|\cdot\|_\infty)$ with $\|\phi\| \le 1$; moreover, by (3.13), for all $F \in C_c(\mathbb{C})$ we have

$$|\phi(F)| \le \|F\|_{\infty,\mathbb{R}}. \tag{3.14}$$

Now, fix $F \in C_c(\mathbb{C})$. For all $\epsilon > 0$ we can find $F_\epsilon \in C_c(\mathbb{C})$ such that F_ϵ restricted to $\mathbb{R}$ belongs to $C_c^1(\mathbb{R})$ and $\|F - F_\epsilon\|_{\infty,\mathbb{R}} \le \epsilon$. Then, for all $\epsilon > 0$ and for all n we have

$$\begin{aligned}
&\left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \phi(F) \right| \\
&\le \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\lambda_j(A_{\boldsymbol{n}})) \right| + \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\lambda_j(A_{\boldsymbol{n}})) - \phi(F_\epsilon) \right| + |\phi(F_\epsilon) - \phi(F)| \\
&\le \|F - F_\epsilon\|_{\infty,\mathbb{R}} + \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F_\epsilon(\lambda_j(A_{\boldsymbol{n}})) - \phi(F_\epsilon) \right| + |\phi(F_\epsilon - F)|.
\end{aligned}$$

Considering that (3.11) holds for F_ϵ by Theorem 3.4 and using (3.14), we have

$$\limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \phi(F) \right| \le \epsilon + \epsilon.$$

Passing to the limit as $\epsilon \to 0$, we obtain

$$\limsup_{n\to\infty} \left| \frac{1}{N(\boldsymbol{n})} \sum_{j=1}^{N(\boldsymbol{n})} F(\lambda_j(A_{\boldsymbol{n}})) - \phi(F) \right| = 0,$$

which means that (3.11) holds for every $F \in C_c(\mathbb{C})$. □

Two important corollaries of Theorems 3.3 and 3.5 are given in the following. They will be used in Section 5.3 to prove the asymptotic spectral properties of GLT sequences.

Corollary 3.1. *Let $\{A_{\boldsymbol{n}}\}_n$ be a sequence of matrices. 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 $f_m : D \subset \mathbb{R}^d \to \mathbb{C}$;*
3. *$|f_m| \to |f|$ in measure over D when $m \to \infty$, being $f : D \to \mathbb{C}$ another measurable function.*[2]

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

Proof. Apply Theorem 3.3 with

$$\phi_m(F) = \frac{1}{\mu_d(D)} \int_D F(|f_m(\mathbf{x})|)d\mathbf{x}, \qquad \phi(F) = \frac{1}{\mu_d(D)} \int_D F(|f(\mathbf{x})|)d\mathbf{x}.$$

Note that $\phi_m(F) \to \phi(F)$ for all $F \in C_c(\mathbb{R})$ by Lemma 2.1. □

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 $f_m : D \subset \mathbb{R}^d \to \mathbb{C}$;*
3. *$f_m \to f$ in measure over D when $m \to \infty$, being $f : D \to \mathbb{C}$ another measurable function.*

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

Proof. Apply Theorem 3.5 with

$$\phi_m(F) = \frac{1}{\mu_d(D)} \int_D F(f_m(\mathbf{x}))d\mathbf{x}, \qquad \phi(F) = \frac{1}{\mu_d(D)} \int_D F(f(\mathbf{x}))d\mathbf{x},$$

and use again Lemma 2.1 to see that $\phi_m(F) \to \phi(F)$ for all $F \in C_c(\mathbb{C})$. □

[2] Note that '$f_m \to f$ in measure' implies '$|f_m| \to |f|$ in measure'.

3.2 The a.c.s. algebra

In this section, we investigate the numerous 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 5.4. We begin with the following observation, whose proof is very simple and is left to the reader.

Remark 3.2. 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 $n\ge n_m$,
$$A_{\boldsymbol{n}}=B_{\boldsymbol{n},m}+R_{\boldsymbol{n},m}+N_{\boldsymbol{n},m},$$
where $\text{rank}(R_{\boldsymbol{n},m})\le c(m)N(\boldsymbol{n})$, $\|N_{\boldsymbol{n},m}\|\le\omega(m)$ and $\lim_{m\to\infty}c(m)=\lim_{m\to\infty}\omega(m)=0$;
- for every m there exists n'_m such that, for $n\ge n'_m$,
$$A'_{\boldsymbol{n}}=B'_{\boldsymbol{n},m}+R'_{\boldsymbol{n},m}+N'_{\boldsymbol{n},m},$$
where $\text{rank}(R'_{\boldsymbol{n},m})\le c'(m)N(\boldsymbol{n})$, $\|N'_{\boldsymbol{n},m}\|\le\omega'(m)$ and $\lim_{m\to\infty}c'(m)=\lim_{m\to\infty}\omega'(m)=0$.

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

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

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 $n\ge n_M$,

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

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

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

3. For every $M>0$ *there exists* n_M *such that, for* $n\ge n_M$,

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

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

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

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 $n\ge n_M$,

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

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

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

($2\Rightarrow 1$) 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 lim sup, 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.

($1\Rightarrow 3$) 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 a 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 $n \ge n_M$,

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

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

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

where $\mathrm{rank}(A_{\boldsymbol{n},M}) \le r(M)N(\boldsymbol{n})$, $\|\tilde{A}_{\boldsymbol{n},M}\| \le M$ and $\lim_{M\to\infty} r(M) = 0$. By the minimax principle for singular values [7, Problem III.6.1], for all $i = 1,\ldots,N(\boldsymbol{n})$ we have

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

Since $\mathrm{rank}(A_{\boldsymbol{n},M}) \le r(M)N(\boldsymbol{n})$, $\hat{A}_{\boldsymbol{n},M}$ has at most $r(M)N(\boldsymbol{n})$ nonzero singular values. Therefore, $A_{\boldsymbol{n}}$ 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.; cf. [51, Proposition 2.7].

Proposition 3.3. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence such that $\{A_{\boldsymbol{n}}\}_n \sim_\sigma f$ for some measurable $f : D \subset \mathbb{R}^d \to \mathbb{C}$. Then $\{A_{\boldsymbol{n}}\}_n$ is s.u.*

Proof. If we could choose $F = \chi_{(M,\infty)}$ as a test function in (2.18), then we would obtain

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

since $\mu_d\{|f| > M\} \to 0$ as $M \to \infty$ (by the dominated convergence theorem [37]), the proof would be over, thanks to Proposition 3.2. Nevertheless, $\chi_{(M,\infty)}$ cannot be chosen as a test function in (2.18), since it does not belong to $C_c(\mathbb{R})$. Therefore, to obtain the thesis, we need a little more work.

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

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

and

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

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

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

and so

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

By Proposition 3.2, $\{A_{\boldsymbol{n}}\}_n$ 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 [41].

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 $n \ge n_m$,

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

where

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

and $\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0$. 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 $n \ge n(M)$,

$$\begin{aligned} A_{\boldsymbol{n}} &= \hat{A}_{\boldsymbol{n},M} + \tilde{A}_{\boldsymbol{n},M}, \\ A'_{\boldsymbol{n}} &= \hat{A}'_{\boldsymbol{n},M} + \tilde{A}'_{\boldsymbol{n},M}, \end{aligned}$$

where

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

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

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

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

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 sometimes 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 of Definition 3.1. In this section, we provide two useful criterions to solve this problem.

The first criterion, expressed in Theorem 3.6 and Corollary 3.3, is formulated in terms of Schatten p-norms; it will be mainly applied with $p = 1$.

Lemma 3.2. *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

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

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 a SVD of C and write

$$C = U\Sigma V^* = U\Sigma^{(1)}V^* + U\Sigma^{(2)}V^*,$$

where $\Sigma^{(1)}$ 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 $\Sigma^{(2)} = \Sigma - \Sigma^{(1)}$ is obtained from Σ by setting to 0 all the singular values that exceed $\epsilon^{\frac{1}{p+1}}$. Then

$$C = R + N,$$

where $R = U\Sigma^{(1)}V^*$ and $N = U\Sigma^{(2)}V^*$ satisfy $\text{rank}(R) \le \epsilon^{\frac{1}{p+1}} s$ and $\|N\| \le \epsilon^{\frac{1}{p+1}}$. □

Theorem 3.6. *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, for every m there exists $\hat{n}_m$ such that, for $n \ge \hat{n}_m$,

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

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

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

Proof. By Lemma 3.2, for every m and every $n \geq n_m$ we have

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

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

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

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

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

Since $\epsilon(m) \to 0$ by assumption, the thesis is proved with $c(m) = \omega(m) = \left(\epsilon(m) + \frac{1}{m}\right)^{\frac{1}{p+1}}$. □

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 $\{A_{\boldsymbol{n}}\}_n$.

The second criterion, expressed in Theorem 3.7 and Corollary 3.4, is formulated in terms of the singular value distribution.

Theorem 3.7. *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 each m, and $g_m \to 0$ in measure. Then, for every m there exists n_m such that, for $n \geq n_m$,*

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

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

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

Proof. For any fixed $k \in \mathbb{N}$, let $F_k \in C_c(\mathbb{R})$ such that $F_k = 1$ over $[0, 1/2k]$, $F = 0$ outside $[-1/k, 1/k]$, and $0 \leq F_k \leq 1$ over $\mathbb{R}$. For every m, k, we have

$$\begin{aligned}
\frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} :\ \sigma_i(C_{\boldsymbol{n},m}) > 1/k\}}{N(\boldsymbol{n})} &= 1 - \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} :\ \sigma_i(C_{\boldsymbol{n},m}) \leq 1/k\}}{N(\boldsymbol{n})} \\
&= 1 - \frac{1}{N(\boldsymbol{n})} \sum_{i=1}^{N(\boldsymbol{n})} \chi_{[0,1/k]}(\sigma_i(C_{\boldsymbol{n},m})) \\
&\leq 1 - \frac{1}{N(\boldsymbol{n})} \sum_{i=1}^{N(\boldsymbol{n})} F_k(\sigma_i(C_{\boldsymbol{n},m})) \overset{n\to\infty}{\longrightarrow} c(m,k), \qquad (3.15)
\end{aligned}$$

where

$$c(m,k) = 1 - \frac{1}{\mu_d(D)} \int_D F_k(|g_m(\mathbf{x})|) d\mathbf{x}.$$

By Lemma 2.1, $c(m,k) \to 0$ as $m \to \infty$ and so there exists a sequence $\{k_m\}_m$ of natural numbers such that

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

From (3.15),

$$\limsup_{n\to\infty} \frac{\#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \ \sigma_i(C_{\boldsymbol{n},m}) > 1/k_m\}}{N(\boldsymbol{n})} \le c(m,k_m). \tag{3.16}$$

Let $C_{\boldsymbol{n},m} = U_{\boldsymbol{n},m}\Sigma_{\boldsymbol{n},m}V_{\boldsymbol{n},m}^*$ be a SVD of $C_{\boldsymbol{n},m}$. Let $\Sigma_{\boldsymbol{n},m}^{(1)}$ be the matrix obtained from $\Sigma_{\boldsymbol{n},m}$ by setting to 0 all the singular vaules that are less than or equal to $1/k_m$, and let $\Sigma_{\boldsymbol{n},m}^{(2)} = \Sigma_{\boldsymbol{n},m} - \Sigma_{\boldsymbol{n},m}^{(1)}$ be the matrix obtained from $\Sigma_{\boldsymbol{n},m}$ by setting to 0 all the singular values that exceed $1/k_m$. Then we can write

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

where $R_{\boldsymbol{n},m} = U_{\boldsymbol{n},m}\Sigma_{\boldsymbol{n},m}^{(1)}V_{\boldsymbol{n},m}^*$ and $N_{\boldsymbol{n},m} = U_{\boldsymbol{n},m}\Sigma_{\boldsymbol{n},m}^{(2)}V_{\boldsymbol{n},m}^*$. By definition, $\|N_{\boldsymbol{n},m}\| \le 1/k_m$. Moreover, Eq. (3.16) says that

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

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

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

This completes the proof. □

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 each m, and $g_m \to 0$ in measure. Then $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.*

4 LT and sLT sequences

In this section, we introduce and analyze the so-called Locally Toeplitz operator. Then, we define the Locally Toeplitz (LT) and separable Locally Toeplitz (sLT) sequences, and we study their properties. The results contained in this section are of fundamental importance for developing the theory of Generalized Locally Toeplitz (GLT) 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

$$\begin{aligned} &LT_{n_1,n_2}^{m_1,m_2}(a,f_1\otimes f_2)\\ &=\operatorname*{diag}_{j_1=1,\ldots,m_1} T_{\lfloor n_1/m_1\rfloor}(f_1)\otimes\Big[\operatorname*{diag}_{j_2=1,\ldots,m_2} a\Big(\frac{j_1}{m_1},\frac{j_2}{m_2}\Big)T_{\lfloor n_2/m_2\rfloor}(f_2)\ \oplus\ O_{n_2\bmod m_2}\Big]\ \oplus\ O_{(n_1\bmod m_1)n_2}.\end{aligned}$$

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 property of 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 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. The proof of Theorem 4.1 is rather technical and may be skipped in a first reading; only the statement of the theorem is relevant for what follows.

