

## BLOCK GENERALIZED LOCALLY TOEPLITZ SEQUENCES: THEORY AND APPLICATIONS IN THE MULTIDIMENSIONAL CASE\*

GIOVANNI BARBARINO<sup>†</sup>, CARLO GARONI<sup>‡</sup>, AND STEFANO SERRA-CAPIZZANO<sup>§</sup>

**Abstract.** In computational mathematics, when dealing with a large linear discrete problem (e.g., a linear system) arising from the numerical discretization of a partial differential equation (PDE), knowledge of the spectral distribution of the associated matrix has proved to be useful information for designing/analyzing appropriate solvers—especially, preconditioned Krylov and multigrid solvers—for the considered problem. Actually, this spectral information is of interest also in itself as long as the eigenvalues of the aforementioned matrix represent physical quantities of interest, which is the case for several problems from engineering and applied sciences (e.g., the study of natural vibration frequencies in an elastic material). The theory of multilevel generalized locally Toeplitz (GLT) sequences is a powerful apparatus for computing the asymptotic spectral distribution of matrices  $A_n$  arising from virtually any kind of numerical discretization of PDEs. Indeed, when the mesh-fineness parameter  $n$  tends to infinity, these matrices  $A_n$  give rise to a sequence  $\{A_n\}_n$ , which often turns out to be a multilevel GLT sequence or one of its “relatives”, i.e., a multilevel block GLT sequence or a (multilevel) reduced GLT sequence. In particular, multilevel block GLT sequences are encountered in the discretization of systems of PDEs as well as in the higher-order finite element or discontinuous Galerkin approximation of scalar/vectorial PDEs. In this work, we systematically develop the theory of multilevel block GLT sequences as an extension of the theories of (unilevel) GLT sequences [Garoni and Serra-Capizzano, *Generalized Locally Toeplitz Sequences: Theory and Applications*. Vol. I., Springer, Cham, 2017], multilevel GLT sequences [Garoni and Serra-Capizzano, *Generalized Locally Toeplitz Sequences: Theory and Applications*. Vol. II., Springer, Cham, 2018], and block GLT sequences [Barbarino, Garoni, and Serra-Capizzano, *Electron. Trans. Numer. Anal.*, 53 (2020), pp. 28–112]. We also present several emblematic applications of this theory in the context of PDE discretizations.

**Key words.** asymptotic distribution of singular values and eigenvalues, multilevel block Toeplitz matrices, multilevel block generalized locally Toeplitz matrices, numerical discretization of partial differential equations, finite differences, finite elements, isogeometric analysis, discontinuous Galerkin methods, tensor products, B-splines

**AMS subject classifications.** 15A18, 15B05, 47B06, 65N06, 65N30, 65N25, 15A60, 15A69, 65D07

<b>1</b>	<b>Introduction</b>	<b>115</b>
<b>2</b>	<b>Mathematical background</b>	<b>116</b>
2.1	Notation and terminology . . . . .	116
2.1.1	General notation and terminology . . . . .	116
2.1.2	Multi-index notation . . . . .	117
2.1.3	Multilevel block matrix-sequences . . . . .	119
2.2	Preliminaries on matrix analysis . . . . .	120
2.2.1	Matrix norms . . . . .	120
2.2.2	Tensor products and direct sums . . . . .	120
2.3	Preliminaries on measure and integration theory . . . . .	123
2.3.1	Measurability . . . . .	123
2.3.2	Essential range of matrix-valued functions . . . . .	124
2.3.3	$L^p$ -norms of matrix-valued functions . . . . .	124
2.3.4	Convergence in measure and the topology $\tau_{\text{measure}}$ . . . . .	125
2.3.5	Multivariate Riemann-integrable functions . . . . .	126

---

\*Received August 12, 2019. Accepted November 6, 2019. Published online on January 30, 2020. Recommended by L. Reichel. Carlo Garoni acknowledges the MIUR Excellence Department Project awarded to the Department of Mathematics, University of Rome “Tor Vergata”, CUP E83C18000100006, and the support obtained by the Beyond Borders Programme of the University of Rome “Tor Vergata” through the project ASTRID, CUP E84I19002250005.

<sup>†</sup>Faculty of Sciences, Scuola Normale Superiore, Italy ([giovanni.barbarino@sns.it](mailto:giovanni.barbarino@sns.it)).

<sup>‡</sup>Department of Mathematics, University of Rome “Tor Vergata”, Italy; Department of Science and High Technology, University of Insubria, Italy ([garoni@mat.uniroma2.it](mailto:garoni@mat.uniroma2.it); [carlo.garoni@uninsubria.it](mailto:carlo.garoni@uninsubria.it)).

<sup>§</sup>Department of Humanities and Innovation, University of Insubria, Italy; Department of Information Technology, Uppsala University, Sweden ([stefano.serrac@uninsubria.it](mailto:stefano.serrac@uninsubria.it); [stefano.serra@it.uu.se](mailto:stefano.serra@it.uu.se)).

2.4	Singular value and eigenvalue distribution of a sequence of matrices . . . . .	127
2.4.1	The notion of singular value and eigenvalue distribution . . . . .	127
2.4.2	Clustering and attraction . . . . .	128
2.4.3	Zero-distributed sequences . . . . .	129
2.4.4	Sparsely unbounded and sparsely vanishing sequences of matrices . . . . .	129
2.4.5	Spectral distribution of sequences of perturbed/compressed/expanded Hermitian matrices . . . . .	131
2.5	Approximating classes of sequences . . . . .	131
2.5.1	Definition of a.c.s. and the a.c.s. topology $\tau_{\text{a.c.s.}}$ . . . . .	131
2.5.2	$\tau_{\text{a.c.s.}}$ and $\tau_{\text{measure}}$ . . . . .	132
2.5.3	The a.c.s. tools for computing singular value and eigenvalue distributions . . . . .	132
2.5.4	The a.c.s. algebra . . . . .	132
2.5.5	A criterion to identify a.c.s. . . . .	133
2.5.6	An extension of the concept of a.c.s. . . . .	133
2.6	Multilevel block Toeplitz matrices . . . . .	134
2.7	Multilevel block diagonal sampling matrices . . . . .	137
<b>3</b>	<b>Multilevel block locally Toeplitz sequences</b> . . . . .	<b>139</b>
3.1	Multilevel block LT operators . . . . .	139
3.1.1	Definition of multilevel block LT operators . . . . .	140
3.1.2	Properties of multilevel block LT operators . . . . .	144
3.2	Definition of multilevel block LT sequences . . . . .	147
3.3	Fundamental examples of multilevel block LT sequences . . . . .	147
3.3.1	Zero-distributed sequences . . . . .	148
3.3.2	Sequences of multilevel block diagonal sampling matrices . . . . .	148
3.3.3	Multilevel block Toeplitz sequences . . . . .	152
3.4	Singular value and spectral distribution of sums of products of multilevel block LT sequences . . . . .	156
3.5	Algebraic properties of multilevel block LT sequences . . . . .	157
3.6	Characterizations and simplified definition of multilevel block LT sequences . . . . .	158
<b>4</b>	<b>Multilevel block generalized locally Toeplitz sequences</b> . . . . .	<b>160</b>
4.1	Equivalent definitions of multilevel block GLT sequences . . . . .	160
4.2	Singular value and spectral distribution of multilevel block GLT sequences . . . . .	161
4.3	Multilevel block GLT sequences and matrix-valued measurable functions . . . . .	163
4.4	The multilevel block GLT algebra . . . . .	164
4.5	Topological density results for multilevel block GLT sequences . . . . .	167
4.6	Characterizations of multilevel block GLT sequences . . . . .	168
4.7	Sequences of multilevel block diagonal sampling matrices . . . . .	169
4.8	Sequences of block matrices with multilevel block GLT blocks . . . . .	170
4.9	Further possible definitions of multilevel block GLT sequences . . . . .	171
<b>5</b>	<b>Summary of the theory</b> . . . . .	<b>172</b>
<b>6</b>	<b>Applications</b> . . . . .	<b>180</b>
6.1	FD discretization of systems of PDEs . . . . .	180
6.2	Higher-order FE discretization of diffusion equations . . . . .	186
6.3	Higher-order FE discretization of convection-diffusion-reaction equations . . . . .	199
6.4	Higher-order FE discretization of systems of PDEs . . . . .	202
6.5	Higher-order isogeometric Galerkin discretization of eigenvalue problems . . . . .	209

**1. Introduction.** The theory of generalized locally Toeplitz (GLT) sequences stems from Tilli's work on locally Toeplitz (LT) sequences [66] and from the spectral theory of Toeplitz matrices [2, 17, 18, 19, 20, 49, 53, 67, 69, 70, 71]. It was then carried forward in [40, 41, 62, 63], and it has been recently extended in [3, 4, 5, 6, 7, 9]. This theory, especially in its multidimensional version [41, 62, 63], is a powerful apparatus for computing the asymptotic spectral distribution of matrices arising from the numerical discretization of continuous problems such as integral equations (IEs) and, especially, partial differential equations (PDEs). Experience reveals that virtually any kind of numerical method for the discretization of PDEs gives rise to structured matrices  $A_n$  whose asymptotic spectral distribution, as the mesh-fineness parameter  $n$  tends to infinity, can be computed through the theory of GLT sequences. We refer the reader to [40, Section 10.5], [41, Section 7.3], and [16, 62, 63] for applications of the theory of GLT sequences in the context of finite difference (FD) discretizations of PDEs; to [40, Section 10.6], [41, Section 7.4], and [11, 16, 33, 55, 63] for the finite element (FE) case; to [13] for the finite volume (FV) case; to [40, Section 10.7], [41, Sections 7.5–7.7], and [27, 35, 36, 37, 38, 56] for the case of isogeometric analysis (IgA) discretizations, both in the collocation and Galerkin frameworks; and to [31] for a further application to fractional differential equations. We also refer the reader to [40, Section 10.4] and [1, 59] for a look at the GLT approach for sequences of matrices arising from IE discretizations.

It is worth emphasizing that the asymptotic spectral distribution of PDE discretization matrices, whose computation is the main objective of the theory of GLT sequences, is not only interesting from a theoretical viewpoint but can also be used for practical purposes. For example, it is known that the convergence properties of mainstream iterative solvers, such as multigrid and preconditioned Krylov methods, strongly depend on the spectral features of the matrices to which they are applied. The spectral distribution can then be exploited to design efficient solvers of this kind and to analyze/predict their performance. In this regard, we recall that noteworthy estimates on the superlinear convergence of the conjugate gradient method obtained by Beckermann and Kuijlaars in [10] are closely related to the asymptotic spectral distribution of the considered matrices. More recently, in the context of Galerkin and collocation IgA discretizations of elliptic PDEs, the spectral distribution computed through the theory of GLT sequences in a series of papers [27, 35, 36, 37, 38] was exploited in [25, 26, 28] to devise and analyze optimal and robust multigrid solvers for IgA linear systems. In addition to the design and analysis of appropriate solvers, the spectral distribution of PDE discretization matrices is of interest also in itself whenever the eigenvalues of such matrices represent relevant physical quantities. This is the case for a broad class of problems arising in engineering and applied sciences such as the study of natural vibration frequencies for an elastic material; see the review [47] and the references therein.

In [8], starting from the original intuition in [63, Section 3.3] and based on the recent contributions [3, 6, 7, 9, 39, 42, 45, 46], the theory of block GLT sequences has been developed in a systematic way as an extension of the theory of GLT sequences. The focus of [8], however, is only on the unidimensional (or unilevel) version of the theory, which allows one to face only unidimensional PDEs (i.e., ordinary differential equations). In this work, we complete [8] by covering the multidimensional (or multilevel) version of the theory, also known as the theory of multilevel block GLT sequences. Such a completion is of the utmost importance in practical applications; in particular, it provides the necessary tools for computing the spectral distribution of multilevel block matrices arising from the discretization of systems of PDEs [63, Section 3.3] and from the higher-order FE or discontinuous Galerkin (DG) approximation of scalar/vectorial PDEs [12, 34, 44, 47]. In addition to developing the theory of multilevel block GLT sequences, we also present some of its most emblematic applications in the context of PDE discretizations.

The present work is structured as a long research article in book form. Chapter 2 collects the necessary preliminaries. Chapters 3 and 4 cover the theory of multilevel block GLT sequences, which is finally summarized in Chapter 5. Chapter 6 is devoted to applications. The exposition in this work is conducted on an abstract level; for motivations and insights we recommend that the reader takes a look at the extended introduction of [8]. Needless to say, the reader who knows [8] will be certainly facilitated in reading this work.

**2. Mathematical background.** This chapter collects the necessary preliminaries for developing the theory of multilevel block GLT sequences.

### 2.1. Notation and terminology.

#### 2.1.1. General notation and terminology.

- A permutation  $\sigma$  of the set  $\{1, 2, \dots, n\}$  is denoted by  $[\sigma(1), \sigma(2), \dots, \sigma(n)]$ .
- $O_m$  and  $I_m$  denote, respectively, the  $m \times m$  zero matrix and the  $m \times m$  identity matrix. Sometimes, when the size  $m$  can be inferred from the context,  $O$  and  $I$  are used instead of  $O_m$  and  $I_m$ . The symbol  $O$  is also used to indicate rectangular zero matrices whose sizes are clear from the context.
- For every  $s \in \mathbb{N}$  and every  $\alpha, \beta = 1, \dots, s$ , we denote by  $E_{\alpha\beta}^{(s)}$  the  $s \times s$  matrix having 1 in position  $(\alpha, \beta)$  and 0 elsewhere.
- The eigenvalues and the singular values of a matrix  $X \in \mathbb{C}^{m \times m}$  are denoted by  $\lambda_j(X)$ ,  $j = 1, \dots, m$ , and  $\sigma_j(X)$ ,  $j = 1, \dots, m$ , respectively. The maximum and minimum singular values of  $X$  are also denoted by  $\sigma_{\max}(X)$  and  $\sigma_{\min}(X)$ , respectively. The spectrum of  $X$  is denoted by  $\Lambda(X)$ .
- If  $1 \leq p \leq \infty$ , the symbol  $|\cdot|_p$  denotes both the  $p$ -norm of vectors and the associated operator norm for matrices:

$$|\mathbf{x}|_p = \begin{cases} (\sum_{i=1}^m |x_i|^p)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \max_{i=1, \dots, m} |x_i|, & \text{if } p = \infty, \end{cases} \quad \mathbf{x} \in \mathbb{C}^m,$$

$$|X|_p = \max_{\substack{\mathbf{x} \in \mathbb{C}^m \\ \mathbf{x} \neq \mathbf{0}}} \frac{|X\mathbf{x}|_p}{|\mathbf{x}|_p}, \quad X \in \mathbb{C}^{m \times m}.$$

The 2-norm  $|\cdot|_2$  is also known as the spectral (or Euclidean) norm; it will be preferably denoted by  $\|\cdot\|$ .

- Given  $X \in \mathbb{C}^{m \times m}$  and  $1 \leq p \leq \infty$ ,  $\|X\|_p$  denotes the Schatten  $p$ -norm of  $X$ , which is defined as the  $p$ -norm of the vector  $(\sigma_1(X), \dots, \sigma_m(X))$ . The Schatten 1-norm is also called the trace-norm. The Schatten 2-norm  $\|X\|_2$  coincides with the classical Frobenius norm  $(\sum_{i,j=1}^m |x_{ij}|^2)^{1/2}$ . The Schatten  $\infty$ -norm  $\|X\|_\infty = \sigma_{\max}(X)$  is the classical 2-norm  $\|X\|$ . For more on Schatten  $p$ -norms, see [14].
- $\Re(X)$  and  $\Im(X)$  are, respectively, the real and imaginary parts of the (square) matrix  $X$ , i.e.,  $\Re(X) = \frac{X+X^*}{2}$  and  $\Im(X) = \frac{X-X^*}{2i}$ , where  $X^*$  is the conjugate transpose of  $X$  and  $i$  is the imaginary unit.
- If  $X, Y \in \mathbb{C}^{m \times \ell}$ , their componentwise (or Hadamard) product  $X \circ Y$  is the  $m \times \ell$  matrix defined by  $(X \circ Y)_{ij} = x_{ij}y_{ij}$ , for  $i = 1, \dots, m$  and  $j = 1, \dots, \ell$ .
- If  $X \in \mathbb{C}^{m \times m}$ , we denote by  $X^\dagger$  the Moore-Penrose pseudoinverse of  $X$ .
- $C_c(\mathbb{C})$  (resp.,  $C_c(\mathbb{R})$ ) is the space of complex-valued (resp., real-valued) continuous functions defined on  $\mathbb{C}$  (resp.,  $\mathbb{R}$ ) and with bounded support.
- Let  $w_i : D_i \rightarrow \mathbb{C}^{r_i \times r_i}$ ,  $i = 1, \dots, d$ , set  $\mathbf{r} = (r_1, \dots, r_d)$  and  $N(\mathbf{r}) = r_1 \cdots r_d$ . We define the tensor-product function  $w_1 \otimes \cdots \otimes w_d : D_1 \times \cdots \times D_d \rightarrow \mathbb{C}^{N(\mathbf{r}) \times N(\mathbf{r})}$  as follows: for every  $(\xi_1, \dots, \xi_d) \in D_1 \times \cdots \times D_d$ ,

$$(w_1 \otimes \cdots \otimes w_d)(\xi_1, \dots, \xi_d) = w_1(\xi_1) \otimes \cdots \otimes w_d(\xi_d),$$

where  $\otimes$  denotes the tensor (Kronecker) product of matrices (see Section 2.2.2).

- If  $z \in \mathbb{C}$  and  $\varepsilon > 0$ , we denote by  $D(z, \varepsilon)$  the open disk with center  $z$  and radius  $\varepsilon$ , i.e.,  $D(z, \varepsilon) = \{w \in \mathbb{C} : |w - z| < \varepsilon\}$ . If  $S \subseteq \mathbb{C}$  and  $\varepsilon > 0$ , we denote by  $D(S, \varepsilon)$  the  $\varepsilon$ -expansion of  $S$ , which is defined as  $D(S, \varepsilon) = \bigcup_{z \in S} D(z, \varepsilon)$ .
- $\chi_E$  is the characteristic (indicator) function of the set  $E$ .
- A concave bounded continuous function  $\varphi : [0, \infty) \rightarrow [0, \infty)$  such that  $\varphi(0) = 0$  and  $\varphi > 0$  on  $(0, \infty)$  is referred to as a gauge function. It can be shown that any gauge function  $\varphi$  is non-decreasing and subadditive, i.e.,  $\varphi(x + y) \leq \varphi(x) + \varphi(y)$  for all  $x, y \in [0, \infty)$ ; see, e.g., [40, Exercise 2.4].
- If  $g : D \rightarrow \mathbb{C}$  is continuous over  $D$ , with  $D \subseteq \mathbb{C}^k$  for some  $k$ , then we denote by  $\omega_g(\cdot)$  the modulus of continuity of  $g$ ,

$$\omega_g(\delta) = \sup_{\substack{\mathbf{x}, \mathbf{y} \in D \\ \|\mathbf{x} - \mathbf{y}\|_\infty \leq \delta}} |g(\mathbf{x}) - g(\mathbf{y})|, \quad \delta > 0.$$

- If  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$  are such that  $x_i \leq y_i$  for all  $i = 1, \dots, d$ , then the symbol  $[\mathbf{x}, \mathbf{y}]$  denotes the hyperrectangle  $[\mathbf{x}, \mathbf{y}] = [x_1, y_1] \times \dots \times [x_d, y_d]$ . Similar meanings have the symbols  $(\mathbf{x}, \mathbf{y})$ ,  $(\mathbf{x}, \mathbf{y})$ ,  $[\mathbf{x}, \mathbf{y}]$ .
- $\mu_k$  denotes the Lebesgue measure in  $\mathbb{R}^k$ . Throughout this work, unless stated otherwise, all the terminology from measure theory (such as “measurable set”, “measurable function”, “a.e.”, etc.) is always referred to the Lebesgue measure.
- Let  $D \subseteq \mathbb{R}^k$ , let  $r \geq 1$  and  $1 \leq p \leq \infty$ . A matrix-valued function  $f : D \rightarrow \mathbb{C}^{r \times r}$  is said to be measurable (resp., continuous, a.e. continuous, bounded, Riemann-integrable, in  $L^p(D)$ , in  $C^\infty(D)$ , etc.) if its components  $f_{\alpha\beta} : D \rightarrow \mathbb{C}$ ,  $\alpha, \beta = 1, \dots, r$ , are measurable (resp., continuous, a.e. continuous, bounded, Riemann-integrable, in  $L^p(D)$ , in  $C^\infty(D)$ , etc.). The space of functions  $f : D \rightarrow \mathbb{C}^{r \times r}$  belonging to  $L^p(D)$  will be denoted by  $L^p(D, r)$  in order to emphasize the dependence on  $r$ . For the space of scalar functions  $L^p(D, 1)$ , we will preferably use the traditional simpler notation  $L^p(D)$ .
- Let  $f_m, f : D \subseteq \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  be measurable. We say that  $f_m$  converges to  $f$  in measure (resp., a.e., in  $L^p(D)$ , etc.) if  $(f_m)_{\alpha\beta}$  converges to  $f_{\alpha\beta}$  in measure (resp., a.e., in  $L^p(D)$ , etc.) for all  $\alpha, \beta = 1, \dots, r$ .
- If  $D$  is any measurable subset of some  $\mathbb{R}^k$  and  $r \in \mathbb{N}$ , then we set

$$\mathfrak{M}_D^{(r)} = \{f : D \rightarrow \mathbb{C}^{r \times r} : f \text{ is measurable}\}.$$

If  $D = [0, 1]^d \times [-\pi, \pi]^d$ , we preferably use the notation  $\mathfrak{M}_d^{(r)}$  instead of  $\mathfrak{M}_D^{(r)}$ :

$$\mathfrak{M}_d^{(r)} = \{\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{r \times r} : \kappa \text{ is measurable}\}.$$

- We use a notation borrowed from probability theory to indicate sets. For example, if  $f, g : D \subseteq \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$ , then  $\{\sigma_{\max}(f) > 0\} = \{\mathbf{x} \in D : \sigma_{\max}(f(\mathbf{x})) > 0\}$ ,  $\mu_k\{\|f - g\| \geq \varepsilon\}$  is the measure of the set  $\{\mathbf{x} \in D : \|f(\mathbf{x}) - g(\mathbf{x})\| \geq \varepsilon\}$ , etc.
- A function of the form  $f(\theta) = \sum_{j=-q}^q f_j e^{ij\theta}$  with  $f_{-q}, \dots, f_q \in \mathbb{C}^{r \times r}$  is said to be a (matrix-valued) trigonometric polynomial. If  $f_{-q} \neq O_r$  or  $f_q \neq O_r$ , then the number  $q$  is referred to as the degree of  $f$ .
- A sequence of matrices is a sequence of the form  $\{A_n\}_n$ , where  $A_n$  is a square matrix of size  $d_n$  such that  $d_n \rightarrow \infty$  as  $n \rightarrow \infty$ .

**2.1.2. Multi-index notation.** A multi-index  $\mathbf{i}$  of size  $d$ , also called a  $d$ -index, is simply a (row) vector in  $\mathbb{Z}^d$ ; its components are denoted by  $i_1, \dots, i_d$ .

- $\mathbf{0}, \mathbf{1}, \mathbf{2}, \dots$  are the vectors of all zeros, all ones, all twos,  $\dots$  (their size will be clear from the context).
- For any  $d$ -index  $\mathbf{m}$ , we set  $N(\mathbf{m}) = \prod_{j=1}^d m_j$ , and we write  $\mathbf{m} \rightarrow \infty$  to indicate that  $\min(\mathbf{m}) \rightarrow \infty$ . The notation  $N(\boldsymbol{\alpha}) = \prod_{j=1}^d \alpha_j$  will be actually used for any vector  $\boldsymbol{\alpha}$  with  $d$  components and not only for  $d$ -indices.
- If  $\mathbf{h}, \mathbf{k}$  are  $d$ -indices, then  $\mathbf{h} \leq \mathbf{k}$  means that  $h_r \leq k_r$  for all  $r = 1, \dots, d$ , while  $\mathbf{h} \not\leq \mathbf{k}$  means that  $h_r > k_r$  for at least one  $r \in \{1, \dots, d\}$ .
- If  $\mathbf{h}, \mathbf{k}$  are  $d$ -indices such that  $\mathbf{h} \leq \mathbf{k}$ , then the multi-index range  $\mathbf{h}, \dots, \mathbf{k}$  (or, more precisely, the  $d$ -index range  $\mathbf{h}, \dots, \mathbf{k}$ ) is the set of cardinality  $N(\mathbf{k} - \mathbf{h} + \mathbf{1})$  given by  $\{\mathbf{j} \in \mathbb{Z}^d : \mathbf{h} \leq \mathbf{j} \leq \mathbf{k}\}$ . We assume for this set the standard lexicographic ordering:

$$(2.1) \quad \left[ \dots \left[ (j_1, \dots, j_d) \right]_{j_d=h_d, \dots, k_d} \right]_{j_{d-1}=h_{d-1}, \dots, k_{d-1}} \dots \right]_{j_1=h_1, \dots, k_1}.$$

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

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

- When a  $d$ -index  $\mathbf{j}$  varies over a  $d$ -index range  $\mathbf{h}, \dots, \mathbf{k}$  (this is often written as  $\mathbf{j} = \mathbf{h}, \dots, \mathbf{k}$ ), it is understood that  $\mathbf{j}$  varies from  $\mathbf{h}$  to  $\mathbf{k}$  following the specific ordering (2.1). For instance, if  $\mathbf{m} \in \mathbb{N}^d$  and we write  $\mathbf{x} = [x_i]_{i=1}^{\mathbf{m}}$ , then  $\mathbf{x}$  is a vector of size  $N(\mathbf{m})$  whose components  $x_i$ ,  $i = 1, \dots, \mathbf{m}$ , are ordered in accordance with (2.1): the first component is  $x_1 = x_{(1, \dots, 1, 1)}$ , the second component is  $x_{(1, \dots, 1, 2)}$ , and so on until the last component, which is  $x_{\mathbf{m}} = x_{(m_1, \dots, m_d)}$ . Similarly, if

$$(2.2) \quad X = [x_{ij}]_{i,j=1}^{\mathbf{m}},$$

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

- If  $\mathbf{h}, \mathbf{k}$  are  $d$ -indices such that  $\mathbf{h} \leq \mathbf{k}$ , then the notation  $\sum_{\mathbf{j}=\mathbf{h}}^{\mathbf{k}}$  indicates the summation over all  $\mathbf{j}$  in  $\mathbf{h}, \dots, \mathbf{k}$ .
- If  $i, j$  are  $d$ -indices, then  $i \preceq j$  means that  $i$  precedes (or equals)  $j$  in the lexicographic ordering (which is a total ordering on  $\mathbb{Z}^d$ ). Moreover, we define

$$i \wedge j = \begin{cases} i, & \text{if } i \preceq j, \\ j, & \text{if } i \succ j. \end{cases}$$

Note that  $i \wedge j$  is the minimum among  $i$  and  $j$  with respect to the lexicographic ordering. In the case where  $i$  and  $j$  are 1-indices (i.e., normal scalar indices), it is clear that  $i \wedge j = \min(i, j)$ .

- Let  $\{a_n\}_{n \in \mathbb{N}^d}$  be a family of numbers parameterized by a  $d$ -index  $\mathbf{n}$ . The limit of  $a_n$  as  $\mathbf{n} \rightarrow \infty$  is defined, as in the case of a traditional sequence  $\{a_n\}_{n \in \mathbb{N}}$ , in the following way:  $\lim_{\mathbf{n} \rightarrow \infty} a_n = a$  if and only if for every  $\varepsilon > 0$  there exists  $\mathbf{N}$  such that  $a_n \in D(a, \varepsilon)$  for  $\mathbf{n} \geq \mathbf{N}$ . Moreover, we define

$$\limsup_{\mathbf{n} \rightarrow \infty} a_n = \lim_{\mathbf{n} \rightarrow \infty} \left( \sup_{\mathbf{m} \geq \mathbf{n}} a_m \right), \quad \liminf_{\mathbf{n} \rightarrow \infty} a_n = \lim_{\mathbf{n} \rightarrow \infty} \left( \inf_{\mathbf{m} \geq \mathbf{n}} a_m \right).$$

- Operations involving  $d$ -indices that have no meaning in the vector space  $\mathbb{Z}^d$  must be interpreted in the componentwise sense. For instance,

$$\begin{aligned} \mathbf{i}\mathbf{j} &= (i_1j_1, \dots, i_dj_d), \\ \alpha\mathbf{i}/\mathbf{j} &= (\alpha i_1/j_1, \dots, \alpha i_d/j_d), \quad \alpha \in \mathbb{C}, \\ \min(\mathbf{i}, \mathbf{j}) &= (\min(i_1, j_1), \dots, \min(i_d, j_d)), \\ \mathbf{i} \bmod \mathbf{m} &= (i_1 \bmod m_1, \dots, i_d \bmod m_d), \end{aligned}$$

etc.