Theorem 4.1. *For any $\boldsymbol{m},\boldsymbol{n}\in\mathbb{N}^d$ there exists a permutation matrix $\Pi_{\boldsymbol{n}}^{\boldsymbol{m}}$, depending only on $\boldsymbol{m}$ and $\boldsymbol{n}$, 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 $d\geq 2$, let $\boldsymbol{n}'=(n_2,\ldots,n_d)$ and $\boldsymbol{m}'=(m_2,\ldots,m_d)$. By definition,

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

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 $N(\boldsymbol{n}',\boldsymbol{m}')=N(\boldsymbol{n}')-N(\boldsymbol{m}')N(\lfloor\boldsymbol{n}'/\boldsymbol{m}'\rfloor)$, we have

$$LT_{\boldsymbol{n}'}^{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big),f_2\otimes\cdots\otimes f_d\Big)=\Pi_{\boldsymbol{n}'}^{\boldsymbol{m}'}\Big[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor\boldsymbol{n}'/\boldsymbol{m}'\rfloor}(f_2\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')}\Big](\Pi_{\boldsymbol{n}'}^{\boldsymbol{m}'})^T.\quad(4.2)$$

Let us now work on the argument of the 'diag operator' in (4.1). By Lemma 2.3, Eq. (4.2) and the properties of tensor products (see Section 2.3.1), we get

$$
\begin{aligned}
&T_{\lfloor n_1/m_1\rfloor}(f_1)\otimes LT_{\boldsymbol{n}'}^{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big), f_2\otimes\cdots\otimes f_d\Big)\\
&=\Pi_{(\lfloor n_1/m_1\rfloor,N(\boldsymbol{n}'));[2,1]}\left\{LT_{\boldsymbol{n}'}^{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big), f_2\otimes\cdots\otimes f_d\Big)\otimes T_{\lfloor n_1/m_1\rfloor}(f_1)\right\}(\Pi_{(\lfloor n_1/m_1\rfloor,N(\boldsymbol{n}'));[2,1]})^T\\
&=\Pi_{(\lfloor n_1/m_1\rfloor,N(\boldsymbol{n}'));[2,1]}\left\{\Pi_{\boldsymbol{n}'}^{\boldsymbol{m}'}\left[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor}(f_2\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')}\right](\Pi_{\boldsymbol{n}'}^{\boldsymbol{m}'})^T\right.\\
&\qquad\qquad\left.\otimes\ T_{\lfloor n_1/m_1\rfloor}(f_1)\right\}(\Pi_{(\lfloor n_1/m_1\rfloor,N(\boldsymbol{n}'));[2,1]})^T\\
&=\Pi_{(\lfloor n_1/m_1\rfloor,N(\boldsymbol{n}'));[2,1]}(\Pi_{\boldsymbol{n}'}^{\boldsymbol{m}'}\otimes I_{\lfloor n_1/m_1\rfloor})\left\{\left[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor}(f_2\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')}\right]\right.\\
&\qquad\qquad\left.\otimes\ T_{\lfloor n_1/m_1\rfloor}(f_1)\right\}(\Pi_{\boldsymbol{n}'}^{\boldsymbol{m}'}\otimes I_{\lfloor n_1/m_1\rfloor})^T(\Pi_{(\lfloor n_1/m_1\rfloor,N(\boldsymbol{n}'));[2,1]})^T. \qquad (4.3)
\end{aligned}
$$

Using Remark 2.1, Lemma 2.3, Lemma 2.6 and the properties of tensor products and direct sums (see Section 2.3.1), we obtain

$$
\begin{aligned}
&\left[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor}(f_2\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')}\right]\otimes T_{\lfloor n_1/m_1\rfloor}(f_1)\\
&=D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor}(f_2\otimes\cdots\otimes f_d)\otimes T_{\lfloor n_1/m_1\rfloor}(f_1)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor}\\
&=\Pi_{(N(\boldsymbol{m}'),\lfloor n_1/m_1\rfloor,N(\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor));[1,3,2]}\left[D_{\boldsymbol{m}'}\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{n}'/\boldsymbol{m}'\rfloor}(f_2\otimes\cdots\otimes f_d)\right]\\
&\quad\cdot\ (\Pi_{(N(\boldsymbol{m}'),\lfloor n_1/m_1\rfloor,N(\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor));[1,3,2]})^T\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor}\\
&=\Pi_{(N(\boldsymbol{m}'),\lfloor n_1/m_1\rfloor,N(\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor));[1,3,2]}\left[D_{\boldsymbol{m}'}\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]\\
&\quad\cdot\ (\Pi_{(N(\boldsymbol{m}'),\lfloor n_1/m_1\rfloor,N(\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor));[1,3,2]})^T\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor}\\
&=(\Pi_{(N(\boldsymbol{m}'),\lfloor n_1/m_1\rfloor,N(\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor));[1,3,2]}\oplus I_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor})\\
&\quad\cdot\ \left[D_{\boldsymbol{m}'}\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{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor}\right]\\
&\quad\cdot\ (\Pi_{(N(\boldsymbol{m}'),\lfloor n_1/m_1\rfloor,N(\lfloor \boldsymbol{n}'/\boldsymbol{m}'\rfloor));[1,3,2]}\oplus I_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor})^T. \qquad (4.4)
\end{aligned}
$$

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

$$
\begin{aligned}
&T_{\lfloor n_1/m_1\rfloor}(f_1)\otimes LT_{\boldsymbol{n}'}^{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big), f_2\otimes\cdots\otimes f_d\Big)\\
&=P_{\boldsymbol{n}}^{\boldsymbol{m}}\left[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor}\right](P_{\boldsymbol{n}}^{\boldsymbol{m}})^T, \qquad (4.5)
\end{aligned}
$$

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

$$
\begin{aligned}
&LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f_1\otimes\cdots\otimes f_d)\\
&=\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)\underset{j_1=1,\ldots,m_1}{\mathrm{diag}}\Big[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor}\Big]\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)^T\\
&\quad\oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d}.
\end{aligned}
$$

From Lemma 2.5,

$$\begin{aligned}
&\operatorname*{diag}_{j_1=1,\ldots,m_1}\left[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor}\right]\\
&=\bigoplus_{j_1=1}^{m_1}\left[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor}\right]\\
&=V_{\boldsymbol{n}}^{\boldsymbol{m}}\left[\bigoplus_{j_1=1}^{m_1}\left[D_{\boldsymbol{m}'}\Big(a\Big(\frac{j_1}{m_1},\cdot\Big)\Big)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\right]\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor m_1}\right](V_{\boldsymbol{n}}^{\boldsymbol{m}})^T\\
&=V_{\boldsymbol{n}}^{\boldsymbol{m}}\left[D_{\boldsymbol{m}}(a)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor m_1}\right](V_{\boldsymbol{n}}^{\boldsymbol{m}})^T,
\end{aligned}$$

where

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

Thus,

$$\begin{aligned}
&LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f_1\otimes\cdots\otimes f_d)\\
&=\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)V_{\boldsymbol{n}}^{\boldsymbol{m}}\left[D_{\boldsymbol{m}}(a)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor m_1}\right](V_{\boldsymbol{n}}^{\boldsymbol{m}})^T\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)^T\\
&\quad\oplus\ O_{(n_1 \bmod m_1)n_2\cdots n_d}\\
&=\left[\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)V_{\boldsymbol{n}}^{\boldsymbol{m}}\ \oplus\ I_{(n_1 \bmod m_1)n_2\cdots n_d}\right]\\
&\quad\cdot\left[D_{\boldsymbol{m}}(a)\otimes T_{\lfloor \boldsymbol{n}/\boldsymbol{m}\rfloor}(f_1\otimes\cdots\otimes f_d)\ \oplus\ O_{N(\boldsymbol{n}',\boldsymbol{m}')\lfloor n_1/m_1\rfloor m_1+(n_1 \bmod m_1)n_2\cdots n_d}\right]\\
&\quad\cdot\left[(V_{\boldsymbol{n}}^{\boldsymbol{m}})^T\Big(\bigoplus_{j_1=1}^{m_1}P_{\boldsymbol{n}}^{\boldsymbol{m}}\Big)^T\ \oplus\ I_{(n_1 \bmod m_1)n_2\cdots n_d}\right].
\end{aligned}$$

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

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. 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. Note also that $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)=LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,g)$ whenever $f=g$ a.e.

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 $(X\otimes Y)^*=X^*\otimes Y^*$, $(X\oplus Y)^*=X^*\oplus Y^*$ (see Section 2.3.1), and from the equality $[T_{\boldsymbol{k}}(f)]^*=T_{\boldsymbol{k}}(\overline{f})$ (see Section 2.5).

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 the definition of $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)$, the invariance of $\|\cdot\|$ and $\|\cdot\|_p$ by unitary transformations (such as permutations), and the equalities (2.11)–(2.12). □

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

By Hölder's inequality [37], 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 $\boldsymbol{n},\boldsymbol{m}\in\mathbb{N}^d$, 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*

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

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 2.7. In particular, for every $\boldsymbol{m}\in\mathbb{N}^d$ there exists $\boldsymbol{n}_{\boldsymbol{m}}\in\mathbb{N}^d$ such that, for $\boldsymbol{n}\ge\boldsymbol{n}_{\boldsymbol{m}}$,

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

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

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,

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

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 from (4.10) and Lemma 3.2. □

The next two theorems 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 3.3 and 3.5, they play a central role in the computation of the singular value and eigenvalue distribution of GLT sequences.

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

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

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

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

plus further $N(\boldsymbol{n}) - N(\boldsymbol{m})N(\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor)$ singular values equal to 0. Note that $\frac{N(\boldsymbol{n})-N(\boldsymbol{m})N(\lfloor \boldsymbol{n}/\boldsymbol{m} \rfloor)}{N(\boldsymbol{n})} \to 0$ as $\boldsymbol{n} \to \infty$. Therefore, for any $F \in C_c(\mathbb{R})$,

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

where the latter equality is due to Theorem 2.5. This proves (4.12).

If $a_1, \ldots, a_p$ 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 $\boldsymbol{\theta} \in [-\pi, \pi]^d$, and so

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

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

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

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

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

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

where the latter equality follows from Theorem 2.5. This proves (4.15).

If $a_1,\dots,a_p$ are Riemann-integrable, then the function $\mathbf{x}\mapsto F\left(\Re\left(\sum_{i=1}^{p}a_i\left(\mathbf{x}\right)f_i(\boldsymbol{\theta})\right)\right)$ 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

First we start with the definitions and basic examples, including Toeplitz and diagonal sampling matrix-sequences, which represent the building blocks for approximating PDE matrix-sequences. Then we illustrate the main properties of LT and sLT sequences.

4.2.1 Definition and basic examples: Toeplitz and diagonal sampling matrix-sequences

Definition 4.3 (**LT sequence**)**.** Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, with $\boldsymbol{n}\in\mathbb{N}^d$. We say that $\{A_{\boldsymbol{n}}\}_n$ is a Locally Toeplitz (LT) sequence if there exist

- a Riemann-integrable function $a:[0,1]^d\to\mathbb{C}$,
- a function $f\in L^1([-\pi,\pi]^d)$,

such that $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A_n\}_n$, in the following sense: for all $\boldsymbol{m}\in\mathbb{N}^d$ there is $\boldsymbol{n}_{\boldsymbol{m}}\in\mathbb{N}^d$ such that, for $\boldsymbol{n}\geq\boldsymbol{n}_{\boldsymbol{m}}$,

$$\begin{gathered}A_{\boldsymbol{n}}=LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)+R_{\boldsymbol{n},\boldsymbol{m}}+N_{\boldsymbol{n},\boldsymbol{m}},\\ \operatorname{rank}(R_{\boldsymbol{n},\boldsymbol{m}})\leq c(\boldsymbol{m})N(\boldsymbol{n}),\qquad \|N_{\boldsymbol{n},\boldsymbol{m}}\|\leq\omega(\boldsymbol{m}),\end{gathered}\tag{4.18}$$

where the quantities $\boldsymbol{n}_{\boldsymbol{m}}$, $c(\boldsymbol{m})$, $\omega(\boldsymbol{m})$ 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 sequence $\{A_{\boldsymbol{n}}\}_n$, a is the *weight function* and f is the *generating function*.[3]

[3]We refer the reder to the introduction of Tilli's paper [54] for the origin and the meaning of this terminology.

Remark 4.2. An equivalent definition of LT sequence is obtained by replacing, in Definition 4.3, 'for all $\boldsymbol{m} \in \mathbb{N}^d$' with 'for all sufficiently large $\boldsymbol{m} \in \mathbb{N}^d$' (i.e., 'for every $\boldsymbol{m}$ greater than or equal to some $\hat{\boldsymbol{m}} \in \mathbb{N}^d$'). Indeed, suppose that the splitting (4.18) 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 $\boldsymbol{n}_{\boldsymbol{m}} = \boldsymbol{1},\ c(\boldsymbol{m}) = 1,\ \omega(\boldsymbol{m}) = 0$ and $R_{\boldsymbol{n},\boldsymbol{m}} = A_{\boldsymbol{n},\boldsymbol{m}} - B_{\boldsymbol{n},\boldsymbol{m}},\ N_{\boldsymbol{n},\boldsymbol{m}} = O$ for all the other values of $\boldsymbol{m}$, we see that they actually hold for every $\boldsymbol{m} \in \mathbb{N}^d$.