- When a multi-index appears as subscript or superscript, we sometimes suppress the brackets to simplify the notation. For instance, the component of the vector  $\mathbf{x} = [x_i]_{i=1}^{\mathbf{m}}$  corresponding to the  $d$ -index  $\mathbf{i}$  is denoted by  $x_{\mathbf{i}}$  or  $x_{i_1, \dots, i_d}$ , and we often avoid the heavy notation  $x_{(i_1, \dots, i_d)}$ .
- For every  $s \in \mathbb{N}$  and  $\mathbf{n} \in \mathbb{N}^d$ , we denote by  $\Pi_{\mathbf{n}, s}$  the permutation matrix given by

$$\Pi_{\mathbf{n}, s} = \begin{bmatrix} I_s \otimes \mathbf{e}_1^T \\ I_s \otimes \mathbf{e}_2^T \\ \vdots \\ I_s \otimes \mathbf{e}_n^T \end{bmatrix} = \sum_{k=1}^n \mathbf{e}_k \otimes I_s \otimes \mathbf{e}_k^T,$$

where  $\otimes$  denotes the tensor (Kronecker) product (see Section 2.2.2) and  $\mathbf{e}_i$ ,  $i = 1, \dots, n$ , are the vectors of the canonical basis of  $\mathbb{C}^{N(\mathbf{n})}$ , which, for convenience, are indexed by a  $d$ -index  $\mathbf{i} = 1, \dots, n$  instead of a linear index  $i = 1, \dots, N(\mathbf{n})$ . For every  $s, r \in \mathbb{N}$  and  $\mathbf{n} \in \mathbb{N}^d$ , we define the permutation matrix

$$(2.3) \quad \Pi_{\mathbf{n}, s, r} = \Pi_{\mathbf{n}, s} \otimes I_r.$$

- For every  $s \in \mathbb{N}^d$  and every  $\ell, k = 1, \dots, s$ , we denote by  $E_{\ell k}^{(s)}$  the  $N(s) \times N(s)$  matrix having 1 in position  $(\ell, k)$  and 0 elsewhere.
- A  $d$ -variate  $r \times r$  matrix-valued trigonometric polynomial is a finite linear combination, with coefficients in  $\mathbb{C}^{r \times r}$ , of the  $d$ -variate Fourier frequencies

$$e^{i\mathbf{k} \cdot \boldsymbol{\theta}} = e^{i(k_1\theta_1 + \dots + k_d\theta_d)}, \quad \mathbf{k} \in \mathbb{Z}^d,$$

that is, a function of the form

$$f(\boldsymbol{\theta}) = \sum_{\mathbf{k}=-\mathbf{N}}^{\mathbf{N}} f_{\mathbf{k}} e^{i\mathbf{k} \cdot \boldsymbol{\theta}}, \quad f_{-\mathbf{N}}, \dots, f_{\mathbf{N}} \in \mathbb{C}^{r \times r}, \quad \mathbf{N} \in \mathbb{N}^d.$$

A number of examples in order to help the reader become familiar with the multi-index notation are presented in [41, Section 2.1.2].

**2.1.3. Multilevel block matrix-sequences.** Recall from Section 2.1.1 that a sequence of matrices is a sequence of the form  $\{A_n\}_n$ , where  $n$  varies in some infinite subset of  $\mathbb{N}$  and  $A_n$  is a square matrix of size  $d_n \rightarrow \infty$ . A  $d$ -level  $r$ -block matrix-sequence is a special sequence of matrices of the form  $\{A_n\}_n$ , where:

- $n$  varies in some infinite subset of  $\mathbb{N}$ ;
- $\mathbf{n} = \mathbf{n}(n)$  is a  $d$ -index with positive components which depends on  $n$  and satisfies  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ ;
- $A_n$  is a square matrix of size  $N(\mathbf{n})r$ .

Recall from Section 2.1.2 that  $\mathbf{n} \rightarrow \infty$  means that  $\min(\mathbf{n}) \rightarrow \infty$ .

## 2.2. Preliminaries on matrix analysis.

**2.2.1. Matrix norms.** For the reader's convenience, we report in this section some matrix-norm inequalities that we shall use throughout this work. Given a matrix  $X \in \mathbb{C}^{m \times m}$ , important bounds for  $\|X\|$  in terms of the components of  $X$  are the following [40, pp. 29–30]:

$$|x_{ij}| \leq \|X\|, \quad i, j = 1, \dots, m, \quad X \in \mathbb{C}^{m \times m},$$

$$\|X\| \leq \sqrt{|X|_1 |X|_\infty} \leq \max(|X|_1, |X|_\infty) \leq \sum_{i,j=1}^m |x_{ij}|, \quad X \in \mathbb{C}^{m \times m}.$$

Since  $\|X\| = \sigma_{\max}(X)$  and  $\text{rank}(X)$  is the number of nonzero singular values of  $X$ , we have

$$\|X\| \leq \|X\|_1 \leq \text{rank}(X) \|X\| \leq m \|X\|, \quad X \in \mathbb{C}^{m \times m}.$$

Another important trace-norm inequality is the following [40, p. 33]:

$$\|X\|_1 \leq \sum_{i,j=1}^m |x_{ij}|, \quad X \in \mathbb{C}^{m \times m}.$$

A bound for the Frobenius norm in terms of the spectral norm and the trace-norm is provided by the following inequality:

$$(2.4) \quad \|X\|_2 = \sqrt{\sum_{i=1}^m \sigma_i(X)^2} \leq \sqrt{\sigma_{\max}(X) \sum_{i=1}^m \sigma_i(X)} = \sqrt{\|X\| \|X\|_1}, \quad X \in \mathbb{C}^{m \times m}.$$

If  $1 \leq p, q \leq \infty$  are conjugate exponents, i.e.,  $1/p + 1/q = 1$ , then the following Hölder-type inequality holds for the Schatten norms [14]:

$$(2.5) \quad \|XY\|_1 \leq \|X\|_p \|Y\|_q, \quad X, Y \in \mathbb{C}^{m \times m}.$$

Moreover, for  $1 \leq p \leq \infty$ , we have

$$(2.6) \quad \|AXB\|_p \leq \|A\| \|X\|_p \|B\|, \quad A, X, B \in \mathbb{C}^{m \times m}.$$

This inequality actually holds for all unitarily invariant norms and not only for the Schatten norms; see [14, Proposition IV.2.4].

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

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

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

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

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



(i) Associativity: for all matrices  $X, Y, Z$ ,

$$\begin{aligned}(X \otimes Y) \otimes Z &= X \otimes (Y \otimes Z), \\ (X \oplus Y) \oplus Z &= X \oplus (Y \oplus Z).\end{aligned}$$

(ii) If  $X_1, X_2$  can be multiplied and  $Y_1, Y_2$  can be multiplied, then

$$\begin{aligned}(X_1 \otimes Y_1)(X_2 \otimes Y_2) &= (X_1 X_2) \otimes (Y_1 Y_2), \\ (X_1 \oplus Y_1)(X_2 \oplus Y_2) &= (X_1 X_2) \oplus (Y_1 Y_2).\end{aligned}$$

(iii) For all matrices  $X, Y$ ,

$$\begin{aligned}(X \otimes Y)^* &= X^* \otimes Y^*, & (X \otimes Y)^T &= X^T \otimes Y^T \\ (X \oplus Y)^* &= X^* \oplus Y^*, & (X \oplus Y)^T &= X^T \oplus Y^T.\end{aligned}$$

(iv) Bilinearity (of tensor products): for each fixed matrix  $X$ , the application

$$Y \mapsto X \otimes Y$$

is linear on  $\mathbb{C}^{\ell_1 \times \ell_2}$  for all  $\ell_1, \ell_2 \in \mathbb{N}$ ; for each fixed matrix  $Y$ , the application

$$X \mapsto X \otimes Y$$

is linear on  $\mathbb{C}^{m_1 \times m_2}$  for all  $m_1, m_2 \in \mathbb{N}$ .

From (i)–(iv), a lot of other properties follow. For example, if  $\mathbf{v}$  is a (column) vector and  $X, Y$  are matrices that can be multiplied, then  $(\mathbf{v} \otimes X)Y = (\mathbf{v} \otimes X)([1] \otimes Y) = \mathbf{v} \otimes (XY)$ . If  $X, Y$  are invertible, then  $X \otimes Y$  is invertible with inverse  $X^{-1} \otimes Y^{-1}$ . If  $X, Y$  are normal (resp., Hermitian, symmetric, unitary), then  $X \otimes Y$  is also normal (resp., Hermitian, symmetric, unitary). If  $X \in \mathbb{C}^{m \times m}$  and  $Y \in \mathbb{C}^{\ell \times \ell}$ , then the eigenvalues and singular values of  $X \otimes Y$  are given by

$$\begin{aligned}\{\lambda_i(X)\lambda_j(Y) : i = 1, \dots, m, j = 1, \dots, \ell\}, \\ \{\sigma_i(X)\sigma_j(Y) : i = 1, \dots, m, j = 1, \dots, \ell\},\end{aligned}$$

and the eigenvalues and singular values of  $X \oplus Y$  are given by

$$\begin{aligned}\{\lambda_i(X), \lambda_j(Y) : i = 1, \dots, m, j = 1, \dots, \ell\}, \\ \{\sigma_i(X), \sigma_j(Y) : i = 1, \dots, m, j = 1, \dots, \ell\};\end{aligned}$$

see [40, Exercise 2.5]. In particular, for all  $X \in \mathbb{C}^{m \times m}$ ,  $Y \in \mathbb{C}^{\ell \times \ell}$ , and  $1 \leq p \leq \infty$ , we have

$$\begin{aligned}\|X \otimes Y\|_p &= \|X\|_p \|Y\|_p, \\ (2.7) \quad \|X \oplus Y\|_p &= |(\|X\|_p, \|Y\|_p)|_p = \begin{cases} (\|X\|_p^p + \|Y\|_p^p)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \max(\|X\|_\infty, \|Y\|_\infty), & \text{if } p = \infty, \end{cases}\end{aligned}$$

and

$$(2.8) \quad \begin{aligned}\text{rank}(X \otimes Y) &= \text{rank}(X)\text{rank}(Y), \\ \text{rank}(X \oplus Y) &= \text{rank}(X) + \text{rank}(Y).\end{aligned}$$

In addition to the properties considered so far, we need to highlight two further properties of tensor products that can be found in [41, Section 2.5]. The first one is the *multi-index formula for tensor products*: if we have  $d$  matrices  $X_k \in \mathbb{C}^{m_k \times \ell_k}$ ,  $k = 1, \dots, d$ , then

$$(2.9) \quad (X_1 \otimes X_2 \otimes \cdots \otimes X_d)_{ij} = (X_1)_{i_1 j_1} (X_2)_{i_2 j_2} \cdots (X_d)_{i_d j_d}$$

for  $i = 1, \dots, \mathbf{m}$  and  $j = 1, \dots, \ell$ , where  $\mathbf{m} = (m_1, m_2, \dots, m_d)$  and  $\ell = (\ell_1, \ell_2, \dots, \ell_d)$ . Note that (2.9) can be rewritten as follows:

$$X_1 \otimes X_2 \otimes \cdots \otimes X_d = [(X_1)_{i_1 j_1} (X_2)_{i_2 j_2} \cdots (X_d)_{i_d j_d}]_{\substack{i=1, \dots, \mathbf{m} \\ j=1, \dots, \ell}}$$

The second property is a natural *upper bound for the rank of the difference of two tensor products* formed by  $d$  factors. More precisely, suppose we have  $2d$  matrices  $X_1, \dots, X_d, Y_1, \dots, Y_d$ , with  $X_i, Y_i \in \mathbb{C}^{m_i \times m_i}$  for all  $i = 1, \dots, d$ , then,

$$\text{rank}(X_1 \otimes \cdots \otimes X_d - Y_1 \otimes \cdots \otimes Y_d) \leq N(\mathbf{m}) \sum_{i=1}^d \frac{\text{rank}(X_i - Y_i)}{m_i},$$

where  $\mathbf{m} = (m_1, \dots, m_d)$ .

Concerning the distributive properties of tensor products with respect to direct sums, it follows directly from the definitions that, for all matrices  $X_1, \dots, X_d, Y$ ,

$$(2.10) \quad (X_1 \oplus X_2 \oplus \cdots \oplus X_d) \otimes Y = (X_1 \otimes Y) \oplus (X_2 \otimes Y) \oplus \cdots \oplus (X_d \otimes Y).$$

**DEFINITION 2.1 (Permutation matrix).** *Let  $\varsigma$  be a permutation of the set  $\{1, \dots, m\}$ . The permutation matrix  $\Pi_\varsigma$  associated with  $\varsigma$  is the  $m \times m$  matrix whose rows are  $\mathbf{e}_{\varsigma(1)}^T, \dots, \mathbf{e}_{\varsigma(m)}^T$  (in this order), where  $\mathbf{e}_1, \dots, \mathbf{e}_m$  are the vectors of the canonical basis of  $\mathbb{C}^m$ .*

**LEMMA 2.2.** *For every  $\mathbf{m} \in \mathbb{N}^d$  and every permutation  $\sigma$  of the set  $\{1, \dots, d\}$ , there exists a unique permutation matrix  $\Pi_{\mathbf{m};\sigma}$  that depends only on  $\mathbf{m}$  and  $\sigma$  and satisfies*

$$(2.11) \quad X_{\sigma(1)} \otimes \cdots \otimes X_{\sigma(d)} = \Pi_{\mathbf{m};\sigma} (X_1 \otimes \cdots \otimes X_d) \Pi_{\mathbf{m};\sigma}^T$$

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

*Proof.* The existence of a matrix  $\Pi_{\mathbf{m};\sigma}$  with the required properties has already been proved in [41, Lemma 2.6]. We prove the uniqueness. Suppose that  $\hat{\Pi}_{\mathbf{m};\sigma}$  is another permutation matrix that depends only on  $\mathbf{m}$  and  $\sigma$  and satisfies (2.11) for all matrices  $X_1 \in \mathbb{C}^{m_1 \times m_1}, \dots, X_d \in \mathbb{C}^{m_d \times m_d}$ . This immediately implies that

$$\Pi_{\mathbf{m};\sigma} (X_1 \otimes \cdots \otimes X_d) \Pi_{\mathbf{m};\sigma}^T = \hat{\Pi}_{\mathbf{m};\sigma} (X_1 \otimes \cdots \otimes X_d) \hat{\Pi}_{\mathbf{m};\sigma}^T$$

for all matrices  $X_1 \in \mathbb{C}^{m_1 \times m_1}, \dots, X_d \in \mathbb{C}^{m_d \times m_d}$ , and hence

$$\Pi_{\mathbf{m};\sigma} X \Pi_{\mathbf{m};\sigma}^T = \hat{\Pi}_{\mathbf{m};\sigma} X \hat{\Pi}_{\mathbf{m};\sigma}^T$$

for all matrices  $X \in \mathbb{C}^{N(\mathbf{m}) \times N(\mathbf{m})}$  because

$$\text{span}(X_1 \otimes \cdots \otimes X_d : X_1 \in \mathbb{C}^{m_1 \times m_1}, \dots, X_d \in \mathbb{C}^{m_d \times m_d}) = \mathbb{C}^{N(\mathbf{m}) \times N(\mathbf{m})}.$$

Indeed,  $\mathbb{C}^{N(\mathbf{m}) \times N(\mathbf{m})} = \text{span}(E_{ij} : i, j = 1, \dots, \mathbf{m})$ , where  $E_{ij} = E_{i_1 j_1} \otimes \cdots \otimes E_{i_d j_d}$  and  $E_{i_k j_k}$  is the  $m_k \times m_k$  matrix having 1 in position  $(i_k, j_k)$  and 0 elsewhere, so that  $E_{ij}$  is the matrix having 1 in position  $(i, j)$  and 0 elsewhere. Thus,

$$(2.12) \quad X = P X P^T, \quad P = \Pi_{\mathbf{m};\sigma}^T \hat{\Pi}_{\mathbf{m};\sigma},$$

for all matrices  $X \in \mathbb{C}^{N(\mathbf{m}) \times N(\mathbf{m})}$ . Considering that  $P$  is the product of two permutation matrices, it is itself a permutation matrix, say  $P^T = \Pi_\tau$  with  $\tau$  being a permutation of  $\{1, \dots, N(\mathbf{m})\}$ . Hence, by definition, the columns of  $P$  are  $\mathbf{e}_{\tau(1)}, \dots, \mathbf{e}_{\tau(N(\mathbf{m}))}$  (in this order), with  $\mathbf{e}_1, \dots, \mathbf{e}_{N(\mathbf{m})}$  being the vectors of the canonical basis of  $\mathbb{C}^{N(\mathbf{m})}$ . Let  $E_{ij}$  be the  $N(\mathbf{m}) \times N(\mathbf{m})$  matrix having 1 in position  $(i, j)$  and 0 elsewhere for all  $i, j = 1, \dots, N(\mathbf{m})$ . By (2.12), for all  $i, j = 1, \dots, N(\mathbf{m})$ , we have

$$\mathbf{e}_i \mathbf{e}_j^T = E_{ij} = P E_{ij} P^T = P \mathbf{e}_i \mathbf{e}_j^T P^T = \mathbf{e}_{\tau(i)} \mathbf{e}_{\tau(j)}^T,$$

which implies that  $\tau(i) = i$  and  $\tau(j) = j$  for all  $i, j = 1, \dots, N(\mathbf{m})$ . Hence,  $\tau$  is the identity permutation,  $P^T = \Pi_\tau = I_{N(\mathbf{m})} = P$ , and  $\hat{\Pi}_{\mathbf{m};\sigma} = \Pi_{\mathbf{m};\sigma}$ .  $\square$

REMARK 2.3 (The permutation matrix  $\Gamma_{\mathbf{m},\mathbf{s}}$ ). Let  $\mathbf{m}, \mathbf{s} \in \mathbb{N}^d$ , and let  $\sigma$  be the permutation of the set  $\{1, \dots, 2d\}$  given by  $\sigma = [1, d+1, 2, d+2, \dots, d, 2d]$ . In what follows, we denote by  $(\mathbf{m}, \mathbf{s})$  the  $2d$ -index  $(m_1, \dots, m_d, s_1, \dots, s_d)$ .

- If  $\mathbf{s} = \mathbf{1}$ , then  $\Pi_{(\mathbf{m},\mathbf{s});\sigma} = I_{N(\mathbf{m})}$ . Indeed, by definition,  $\Pi_{(\mathbf{m},\mathbf{1});\sigma}$  is the unique permutation matrix that satisfies

$$\begin{aligned} X_1 \otimes X_{d+1} \otimes X_2 \otimes X_{d+2} \otimes \dots \otimes X_d \otimes X_{2d} \\ = \Pi_{(\mathbf{m},\mathbf{1});\sigma} (X_1 \otimes X_2 \otimes \dots \otimes X_d \otimes X_{d+1} \otimes X_{d+2} \otimes \dots \otimes X_{2d}) \Pi_{(\mathbf{m},\mathbf{1});\sigma}^T \end{aligned}$$

for all  $X_1 \in \mathbb{C}^{m_1 \times m_1}, \dots, X_d \in \mathbb{C}^{m_d \times m_d}, X_{d+1} \in \mathbb{C}^{1 \times 1}, \dots, X_{2d} \in \mathbb{C}^{1 \times 1}$ . Considering that  $X_{d+1}, \dots, X_{2d}$  are scalars, the previous equation becomes

$$\begin{aligned} (X_{d+1} \dots X_{2d}) (X_1 \otimes X_2 \otimes \dots \otimes X_d) \\ = (X_{d+1} \dots X_{2d}) \Pi_{(\mathbf{m},\mathbf{1});\sigma} (X_1 \otimes X_2 \otimes \dots \otimes X_d) \Pi_{(\mathbf{m},\mathbf{1});\sigma}^T \end{aligned}$$

for all  $X_1 \in \mathbb{C}^{m_1 \times m_1}, \dots, X_d \in \mathbb{C}^{m_d \times m_d}, X_{d+1} \in \mathbb{C}^{1 \times 1}, \dots, X_{2d} \in \mathbb{C}^{1 \times 1}$ , which is equivalent to

$$(X_1 \otimes X_2 \otimes \dots \otimes X_d) = \Pi_{(\mathbf{m},\mathbf{1});\sigma} (X_1 \otimes X_2 \otimes \dots \otimes X_d) \Pi_{(\mathbf{m},\mathbf{1});\sigma}^T$$

for all  $X_1 \in \mathbb{C}^{m_1 \times m_1}, \dots, X_d \in \mathbb{C}^{m_d \times m_d}$ . Since  $I_{N(\mathbf{m})}$  satisfies the previous equation, we conclude that  $\Pi_{(\mathbf{m},\mathbf{1});\sigma} = I_{N(\mathbf{m})}$ .

- If  $d = 1$ , then  $\Pi_{(\mathbf{m},\mathbf{s});\sigma} = \Pi_{(\mathbf{m},\mathbf{s});[1,2]} = I_{m_s}$ . This is obvious because  $I_{m_s}$  satisfies the same equation satisfied by  $\Pi_{(\mathbf{m},\mathbf{s});[1,2]}$ , due to the fact that  $\sigma = [1, 2]$  is the identity. In what follows, the matrix  $\Pi_{(\mathbf{m},\mathbf{s});\sigma}$  will be denoted by  $\Gamma_{\mathbf{m},\mathbf{s}}$ .

LEMMA 2.4. For every  $\mathbf{m} \in \mathbb{N}^d$  and every permutation  $\sigma$  of the set  $\{1, \dots, d\}$ , there exists a permutation matrix  $V_{\mathbf{m};\sigma}$  of size  $m_1 + \dots + m_d$  such that

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

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

*Proof.* See [41, Lemma 2.7].  $\square$

### 2.3. Preliminaries on measure and integration theory.

**2.3.1. Measurability.** The following lemma is derived from the results in [14, Section VI.1]. It will be used essentially everywhere in this work, either explicitly or implicitly.

LEMMA 2.5. Let  $f : D \subseteq \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  be measurable and  $g : \mathbb{C}^r \rightarrow \mathbb{C}$  be continuous and symmetric in its  $r$  arguments, i.e.,  $g(\lambda_1, \dots, \lambda_r) = g(\lambda_{\rho(1)}, \dots, \lambda_{\rho(r)})$  for all permutations  $\rho$  of  $\{1, \dots, r\}$ . Then, the function  $\mathbf{x} \mapsto g(\lambda_1(f(\mathbf{x})), \dots, \lambda_r(f(\mathbf{x})))$  is well-defined (independently of the ordering of the eigenvalues of  $f(\mathbf{x})$ ) and measurable. As a consequence:

- the function  $\mathbf{x} \mapsto g(\sigma_1(f(\mathbf{x})), \dots, \sigma_r(f(\mathbf{x})))$  is measurable;
- the functions  $\mathbf{x} \mapsto \sum_{i=1}^r F(\lambda_i(f(\mathbf{x})))$  and  $\mathbf{x} \mapsto \sum_{i=1}^r F(\sigma_i(f(\mathbf{x})))$  are measurable for all continuous  $F : \mathbb{C} \rightarrow \mathbb{C}$ ;
- the function  $\mathbf{x} \mapsto \|f(\mathbf{x})\|_p$  is measurable for all  $p \in [1, \infty]$ .

REMARK 2.6 (Existence of an ordering for the eigenvalues  $\lambda_i(f(\mathbf{x}))$ ). Let the matrix-valued function  $f : D \subseteq \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  be measurable. In the case where all the eigenvalues of the matrix  $f(\mathbf{x})$  are real for almost every  $\mathbf{x} \in D$ , one can define the eigenvalue function  $\lambda_i(f(\mathbf{x}))$  as a measurable function taking the value of the  $i$ th largest eigenvalue of  $f(\mathbf{x})$ . In general, even if  $f$  is continuous, we are not able to find  $r$  continuous functions acting as eigenvalue functions; see [14, Example VI.1.3]. Thus, a convenient ordering on the eigenvalues  $\lambda_i(f(\mathbf{x}))$  cannot be prescribed beforehand. In such cases,  $\lambda_i(f(\mathbf{x}))$  has not to be intended as a function in  $\mathbf{x}$  but as an element of the spectrum  $\Lambda(f(\mathbf{x}))$  ordered in an arbitrary way. Lemma 2.5 is then important as it allows us to work with the spectrum as a whole, without having to specify which ordering we are imposing on the eigenvalues  $\lambda_i(f(\mathbf{x}))$ . In what follows, when we talk about the  $i$ th eigenvalue function  $\lambda_i(f(\mathbf{x}))$ , we are implicitly assuming that this function exists as a measurable function; more precisely, we are assuming that there exist  $r$  measurable functions  $\lambda_i(f(\mathbf{x}))$ ,  $i = 1, \dots, r$ , from  $D$  to  $\mathbb{C}$  such that, for each fixed  $\mathbf{x} \in D$ , the eigenvalues of  $f(\mathbf{x})$  are given by  $\lambda_1(f(\mathbf{x})), \dots, \lambda_r(f(\mathbf{x}))$ .

**2.3.2. Essential range of matrix-valued functions.** If  $f : D \subseteq \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  is a measurable matrix-valued function, then the essential range of  $f$  is denoted by  $\mathcal{ER}(f)$  and is defined as follows:

$$\begin{aligned} \mathcal{ER}(f) &= \{z \in \mathbb{C} : \mu_k\{\exists j \in \{1, \dots, r\} : \lambda_j(f) \in D(z, \varepsilon)\} > 0 \text{ for all } \varepsilon > 0\} \\ &= \{z \in \mathbb{C} : \mu_k\{\min_{j=1, \dots, r} |\lambda_j(f) - z| < \varepsilon\} > 0 \text{ for all } \varepsilon > 0\}, \end{aligned}$$

where we recall that, according to our notation,

$$\{\exists j \in \{1, \dots, r\} : \lambda_j(f) \in D(z, \varepsilon)\} = \{\mathbf{x} \in D : \exists j \in \{1, \dots, r\} : \lambda_j(f(\mathbf{x})) \in D(z, \varepsilon)\}$$

and

$$\{\min_{j=1, \dots, r} |\lambda_j(f) - z| < \varepsilon\} = \{\mathbf{x} \in D : \min_{j=1, \dots, r} |\lambda_j(f(\mathbf{x})) - z| < \varepsilon\}.$$

Note that  $\mathcal{ER}(f)$  is well-defined because the function  $\mathbf{x} \mapsto \min_{j=1, \dots, r} |\lambda_j(f(\mathbf{x})) - z|$  is measurable by Lemma 2.5. It can be shown that  $\mathcal{ER}(f)$  is closed and  $\Lambda(f) \subseteq \mathcal{ER}(f)$  a.e. [8, Lemma 2.2]. In the case where the eigenvalue functions  $\lambda_j(f) : D \rightarrow \mathbb{C}$ ,  $j = 1, \dots, r$ , are measurable, we have

$$\mathcal{ER}(f) = \bigcup_{j=1}^r \mathcal{ER}(\lambda_j(f)).$$

**2.3.3.  $L^p$ -norms of matrix-valued functions.** Let  $D$  be any measurable subset of some  $\mathbb{R}^k$ , let  $r \geq 1$ , and let  $1 \leq p \leq \infty$ . For any measurable function  $f : D \rightarrow \mathbb{C}^{r \times r}$  we define

$$\|f\|_{L^p} = \begin{cases} \left( \int_D \|f(\mathbf{x})\|_p^p d\mathbf{x} \right)^{1/p}, & \text{if } 1 \leq p < \infty, \\ \text{ess sup}_{\mathbf{x} \in D} \|f(\mathbf{x})\|, & \text{if } p = \infty. \end{cases}$$

Note that this definition is well-posed by Lemma 2.5. In the case where  $r = 1$ , it reduces to the classical definition of  $L^p$ -norms for scalar functions. As highlighted in [29, p. 164], for

every  $p \in [1, \infty]$  there exist constants  $A_p, B_p > 0$  such that, for all  $f \in L^p(D, r)$ ,

$$\begin{aligned}
 A_p \|f\|_{L^p}^p &\leq \sum_{\alpha, \beta=1}^r \|f_{\alpha\beta}\|_{L^p}^p \leq B_p \|f\|_{L^p}^p, & \text{if } 1 \leq p < \infty, \\
 A_\infty \|f\|_{L^\infty} &\leq \max_{\alpha, \beta=1, \dots, r} \|f_{\alpha\beta}\|_{L^\infty} \leq B_\infty \|f\|_{L^\infty}, & \text{if } p = \infty.
 \end{aligned}$$

This means that  $L^p(D, r)$ , which we defined in Section 2.1.1 as the set of functions  $f : D \rightarrow \mathbb{C}^{r \times r}$  such that each component  $f_{\alpha\beta}$  belongs to  $L^p(D)$ , can also be defined as the set of measurable functions  $f : D \rightarrow \mathbb{C}^{r \times r}$  such that  $\|f\|_{L^p} < \infty$ . Moreover, if we identify two functions  $f, g \in L^p(D, r)$  whenever  $f(\mathbf{x}) = g(\mathbf{x})$  for almost every  $\mathbf{x} \in D$ , then the map  $f \mapsto \|f\|_{L^p}$  is a norm on  $L^p(D, r)$  which induces on  $L^p(D, r)$  the componentwise  $L^p$  convergence, that is,  $f_m \rightarrow f$  in  $L^p(D, r)$  according to the norm  $\|\cdot\|_{L^p}$  if and only if  $(f_m)_{\alpha\beta} \rightarrow f_{\alpha\beta}$  in  $L^p(D)$  for all  $\alpha, \beta = 1, \dots, r$ .