Remark 4.3. If $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ in the sense of Definition 4.3, then $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ (in the sense of the classical Definition 3.1) for all sequences $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m$ such that $\boldsymbol{m} \to \infty$ when $m \to \infty$.

Remark 4.4. Suppose that $\{\{B_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ (in the sense of Definition 4.3) and $\{\{B'_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A'_{\boldsymbol{n}}\}_n$ (in the sense of Definition 4.3). Then:

1. $\{\{B^*_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A^*_{\boldsymbol{n}}\}_n$ (in the sense of Definition 4.3);
2. $\{\{\alpha B_{\boldsymbol{n},\boldsymbol{m}} + \beta B'_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{\alpha A_{\boldsymbol{n}} + \beta A'_{\boldsymbol{n}}\}_n$ (in the sense of Definition 4.3) for all $\alpha, \beta \in \mathbb{C}$;
3. if $\{A_{\boldsymbol{n}}\}_n$ and $\{A'_{\boldsymbol{n}}\}_n$ are s.u., then $\{\{B_{\boldsymbol{n},\boldsymbol{m}} B'_{\boldsymbol{n},\boldsymbol{m}}\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A_{\boldsymbol{n}} A'_{\boldsymbol{n}}\}_n$ (in the sense of Definition 4.3).

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 3.2 and Propositions 3.1–3.4.

Definition 4.4 (sLT sequence). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, with $\boldsymbol{n} \in \mathbb{N}^d$. We say that $\{A_{\boldsymbol{n}}\}_n$ is a separable Locally Toeplitz (sLT) sequence if there exist

- a Riemann-integrable function $a : [0, 1]^d \to \mathbb{C}$,
- 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 f. From now on, if a matrix-sequence $\{A_{\boldsymbol{n}}\}_n$ is given (with $\boldsymbol{n} \in \mathbb{N}^d$) and 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 4.5. Suppose that $\{A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} a \otimes f_i,\ i = 1, \ldots, r$. Then $\{\sum_{i=1}^r A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} a \otimes (\sum_{i=1}^r f_i)$. Suppose that $\{A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} a_i \otimes f,\ i = 1, \ldots, r$. 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: sequences distributed in the singular value sense as the zero function, 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.

Theorem 4.4. *Assume that a sequence $\{A_{\boldsymbol{n}}\}_n$ is given, $A_{\boldsymbol{n}}$ of size d_n. Then the following are equivalent*

- $\{A_{\boldsymbol{n}}\}_n \sim_\sigma 0$*;*
- $\{\{O_{\boldsymbol{n}}\}_n\}_{\boldsymbol{m}}$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$*.*

Under the additional assumption that $d_n = N(\boldsymbol{n})$, the previous statements are equivalent to the following

- $\{A_{\boldsymbol{n}}\}_n \sim_{\rm sLT} a \otimes f$ *with* $a \otimes f \equiv 0$*;*
- $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$ *with* $a \otimes f \equiv 0$*;*

Proof. Let $m > 0$ and let $\phi(m, n)$ the cardinality of the singular values of $A_{\boldsymbol{n}}$ which are bounded by $1/m$. Since $\{A_{\boldsymbol{n}}\}_n \sim_\sigma 0$ there exist a function $c(m)$ such that $\lim_{m\to 0} c(m) = 0$ and $\phi(n, m) \le c(m)d_n$ where d_n the size of $A_{\boldsymbol{n}}$. Now take the singular value decomposition [30] of $A_{\boldsymbol{n}}$. We have

$$A_{\boldsymbol{n}} = U_{\boldsymbol{n}}\Sigma_{\boldsymbol{n}}V_{\boldsymbol{n}}$$

with $U_{\boldsymbol{n}}, V_{\boldsymbol{n}}$ unitary and $\Sigma_{\boldsymbol{n}}$ diagonal containing the singular values. We split $\Sigma_{\boldsymbol{n}}$ as $\Sigma_{\boldsymbol{n},m,<1/m} + \Sigma_{\boldsymbol{n},m,\ge 1/m}$, where the diagonal matrix $\Sigma_{\boldsymbol{n},m,<1/m}$ contains the singular values less than $1/m$ in the same position as $\Sigma_{\boldsymbol{n}}$ and zero otherwise. Because $U_{\boldsymbol{n}}, V_{\boldsymbol{n}}$ are unitary we deduce

$$A_{\boldsymbol{n}} = N_{\boldsymbol{n},m} + R_{\boldsymbol{n},m}, \;\; N_{\boldsymbol{n},m} = U_{\boldsymbol{n}}\Sigma_{\boldsymbol{n},m,<1/m}V_{\boldsymbol{n}}, \; R_{\boldsymbol{n},m} = U_{\boldsymbol{n}}\Sigma_{\boldsymbol{n},m,\ge 1/m}V_{\boldsymbol{n}}$$

with

$$\|N_{\boldsymbol{n},m}\| < 1/m, \;\; \text{rank}(R_{\boldsymbol{n},m}) \le c(m)d_n.$$

By definition of a.c.s., the latter shows that $\{\{O_{\boldsymbol{n}}\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Now consider $d_n = N(\boldsymbol{n})$. By observing that $a \otimes f \equiv 0$ is equivalent to the fact that $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a, f)$ is the null matrix, the equivalence of latter two statements to the previous is plain. □

Remark 4.6. Taking into consideration the standard notion of clustering [56], the statements in Lemma 4.4 are also equivalent to write that the sequence $\{A_{\boldsymbol{n}}\}_n$ is clustered at zero.

For proving Theorems 4.5–4.6, we will need the following technical lemma, whose proof may be skipped in a first reading.

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 $m(\boldsymbol{h}) \to \infty$ and $x(m(\boldsymbol{h}), \boldsymbol{h}) \to \xi$ when $\boldsymbol{h} \to \infty$.

Proof. Let $\hat{\mathbb{N}} = \{m_1, m_2, \ldots\}$; we denote by m_+ (m_-) the successor (predecessor) of 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 $m \ge m_r$. We construct an injective function $\boldsymbol{h}(\cdot) : \{m_r, m_{r+1}, \ldots\} \to \mathbb{N}^d$ as follows: we set $\boldsymbol{h}(m_r) = \boldsymbol{1}$ and, for $m \in \hat{\mathbb{N}} \backslash \{m_r\}$, we choose $\boldsymbol{h}(m) > \boldsymbol{h}(m_-)$ such that, for $\boldsymbol{h} \ge \boldsymbol{h}(m)$,

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

Hence, we have a sequence

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

and, if $m \ge m_r$ and $\boldsymbol{h} \ge \boldsymbol{h}(m)$,

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

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

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

Then, when $\boldsymbol{h}\to\infty$, we have $m(\boldsymbol{h})\to\infty$. Moreover, noting that $m(\boldsymbol{h})\ge m_r$ and $\boldsymbol{h}\ge\boldsymbol{h}(m(\boldsymbol{h}))$ for all $\boldsymbol{h}\in\mathbb{N}^d$, by (4.19) we also have $|x(m(\boldsymbol{h}),\boldsymbol{h})-\xi|\le|x(m(\boldsymbol{h}))-\xi|+1/m(\boldsymbol{h})\to 0$. □

Theorem 4.5. *Let $a:[0,1]^d\to\mathbb{C}$ be Riemann-integrable and consider the sequence of matrices $\{D_{\boldsymbol{n}}(a)\}_n$, where $\boldsymbol{n}\in\mathbb{N}^d$ and, 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

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

where ω_a is the modulus of continuity of a:

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

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 take, in Definition 4.3, $\boldsymbol{n}_{\boldsymbol{m}}=\boldsymbol{m}^2,\ c(\boldsymbol{m})=\sum_{i=1}^d(1/m_i)$ and $\omega(\boldsymbol{m})=\sum_{i=1}^d\omega_a(2/m_i)$.

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

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

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

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

because

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

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

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

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

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

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

For the case $d > 1$, $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,1)$ is a $N(\boldsymbol{n}) \times N(\boldsymbol{n})$ 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

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

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, as in (4.21), one can show 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

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

see (4.22) and (2.2). Hence,

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

where

$$R_{\boldsymbol{n},\boldsymbol{m}} = \mathop{\mathrm{diag}}_{j_1=1,\ldots,m_1} \left[\mathop{\mathrm{diag}}_{i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor} R_{n_2,\ldots,n_d,m_2,\ldots,m_d}^{[j_1/m_1]}\right] \oplus \mathop{\mathrm{diag}}_{i_1=m_1\lfloor n_1/m_1\rfloor+1,\ldots,n_1} D_{n_2,\ldots,n_d}\Big(a\Big(\frac{i_1}{n_1},\cdot\Big)\Big),$$

$$N_{\boldsymbol{n},\boldsymbol{m}} = \mathop{\mathrm{diag}}_{j_1=1,\ldots,m_1} \left[\mathop{\mathrm{diag}}_{i_1=(j_1-1)\lfloor n_1/m_1\rfloor+1,\ldots,j_1\lfloor n_1/m_1\rfloor} N_{\boldsymbol{n},\boldsymbol{m}}^{[j_1/m_1,\, i_1/n_1]}\right] \oplus O_{(n_1 \bmod m_1)n_2\cdots n_d}.$$

By (4.23) and (2.11), (2.13), we have

$$\begin{aligned}
\mathrm{rank}(R_{\boldsymbol{n},\boldsymbol{m}}) &\le m_1 \left\lfloor \frac{n_1}{m_1}\right\rfloor n_2\cdots n_d \sum_{k=2}^{d} \frac{m_k}{n_k} + (n_1 \bmod m_1) n_2\cdots n_d\\
&\le n_1 n_2\cdots n_d \sum_{k=2}^{d} \frac{m_k}{n_k} + m_1 n_2\cdots n_d = N(\boldsymbol{n})\sum_{k=1}^{d}\frac{m_k}{n_k},\\
\|N_{\boldsymbol{n},\boldsymbol{m}}\| &\le \sum_{k=1}^{d} \omega_a\Big(\frac{1}{m_k} + \frac{m_k}{n_k}\Big),
\end{aligned}$$

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 [37]. By the first part of the proof, $\{D_{\boldsymbol{n}}(a_m)\}_n \sim_{\mathrm{sLT}} a_m \otimes 1$. Hence, for each m and each $\boldsymbol{h} \in \mathbb{N}^d$ there is $\boldsymbol{n}_{m,\boldsymbol{h}}$ such that, for $\boldsymbol{n} \ge \boldsymbol{n}_{m,\boldsymbol{h}}$,

$$D_{\boldsymbol{n}}(a_m) = LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m,1) + R_{\boldsymbol{n},m,\boldsymbol{h}} + N_{\boldsymbol{n},m,\boldsymbol{h}},$$

$$\mathrm{rank}(R_{\boldsymbol{n},m,\boldsymbol{h}}) \le c(m,\boldsymbol{h})N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},m,\boldsymbol{h}}\| \le \omega(m,\boldsymbol{h}),$$

where

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

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

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

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 $\boldsymbol{n}_m$ such that, for $\boldsymbol{n}\geq\boldsymbol{n}_m$,

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

where

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

It follows that, for every m, every $\boldsymbol{h}\in\mathbb{N}^d$ and every $\boldsymbol{n}\geq\max(\boldsymbol{n}_m,\boldsymbol{n}_{m,\boldsymbol{h}})$,

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

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

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

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

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

The application of Lemma 3.2 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 matrices $\{T_{\boldsymbol{n}}(f)\}_n$, where $\boldsymbol{n}\in\mathbb{N}^d$ and, of course, $\boldsymbol{n}\to\infty$ as $n\to\infty$. Then $\{T_{\boldsymbol{n}}(f)\}_n\sim_{\mathrm{LT}} 1\otimes f$.*

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 in its full generality, when f is only assumed to be 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

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

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 take, in Definition 4.3, $\boldsymbol{n_m} = \boldsymbol{m}^2,\ c(\boldsymbol{m}) = \sum_{i=1}^{d}(2r_i + 1)/m_i$ and $\omega(\boldsymbol{m}) = 0$.