**2.3.4. Convergence in measure and the topology  $\tau_{\text{measure}}$ .** The convergence in measure plays a central role in the theory of multilevel block GLT sequences. A basic lemma about this convergence is reported below [15, Corollary 2.2.6].

LEMMA 2.7. *Let  $f_m, g_m, f, g : D \subseteq \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  be measurable functions.*

- *If  $f_m \rightarrow f$  in measure and  $g_m \rightarrow g$  in measure, then  $\alpha f_m + \beta g_m \rightarrow \alpha f + \beta g$  in measure for all  $\alpha, \beta \in \mathbb{C}$ .*
- *If  $f_m \rightarrow f$  in measure,  $g_m \rightarrow g$  in measure, and  $\mu_k(D) < \infty$ , then  $f_m g_m \rightarrow f g$  in measure.*

Let  $\varphi : [0, \infty) \rightarrow [0, \infty)$  be a gauge function, let  $D \subset \mathbb{R}^k$  be a measurable set with  $0 < \mu_k(D) < \infty$ , and let

$$\mathfrak{M}_D^{(r)} = \{f : D \rightarrow \mathbb{C}^{r \times r} : f \text{ is measurable}\}.$$

Suppose first that  $r = 1$ . If we define

$$\begin{aligned}
 p_{\text{measure}}^\varphi(f) &= \frac{1}{\mu_k(D)} \int_D \varphi(|f|), & f \in \mathfrak{M}_D^{(1)}, \\
 d_{\text{measure}}^\varphi(f, g) &= p_{\text{measure}}^\varphi(f - g), & f, g \in \mathfrak{M}_D^{(1)},
 \end{aligned}$$

then  $d_{\text{measure}}^\varphi$  is a complete pseudometric on  $\mathfrak{M}_D^{(1)}$  such that a sequence  $\{f_m\}_m \subset \mathfrak{M}_D^{(1)}$  converges to  $f \in \mathfrak{M}_D^{(1)}$  according to  $d_{\text{measure}}^\varphi$  if and only if  $f_m \rightarrow f$  in measure. In particular,  $d_{\text{measure}}^\varphi(f, g) = 0$  if and only if  $f \rightarrow g$  in measure, that is, if and only if  $f = g$  a.e. The topology induced on  $\mathfrak{M}_D^{(1)}$  by  $d_{\text{measure}}^\varphi$  is the same for all gauge functions  $\varphi$ ; it is denoted by  $\tau_{\text{measure}}$ , and it is referred to as the topology of convergence in measure on  $\mathfrak{M}_D^{(1)}$ .

Suppose now that  $r \geq 1$ . If we define

$$\begin{aligned}
 \hat{p}_{\text{measure}}^\varphi(f) &= \max_{\alpha, \beta=1, \dots, r} p_{\text{measure}}^\varphi(f_{\alpha\beta}), & f \in \mathfrak{M}_D^{(r)}, \\
 \hat{d}_{\text{measure}}^\varphi(f, g) &= \hat{p}_{\text{measure}}^\varphi(f - g), & f, g \in \mathfrak{M}_D^{(r)},
 \end{aligned}$$

then  $\hat{d}_{\text{measure}}^\varphi$  is a complete pseudometric on  $\mathfrak{M}_D^{(r)}$  such that a sequence  $\{f_m\}_m \subset \mathfrak{M}_D^{(r)}$  converges to  $f \in \mathfrak{M}_D^{(r)}$  according to  $\hat{d}_{\text{measure}}^\varphi$  if and only if  $f_m \rightarrow f$  in measure. In particular,  $\hat{d}_{\text{measure}}^\varphi(f, g) = 0$  if and only if  $f \rightarrow g$  in measure, that is, if and only if  $f = g$  a.e. The topology induced on  $\mathfrak{M}_D^{(r)}$  by  $\hat{d}_{\text{measure}}^\varphi$  is the same for all gauge functions  $\varphi$ ; it is denoted by  $\tau_{\text{measure}}$ , and it is referred to as the topology of convergence in measure on  $\mathfrak{M}_D^{(r)}$ .

Now, let

$$p_{\text{measure}}^{\varphi}(f) = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^r \varphi(\sigma_i(f))}{r}, \quad f \in \mathfrak{M}_D^{(r)},$$

$$d_{\text{measure}}^{\varphi}(f, g) = p_{\text{measure}}^{\varphi}(f - g), \quad f, g \in \mathfrak{M}_D^{(r)}.$$

As highlighted in [8, Section 2.3.4],  $d_{\text{measure}}^{\varphi}$  is another pseudometric on  $\mathfrak{M}_D^{(r)}$ , which is metrically equivalent to  $\hat{d}_{\text{measure}}^{\varphi}$  because

$$\frac{1}{r} \hat{p}_{\text{measure}}^{\varphi}(f) \leq p_{\text{measure}}^{\varphi}(f) \leq r^2 \hat{p}_{\text{measure}}^{\varphi}(f).$$

In particular,  $d_{\text{measure}}^{\varphi}$  induces on  $\mathfrak{M}_D^{(r)}$  the topology  $\tau_{\text{measure}}$  of convergence in measure, and it is complete on  $\mathfrak{M}_D^{(r)}$ , just as  $\hat{d}_{\text{measure}}^{\varphi}$ . Throughout this work, we will use the notations

$$p_{\text{measure}} = p_{\text{measure}}^{\psi}, \quad d_{\text{measure}} = d_{\text{measure}}^{\psi}, \quad \psi(x) = \frac{x}{1+x}.$$

For the next lemma, see [8, Lemma 2.4].

LEMMA 2.8. *Let  $g_m, g : D \subset \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  be measurable functions defined on a set  $D$  with  $0 < \mu_k(D) < \infty$ . If*

$$\lim_{m \rightarrow \infty} \frac{1}{\mu_k(D)} \int_D \frac{\sum_{j=1}^r F(\sigma_j(g_m - g))}{r} = F(0), \quad \forall F \in C_c(\mathbb{R}),$$

then  $g_m \rightarrow g$  in measure.

REMARK 2.9. Let  $f : D \rightarrow \mathbb{C}^{r \times r}$  be a measurable function defined on a set  $D \subset \mathbb{R}^k$  with  $0 < \mu_k(D) < \infty$ , and assume that

$$\frac{1}{\mu_k(D)} \int_D \frac{\sum_{j=1}^r F(\sigma_j(f))}{r} = F(0), \quad \forall F \in C_c(\mathbb{R}).$$

Then  $f = O_r$  a.e. Indeed, by Lemma 2.8, the previous equation implies that  $f \rightarrow O_r$  in measure, i.e.,  $f = O_r$  a.e.

**2.3.5. Multivariate Riemann-integrable functions.** A function  $a : [0, 1]^d \rightarrow \mathbb{C}$  is said to be Riemann-integrable if its real and imaginary parts  $\Re(a), \Im(a) : [0, 1]^d \rightarrow \mathbb{R}$  are Riemann-integrable in the classical sense. Recall that any Riemann-integrable function is *bounded* by definition. We report below a list of properties possessed by Riemann-integrable functions that will be used in this paper, either explicitly or implicitly.

- If  $\alpha, \beta \in \mathbb{C}$  and  $a, b : [0, 1]^d \rightarrow \mathbb{C}$  are Riemann-integrable, then  $\alpha a + \beta b$  is Riemann-integrable.
- If  $a, b : [0, 1]^d \rightarrow \mathbb{C}$  are Riemann-integrable, then  $ab$  is Riemann-integrable.
- If  $a : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable and  $F : \mathbb{C} \rightarrow \mathbb{C}$  is continuous, then  $F(a) : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable.
- If  $a : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable, then  $a$  belongs to  $L^\infty([0, 1]^d)$  and its Lebesgue and Riemann integrals over  $[0, 1]^d$  coincide.
- If  $a : [0, 1]^d \rightarrow \mathbb{C}$  is bounded, then  $a$  is Riemann-integrable if and only if  $a$  is continuous a.e.

Note that the last two properties imply the first three. The proof of the second-to-last property can be found in [57, pp. 73–74], while the last property is Lebesgue's characterization theorem of Riemann-integrable functions [57, p. 104]. Note that the proofs in [57] are made for the

case  $d = 1$  only, but the generalization to the case  $d > 1$  is straightforward. A further property of Riemann-integrable functions that will be used in this paper is reported in the next lemma [41, Lemma 2.5].

LEMMA 2.10. *Let  $a : [0, 1]^d \rightarrow \mathbb{R}$  be Riemann-integrable. For each  $\mathbf{n} \in \mathbb{N}^d$ , consider the partition of  $(0, 1]^d$  given by the  $d$ -dimensional hyperrectangles*

$$I_{\mathbf{i}, \mathbf{n}} = \left( \frac{\mathbf{i} - \mathbf{1}}{\mathbf{n}}, \frac{\mathbf{i}}{\mathbf{n}} \right] = \left( \frac{i_1 - 1}{n_1}, \frac{i_1}{n_1} \right] \times \cdots \times \left( \frac{i_d - 1}{n_d}, \frac{i_d}{n_d} \right], \quad \mathbf{i} = \mathbf{1}, \dots, \mathbf{n},$$

and let

$$a_{\mathbf{i}, \mathbf{n}} \in \left[ \inf_{\mathbf{x} \in I_{\mathbf{i}, \mathbf{n}}} a(\mathbf{x}), \sup_{\mathbf{x} \in I_{\mathbf{i}, \mathbf{n}}} a(\mathbf{x}) \right], \quad \mathbf{i} = \mathbf{1}, \dots, \mathbf{n}.$$

Then

$$\sum_{\mathbf{i}=1}^{\mathbf{n}} a_{\mathbf{i}, \mathbf{n}} \chi_{I_{\mathbf{i}, \mathbf{n}}} \rightarrow a \text{ a.e. in } [0, 1]^d \text{ as } \mathbf{n} \rightarrow \infty, \quad \lim_{\mathbf{n} \rightarrow \infty} \frac{1}{N(\mathbf{n})} \sum_{\mathbf{i}=1}^{\mathbf{n}} a_{\mathbf{i}, \mathbf{n}} = \int_{[0, 1]^d} a(\mathbf{x}) d\mathbf{x}.$$

## 2.4. Singular value and eigenvalue distribution of a sequence of matrices.

### 2.4.1. The notion of singular value and eigenvalue distribution.

We here introduce the fundamental definitions of singular value and eigenvalue (or spectral) distribution for a given sequence of matrices.

DEFINITION 2.11 (Singular value and eigenvalue distribution of a sequence of matrices). *Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ , and let  $f : D \subset \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  be a measurable matrix-valued function defined on a set  $D$  with  $0 < \mu_k(D) < \infty$ .*

- We say that  $\{A_n\}_n$  has a (asymptotic) singular value distribution described by  $f$ , and we write  $\{A_n\}_n \sim_\sigma f$ , if

$$(2.13) \quad \lim_{n \rightarrow \infty} \frac{1}{d_n} \sum_{j=1}^{d_n} F(\sigma_j(A_n)) = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^r F(\sigma_i(f(\mathbf{x})))}{r} d\mathbf{x}, \quad \forall F \in C_c(\mathbb{R}).$$

In this case, the function  $f$  is referred to as the singular value symbol of  $\{A_n\}_n$ .

- We say that  $\{A_n\}_n$  has an (asymptotic) eigenvalue (or spectral) distribution described by  $f$ , and we write  $\{A_n\}_n \sim_\lambda f$ , if

$$(2.14) \quad \lim_{n \rightarrow \infty} \frac{1}{d_n} \sum_{j=1}^{d_n} F(\lambda_j(A_n)) = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^r F(\lambda_i(f(\mathbf{x})))}{r} d\mathbf{x}, \quad \forall F \in C_c(\mathbb{C}).$$

In this case, the function  $f$  is referred to as the eigenvalue (or spectral) symbol of  $\{A_n\}_n$ .

Note that Definition 2.11 is well-posed by Lemma 2.5, which ensures that the functions  $\mathbf{x} \mapsto \sum_{i=1}^r F(\sigma_i(f(\mathbf{x})))$  and  $\mathbf{x} \mapsto \sum_{i=1}^r F(\lambda_i(f(\mathbf{x})))$  are measurable. In this work, whenever we write a relation such as  $\{A_n\}_n \sim_\sigma f$  or  $\{A_n\}_n \sim_\lambda f$ , it is understood that  $f$  is as in Definition 2.11, that is,  $f$  is a measurable function taking values in  $\mathbb{C}^{r \times r}$  for some  $r \geq 1$  and is defined on a subset  $D$  of some  $\mathbb{R}^k$  with  $0 < \mu_k(D) < \infty$ .

REMARK 2.12 (Informal meaning of the singular value and eigenvalue distribution). The informal meaning behind the spectral distribution (2.14) is the following: assuming that  $f$  possesses  $r$  a.e. continuous eigenvalue functions  $\lambda_i(f(\mathbf{x}))$ ,  $i = 1, \dots, r$ , then the eigenvalues of  $A_n$ , except possibly for  $o(d_n)$  outliers, can be subdivided into  $r$  different subsets of approximately the same cardinality, and, for  $n$  large enough, the eigenvalues belonging to

the  $i$ th subset are approximately equal to the samples of the  $i$ th eigenvalue function  $\lambda_i(f(\mathbf{x}))$  over a uniform grid in the domain  $D$ . For instance, if  $k = 1$ ,  $d_n = nr$ , and  $D = [a, b]$ , then, assuming we have no outliers, the eigenvalues of  $A_n$  are approximately equal to

$$\lambda_i\left(f\left(a + j\frac{b-a}{n}\right)\right), \quad j = 1, \dots, n, \quad i = 1, \dots, r,$$

for  $n$  large enough. Similarly, if  $k = 2$ ,  $d_n = n^2r$ , and  $D = [a_1, b_1] \times [a_2, b_2]$ , then, assuming we have no outliers, the eigenvalues of  $A_n$  are approximately equal to

$$\lambda_i\left(f\left(a_1 + j_1\frac{b_1-a_1}{n}, a_2 + j_2\frac{b_2-a_2}{n}\right)\right), \quad j_1, j_2 = 1, \dots, n, \quad i = 1, \dots, r,$$

for  $n$  large enough; and so on for  $k \geq 3$ . A completely analogous meaning can also be given for the singular value distribution (2.13).

**REMARK 2.13 (Rearrangement).** Let  $D = [a_1, b_1] \times \dots \times [a_k, b_k] \subset \mathbb{R}^k$ , and let  $f : D \rightarrow \mathbb{C}^{r \times r}$  be a measurable function possessing  $r$  real-valued a.e. continuous eigenvalue functions  $\lambda_i(f(\mathbf{x}))$ ,  $i = 1, \dots, r$ . Compute for each  $\rho \in \mathbb{N}$  the uniform samples

$$\lambda_i\left(f\left(a_1 + j_1\frac{b_1-a_1}{\rho}, \dots, a_k + j_k\frac{b_k-a_k}{\rho}\right)\right), \quad j_1, \dots, j_k = 1, \dots, \rho, \quad i = 1, \dots, r,$$

sort them in non-decreasing order, and put them into a vector  $(s_1, s_2, \dots, s_{r\rho^k})$ . Let  $\phi_\rho : [0, 1] \rightarrow \mathbb{R}$  be the piecewise linear non-decreasing function that interpolates the samples  $(s_0 = s_1, s_1, s_2, \dots, s_{r\rho^k})$  at the nodes  $(0, \frac{1}{r\rho^k}, \frac{2}{r\rho^k}, \dots, 1)$ , i.e.,

$$\begin{cases} \phi_\rho\left(\frac{i}{r\rho^k}\right) = s_i, & i = 0, \dots, r\rho^k, \\ \phi_\rho \text{ linear on } \left[\frac{i}{r\rho^k}, \frac{i+1}{r\rho^k}\right], & i = 0, \dots, r\rho^k - 1. \end{cases}$$

When  $\rho \rightarrow \infty$ , the function  $\phi_\rho$  converges a.e. to a function  $\phi$ , which is non-decreasing on  $(0, 1)$  and satisfies

$$(2.15) \quad \int_0^1 F(\phi(t))dt = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^r F(\lambda_i(f(\mathbf{x})))}{r} d\mathbf{x}, \quad \forall F \in C_c(\mathbb{C}).$$

The proof of this result is omitted because it is rather technical; it involves arguments from [40, solution of Exercise 3.1] and [6]. The function  $\phi$  is referred to as the canonical rearranged version of  $f$ . What is interesting about  $\phi$  is that, by (2.15), if  $\{A_n\}_n \sim_\lambda f$ , then  $\{A_n\}_n \sim_\lambda \phi$ , i.e., if  $f$  is a spectral symbol of  $\{A_n\}_n$ , then the same is true for  $\phi$ . Moreover,  $\phi$  is a univariate non-decreasing scalar function, and hence it is much easier to handle than  $f$ . According to Remark 2.12, if we have  $\{A_n\}_n \sim_\lambda f$  (and hence also  $\{A_n\}_n \sim_\lambda \phi$ ), then, for  $n$  large enough, the eigenvalues of  $A_n$ , with the possible exception of  $o(d_n)$  outliers, are approximately equal to the samples of  $\phi$  over a uniform grid in  $[0, 1]$ .

**2.4.2. Clustering and attraction.** In what follows, if  $S \subseteq \mathbb{C}$  and  $\varepsilon > 0$ , we denote by  $D(S, \varepsilon)$  the  $\varepsilon$ -expansion of  $S$ , which is defined as  $D(S, \varepsilon) = \bigcup_{z \in S} D(z, \varepsilon)$ .

**DEFINITION 2.14 (Clustering of a sequence of matrices).** Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ , and let  $S \subseteq \mathbb{C}$  be a nonempty subset of  $\mathbb{C}$ .

- We say that  $\{A_n\}_n$  is strongly clustered at  $S$  (in the sense of the eigenvalues), or equivalently, that the eigenvalues of  $\{A_n\}_n$  are strongly clustered at  $S$ , if, for every  $\varepsilon > 0$ , the number of eigenvalues of  $A_n$  lying outside  $D(S, \varepsilon)$  is bounded by a constant  $C_\varepsilon$  independent of  $n$ , that is, for every  $\varepsilon > 0$ ,

$$(2.16) \quad \#\{j \in \{1, \dots, d_n\} : \lambda_j(A_n) \notin D(S, \varepsilon)\} = O(1).$$



- We say that  $\{A_n\}_n$  is weakly clustered at  $S$  (in the sense of the eigenvalues), or equivalently, that the eigenvalues of  $\{A_n\}_n$  are weakly clustered at  $S$ , if, for every  $\varepsilon > 0$ ,

$$(2.17) \quad \#\{j \in \{1, \dots, d_n\} : \lambda_j(A_n) \notin D(S, \varepsilon)\} = o(d_n).$$

By replacing “eigenvalues” with “singular values” and  $\lambda_j(A_n)$  with  $\sigma_j(A_n)$  in (2.16)–(2.17), we obtain the definitions of a sequence of matrices strongly or weakly clustered at a nonempty subset of  $\mathbb{C}$  in the sense of the singular values.

Throughout this work, when we speak of strong/weak cluster, sequence of matrices strongly/weakly clustered, etc., without further specifications, it is understood “in the sense of the eigenvalues”. When the clustering is intended in the sense of the singular values, this is specified every time.

**DEFINITION 2.15** (Spectral attraction). *Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ , and let  $z \in \mathbb{C}$ . We say that  $z$  strongly attracts the spectrum  $\Lambda(A_n)$  with infinite order if, once we have ordered the eigenvalues of  $A_n$  according to their distance from  $z$ ,*

$$|\lambda_1(A_n) - z| \leq |\lambda_2(A_n) - z| \leq \dots \leq |\lambda_{d_n}(A_n) - z|,$$

the following limit relation holds for each fixed  $j \geq 1$ :

$$\lim_{n \rightarrow \infty} |\lambda_j(A_n) - z| = 0.$$

For the next theorem and its corollary, see [8, Theorem 2.12 and Corollary 2.13].

**THEOREM 2.16.** *If  $\{A_n\}_n \sim_\lambda f$ , then  $\{A_n\}_n$  is weakly clustered at the essential range  $\mathcal{ER}(f)$  and every point of  $\mathcal{ER}(f)$  strongly attracts the spectrum  $\Lambda(A_n)$  with infinite order.*

**COROLLARY 2.17.** *If  $\{A_n\}_n \sim_\lambda f$  and  $\Lambda(A_n)$  is contained in  $S \subseteq \mathbb{C}$  for all  $n$ , then  $\mathcal{ER}(f)$  is contained in the closure  $\bar{S}$ .*

**2.4.3. Zero-distributed sequences.** A sequence of matrices  $\{Z_n\}_n$  with  $Z_n$  of size  $d_n$  is said to be zero-distributed if  $\{Z_n\}_n \sim_\sigma 0$ , i.e.,

$$\lim_{n \rightarrow \infty} \frac{1}{d_n} \sum_{j=1}^{d_n} F(\sigma_j(Z_n)) = F(0), \quad \forall F \in C_c(\mathbb{R}).$$

It is clear that, for any  $r \geq 1$ ,  $\{Z_n\}_n \sim_\sigma 0$  is equivalent to  $\{Z_n\}_n \sim_\sigma O_r$ . Theorem 2.18 provides a characterization of zero-distributed sequences together with a sufficient condition for detecting such sequences. For the related proof, see [40, Theorems 3.2 and 3.3].

**THEOREM 2.18.** *Let  $\{Z_n\}_n$  be a sequence of matrices with  $Z_n$  of size  $d_n$ .*

1.  $\{Z_n\}_n \sim_\sigma 0$  if and only if  $Z_n = R_n + N_n$  with  $\lim_{n \rightarrow \infty} (\text{rank}(R_n)/d_n) = \lim_{n \rightarrow \infty} \|N_n\| = 0$ .
2.  $\{Z_n\}_n \sim_\sigma 0$  if there exists  $p \in [1, \infty)$  such that  $\lim_{n \rightarrow \infty} (\|Z_n\|_p^p/d_n) = 0$ .

**2.4.4. Sparsely unbounded and sparsely vanishing sequences of matrices.** The notions of sparsely unbounded and sparsely vanishing sequences of matrices play an important role within the framework of the theory of multilevel block GLT sequences.

**DEFINITION 2.19** (Sparsely unbounded sequence of matrices). *A sequence of matrices  $\{A_n\}_n$  with  $A_n$  of size  $d_n$  is said to be sparsely unbounded (s.u.), if for every  $M > 0$  there exists  $n_M$  such that, for  $n \geq n_M$ ,*

$$\frac{\#\{i \in \{1, \dots, d_n\} : \sigma_i(A_n) > M\}}{d_n} \leq r(M),$$

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

For the proofs of the next three propositions, we refer the reader to Proposition 5.3 of [40] and Propositions 2.2 and 2.3 of [45]. Note that the proof in [40] is made for  $d_n = n$  and the proofs in [45] are made for  $d_n = sn$  for a fixed  $s \in \mathbb{N}$ , but the extension to the case of a general  $d_n$  tending to infinity is straightforward.

PROPOSITION 2.20. *Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ . The following statements are equivalent.*

1.  $\{A_n\}_n$  is s.u.
2.  $\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\#\{i \in \{1, \dots, d_n\} : \sigma_i(A_n) > M\}}{d_n} = 0$ .
3. For every  $M > 0$  there exists  $n_M$  such that, for  $n \geq n_M$ ,

$$A_n = \hat{A}_{n,M} + \tilde{A}_{n,M}, \quad \text{rank}(\hat{A}_{n,M}) \leq r(M)d_n, \quad \|\tilde{A}_{n,M}\| \leq M,$$

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

PROPOSITION 2.21. *If  $\{A_n\}_n, \{A'_n\}_n$  are s.u., then  $\{A_n A'_n\}_n$  is s.u.*

PROPOSITION 2.22. *If  $\{A_n\}_n \sim_\sigma f$ , then  $\{A_n\}_n$  is s.u.*

REMARK 2.23. Let  $\{A_n\}_n$  be an s.u. sequence of Hermitian matrices with  $A_n$  of size  $d_n$ . Then, the following stronger version of condition 3 in Proposition 2.20 is satisfied [8, Remark 2.19]: for every  $M > 0$  there exists  $n_M$  such that, for  $n \geq n_M$ ,

$$A_n = \hat{A}_{n,M} + \tilde{A}_{n,M}, \quad \text{rank}(\hat{A}_{n,M}) \leq r(M)d_n, \quad \|\tilde{A}_{n,M}\| \leq M,$$

where  $\lim_{M \rightarrow \infty} r(M) = 0$ , the matrices  $\hat{A}_{n,M}$  and  $\tilde{A}_{n,M}$  are Hermitian, and for all functions  $g : \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $g(0) = 0$ , we have

$$g(\hat{A}_{n,M} + \tilde{A}_{n,M}) = g(\hat{A}_{n,M}) + g(\tilde{A}_{n,M}).$$

Strictly related to the notion of sparsely unbounded sequences of matrices is the notion of sparsely vanishing sequences of matrices.

DEFINITION 2.24 (Sparsely vanishing sequence of matrices). *A sequence of matrices  $\{A_n\}_n$  with  $A_n$  of size  $d_n$  is said to be sparsely vanishing (s.v.), if for every  $M > 0$  there exists  $n_M$  such that, for  $n \geq n_M$ ,*

$$\frac{\#\{i \in \{1, \dots, d_n\} : \sigma_i(A_n) < 1/M\}}{d_n} \leq r(M),$$

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

REMARK 2.25. If  $\{A_n\}_n$  is s.v., then  $\{A_n^\dagger\}_n$  is s.u. This follows from the fact that the singular values of  $A^\dagger$  are  $1/\sigma_1(A), \dots, 1/\sigma_r(A), 0, \dots, 0$ , where  $\sigma_1(A), \dots, \sigma_r(A)$  are the nonzero singular values of  $A$  ( $r = \text{rank}(A)$ ).

REMARK 2.26. A sequence of matrices  $\{A_n\}_n$  with  $A_n$  of size  $d_n$  is s.v. if and only if

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\#\{i \in \{1, \dots, d_n\} : \sigma_i(A_n) < 1/M\}}{d_n} = 0;$$

see [40, Remark 8.6].

Proposition 2.27 is the analog of Proposition 2.22 for s.v. sequences of matrices [46, Proposition 2.3].

PROPOSITION 2.27. *If  $\{A_n\}_n \sim_\sigma f$ , then  $\{A_n\}_n$  is s.v. if and only if  $f$  is invertible a.e.*

**2.4.5. Spectral distribution of sequences of perturbed/compressed/expanded Hermitian matrices.** Theorem 2.28 reports from [9] a recent important result about the spectral distribution of sequences of perturbed Hermitian matrices. It nicely extends previous results obtained in [43, 48].

**THEOREM 2.28.** *Let  $\{X_n\}_n, \{Y_n\}_n$  be sequences of matrices with  $X_n, Y_n$  of size  $d_n$ , and set  $A_n = X_n + Y_n$ . Assume that the following conditions are met.*

1. *Every  $X_n$  is Hermitian and  $\{X_n\}_n \sim_\lambda f$ .*
  2.  *$\|Y_n\|_2 = o(\sqrt{d_n})$ .*
- Then  $\{A_n\}_n \sim_\lambda f$ .*

**REMARK 2.29.** If  $\|Y_n\| \leq C$  for some constant  $C$  independent of  $n$  and  $\|Y_n\|_1 = o(d_n)$ , then  $Y_n$  satisfies the second assumption in Theorem 2.28 by (2.4).

Theorem 2.30 concerns the singular value and spectral distribution of sequences of matrices obtained as a compression (or expansion) of another sequence of matrices. For the proof, we refer the reader to [52, Theorem 4.3 and Corollary 4.4].

**THEOREM 2.30.** *Let  $\{X_n\}_n$  be a sequence of matrices with  $X_n$  of size  $d_n$ , and let  $\{P_n\}_n$  be a sequence such that  $P_n \in \mathbb{C}^{d_n \times \delta_n}$ ,  $P_n^* P_n = I_{\delta_n}$ ,  $\delta_n \leq d_n$ , and  $\delta_n/d_n \rightarrow 1$  as  $n \rightarrow \infty$ .*

- *We have  $\{X_n\}_n \sim_\sigma f$  if and only if  $\{P_n^* X_n P_n\}_n \sim_\sigma f$ .*
- *In the case where the matrices  $X_n$  are Hermitian, we have  $\{X_n\}_n \sim_\lambda f$  if and only if  $\{P_n^* X_n P_n\}_n \sim_\lambda f$ .*

## 2.5. Approximating classes of sequences.

### 2.5.1. Definition of a.c.s. and the a.c.s. topology $\tau_{\text{a.c.s.}}$

**DEFINITION 2.31** (Approximating class of sequences). *Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ , and let  $\{\{B_{n,m}\}_n\}_m$  be a sequence of sequences of matrices with  $B_{n,m}$  of size  $d_n$ . We say that  $\{\{B_{n,m}\}_n\}_m$  is an approximating class of sequences (a.c.s.) for  $\{A_n\}_n$  if the following condition is met: for every  $m$  there exists  $n_m$  such that, for  $n \geq n_m$ ,*

$$A_n = B_{n,m} + R_{n,m} + N_{n,m}, \quad \text{rank}(R_{n,m}) \leq c(m)d_n, \quad \|N_{n,m}\| \leq \omega(m),$$

where  $n_m, c(m), \omega(m)$  depend only on  $m$  and  $\lim_{m \rightarrow \infty} c(m) = \lim_{m \rightarrow \infty} \omega(m) = 0$ .