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

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

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

For the case $d > 1$, by induction hypothesis we have

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

From the definition of $LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1, f)$ and the properties of tensor products and direct sums (see Section 2.3.1), we obtain

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

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

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

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 $(2r_1 + 1)m_1$. Thus, recalling Lemma 2.6, we arrive at

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

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

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

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 $f = \sum_{\boldsymbol{j}=-\boldsymbol{r}}^{\boldsymbol{r}} f_{\boldsymbol{j}} \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}$ 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, $\{T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})\}_n \sim_{LT} 1 \otimes \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}}$, hence $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1, \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})\}_n$ in the sense of Definition 4.3. It follows that $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(1, f)\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{T_{\boldsymbol{n}}(f)\}_n$ in the sense of Definition 4.3; see Remark 4.4. 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 $\boldsymbol{n}_{m,\boldsymbol{h}}$ such that, for $\boldsymbol{n} \geq \boldsymbol{n}_{m,\boldsymbol{h}}$,

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

where

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

Moreover, by Theorem 2.6,

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

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

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

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 $\boldsymbol{n} \geq \max(\boldsymbol{n}_m, \boldsymbol{n}_{m,\boldsymbol{h}})$,

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

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 $x(m,\boldsymbol{h}) = c(m,\boldsymbol{h}) + \omega(m,\boldsymbol{h})$. Then, for every $\boldsymbol{h} \in \mathbb{N}^d$ and every $\boldsymbol{n} \geq \max(\boldsymbol{n}_{m(\boldsymbol{h})}, \boldsymbol{n}_{m(\boldsymbol{h}),\boldsymbol{h}})$,

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

The application of Lemma 3.2 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 of LT and sLT sequences

We begin with a basic spectral result for sLT sequences. It is stated as a lemma, both because it is needed for the proof of Theorem 4.7 and because it will be generalized afterwards, in the more general context of GLT sequences (so, we do not need to remember it). To simplify the presentation, from now on, until the end of this section, it is understood that multi-indices are d-indices (i.e., they have length d).

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 $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m$ 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, for all $F \in C_c(\mathbb{R})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{r=1}^{N(\boldsymbol{n})} F\left(\sigma_r(LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f))\right) = \phi_{\boldsymbol{m}}(F),$$

where

$$\lim_{m\to\infty} \phi_{\boldsymbol{m}}(F) = \phi(F) = \frac{1}{(2\pi)^d} \int_{[0,1]^d\times[-\pi,\pi]^d} F(a(\mathbf{x})f(\boldsymbol{\theta}))d\mathbf{x}d\boldsymbol{\theta}.$$

Therefore, by Theorem 3.3, $\{A_{\boldsymbol{n}}\}_n \sim_\sigma a \otimes f$. □

As a consequence of Lemma 4.2 and Proposition 3.3, every LT sequence is s.u. in the sense of 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. By Lemma 4.2 and Proposition 3.3, every LT sequence is s.u., so in particular $\{A_{\boldsymbol{n}}\}_n$ and $\{\tilde{A}_{\boldsymbol{n}}\}_n$ are s.u. Since $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$ (in the sense of Definition 4.3) and $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde{a},\tilde{f})\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{\tilde{A}_{\boldsymbol{n}}\}_n$ (in the sense of Definition 4.3), the product $\{\{LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a,f)LT_{\boldsymbol{n}}^{\boldsymbol{m}}(\tilde{a},\tilde{f})\}_n\}_{\boldsymbol{m}}$ is an a.c.s. for $\{A_{\boldsymbol{n}}\tilde{A}_{\boldsymbol{n}}\}_n$ (in the sense of Definition 4.3); see Remark 4.4. The thesis now follows from Definition 4.3 and 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 $\boldsymbol{n} \in \mathbb{N}^d$ and, 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 $A_{\boldsymbol{n}} = D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)$. Theorems 4.9–4.10 show that the sequences of the form $\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n$ play a special 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. More precisely, any LT sequence $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$ admits $\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n$ 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. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a\otimes f$. Then, for any $\{a_m\}$, $\{f_m\}$, $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ 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 $\{A_{\boldsymbol{n}}\}_n$. In particular, for any $\{a_m\}$, $\{f_m\}$ with the above properties, $\{\{D_{\boldsymbol{n}}(a_m)T_{\boldsymbol{n}}(f_m)\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Taking $a_m = a$ and $f_m = f$, we see that $\{\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.

Proof. Let $\{a_m\}$, $\{f_m\}$, $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ as in the statement of the theorem. Then, for each m and each $\boldsymbol{h}\in\mathbb{N}^d$ there is $\boldsymbol{n}_{m,\boldsymbol{h}}$ such that, for $\boldsymbol{n}\ge\boldsymbol{n}_{m,\boldsymbol{h}}$,

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

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 $\boldsymbol{n}_{\boldsymbol{h}}$ such that, for $\boldsymbol{n}\ge\boldsymbol{n}_{\boldsymbol{h}}$,

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

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 $\boldsymbol{n}\ge\max(\boldsymbol{n}_{m,\boldsymbol{h}},\boldsymbol{n}_{\boldsymbol{h}})$,

$$A_{\boldsymbol{n}} = A_{\boldsymbol{n}}^{(m)} + \big[LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m,f_m)\big] + (R_{\boldsymbol{n},\boldsymbol{h}} - R_{\boldsymbol{n},m,\boldsymbol{h}}) + (N_{\boldsymbol{n},\boldsymbol{h}} - N_{\boldsymbol{n},m,\boldsymbol{h}}), \tag{4.26}$$
$$\mathrm{rank}(R_{\boldsymbol{n},\boldsymbol{h}} - R_{\boldsymbol{n},m,\boldsymbol{h}}) \le (c(\boldsymbol{h}) + c(m,\boldsymbol{h}))N(\boldsymbol{n}), \qquad \|N_{\boldsymbol{n},\boldsymbol{h}} - N_{\boldsymbol{n},m,\boldsymbol{h}}\| \le \omega(\boldsymbol{h}) + \omega(m,\boldsymbol{h}). \tag{4.27}$$

Thanks to Propositions 4.1–4.2 and to Theorem 2.6, we have

$$\begin{aligned}
&\|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a_m,f_m)\|_1 \le \|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a,f-f_m)\|_1 + \|LT_{\boldsymbol{n}}^{\boldsymbol{h}}(a-a_m,f_m)\|_1\\
&= \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}}\Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big|\,\|T_{\lfloor\boldsymbol{n}/\boldsymbol{h}\rfloor}(f-f_m)\|_1 + \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}}\Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big|\,\|T_{\lfloor\boldsymbol{n}/\boldsymbol{h}\rfloor}(f_m)\|_1\\
&\le N(\boldsymbol{n})\|a\|_\infty\|f-f_m\|_{L^1} + \|f_m\|_{L^1}\frac{N(\boldsymbol{n})}{N(\boldsymbol{h})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}}\Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big|\\
&\le \left[\|a\|_\infty\|f-f_m\|_{L^1} + \sup_k\|f_k\|_{L^1}\frac{1}{N(\boldsymbol{h})}\sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}}\Big|a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big|\right]N(\boldsymbol{n});
\end{aligned} \tag{4.28}$$

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

$$\varepsilon(m,\boldsymbol{h}) = \|a\|_\infty \|f-f_m\|_{L^1} + \sup_k \|f_k\|_{L^1} \frac{1}{N(\boldsymbol{h})} \sum_{\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{h}} \Big| a\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big) - a_m\Big(\frac{\boldsymbol{j}}{\boldsymbol{h}}\Big)\Big| \tag{4.29}$$

satisfies

$$\lim_{m\to\infty} \lim_{\boldsymbol{h}\to\infty} \varepsilon(m,\boldsymbol{h}) = \lim_{m\to\infty} \left(\|a\|_\infty \|f-f_m\|_{L^1} + \sup_k \|f_k\|_{L^1} \int_{[0,1]^d} |a(\mathbf{x}) - a_m(\mathbf{x})| d\mathbf{x} \right) = 0. \tag{4.30}$$

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)–(4.28), for every m and every $\boldsymbol{n} \geq \max(\boldsymbol{n}_{m,\boldsymbol{h}(m)}, \boldsymbol{n}_{\boldsymbol{h}(m)})$,

$$A_{\boldsymbol{n}} = A_{\boldsymbol{n}}^{(m)} + \left[LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a,f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a_m,f_m)\right] + (R_{\boldsymbol{n},\boldsymbol{h}(m)} - R_{\boldsymbol{n},m,\boldsymbol{h}(m)}) + (N_{\boldsymbol{n},\boldsymbol{h}(m)} - N_{\boldsymbol{n},m,\boldsymbol{h}(m)}),$$

$$\mathrm{rank}(R_{\boldsymbol{n},\boldsymbol{h}(m)} - R_{\boldsymbol{n},m,\boldsymbol{h}(m)}) \leq [c(\boldsymbol{h}(m)) + c(m,\boldsymbol{h}(m))]\, N(\boldsymbol{n}),$$

$$\|N_{\boldsymbol{n},\boldsymbol{h}(m)} - N_{\boldsymbol{n},m,\boldsymbol{h}(m)}\| \leq \omega(\boldsymbol{h}(m)) + \omega(m,\boldsymbol{h}(m)),$$

$$\left\|LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a,f) - LT_{\boldsymbol{n}}^{\boldsymbol{h}(m)}(a_m,f_m)\right\|_1 \leq \varepsilon(m,\boldsymbol{h}(m))N(\boldsymbol{n}).$$

Using Lemma 3.2, 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. □

Theorem 4.10. *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 are equivalent.*

1. $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{LT}} a \otimes f$.
2. *There exist sequences $\{a_m\}$, $\{f_m\}$ such that:*
   - $a_m : [0,1]^d \to \mathbb{C}$ *is continuous,* $\|a_m\|_\infty \leq \|a\|_{L^\infty}$ *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* $\{A_{\boldsymbol{n}}\}_n$.
3. *There exist* $\{a_m\}$, $\{f_m\}$, $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ *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* $L^1([-\pi,\pi]^d)$*;*
   - $\{A_{\boldsymbol{n}}^{(m)}\}_n \sim_{\mathrm{LT}} a_m \otimes f_m$ *and* $\{\{A_{\boldsymbol{n}}^{(m)}\}_n\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$.
4. $\{\{D_{\boldsymbol{n}}(a)T_{\boldsymbol{n}}(f)\}_n\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$.

Proof. ($1 \Rightarrow 2$) Since any Riemann-integrable function is bounded, we have $a \in L^\infty([0,1]^d)$. Hence, by the Lusin theorem [37], there exists a sequence of continuous functions $\hat{a}_m : [0,1]^d \to \mathbb{C}$ such that $\|\hat{a}_m\|_\infty \le \|a\|_{L^\infty}$ for all m and $\hat{a}_m \to a$ in measure. This implies that $\hat{a}_m \to a$ also in $L^1([0,1]^d)$, because of 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 2; note also that $a_m \to a$ in $L^1([0,1]^d)$ by the dominated convergence theorem.

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 2.

The application of Theorem 4.9 concludes the proof.

($2 \Rightarrow 3$) Obvious; we just recall that, under the assumptions in item 2, $a_m \to a$ in $L^1([0,1]^d)$ by the dominated convergence theorem; moreover, $\{\{D_{\boldsymbol{n}}(a_m)T_{\boldsymbol{n}}(f_m)\}_n\}_m \sim_{\rm LT} a_m \otimes f_m$ by Theorem 4.8.

($3 \Rightarrow 1$) By assumption, for each m and each $\boldsymbol{h} \in \mathbb{N}^d$ there is $\boldsymbol{n}_{m,\boldsymbol{h}}$ such that, for $\boldsymbol{n} \ge \boldsymbol{n}_{m,\boldsymbol{h}}$,

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

where

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

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

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

where

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

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

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

where in the last inequalities we used (4.28); the quantity $\varepsilon(m,\boldsymbol{h})$ is defined in (4.29) and satisfies (4.30). 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 $x(m,\boldsymbol{h}) = \varepsilon(m,\boldsymbol{h}) + c(m,\boldsymbol{h}) + \omega(m,\boldsymbol{h})$. Then, for every $\boldsymbol{h} \in \mathbb{N}^d$ and every $\boldsymbol{n} \ge \max(\boldsymbol{n}_{m(\boldsymbol{h})}, \boldsymbol{n}_{m(\boldsymbol{h}),\boldsymbol{h}})$,

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

The application of Lemma 3.2 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 $3 \Rightarrow 1$.

To conclude the proof of the theorem, we note that $1 \Rightarrow 4$ (by Theorem 4.9) and $4 \Rightarrow 3$ (obviously). □

Remark 4.7. Theorem 4.10 continues to hold if f is assumed to be separable and we add in item 2 the requirement that each f_m is separable. The proof is left as an exercise for the reader. Note that, if f is separable, we can replace item 1 with '$\{A_{\boldsymbol{n}}\}_n \sim_{\rm sLT} a \otimes f$'.

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 [45, 46] in terms of sLT sequences, can be equivalently formulated in terms of LT sequences.

Proposition 4.4. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\rm LT} a \otimes f$. Then, for any $m \in \mathbb{N}$ there exist matrix-sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm sLT} a \otimes f_{i,m}$, $i = 1, \ldots, N_m$, 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 $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm sLT} a \otimes f_{i,m}$, $i = 1, \ldots, N_m$. For example, in view of Theorem 4.8, one can choose $A_{\boldsymbol{n}}^{(i,m)} = D_{\boldsymbol{n}}(a) T_{\boldsymbol{n}}(f_{i,m})$. By Remark 4.5, $\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm LT} a \otimes (\sum_{i=1}^{N_m} f_{i,m}) = a \otimes f_m$. Hence, by Theorem 4.9, $\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. □

5 GLT sequences

In this section we develop the GLT theory. In summary we prove all the statements contained in items **GLT1-GLT5** and, in Theorem 5.6, we prove that f(Hermitian GLT) =Hermitian GLT, under mild assumptions of f.