Roughly speaking,  $\{\{B_{n,m}\}_n\}_m$  is an a.c.s. for  $\{A_n\}_n$  if, for all sufficiently large  $m$ , the sequence  $\{B_{n,m}\}_n$  approximates (asymptotically) the sequence  $\{A_n\}_n$  in the sense that  $A_n$  is eventually equal to  $B_{n,m}$  plus a small-rank matrix (with respect to the matrix size  $d_n$ ) plus a small-norm matrix.

It turns out that, for each fixed sequence of positive integers  $d_n$  such that  $d_n \rightarrow \infty$ , the notion of a.c.s. is a notion of convergence in the space of all sequences of matrices corresponding to  $\{d_n\}_n$ , i.e.,

$$(2.18) \quad \mathcal{E} = \{\{A_n\}_n : A_n \in \mathbb{C}^{d_n \times d_n} \text{ for every } n\}.$$

To be precise, for every  $\varphi : [0, \infty) \rightarrow [0, \infty)$  and every square matrix  $A \in \mathbb{C}^{\ell \times \ell}$ , let

$$p^\varphi(A) = \frac{1}{\ell} \sum_{i=1}^{\ell} \varphi(\sigma_i(A))$$

and define

$$(2.19) \quad \begin{aligned} p_{\text{a.c.s.}}^\varphi(\{A_n\}_n) &= \limsup_{n \rightarrow \infty} p^\varphi(A_n), & \{A_n\}_n \in \mathcal{E}, \\ d_{\text{a.c.s.}}^\varphi(\{A_n\}_n, \{B_n\}_n) &= p_{\text{a.c.s.}}^\varphi(\{A_n - B_n\}_n), & \{A_n\}_n, \{B_n\}_n \in \mathcal{E}. \end{aligned}$$

**THEOREM 2.32.** *Let  $\varphi : [0, \infty) \rightarrow [0, \infty)$  be a gauge function. Fix a sequence of positive integers  $d_n$  such that  $d_n \rightarrow \infty$ , and let  $\mathcal{E}$  be the space (2.18). The following properties hold.*

1.  $d_{\text{a.c.s.}}^\varphi$  in (2.19) is a complete pseudometric on  $\mathcal{E}$  such that  $d_{\text{a.c.s.}}^\varphi(\{A_n\}_n, \{B_n\}_n) = 0$  if and only if  $\{A_n - B_n\}_n$  is zero-distributed.
2. Suppose  $\{A_n\}_n \in \mathcal{E}$  and  $\{\{B_{n,m}\}_m\}_n \subset \mathcal{E}$ . Then,  $\{\{B_{n,m}\}_m\}_n$  is an a.c.s. for  $\{A_n\}_n$  if and only if we have  $d_{\text{a.c.s.}}^\varphi(\{A_n\}_n, \{\{B_{n,m}\}_m\}_n) \rightarrow 0$  as  $m \rightarrow \infty$ .

Theorem 2.32 was proved in [7]. It justifies the convergence notation  $\{\{B_{n,m}\}_m\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ , which will be used to indicate that  $\{\{B_{n,m}\}_m\}_n$  is an a.c.s. for  $\{A_n\}_n$ . The topology induced on  $\mathcal{E}$  by the pseudometric  $d_{\text{a.c.s.}}^\varphi$  is the same for all gauge functions  $\varphi$ , it is denoted by  $\tau_{\text{a.c.s.}}$ , and it is referred to as the a.c.s. topology. Throughout this work, we will use the notations

$$p_{\text{a.c.s.}} = p_{\text{a.c.s.}}^\psi, \quad d_{\text{a.c.s.}} = d_{\text{a.c.s.}}^\psi, \quad \psi(x) = \frac{x}{1+x}.$$

**2.5.2.  $\tau_{\text{a.c.s.}}$  and  $\tau_{\text{measure}}$ .** Theorem 2.33 highlights important connections between  $\tau_{\text{a.c.s.}}$  and  $\tau_{\text{measure}}$  or, to be more precise, between the pseudometrics  $d_{\text{a.c.s.}}^\varphi$  and  $d_{\text{measure}}^\varphi$  inducing these two topologies. Actually, the connections between  $\tau_{\text{a.c.s.}}$  and  $\tau_{\text{measure}}$  are so deep that they may lead to a “bridge”, in the precise mathematical sense established in [22], between measure theory and the asymptotic linear algebra theory underlying the notion of a.c.s.; a bridge that could be exploited to obtain matrix theory results from measure theory results and vice versa. For deeper insights on this topic, we suggest reading [7, Section 1]. For the proof of Theorem 2.33, see [8, Section 2.5.2].

**THEOREM 2.33.** *If  $\{A_n\}_n \sim_\sigma f$ , then  $p_{\text{a.c.s.}}^\varphi(\{A_n\}_n) = p_{\text{measure}}^\varphi(f)$  for every bounded continuous function  $\varphi : [0, \infty) \rightarrow [0, \infty)$ .*

**2.5.3. The a.c.s. tools for computing singular value and eigenvalue distributions.** The importance of the a.c.s. notion resides in Theorems 2.34 and 2.35, for which we refer the reader to [45, Theorems 3.1 and 3.2].

**THEOREM 2.34.** *Let  $\{A_n\}_n, \{B_{n,m}\}_n$  be sequences of matrices and  $f, f_m : D \rightarrow \mathbb{C}^{r \times r}$  be measurable functions defined on a set  $D \subset \mathbb{R}^k$  with  $0 < \mu_k(D) < \infty$ . Assume that:*

1.  $\{B_{n,m}\}_n \sim_\sigma f_m$  for every  $m$ ;
2.  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ ;
3.  $f_m \rightarrow f$  in measure.

Then  $\{A_n\}_n \sim_\sigma f$ .

**THEOREM 2.35.** *Let  $\{A_n\}_n, \{B_{n,m}\}_n$  be sequences of Hermitian matrices, and let  $f, f_m : D \rightarrow \mathbb{C}^{r \times r}$  be measurable functions defined on a set  $D \subset \mathbb{R}^k$  with  $0 < \mu_k(D) < \infty$ . Assume that:*

1.  $\{B_{n,m}\}_n \sim_\lambda f_m$  for every  $m$ ;
2.  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ ;
3.  $f_m \rightarrow f$  in measure.

Then  $\{A_n\}_n \sim_\lambda f$ .

**REMARK 2.36.** Let  $\{A_n\}_n$  and  $\{B_n\}_n$  be sequences of matrices with  $A_n$  and  $B_n$  of size  $d_n$ , and suppose that  $d_{\text{a.c.s.}}(\{A_n\}_n, \{B_n\}_n) = 0$  (which is equivalent to  $\{A_n - B_n\}_n \sim_\sigma 0$  by Theorem 2.32). By Theorems 2.34 and 2.35,

- $\{A_n\}_n \sim_\sigma f \iff \{B_n\}_n \sim_\sigma f$ ;
- if the matrices  $A_n$  and  $B_n$  are Hermitian, then  $\{A_n\}_n \sim_\lambda f \iff \{B_n\}_n \sim_\lambda f$ .

**2.5.4. The a.c.s. algebra.** Theorem 2.37 collects important algebraic properties possessed by the a.c.s. For the proof, we refer the reader to [45, Theorem 2.3].

**THEOREM 2.37.** *Let  $\{A_n\}_n, \{A'_n\}_n, \{B_{n,m}\}_n, \{B'_{n,m}\}_n$  be sequences of matrices such that  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  and  $\{B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A'_n\}_n$ . Then, the following properties hold.*

- $\{B_{n,m}^*\}_n \xrightarrow{\text{a.c.s.}} \{A_n^*\}_n$ .

- $\{\alpha B_{n,m} + \beta B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{\alpha A_n + \beta A'_n\}_n$  for all  $\alpha, \beta \in \mathbb{C}$ .
- If  $\{A_n\}_n, \{A'_n\}_n$  are s.u., then  $\{B_{n,m}B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_nA'_n\}_n$ .
- If  $\{C_n\}_n$  is s.u., then  $\{B_{n,m}C_n\}_n \xrightarrow{\text{a.c.s.}} \{A_nC_n\}_n$ .

Another important algebraic property of a.c.s. is stated in the next theorem [39, Lemma 1].

**THEOREM 2.38.** *Let  $s \in \mathbb{N}$ , let  $\{A_n = [A_{n,ij}]_{i,j=1}^s\}_n$  and  $\{B_n^{(m)} = [B_{n,ij}^{(m)}]_{i,j=1}^s\}_n$  be sequences of block matrices, and suppose that*

$$\{B_{n,ij}^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{A_{n,ij}\}_n, \quad i, j = 1, \dots, s.$$

Then  $\{B_n^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

**2.5.5. A criterion to identify a.c.s.** In practical applications, it often happens that a sequence of sequences of matrices  $\{\{B_{n,m}\}_m\}_n$  is given together with another sequence of matrices  $\{A_n\}_n$ , and one would like to show that  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  without resorting to Definition 2.31. A way for solving this problem consists in choosing a suitable gauge function  $\varphi$  and proving that  $d_{\text{a.c.s.}}^{\varphi}(\{B_{n,m}\}_n, \{A_n\}_n) \rightarrow 0$  as  $m \rightarrow \infty$ . Another criterion is provided in the next theorem [40, Corollary 5.3].

**THEOREM 2.39.** *Let  $\{A_n\}_n, \{B_{n,m}\}_n$  be sequences of matrices with  $A_n, B_{n,m}$  of size  $d_n$ , and let  $1 \leq p < \infty$ . Suppose that for every  $m$  there exists  $n_m$  such that, for  $n \geq n_m$ ,*

$$\|A_n - B_{n,m}\|_p^p \leq \varepsilon(m, n)d_n,$$

where  $\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \varepsilon(m, n) = 0$ . Then  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

**2.5.6. An extension of the concept of a.c.s.** We now provide a natural extension of the a.c.s. notion. The underlying idea is that, in Definition 2.31, one could choose to approximate  $\{A_n\}_n$  by a class of sequences  $\{\{B_{n,\alpha}\}_n\}_{\alpha \in \mathcal{A}}$  parameterized by a not necessarily integer parameter  $\alpha$ . For example, one may want to use a parameter  $\varepsilon > 0$  and claim that a given class of sequences  $\{\{B_{n,\varepsilon}\}_n\}_{\varepsilon > 0}$  is an a.c.s. for  $\{A_n\}_n$  as  $\varepsilon \rightarrow 0$ . Intuitively, this assertion should have the following meaning: for every  $\varepsilon > 0$  there exists  $n_\varepsilon$  such that, for  $n \geq n_\varepsilon$ ,

$$A_n = B_{n,\varepsilon} + R_{n,\varepsilon} + N_{n,\varepsilon}, \quad \text{rank}(R_{n,\varepsilon}) \leq c(\varepsilon)d_n, \quad \|N_{n,\varepsilon}\| \leq \omega(\varepsilon),$$

where  $n_\varepsilon, c(\varepsilon), \omega(\varepsilon)$  depend only on  $\varepsilon$  and both  $c(\varepsilon)$  and  $\omega(\varepsilon)$  tend to 0 as  $\varepsilon \rightarrow 0$ . This is in fact the correct meaning.

For the definition of multilevel block LT sequences (Definition 3.10), we need the concept of a.c.s. parameterized by a multi-index  $\mathbf{m} \rightarrow \infty$ . In what follows, a multi-index sequence of sequences of matrices is any class of sequences of the form  $\{\{B_{n,\mathbf{m}}\}_n\}_{\mathbf{m} \in \mathcal{M}}$  which satisfies the following two properties.

1.  $\mathcal{M} \subseteq \mathbb{N}^q$  for some  $q \geq 1$  and  $\mathcal{M} \cap \{\mathbf{i} \in \mathbb{N}^q : \mathbf{i} \geq \mathbf{k}\}$  is nonempty for every  $\mathbf{k} \in \mathbb{N}^q$ . We express the latter condition by saying that  $\infty$  is an *accumulation point* for  $\mathcal{M}$ . This is required to ensure that  $\mathbf{m}$  can tend to  $\infty$  inside  $\mathcal{M}$ .
2. For every  $\mathbf{m} \in \mathcal{M}$ ,  $\{B_{n,\mathbf{m}}\}_n$  is a sequence of matrices as defined at the end of Section 2.1.1.

**DEFINITION 2.40** (Approximating class of sequences as  $\mathbf{m} \rightarrow \infty$ ). *Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ , and let  $\{\{B_{n,\mathbf{m}}\}_n\}_{\mathbf{m} \in \mathcal{M}}$  be a multi-index sequence of sequences of matrices with  $B_{n,\mathbf{m}}$  of size  $d_n$ . We say that  $\{\{B_{n,\mathbf{m}}\}_n\}_{\mathbf{m} \in \mathcal{M}}$  is an a.c.s. for  $\{A_n\}_n$  as  $\mathbf{m} \rightarrow \infty$  if the following property holds: for every  $\mathbf{m} \in \mathcal{M}$  there exists  $n_{\mathbf{m}}$  such that, for  $n \geq n_{\mathbf{m}}$ ,*

$$(2.20) \quad A_n = B_{n,\mathbf{m}} + R_{n,\mathbf{m}} + N_{n,\mathbf{m}}, \quad \text{rank}(R_{n,\mathbf{m}}) \leq c(\mathbf{m})d_n, \quad \|N_{n,\mathbf{m}}\| \leq \omega(\mathbf{m}),$$

where the quantities  $n_m, c(\mathbf{m}), \omega(\mathbf{m})$  depend only on  $\mathbf{m}$  and  $\lim_{\mathbf{m} \rightarrow \infty} c(\mathbf{m}) = \lim_{\mathbf{m} \rightarrow \infty} \omega(\mathbf{m}) = 0$ .

Note that an equivalent definition is obtained by replacing, in Definition 2.40, “for all  $\mathbf{m} \in \mathcal{M}$ ” with “for all sufficiently large  $\mathbf{m} \in \mathcal{M}$ ” (i.e., “for every  $\mathbf{m} \in \mathcal{M}$  that is greater than or equal to some  $\hat{\mathbf{m}}$ ”). Indeed, suppose the splitting (2.20) holds for  $\mathbf{m} \geq \hat{\mathbf{m}}$ . For the other values of  $\mathbf{m}$ , define  $n_m = 1, c(\mathbf{m}) = 1, \omega(\mathbf{m}) = 0$  and  $R_{n,\mathbf{m}} = A_n - B_{n,\mathbf{m}}, N_{n,\mathbf{m}} = O_{d_n}$ . Then, we see that (2.20) holds for every  $\mathbf{m} \in \mathcal{M}$ .