5.1 Definition and characterizations

We first report a 'corrected' version of the original definition of GLT sequences; cf. [45, Definition 2.3] and [46, Definition 1.5]. Then, we will show that such a definition admits some useful equivalent characterizations.

Definition 5.1 (GLT sequence). Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence, with $\boldsymbol{n} \in \mathbb{N}^d$, 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 GLT sequence with symbol κ, and we write $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$, if:

- for any $\epsilon > 0$ there exist matrix-sequences $\{A_{\boldsymbol{n}}^{(i,\epsilon)}\}_n \sim_{\rm sLT} a_{i,\epsilon} \otimes f_{i,\epsilon}$, $i = 1, \ldots, N_\epsilon$;
- $\sum_{i=1}^{N_\epsilon} a_{i,\epsilon} \otimes f_{i,\epsilon} \to \kappa$ in measure over $[0,1]^d \times [-\pi,\pi]^d$ when $\epsilon \to 0$;
- $\big\{\{\sum_{i=1}^{N_\epsilon} A_{\boldsymbol{n}}^{(i,\epsilon)}\}_n\big\}_m$, with $\epsilon = (m+1)^{-1}$, is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.

From now on, until the end of Section 5, it will be always understood that multi-indices are actually d-indices (as in Section 4.2.2). Moreover, if a matrix-sequence $\{A_{\boldsymbol{n}}\}_n$ is given and if we write $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$, it is implicitly assumed 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. More precisely, if $\{A_{\boldsymbol{n}}\}_n \sim_{\rm sLT} a \otimes f$ then $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} a \otimes f$. 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:*

- *there exist matrix-sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\mathrm{LT}} a_{i,m} \otimes f_{i,m},\ i = 1, \ldots, N_m$;*
- *$\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ 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 three conditions 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 $a_{i,\epsilon},\ f_{i,\epsilon},\ \{A_{\boldsymbol{n}}^{(i,\epsilon)}\}_n,\ \epsilon(m) = (m+1)^{-1}$ are as in Definition 5.1. Conversely, suppose that the three conditions of the proposition hold. Then, the three conditions 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 $\{m(\epsilon)\}_{\epsilon>0}$ 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, $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$. □

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 from Proposition 5.1 and Proposition 4.4. □

The next proposition shows that the functions $a_{i,m}$ in Proposition 5.1 may be supposed to be continuous, the functions $f_{i,m}$ may be supposed to be separable trigonometric polynomials, and the matrix-sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n$ may be choosen as $A_{\boldsymbol{n}}^{(i,m)} = D_{\boldsymbol{n}}(a_{i,m})T_{\boldsymbol{n}}(f_{i,m})$.

Proposition 5.2. *We have $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$ if and only if:*

- *there exist functions $a_{i,m},\ f_{i,m},\ i = 1, \ldots, N_m$, where each $a_{i,m} : [0,1]^d \to \mathbb{C}$ is continuous and each $f_{i,m} : [-\pi,\pi]^d \to \mathbb{C}$ is a separable trigonometric polynomial;*
- *$\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ in measure over $[0,1]^d \times [-\pi,\pi]^d$ when $m \to \infty$;*
- *$\big\{\{\sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m})T_{\boldsymbol{n}}(f_{i,m})\}_n\big\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$.*

Proof. If the three conditions of the proposition are met, then also the three conditions of Proposition 5.1 are met, because $\{D_{\boldsymbol{n}}(a_{i,m})T_{\boldsymbol{n}}(f_{i,m})\}_n \sim_{\mathrm{sLT}} a_{i,m} \otimes f_{i,m}$ (Theorem 4.8). Hence, $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$.

Conversely, suppose that $\{A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \kappa$. We show that the three conditions of Proposition 5.2 are met. From Proposition 5.1 we know that:

- there exist $\{\{A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m,\ i = 1, \ldots, N_m$, such that $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\mathrm{sLT}} \hat{a}_{i,m} \otimes \hat{f}_{i,m}$;
- $\sum_{i=1}^{N_m} \hat{a}_{i,m} \otimes \hat{f}_{i,m} \to \kappa$ 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 $\{A_{\boldsymbol{n}}\}_n$.

Let:

* $a_{i,m}^{(k)} : [0,1]^d \to \mathbb{C}$ a continuous function such that $a_{i,m}^{(k)} \to \hat{a}_{i,m}$ in $L^1([0,1]^d)$ and a.e. when $k \to \infty$;
* $f_{i,m}^{(k)} : [-\pi,\pi]^d \to \mathbb{C}$ a separable trigonometric polynomial such that $f_{i,m}^{(k)} \to \hat{f}_{i,m}$ in $L^1([-\pi,\pi]^d)$ and a.e. when $k \to \infty$. To construct $f_{i,m}^{(k)}$, take into account that $\hat{f}_{i,m}$ is separable, say $\hat{f}_{i,m} = \hat{f}_{i,m,1} \otimes \cdots \otimes \hat{f}_{i,m,d}$, $\hat{f}_{i,m,j} \in L^1([-\pi,\pi]^d)$ for all $j = 1,\ldots,d$; choose $f_{i,m}^{(k)} = f_{i,m,1}^{(k)} \otimes \cdots \otimes f_{i,m,d}^{(k)}$, where each $f_{i,m,j}^{(k)}$ is a univariate trigonometric polynomial that converges to $\hat{f}_{i,m,j}$ in $L^1([-\pi,\pi])$ and a.e. when $k \to \infty$.

Since $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm sLT} \hat{a}_{i,m} \otimes \hat{f}_{i,m}$ and $\{\{D_{\boldsymbol{n}}(a_{i,m}^{(k)})T_{\boldsymbol{n}}(f_{i,m}^{(k)})\}_n\}_k \sim_{\rm sLT} a_{i,m}^{(k)} \otimes f_{i,m}^{(k)}$ by Theorem 4.8, it follows from Theorem 4.9 that $\{\{D_{\boldsymbol{n}}(a_{i,m}^{(k)})T_{\boldsymbol{n}}(f_{i,m}^{(k)})\}_n\}_k$ is an a.c.s. for $\{A_{\boldsymbol{n}}^{(i,m)}\}_n$ for each fixed i and m. Hence, for all k there exists $n_{i,m}^{(k)}$ such that, for $n \ge n_{i,m}^{(k)}$,

$$A_{\boldsymbol{n}}^{(i,m)} = D_{\boldsymbol{n}}(a_{i,m}^{(k)})T_{\boldsymbol{n}}(f_{i,m}^{(k)}) + R_{n,k}^{(i,m)} + N_{n,k}^{(i,m)}, \quad \text{rank}(R_{n,k}^{(i,m)}) \le c(i,m,k)N(\boldsymbol{n}), \quad \|N_{n,k}^{(i,m)}\| \le \omega(i,m,k), \tag{5.2}$$

where $\lim_{k\to\infty} c(i,m,k) = \lim_{k\to\infty} \omega(i,m,k) = 0$. For every $\delta > 0$,

$$\mu(m,k,\delta) = \mu_{2d}\left\{\left|\sum_{i=1}^{N_m} a_{i,m}^{(k)} \otimes f_{i,m}^{(k)} - \kappa\right| \ge \delta\right\} \to \mu(m,\delta) = \mu_{2d}\left\{\left|\sum_{i=1}^{N_m} \hat{a}_{i,m} \otimes \hat{f}_{i,m} - \kappa\right| \ge \delta\right\} \qquad \text{as } k \to \infty;$$

this can be seen by writing the measures of sets as integrals of indicator functions and by applying the dominated convergence theorem, taking into account that $\sum_{i=1}^{N_m} a_{i,m}^{(k)} \otimes f_{i,m}^{(k)} \to \sum_{i=1}^{N_m} \hat{a}_{i,m} \otimes \hat{f}_{i,m}$ a.e. when $k \to \infty$. Since $\mu(m,\delta) \to 0$ as $m \to \infty$, for every $\delta > 0$, we can choose $\delta_m \searrow 0$ such that $\mu(m,\delta_m) \to 0$ as $m \to \infty$. Thus, $\mu(m,k,\delta_m) \to \mu(m,\delta_m)$ as $k \to \infty$ and $\mu(m,\delta_m) \to 0$ as $m \to \infty$. In view of the limit relations $\lim_{k\to\infty} c(i,m,k) = \lim_{k\to\infty} \omega(i,m,k) = 0$, $\lim_{k\to\infty} \mu(m,k,\delta_m) = \mu(m,\delta_m)$, it is possible to choose $k_m \nearrow \infty$ such that the following conditions are satisfied for every m:

$$\max_{i=1,\ldots,N_m} c(i,m,k_m) \le \frac{1}{mN_m}, \qquad \max_{i=1,\ldots,N_m} \omega(i,m,k_m) \le \frac{1}{mN_m}, \qquad \mu(m,k_m,\delta_m) \le \frac{1}{m} + \mu(m,\delta_m).$$

Now, $\{\{\sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$: for all m there exists n_m such that, for $n \ge n_m$,

$$A_{\boldsymbol{n}} = \sum_{i=1}^{N_m} A_{\boldsymbol{n}}^{(i,m)} + R_{n,m} + N_{n,m}, \quad \text{rank}(R_{n,m}) \le c(m)N(\boldsymbol{n}), \quad \|N_{n,m}\| \le \omega(m), \tag{5.3}$$

and $\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0$. Combining (5.2)–(5.3), for all m and $n \ge \max(n_m, \max_{i=1,\ldots,N_m} n_{i,m}^{(k_m)})$ we have

$$A_{\boldsymbol{n}} = \sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m}^{(k_m)})T_{\boldsymbol{n}}(f_{i,m}^{(k_m)}) + R_{n,m} + \sum_{i=1}^{N_m} R_{n,k_m}^{(i,m)} + N_{n,m} + \sum_{i=1}^{N_m} N_{n,k_m}^{(i,m)},$$

$$\text{rank}\Big(R_{n,m} + \sum_{i=1}^{N_m} R_{n,k_m}^{(i,m)}\Big) \le \Big(c(m) + \frac{1}{m}\Big)N(\boldsymbol{n}), \qquad \Big\|N_{n,m} + \sum_{i=1}^{N_m} N_{n,k_m}^{(i,m)}\Big\| \le \omega(m) + \frac{1}{m}.$$

Therefore, $\{\{\sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m}^{(k_m)})T_{\boldsymbol{n}}(f_{i,m}^{(k_m)})\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. To finish the proof, we note that

$$\sum_{i=1}^{N_m} a_{i,m}^{(k_m)} \otimes f_{i,m}^{(k_m)} \to \kappa$$

in measure when $m \to \infty$; this is true because, for any $\delta > 0$, we have

$$\mu_{2d}\left\{\left|\sum_{i=1}^{N_m} a_{i,m}^{(k_m)} \otimes f_{i,m}^{(k_m)} - \kappa\right| \ge \delta\right\} = \mu(m,k_m,\delta),$$

which tends to 0, because it is eventually less than $\mu(m,k_m,\delta_m) \to 0$. □

5.2 Approximation results for GLT sequences

The following result is analogous to Corollaries 3.1–3.2, where '$\sim_\lambda$' and '$\sim_\sigma$' are replaced by '$\sim_{\rm GLT}$'.

Theorem 5.1. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. Suppose that:*

- $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ *is an a.c.s. for* $\{A_{\boldsymbol{n}}\}_n$*;*
- $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$ *for every* m*;*
- $\kappa_m \to \kappa$ *in measure over* $[0,1]^d \times [-\pi,\pi]^d$ *when* $m \to \infty$*, being* κ *some measurable function.*

Then $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$*.*

Proof. For every m we have $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$, hence:

- for every k there exist matrix-sequences $\{A^{(i,k)}_{\boldsymbol{n},m}\}_n \sim_{\rm sLT} a_{i,k,m} \otimes f_{i,k,m},\ i = 1,\ldots,N_{k,m}$;
- $\sum_{i=1}^{N_{k,m}} a_{i,k,m} \otimes f_{i,k,m} \to \kappa_m$ in measure over $[0,1]^d \times [-\pi,\pi]^d$ when $k \to \infty$;
- $\big\{\{\sum_{i=1}^{N_{k,m}} A^{(i,k)}_{\boldsymbol{n},m}\}_n\big\}_k$ is an a.c.s. for $\{B_{\boldsymbol{n},m}\}_n$: for every k there exists $n_{k,m}$ such that, for $n \ge n_{k,m}$,

$$B_{\boldsymbol{n},m} = \sum_{i=1}^{N_{k,m}} A^{(i,k)}_{\boldsymbol{n},m} + R_{n,k,m} + N_{n,k,m}, \qquad \text{rank}(R_{n,k,m}) \le c(k,m)N(\boldsymbol{n}), \qquad \|N_{n,k,m}\| \le \omega(k,m),$$

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

Let $\delta_m \searrow 0$. Since $\sum_{i=1}^{N_{k,m}} a_{i,k,m} \otimes f_{i,k,m} \to \kappa_m$ in measure when $k \to \infty$,

$$\mu(m,k,\delta_m) = \mu_{2d}\left\{\left|\sum_{i=1}^{N_{k,m}} a_{i,k,m} \otimes f_{i,k,m} - \kappa_m\right| \ge \delta_m\right\} \to 0 \qquad \text{as } k \to \infty.$$

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 $n \ge n_m$,

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

where $\lim_{m\to\infty} c(m) = \lim_{m\to\infty} \omega(m) = 0$. It follows that, for every m, every k, and every $n \ge \max(n_m, n_{k,m})$,

$$A_{\boldsymbol{n}} = \sum_{i=1}^{N_{k,m}} A^{(i,k)}_{\boldsymbol{n},m} + (R_{n,k,m} + R_{n,m}) + (N_{n,k,m} + N_{n,m}),$$

$$\text{rank}(R_{n,k,m} + R_{n,m}) \le (c(k,m) + c(m))\, N(\boldsymbol{n}), \qquad \|N_{n,k,m} + N_{n,m}\| \le \omega(k,m) + \omega(m).$$

Choose $k_m \nearrow \infty$ such that

$$\lim_{m\to\infty} c(k_m,m) = \lim_{m\to\infty} \omega(k_m,m) = \lim_{m\to\infty} \mu(m,k_m,\delta_m) = 0.$$

Then, for every m and every $n \ge \max(n_m, n_{k_m,m})$,

$$A_{\boldsymbol{n}} = \sum_{i=1}^{N_{k_m,m}} A^{(i,k_m)}_{\boldsymbol{n},m} + (R_{n,k_m,m} + R_{n,m}) + (N_{n,k_m,m} + N_{n,m}),$$

$$\text{rank}(R_{n,k_m,m} + R_{n,m}) \le c(k_m, m) + c(m), \qquad \|N_{n,k_m,m} + N_{n,m}\| \le \omega(k_m, m) + \omega(m).$$

It follows that $\{\{\sum_{i=1}^{N_{k_m,m}} A_{\boldsymbol{n},m}^{(i,k_m)}\}_n\}_m$ is an a.c.s. for $\{A_{\boldsymbol{n}}\}_n$. Moreover, $\{A_{\boldsymbol{n},m}^{(i,k_m)}\}_n \sim_{\rm sLT} a_{i,k_m,m} \otimes f_{i,k_m,m}$ for $i = 1, \ldots, N_{k_m,m}$, and

$$\sum_{i=1}^{N_{k_m,m}} a_{i,k_m,m} \otimes f_{i,k_m,m} \to \kappa$$

in measure over $[0,1]^d \times [-\pi, \pi]^d$ when $m \to \infty$. Indeed, for any $\delta > 0$,

$$\mu_{2d}\left\{\left|\sum_{i=1}^{N_{k_m,m}} a_{i,k_m,m} \otimes f_{i,k_m,m} - \kappa\right| \ge \delta\right\} \le \mu_{2d}\left\{\left|\sum_{i=1}^{N_{k_m,m}} a_{i,k_m,m} \otimes f_{i,k_m,m} - \kappa_m\right| \ge \delta/2\right\} + \mu_{2d}\left\{|\kappa_m - \kappa| \ge \delta/2\right\},$$

$\mu_{2d}\left\{|\kappa_m - \kappa| \ge \delta/2\right\} \to 0$ by assumption (since $\kappa_m \to \kappa$ in measure), and

$$\mu_{2d}\left\{\left|\sum_{i=1}^{N_{k_m,m}} a_{i,k_m,m} \otimes f_{i,k_m,m} - \kappa_m\right| \ge \delta/2\right\} = \mu(m, k_m, \delta/2)$$

tends to 0, because it is eventually less than $\mu(m, k_m, \delta_m) \to 0$. Thus, $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ by Proposition 5.1. □

As a first application of Theorem 5.1, we show in Proposition 5.4 that GLT sequences could be defined in terms of LT sequences instead of sLT sequences. In particular, Proposition 5.4 is the same as Proposition 5.1 with 'sLT' replaced by 'LT'. For the proof of Proposition 5.4 we need to 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; such properties will be investigated in Section 5.4 and give rise to the so-called GLT algebra.

Proposition 5.3. *If*

$$\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa, \qquad \{B_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \xi,$$

then

$$\{\alpha A_{\boldsymbol{n}} + \beta B_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \alpha\kappa + \beta\xi$$

for all $\alpha, \beta \in \mathbb{C}$.

The proof of Proposition 5.3 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 5.4. *We have* $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ *if and only if:*

- *there exist matrix-sequences* $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm LT} a_{i,m} \otimes f_{i,m},\ i = 1, \ldots, N_m$*;*
- $\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ *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. It is clear that, if $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$, then the three conditions hold by Proposition 5.1. Conversely, suppose the three conditions hold. 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 5.3, and $\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m} \to \kappa$ in measure, by hypothesis. The thesis follows from Theorem 5.1. □

The approximation result for GLT sequences stated in Theorem 5.1 admits the following converse, which can be seen as another approximation result for GLT sequences. It looks like a characterization of GLT sequences in terms of a.c.s.

Theorem 5.2. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$. Suppose that:*

- $\{B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa_m$ *for each* m;
- $\kappa_m \to \kappa$ *in measure.*

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

Proof. By Proposition 5.3, $\{A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\}_n \sim_{\rm GLT} \kappa - \kappa_m$ for each m. Hence, by Theorem 5.3 below, $\{A_{\boldsymbol{n}} - B_{\boldsymbol{n},m}\}_n \sim_\sigma \kappa - \kappa_m$, 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 $\{A_{\boldsymbol{n}}\}_n$. □

Corollary 5.2. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$. Then, for any $\{a_{i,m}\}_m$, $\{f_{i,m}\}$, $i = 1, \ldots, N_m$, 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 over* $[0,1]^d \times [-\pi,\pi]^d$;

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

$$\left\{\Big\{\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m} D_{\boldsymbol{n}}(a_{\boldsymbol{j}}^{(m)})T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})\Big\}_n\right\}_m, \qquad a_{\boldsymbol{j}}^{(m)} \in C^\infty([0,1]^d), \qquad \boldsymbol{N}_m \in \mathbb{N}^d. \tag{5.4}$$

Proof. By Theorem 4.8 and Corollary 5.1, we have $\{D_{\boldsymbol{n}}(a_{i,m})T_{\boldsymbol{n}}(f_{i,m})\}_n \sim_{\rm GLT} a_{i,m} \otimes f_{i,m}$. Hence, by Proposition 5.3, $\{\sum_{i=1}^{N_m} D_{\boldsymbol{n}}(a_{i,m})T_{\boldsymbol{n}}(f_{i,m})\}_n \sim_{\rm GLT} \sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m}$. Therefore, the thesis follows from Theorem 5.2 and Lemma 2.2. □

5.3 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 5.3).

Lemma 5.1. *Let $\{A_{\boldsymbol{n}}^{(i)}\}_n \sim_{\rm LT} a_i \otimes f_i$, $i = 1, \ldots, p$. Then $\{\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)}\}_n \sim_\sigma \sum_{i=1}^p a_i \otimes f_i$.*

Proof. Choose any sequence $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m$ 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, for every $F \in C_c(\mathbb{R})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{r=1}^{N(\boldsymbol{n})} F\Big(\sigma_r\Big(\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i, f_i)\Big)\Big) = \phi_{\boldsymbol{m}}(F)$$

and

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

Hence, by Theorem 3.3 and Definition 2.2, $\{\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)}\}_n \sim_\sigma \sum_{i=1}^p a_i \otimes f_i$. □

Theorem 5.3. *If $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ then $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \kappa$.*

Proof. By Proposition 5.4, 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 $\{A_{\boldsymbol{n}}\}_n$. 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 5.3, every GLT sequence is s.u. in the sense of Definition 3.2 (see Proposition 3.3). Using Theorem 5.3, we show in Proposition 5.5 that the symbol of a GLT sequence is unique.

Proposition 5.5. *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 5.3, $\{O_{N(\boldsymbol{n})}\}_n \sim_{\rm GLT} \kappa - \xi$. Therefore, by Theorem 5.3, for all test functions $F \in C_c(\mathbb{R})$ we have

$$F(0) = \frac{1}{(2\pi)^d} \int_{[0,1]^d\times[-\pi,\pi]^d} F(|\kappa(\mathbf{x},\boldsymbol{\theta}) - \xi(\mathbf{x},\boldsymbol{\theta})|) d\mathbf{x} d\boldsymbol{\theta}. \tag{5.5}$$

If we assume by contradiction that κ is not a.e. equal to ξ, then $|\kappa - \xi|$ is not a.e. equal to 0 and we can find an interval $[a,b] \subset (0,\infty)$ such that $\mu_{2d}\{|\kappa - \xi| \in [a,b]\} > 0$. Choosing $F \in C_c(\mathbb{R})$ such that $F \geq 0$ over $\mathbb{R}$, $F = 1$ over $[a,b]$ and $F = 0$ over $(-\infty, 0]$, it is clear that (5.5) does not hold for this test function F. This contradiction concludes the proof. □

Remark 5.1. If $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ then $\{A_{\boldsymbol{n}}^*\}_n \sim_{\rm GLT} \overline{\kappa}$. The proof is easy and is left as an exercise for the reader.

Proposition 5.6. *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 Remark 5.1 we have $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \overline{\kappa}$. Thus, by Proposition 5.5, $\kappa = \overline{\kappa}$ a.e., i.e., $\kappa \in \mathbb{R}$ a.e. □

The next lemma concerns the eigenvalue 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 5.4).

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 any sequence $\{\boldsymbol{m} = \boldsymbol{m}(m)\}_m$ such that $\boldsymbol{m} \to \infty$ when $m \to \infty$. From the properties of a.c.s., see Remark 3.2 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, for every $F \in C_c(\mathbb{C})$,

$$\lim_{n\to\infty} \frac{1}{N(\boldsymbol{n})} \sum_{r=1}^{N(\boldsymbol{n})} F\Big(\lambda_r\Big(\Re\Big(\sum_{i=1}^p LT_{\boldsymbol{n}}^{\boldsymbol{m}}(a_i, f_i)\Big)\Big)\Big) = \phi_{\boldsymbol{m}}(F)$$

and

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

Hence, by Theorem 3.5 and Definition 2.2, $\{\Re(\sum_{i=1}^p A_{\boldsymbol{n}}^{(i)})\}_n \sim_\lambda \Re(\sum_{i=1}^p a_i \otimes f_i)$. □

Theorem 5.4. *Let $\{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa$ and suppose that the matrices $A_{\boldsymbol{n}}$ are Hermitian. Then $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \kappa$.*

Proof. By Proposition 5.4, 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 $A_{\boldsymbol{n}} = \Re(A_{\boldsymbol{n}})$, 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 5.6, 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 satisfied, and it follows that $\{A_{\boldsymbol{n}}\}_n \sim_\lambda \kappa$. □

Remark 5.2. By Proposition 5.3 and Remark 5.1, Lemmas 4.2 and 5.1 are particular cases of Theorem 5.3, and Lemma 5.2 is a particular case of Theorem 5.4.

5.4 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 $A_{\boldsymbol{n}} = \text{ops}(A_{\boldsymbol{n}}^{(1)}, \ldots, A_{\boldsymbol{n}}^{(r)})$ 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 $\kappa = \text{ops}(\kappa_1, \ldots, \kappa_r)$ obtained by performing the same operations on the symbols $\kappa_1, \ldots, \kappa_r$.

Theorem 5.5. *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 Remark 5.1 and Proposition 5.3. We prove the third statement. By assumption and Proposition 5.4, there exist matrix-sequences $\{A_{\boldsymbol{n}}^{(i,m)}\}_n \sim_{\rm LT} a_{i,m} \otimes f_{i,m},\ i = 1, \ldots, N_m$, 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 Proposition 5.2, 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 separable trigonometric polynomials, the functions $a_{i,m},\ b_{j,m}$ might be supposed to be continuous, 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 5.3, $\{A_{\boldsymbol{n}}\}_n \sim_\sigma \kappa$ 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 $\{A_{\boldsymbol{n}} B_{\boldsymbol{n}}\}_n$. By Theorem 4.7 and by the fact that the functions $f_{i,m},\ g_{j,m}$ belong to $L^\infty([-\pi,\pi]^d)$, 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, it is clear that

$$\sum_{i=1}^{N_m}\sum_{j=1}^{M_m} a_{i,m}b_{j,m} \otimes f_{i,m}g_{j,m} = \left(\sum_{i=1}^{N_m} a_{i,m} \otimes f_{i,m}\right)\left(\sum_{j=1}^{M_m} b_{j,m} \otimes g_{j,m}\right) \to \kappa\xi$$

in measure, 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 an algebra over the complex field $\mathbb{C}$. More precisely, fix any sequence of d-indices $\{\boldsymbol{n} = \boldsymbol{n}(n)\}_n$ such that $\boldsymbol{n} \to \infty$ when $n \to \infty$; then,

$$\mathscr{A} = \big\{\{A_{\boldsymbol{n}}\}_n :\ \{A_{\boldsymbol{n}}\}_n \sim_{\rm GLT} \kappa \text{ for some measurable function } \kappa : [0,1]^d \times [-\pi,\pi]^d \to \mathbb{C}\big\} \qquad (5.6)$$

is an algebra over $\mathbb{C}$, with respect to the natural operations of addition, scalar-multiplication and product of matrix-sequences. We call $\mathscr{A}$ the GLT algebra. We are going to see in Theorems 5.6–5.7 that the GLT algebra enjoys other nice properties, in addition to those of Theorem 5.5, which make it look like a 'big container', closed under any type of 'regular' operation.

Theorem 5.6 provides a positive answer to a question raised in [47]. Incidentally, we note that in [47] 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 5.6. *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 $f:\mathbb{R}\to\mathbb{R}$.[4]

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

$$f(A_{\boldsymbol{n}})=p_{m,M}(A_{\boldsymbol{n}})+f(A_{\boldsymbol{n}})-p_{m,M}(A_{\boldsymbol{n}}). \tag{5.7}$$

Since any GLT sequence is s.u. (by Theorem 5.3 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 $n\ge n_M$,

$$A_{\boldsymbol{n}}=\hat A_{\boldsymbol{n},M}+\tilde A_{\boldsymbol{n},M},\qquad \mathrm{rank}(\hat A_{\boldsymbol{n},M})\le r(M)N(\boldsymbol{n}),\qquad \|\tilde A_{\boldsymbol{n},M}\|\le M, \tag{5.8}$$

where $\lim_{M\to\infty}r(M)=0$. However, for the purpose of this proof we need a splitting of the form (5.8) 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 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 $n\ge n_M$,

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

where $\lim_{M\to\infty}r(M)=0$. Let $A_{\boldsymbol{n}}=U_{\boldsymbol{n}}\Lambda_{\boldsymbol{n}}U_{\boldsymbol{n}}^*$ 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 $n\ge n_M$,

$$A_{\boldsymbol{n}}=U_{\boldsymbol{n}}\Lambda_{\boldsymbol{n}}U_{\boldsymbol{n}}^*=U_{\boldsymbol{n}}\hat\Lambda_{\boldsymbol{n},M}U_{\boldsymbol{n}}^*+U_{\boldsymbol{n}}\tilde\Lambda_{\boldsymbol{n},M}U_{\boldsymbol{n}}^*=\hat A_{\boldsymbol{n},M}+\tilde A_{\boldsymbol{n},M},$$

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 (5.8) 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 $g:\mathbb{R}\to\mathbb{R}$.

[4] Recall from Proposition 5.6 that $\kappa\in\mathbb{R}$ a.e., because every $A_{\boldsymbol{n}}$ is Hermitian. Hence, $f(\kappa)$ is well-defined.

Going back to (5.7), for every $M > 0$, every m and every $n \geq n_M$ we can write

$$\begin{aligned} f(A_{\boldsymbol{n}}) &= p_{m,M}(A_{\boldsymbol{n}}) + f(\hat{A}_{\boldsymbol{n},M}) + f(\tilde{A}_{\boldsymbol{n},M}) - p_{m,M}(\hat{A}_{\boldsymbol{n},M}) - p_{m,M}(\tilde{A}_{\boldsymbol{n},M}) \\ &= p_{m,M}(A_{\boldsymbol{n}}) + (f - p_{m,M})(\hat{A}_{\boldsymbol{n},M}) + (f - p_{m,M})(\tilde{A}_{\boldsymbol{n},M}). \end{aligned} \tag{5.9}$$

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 $\text{rank}(R_{\boldsymbol{n},m,M}) = \text{rank}(\hat{A}_{\boldsymbol{n},M})$; 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 (5.9), for every $M > 0$, every m and every $n \geq n_M$ we have

$$f(A_{\boldsymbol{n}}) = p_{m,M}(A_{\boldsymbol{n}}) + R_{\boldsymbol{n},m,M} + N_{\boldsymbol{n},m,M}, \tag{5.10}$$

and, by our construction,

$$\begin{aligned} &\text{rank}(R_{\boldsymbol{n},m,M}) = \text{rank}(\hat{A}_{\boldsymbol{n},M}) \leq r(M)N(\boldsymbol{n}), \\ &\|N_{\boldsymbol{n},m,M}\| \leq |f(0) - p_{m,M}(0)| + \|f - p_{m,M}\|_{\infty,[-M,M]} \leq 2\|f - p_{m,M}\|_{\infty,[-M,M]}. \end{aligned} \tag{5.11}$$

Choose a sequence $\{M_m\}_m$ such that, when $m \to \infty$,

$$M_m \to \infty, \qquad \|f - p_{m,M_m}\|_{\infty,[-M_m,M_m]} \to 0. \tag{5.12}$$

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, $\{p_{m,M_m}(A_{\boldsymbol{n}})\}_n \sim_{\text{GLT}} p_{m,M_m}(\kappa)$ by Theorem 5.5. Finally, $p_{m,M_m}(\kappa) \to f(\kappa)$ a.e. in $[0,1]^d \times [-\pi,\pi]^d$, due to (5.12). In conclusion, all the hypotheses of Theorem 5.1 are satisfied and so $\{f(A_{\boldsymbol{n}})\}_n \sim_{\text{GLT}} f(\kappa)$. $\square$

The last issue we are interested in is to know if $\{A_{\boldsymbol{n}}^{-1}\}_n \sim_{\text{GLT}} \kappa^{-1}$ in the case where $\{A_{\boldsymbol{n}}\}_n \sim_{\text{GLT}} \kappa$, 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 $\{A_{\boldsymbol{n}}^\dagger\}_n \sim_{\text{GLT}} \kappa^{-1}$ when $\{A_{\boldsymbol{n}}\} \sim_{\text{GLT}} \kappa$ 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 5.6, because the matrices $A_{\boldsymbol{n}}$ may fail to be Hermitian and, moreover, $f(x) = x^{-1}$ is not a continuous function 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 $\{A_{\boldsymbol{n}}^\dagger\}_n$ 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 ($r = \text{rank}(A)$).

Remark 5.3. Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence. Then, $\{A_{\boldsymbol{n}}\}_n$ is s.v. if and only if

$$\lim_{M\to\infty}\limsup_{n\to\infty}\frac{\#\{i\in\{1,\ldots,N(\boldsymbol{n})\}:\,\sigma_i(A_{\boldsymbol{n}})<1/M\}}{N(\boldsymbol{n})}=0. \tag{5.13}$$

The proof of this equivalence is easy and follows the same line as the proof of the equivalence (1 ⇔ 2) in Proposition 3.2; the details are left to the reader. Note that (5.13) 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 5.7. *Let $\{A_{\boldsymbol{n}}\}_n$ be a matrix-sequence such that $\{A_{\boldsymbol{n}}\}_n\sim_\sigma f$ for some measurable $f:D\subset\mathbb{R}^k\to\mathbb{C}$. Then $\{A_{\boldsymbol{n}}\}_n$ is s.v. if and only if $f\neq 0$ a.e.*

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

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

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 5.7. *Let $\{A_{\boldsymbol{n}}\}_n\sim_{\rm GLT}\kappa$ with $\kappa\neq 0$ a.e., then $\{A_{\boldsymbol{n}}^\dagger\}_n\sim_{\rm GLT}\kappa^{-1}$.*

Proof. Take a sequence of matrix-sequences $\{\{B_{\boldsymbol{n},m}\}_n\}_m$ such that $\{B_{\boldsymbol{n},m}\}_m\sim_{\rm GLT}\xi_m$ for each m, and $\xi_m\to\kappa^{-1}$ in measure. This can be done because, by Lemma 2.2, there exists $\{\xi_m\}_m$, with ξ_m of the form

$$\xi_m=\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m}a_{\boldsymbol{j}}^{(m)}\otimes \mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}},\qquad a_{\boldsymbol{j}}^{(m)}\in C([0,1]^d),\qquad \boldsymbol{N}_m\in\mathbb{N}^d,$$

such that $\xi_m\to\kappa^{-1}$ in measure; hence, it suffices to take $B_{\boldsymbol{n},m}=\sum_{\boldsymbol{j}=-\boldsymbol{N}_m}^{\boldsymbol{N}_m}D_{\boldsymbol{n}}(a_{\boldsymbol{j}}^{(m)})T_{\boldsymbol{n}}(\mathrm{e}^{\mathrm{i}\boldsymbol{j}\cdot\boldsymbol{\theta}})$, which is a GLT sequence with symbol ξ_m (see Theorem 4.8, Corollary 5.1 and Theorem 5.5).

For every m, by Theorem 5.5 we have $\{B_{\boldsymbol{n},m}A_{\boldsymbol{n}}-I_{N(\boldsymbol{n})}\}_n\sim_{\rm GLT}\xi_m\kappa-1$, and $\xi_m\kappa-1\to 0$ in measure. Therefore, by Proposition 3.7, for every m there exists n_m such that, for $n\ge n_m$,

$$B_{\boldsymbol{n},m}A_{\boldsymbol{n}}=I_{N(\boldsymbol{n})}+R_{\boldsymbol{n},m}+N_{\boldsymbol{n},m}, \tag{5.14}$$

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

where $\lim_{m\to\infty} c(m) \lim_{m\to\infty} \omega(m) = 0$. Multiplying (5.14) by $A_{\boldsymbol{n}}^\dagger$, we obtain that, for every m and every $n \ge n_m$,

$$B_{\boldsymbol{n},m}\, A_{\boldsymbol{n}} A_{\boldsymbol{n}}^\dagger = A_{\boldsymbol{n}}^\dagger + (R_{\boldsymbol{n},m} + N_{\boldsymbol{n},m})A_{\boldsymbol{n}}^\dagger. \tag{5.15}$$

Since $\{A_{\boldsymbol{n}}\}_n$ is s.v. (by Theorem 5.3 and Proposition 5.7), $\{A_{\boldsymbol{n}}^\dagger\}_n$ is s.u. Hence, 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 (5.15) we see that, for every m and every $n \ge \max(n_m, \overline{n}_{M_m})$,

$$C_{\boldsymbol{n},m}\, A_{\boldsymbol{n}} A_{\boldsymbol{n}}^\dagger = A_{\boldsymbol{n}}^\dagger + R'_{\boldsymbol{n},m} + N'_{\boldsymbol{n},m}, \tag{5.16}$$

$$\text{rank}(R'_{\boldsymbol{n},m}) \le \big(c(m) + r(M_m)\big)N(\boldsymbol{n}), \qquad \|N'_{\boldsymbol{n},m}\| \le [\omega(m)]^{1/2}.$$

If the matrices $A_{\boldsymbol{n}}$ were invertible, then $A_{\boldsymbol{n}}^\dagger = A_{\boldsymbol{n}}^{-1}$ and (5.16) 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 5.1), would conclude the proof. In the general case where the matrices $A_{\boldsymbol{n}}$ are not invertible will follow again from (5.16) and Theorem 5.1 as soon as we have proved the following: for every m there is $\hat{n}_m$ 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 $\text{rank}(S_{\boldsymbol{n}}) = \#\{i \in \{1,\ldots,N(\boldsymbol{n})\} : \sigma_i(A_{\boldsymbol{n}}) = 0\}$. Hence, the previous claim follows directly from Definition 5.2 and from the fact that $\{A_{\boldsymbol{n}}\}_n$ is s.v. □

5.4.1 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 discuss 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

$$\mathscr{T} = \left\{ \Big\{ \sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij}) \Big\}_n \ :\ r, q_1, \ldots, q_r \in \mathbb{N},\ g_{ij} \in L^1([-\pi,\pi]^d) \text{ for all } i = 1,\ldots,r \text{ and } j = 1,\ldots,q_i \right\}. \tag{5.17}$$

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 (5.6). Indeed, according to Corollary 5.3,

$$\Big\{ \sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij}) \Big\}_n \sim_{\rm GLT} \sum_{i=1}^r \prod_{j=1}^{q_i} 1 \otimes g_{ij} = 1 \otimes \sum_{i=1}^r \prod_{j=1}^{q_i} g_{ij}.$$

Since $\int_{[0,1]^d\times[-\pi,\pi]^d} (1 \otimes g) = \int_{[-\pi,\pi]^d} g$, Theorem 5.3 and Definition 2.2 immediately give

$$\Big\{ \sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij}) \Big\}_n \sim_\sigma \sum_{i=1}^r \prod_{j=1}^{q_i} g_{ij}.$$

Similarly, if the matrices $\sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij})$ are Hermitian, Theorem 5.4 gives

$$\Big\{ \sum_{i=1}^r \prod_{j=1}^{q_i} T_{\boldsymbol{n}}(g_{ij}) \Big\}_n \sim_\lambda \sum_{i=1}^r \prod_{j=1}^{q_i} g_{ij}. \tag{5.18}$$

The extension of the spectral distribution relation (5.18) 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; see [17, Theorem 9]. Note that, if we remove the hypothesis of 'Hermitianity', then we necessarily have to add some additional assumption. Indeed, (5.18) 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 [17] is a topological assumption on the range of the functions g_{ij}. A completely analogous hypothesis was already used in [16, 19] and, especially, in the pioneering work by Tilli [55], in order to extend the spectral distribution relation expressed in Theorem 2.5 to the case where the generating function f is not real (and hence the related Toeplitz matrices $T_{\boldsymbol{n}}(f)$ are not Hermitian).

6 Conclusions: GLT as a Generalized Fourier Analysis

In this revue we have reported unitarly the main features of the GLT sequences, by improving somehow the technical construction, by making the tools more easily usable in applications, and by proving a new key result concerning the stability of Hermitian GLT sequence under the action of a continuous function.

Furthermore we have clearly stated and proved the properties **GLT1**-**GLT5** reportd in the Introduction. More specifically:

- item **GLT1** is contained in Theorem 5.3 and Theorem 5.4;
- item **GLT2** is contained in Theorem 5.5 and Theorem 5.7;
- item **GLT3** is obtained by putting together Lemma 4.2, Theorem 4.6, and Corollary 5.1;
- item **GLT4** is obtained by combining Lemma 4.2, Theorem 4.5, and Corollary 5.1;
- item **GLT5** is proved in Theorem 4.4, taking into account the inclusion LT⊂GLT given Corollary 5.1.

In is worth observing, as already stated in [46], that the GLT theory and in particular items **GLT1**-**GLT5** represent a powerful tool for generalizing the local Fourier Analysis [5, 10] to much more general contexts in approximating PDEs and IEs; see items **GLT6**, **GLT7**.

6.1 The algebra generated by diagonal sampling matrix-sequences, by zero distributed sequences, and by Toeplitz sequences: a tool for computing the symbol of PDE approximations

As a direct consequence, of items **GLT1**-**GLT5**, the GLT algebra is a super-algebra of sequences containing the algebra generated by diagonal sampling matrix-sequences, by zero distributed sequences, and by Toeplitz sequences. As a consequence (see Subsection 1.1, [45, 46, 6, 13, 14, 15, 23] and references therein), taking a general variable coefficients PDE on a $[0,1]^d$, we deduce that any reasonable approximation of the considered PDE by local methods (Finite Differences, Finite Elements, Collocation/Galerkin IgA, Finite Volumes etc) leads to matrix-sequences that can be approximated by linear combinations of products involving diagonal sampling and Toeplitz matrices.

The case of a general domain can be recovered, especially in the IgA setting, via the use of a geometric map [6, 13, 14, 15], or by using the notion of *reduced GLT sequence* as stated in [46][Section 3.1.4], [45][Remark 2.1]: see also [45][pp 398-399, formula (59)] for a specific example of use, where it turns out that the domain where the PDE is defined has to be simply Peano-Jordan measurable, i.e., its characteristic function is Riemann integrable.

6.2 Future work: reduced GLT, block GLT, and an automatic (symbolic) calculus of the symbol

A future work should concern revisiting the work in [46][Section 3.1.4, Section 3.3] for giving a more applicable GLT theory with general domains and matrix-valued symbols: the latter case is not academic because of vector PDEs (one of the simplest is the linear elasticity in saddle-point form) and because of quadrilateral Finite Elements of degree p over a d dimensional domain, which leads to matrix-sequences having a symbol of size p^d, even in the case of a scalar equation.

In [26] we have treated such a case, but only for a constant coefficient equation, while a proper revisiting of a block GLT theory would represent the ideal framework for dealing with general variable coefficients PDEs and Finite Elements of degree p, regularity $k \in \{0, \ldots, p-1\}$, and dimensionality d.

Finally, a possible idea to be explored is to design an automatic procedure for obtaining the symbol of an approximated PDE, as a function of the principal operator, of the PDE coefficients, of the domain, and of the used approximation technique. Some hints are given in [45][Section 2] and [46][Question 3.1] in connection with the Hörmander calculus [32].

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. **21-6** (2014) 722–743.

[2] ARICÓ A., DONATELLI M. *A V-cycle multigrid for multilevel matrix algebras: proof of optimality.* Numer. Math. **105-4** (2007) 511–547.

[3] ARICÓ A., DONATELLI M., SERRA-CAPIZZANO, S. *V-cycle optimal convergence for certain (multilevel) structured linear systems.* SIAM J. Matrix Anal. Appl. **26-1** (2004) 186–214.

[4] AXELSSON O.BARKER V. *Finite Element Solution of Boundary Value Problems, Theory and Computation.* Academic Press Inc., New York (1984).

[5] BRANDT A. *Guide to multigrid development*. In Multigrid Methods, W. Hackbusch and U. Trottenberg Eds., LNM **960** Springer-Verlag, Berlin (1982) 220–312.

[6] BECKERMANN B., SERRA-CAPIZZANO S. *On the asymptotic spectrum of Finite Element matrix sequences.* SIAM J. Numer. Anal. **45** (2007) 746–769.

[7] BHATIA R. *Matrix Analysis.* Springer-Verlag, New York (1997).

[8] BÖTTCHER A., SILBERMANN B. *Introduction to large truncated Toeplitz matrices.* Springer-Verlag, New York (1999).

[9] BREZZI F., FORTIN M. *Mixed and Hybrid Finite Element Methods.* Springer-Verlag, New York (1991).

[10] CHAN T.F., ELMAN H.C. *Fourier Analysis of iterative methods for Elliptic Problems.* SIAM Review **31-1** (1989) 20–49.

[11] CIARLET P.G. *The Finite Element Method for Elliptic Problems.* SIAM (2002).

[12] COTTRELL J. A., HUGHES T. J. R., BAZILEVS Y. *Isogeometric Analysis: toward integration of CAD and FEA.* John Wiley & Sons (2009).

[13] DONATELLI M., GARONI C., MANNI C., SERRA-CAPIZZANO S., SPELEERS H. *Robust and optimal multi-iterative techniques for IgA Galerkin linear systems.* Comp. Meth. Appl. Mech. Eng. **284** (2015) 230–264.

[14] DONATELLI M., GARONI C., MANNI C., SERRA-CAPIZZANO S., SPELEERS H. *Robust and optimal multi-iterative techniques for IgA collocation linear systems.* Comp. Meth. Appl. Mech. Eng. **284** (2015) 1120–1146.

[15] Donatelli M., Garoni C., Manni C., Serra-Capizzano S., Speleers H. *Spectral analysis and spectral symbol of matrices in isogeometric collocation methods.* Math. Comp. (2015) to appear.

[16] 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.

[17] 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. (2015) to appear.

[18] Donatelli, M., Luati A., Martinelli A. *Spectral filtering for trend estimation.* Linear Algebra Appl. **473** (2015), 217–235.

[19] 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.

[20] Donatelli M., Serra-Capizzano S. *On the regularizing power of multigrid-type algorithms*. SIAM J. Sci. Comput. **27-6** (2006) 2053–2076.

[21] 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 (2015) http://www.it.uu.se/research/publications/reports/2015-008/

[22] 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) http://hdl.handle.net/10277/568

[23] Garoni C., Manni C., Pelosi F., Serra-Capizzano S., Speleers H. *On the spectrum of stiffness matrices arising from isogeometric analysis applied to second order elliptic problems*, Numer. Math. **127-4** (2014) 751–799.

[24] 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 (2015) http://www.it.uu.se/research/publications/reports/2015-005/

[25] 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) http://www.it.uu.se/research/publications/reports/2015-012/

[26] Garoni C., Serra-Capizzano S., Sesana D. *Spectral analysis and spectral symbol of d-variate $\mathbb{Q}_p$ Lagrangian FEM stiffness matrices.* SIAM J. Matrix Anal. Appl.(2015) to appear. Technical Report 2015-021, Department of Information Technology, Uppsala University, Sweden (2015) http://www.it.uu.se/research/publications/reports/2014-021/

[27] Garoni C., Serra-Capizzano S., Sesana D. *Tools for the asymptotic spectrum of non-Hermitian perturbations of Hermitian matrix-sequences and applications.* Integral Eq. Oper. Theory (2014) 10.1007/s00020-014-2157-6.

[28] Garoni C., Serra-Capizzano S., Vassalos P. *A general tool for determining the asymptotic spectral distribution of Hermitian matrix-sequences.* Oper. Matrices (2015) to appear.

[29] Golinskii L., Serra-Capizzano S. *The asymptotic properties of the spectrum of nonsymmetrically perturbed Jacobi matrix sequences.* J. Approx. Theory **144** (2007) 84–102.

[30] Golub G., Van Loan C. *Matrix Computations.* Third Edition, Johns Hopkins University Press, Baltimora (1996).

[31] Grenander U., Szegö G. *Toeplitz forms and their applications.* Second Edition, Chelsea, New York (1984).

[32] Hörmander L. *Pseudo-differential operators and non-elliptic boundary problems.* Annals of Math. **2-83** (1966) 129–209.

[33] Hughes T. J. R., Cottrell J. A., Bazilevs Y. *Isogeometric Analysis: CAD, Finite Elements, NURBS, exact geometry and mesh refinement.* Comput. Meth. Appl. Mech. Engrg. **194** (2005) 4135–4195.

[34] Lax P. *On the stability of difference approximations to solutions of hyperbolic equations with variable coefficients.* Comm. Pure Appl. Math. **14** (1961) 497–520.

[35] LAX P., WENDROFF B. *On the stability of difference schemes.* Comm. Pure Appl. Math. **15** (1962) 363–371.

[36] PARTER S. *On the eigenvalues of certain generalizations of Toeplitz matrices.* Arch. Rat. Math. Mech. **3** (1962) 244–257.

[37] RUDIN W. *Real and complex analysis.* Third Edition, McGraw-Hill, Singapore (1987).

[38] SALINELLI E., SERRA-CAPIZZANO S., SESANA D. *Eigenvalue-eigenvector structure of Schoenmakers-Coffey matrices via Toeplitz technology and applications* Linear Algebra Appl. (2015) 10.1016/j.laa.2015.03.017.

[39] SERRA-CAPIZZANO S. *Locally X matrices, spectral distributions, preconditioning and applications*, SIAM J. Matrix Anal. Appl. **21-4** (2000), 1354–1388.

[40] SERRA-CAPIZZANO S. *A note on the asymptotic spectra of finite difference discretizations of second order elliptic partial differential equations.* Asian J. Math. **4** (2000) 499–514.

[41] SERRA-CAPIZZANO S. *Distribution results on the algebra generated by Toeplitz sequences: a finite dimensional approach.* Linear Algebra Appl. **328** (2001) 121–130.

[42] SERRA-CAPIZZANO S. *Spectral behaviour of matrix sequences and discretized boundary value problems.* Linear Algebra Appl. **337-1/3** (2001) 37–78.

[43] SERRA CAPIZZANO S. *Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences*. Numer. Math. **92-3** (2002) 433–465.

[44] SERRA-CAPIZZANO S. *More inequalities and asymptotics for matrix valued linear positive operators: the noncommutative case.* Oper. Theory Adv. Appl. **135** (2002) 293–315.

[45] SERRA-CAPIZZANO S. *Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations.* Linear Algebra Appl. **366** (2003) 371–402.

[46] SERRA-CAPIZZANO S. *The GLT class as a generalized Fourier Analysis and applications.* Linear Algebra Appl. **419** (2006) 180–233.

[47] SERRA-CAPIZZANO S., SESANA D. *Approximating class of sequences: the Hermitian case.* Linear Algebra Appl. **434** (2011) 1163–1170.

[48] SERRA-CAPIZZANO S., TABLINO-POSSIO C. *Spectral and structural analysis of high precision Finite Difference matrices for Elliptic Operators.* Linear Algebra Appl. **293** (1999) 85–131.

[49] SERRA-CAPIZZANO S., TABLINO-POSSIO C. *Analysis of preconditioning strategies for collocation linear systems.* Linear Algebra Appl. **369** (2003) 41–75.

[50] SERRA-CAPIZZANO S., TABLINO-POSSIO C. *Multigrid methods for multilevel circulant matrices.* SIAM J. Sci. Comput. **26-1** (2004) 55–85.

[51] SESANA D. *Spectral distributions of structured matrix-sequences: tools and applications.* Ph.D. Thesis in Mathematics of Computation, University of Insubria, Como, Italy (2010) http://hdl.handle.net/10277/318.

[52] STRIKWERDA J.C. *Finite Difference Schemes and Partial Differential Equations.* SIAM (2004).

[53] TILLI P. *A note on the spectral distribution of Toeplitz matrices.* Linear Multilinear Algebra **45** (1998) 147–159.

[54] TILLI P. *Locally Toeplitz sequences: spectral properties and applications.* Linear Algebra Appl. **278** (1998) 91–120.

[55] TILLI P. *Some results on complex Toeplitz eigenvalues.* J. Math. Anal. Appl. **239** (1999) 390–401.

[56] TYRTYSHNIKOV E. E. *A unifying approach to some old and new theorems on distribution and clustering.* Linear Algebra Appl. **232** (1996) 1–43.