REMARK 2.41. Definition 2.40 extends the classical definition of a.c.s. (Definition 2.31). Indeed, a classical a.c.s.  $\{\{B_{n,m}\}_n\}_m$  for  $\{A_n\}_n$  is an a.c.s. also in the sense of Definition 2.40 (take  $\mathcal{M}$  as the infinite subset of  $\mathbb{N}$  where  $m$  varies). In addition, if  $\{\{B_{n,m}\}_n\}_{m \in \mathcal{M}}$  is an a.c.s. for  $\{A_n\}_n$  in the sense of Definition 2.40, then  $\{\{B_{n,m}\}_n\}_m$  is an a.c.s. for  $\{A_n\}_n$  (in the sense of the classical Definition 2.31) for all sequences of multi-indices  $\{\mathbf{m} = \mathbf{m}(m)\}_m \subseteq \mathcal{M}$  such that  $\mathbf{m} \rightarrow \infty$  as  $m \rightarrow \infty$ .

REMARK 2.42. Let  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  and  $\{B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A'_n\}_n$  as  $\mathbf{m} \rightarrow \infty$ . The following properties hold.

- $\{B_{n,m}^*\}_n \xrightarrow{\text{a.c.s.}} \{A_n^*\}_n$  as  $\mathbf{m} \rightarrow \infty$ .
- $\{\alpha B_{n,m} + \beta B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{\alpha A_n + \beta A'_n\}_n$  as  $\mathbf{m} \rightarrow \infty$  for all  $\alpha, \beta \in \mathbb{C}$ .
- If  $\{A_n\}_n$  and  $\{A'_n\}_n$  are s.u., then  $\{B_{n,m} B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n A'_n\}_n$ .

The proof of these results is essentially the same as the proof of the analogous results for standard a.c.s.; see Theorem 2.37. Based on the topological results of Section 2.5.1, we can give the following elegant characterization of a.c.s. parameterized by a multi-index  $\mathbf{m} \rightarrow \infty$ .

A multi-index sequence of sequences of matrices  $\{\{B_{n,m}\}_n\}_{m \in \mathcal{M}}$  is an a.c.s. for  $\{A_n\}_n$  as  $\mathbf{m} \rightarrow \infty$  if and only if  $d_{\text{a.c.s.}}(\{A_n\}_n, \{B_{n,m}\}_n) \rightarrow 0$  as  $\mathbf{m} \rightarrow \infty$ .

Throughout this paper, we write “ $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  as  $\mathbf{m} \rightarrow \infty$ ” to indicate that  $\{\{B_{n,m}\}_n\}_{m \in \mathcal{M}}$  is an a.c.s. for  $\{A_n\}_n$  as  $\mathbf{m} \rightarrow \infty$ .

**2.6. Multilevel block Toeplitz matrices.** Given  $\mathbf{n} \in \mathbb{N}^d$ , a matrix of the form

$$[A_{i-j}]_{i,j=1}^{\mathbf{n}} \in \mathbb{C}^{sN(\mathbf{n}) \times sN(\mathbf{n})},$$

with blocks  $A_k \in \mathbb{C}^{s \times s}$  for  $k = -(\mathbf{n} - 1), \dots, \mathbf{n} - 1$ , is called a  $d$ -level  $s$ -block Toeplitz matrix. If  $s = 1$  (resp.,  $d = 1$ ), it is simply referred to as a  $d$ -level (resp.,  $s$ -block) Toeplitz matrix. Given a function  $f \in L^1([-\pi, \pi]^d, s)$ , its Fourier coefficients are denoted by

$$f_{\mathbf{k}} = \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} f(\boldsymbol{\theta}) e^{-i\mathbf{k} \cdot \boldsymbol{\theta}} d\boldsymbol{\theta} \in \mathbb{C}^{s \times s}, \quad \mathbf{k} \in \mathbb{Z}^d,$$

where the integrals are computed componentwise. The  $\mathbf{n}$ th ( $d$ -level  $s$ -block) Toeplitz matrix associated with  $f$  is defined as

$$T_{\mathbf{n}}(f) = [f_{i-j}]_{i,j=1}^{\mathbf{n}} \in \mathbb{C}^{sN(\mathbf{n}) \times sN(\mathbf{n})}.$$

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

For each  $s \in \mathbb{N}$  and  $\mathbf{n} \in \mathbb{N}^d$ , the map  $T_{\mathbf{n}}(\cdot) : L^1([-\pi, \pi]^d, s) \rightarrow \mathbb{C}^{sN(\mathbf{n}) \times sN(\mathbf{n})}$  is linear, i.e.,

$$(2.21) \quad T_{\mathbf{n}}(\alpha f + \beta g) = \alpha T_{\mathbf{n}}(f) + \beta T_{\mathbf{n}}(g), \quad \alpha, \beta \in \mathbb{C}, \quad f, g \in L^1([-\pi, \pi]^d, s).$$

Moreover, it is clear from the definition that  $T_{\mathbf{n}}(I_s) = I_{sN(\mathbf{n})}$ . If  $f \in L^1([-\pi, \pi]^d, s)$ , let  $f^*$  be its conjugate transpose. It can be shown that

$$(2.22) \quad T_{\mathbf{n}}(f)^* = T_{\mathbf{n}}(f^*), \quad f \in L^1([-\pi, \pi]^d, s), \quad s \in \mathbb{N}, \quad \mathbf{n} \in \mathbb{N}^d.$$

In particular, if  $f$  is Hermitian a.e., then the matrices  $T_n(f)$  are Hermitian.

Theorem 2.43 is a fundamental result about multilevel block Toeplitz matrices. It provides the singular value distribution of multilevel block Toeplitz sequences generated by a matrix-valued function  $f \in L^1([-\pi, \pi]^d, s)$  and the spectral distribution of multilevel block Toeplitz sequences generated by a Hermitian matrix-valued function  $f \in L^1([-\pi, \pi]^d, s)$ . For the eigenvalues it goes back to Szegő [49], and for the singular values it was established by Avram [2] and Parter [53]. They assumed  $d = 1$ ,  $s = 1$ , and  $f \in L^\infty([-\pi, \pi]^d, s)$ ; see [19, Section 5] and [20, Section 10.14] for more on the subject in the case of  $L^\infty$  generating functions. The extension to  $d \geq 1$ ,  $s = 1$ , and  $f \in L^1([-\pi, \pi]^d, s)$  was performed by Tyrtshnikov and Zamarashkin [69, 70, 71], and the final generalization to  $f \in L^1([-\pi, \pi]^d, s)$  with arbitrary  $s, d \geq 1$  is due to Tilli [67]. We also refer the reader to [40, Section 6.5] for a proof of Theorem 2.43 based on the notion of approximating classes of sequences; the proof in [40, Section 6.5] is made only for  $d = s = 1$ , but the argument is general and can be extended to the case  $d, s \geq 1$ . Note that the extension to the case  $s = 1$  and  $d \geq 1$  was performed in [41, Section 3.5].

**THEOREM 2.43.** *If  $f \in L^1([-\pi, \pi]^d, s)$  and  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  is any sequence such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ , then  $\{T_n(f)\}_n \sim_\sigma f$ . If moreover  $f$  is Hermitian a.e., then  $\{T_n(f)\}_n \sim_\lambda f$ .*

Important inequalities involving Toeplitz matrices and Schatten  $p$ -norms originally appeared in [64, Corollary 4.2]. They have been generalized to multilevel block Toeplitz matrices in [61, Corollary 3.5]. We report them in the next theorem for future use.

**THEOREM 2.44.** *Let  $f \in L^p([-\pi, \pi]^d, s)$  and  $\mathbf{n} \in \mathbb{N}^d$ , then, using the natural convention  $1/\infty = 0$ , the inequality*

$$\|T_n(f)\|_p \leq \left( \frac{N(\mathbf{n})}{(2\pi)^d} \right)^{1/p} \|f\|_{L^p}$$

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

The next result provides an important relation between tensor products and multilevel block Toeplitz matrices. Observe that, if  $f_i \in L^1([-\pi, \pi]^d, s_i)$ ,  $i = 1, \dots, d$ , and  $\mathbf{s} = (s_1, \dots, s_d)$ , then  $f_1 \otimes \dots \otimes f_d \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$  by Fubini's theorem.

**LEMMA 2.45.** *For every  $\mathbf{n}, \mathbf{s} \in \mathbb{N}^d$ , we have*

$$T_n(f_1 \otimes \dots \otimes f_d) = \Gamma_{\mathbf{n}, \mathbf{s}}^T [T_{n_1}(f_1) \otimes \dots \otimes T_{n_d}(f_d)] \Gamma_{\mathbf{n}, \mathbf{s}}$$

for all functions  $f_1 \in L^1([-\pi, \pi], s_1), \dots, f_d \in L^1([-\pi, \pi], s_d)$ , where  $\Gamma_{\mathbf{n}, \mathbf{s}}$  is defined in Remark 2.3.

*Proof.* The result has already been proved in the case where  $f_1, \dots, f_d$  are matrix-valued trigonometric polynomials; see [44, Lemma 4]. For general functions  $f_1 \in L^1([-\pi, \pi], s_1), \dots, f_d \in L^1([-\pi, \pi], s_d)$ , simply take an  $s_i \times s_i$  matrix-valued trigonometric polynomial  $p_i$  such that  $T_{n_i}(f_i) = T_{n_i}(p_i)$ , and observe that  $T_n(p_1 \otimes \dots \otimes p_d) = T_n(f_1 \otimes \dots \otimes f_d)$ ; for example, one can take  $p_i(\theta) = \sum_{j=-(n_i-1)}^{n_i-1} (f_i)_j e^{ij\theta}$ .  $\square$

The next theorem shows that the product of multilevel block Toeplitz matrices generated by functions in  $L^\infty([-\pi, \pi]^d, s)$  is “close” to the multilevel block Toeplitz matrix generated by the product of the generating functions.

**THEOREM 2.46.** *Let  $f_i \in L^\infty([-\pi, \pi]^d, s)$  for  $i = 1, \dots, q$ . Then,*

$$\lim_{\mathbf{n} \rightarrow \infty} \frac{\|\prod_{i=1}^q T_n(f_i) - T_n(\prod_{i=1}^q f_i)\|_1}{N(\mathbf{n})} = 0.$$

*Proof.* For  $q = 2$  the result is proved in [29, Proposition 2]. In the general case we proceed by induction. Fix  $p \geq 3$  and suppose that the result holds for  $q = p - 1$ . If  $q = p$ , using (2.5) and Theorem 2.44, we obtain

$$\begin{aligned}
 & \frac{1}{N(\mathbf{n})} \left\| \prod_{i=1}^p T_{\mathbf{n}}(f_i) - T_{\mathbf{n}} \left( \prod_{i=1}^p f_i \right) \right\|_1 \\
 &= \frac{1}{N(\mathbf{n})} \left\| \prod_{i=1}^p T_{\mathbf{n}}(f_i) - \left( \prod_{i=1}^{p-2} T_{\mathbf{n}}(f_i) \right) T_{\mathbf{n}}(f_{p-1}f_p) \right. \\
 & \quad \left. + \left( \prod_{i=1}^{p-2} T_{\mathbf{n}}(f_i) \right) T_{\mathbf{n}}(f_{p-1}f_p) - T_{\mathbf{n}} \left( \prod_{i=1}^p f_i \right) \right\|_1 \\
 &\leq \frac{1}{N(\mathbf{n})} \left\| \left( \prod_{i=1}^{p-2} T_{\mathbf{n}}(f_i) \right) (T_{\mathbf{n}}(f_{p-1})T_{\mathbf{n}}(f_p) - T_{\mathbf{n}}(f_{p-1}f_p)) \right\|_1 \\
 & \quad + \frac{1}{N(\mathbf{n})} \left\| \left( \prod_{i=1}^{p-2} T_{\mathbf{n}}(f_i) \right) T_{\mathbf{n}}(f_{p-1}f_p) - T_{\mathbf{n}} \left( \prod_{i=1}^p f_i \right) \right\|_1 \\
 &\leq \frac{1}{N(\mathbf{n})} \left( \prod_{i=1}^{p-2} \|f_i\|_{L^\infty} \right) \|T_{\mathbf{n}}(f_{p-1})T_{\mathbf{n}}(f_p) - T_{\mathbf{n}}(f_{p-1}f_p)\|_1 \\
 & \quad + \frac{1}{N(\mathbf{n})} \left\| \left( \prod_{i=1}^{p-2} T_{\mathbf{n}}(f_i) \right) T_{\mathbf{n}}(f_{p-1}f_p) - T_{\mathbf{n}} \left( \left( \prod_{i=1}^{p-2} f_i \right) (f_{p-1}f_p) \right) \right\|_1.
 \end{aligned}$$

Now, the first term on the right-hand side tends to zero as  $\mathbf{n} \rightarrow \infty$  by [29, Proposition 2], and the second term tends to zero as  $\mathbf{n} \rightarrow \infty$  by the induction hypothesis.  $\square$

We end this section with a result highlighting the connection between multilevel block Toeplitz matrices and block matrices with multilevel block Toeplitz blocks. It generalizes [39, Lemma 3]. Recall that  $\Pi_{\mathbf{n},s,r}$  denotes the special permutation matrix (2.3).

**THEOREM 2.47.** *Let  $\mathbf{n} \in \mathbb{N}^d$ , let  $f_{ij} : [-\pi, \pi]^d \rightarrow \mathbb{C}^{r \times r}$  be in  $L^1([-\pi, \pi]^d, r)$ , for  $i, j = 1, \dots, s$ , and set  $f = [f_{ij}]_{i,j=1}^s$ . The block matrix  $T_{\mathbf{n}} = [T_{\mathbf{n}}(f_{ij})]_{i,j=1}^s$  is similar via the permutation (2.3) to the multilevel block Toeplitz matrix  $T_{\mathbf{n}}(f)$ , that is,  $\Pi_{\mathbf{n},s,r} T_{\mathbf{n}} \Pi_{\mathbf{n},s,r}^T = T_{\mathbf{n}}(f)$ .*

*Proof.* Since  $T_{\mathbf{n}} = \sum_{i,j=1}^s E_{ij}^{(s)} \otimes T_{\mathbf{n}}(f_{ij})$  and  $T_{\mathbf{n}}(f) = \sum_{i,j=1}^s T_{\mathbf{n}}(E_{ij}^{(s)} \otimes f_{ij})$  by the linearity of the map  $T_{\mathbf{n}}(\cdot)$ , it is enough to show that

$$\Pi_{\mathbf{n},s,r}(E \otimes T_{\mathbf{n}}(g))\Pi_{\mathbf{n},s,r}^T = T_{\mathbf{n}}(E \otimes g), \quad \forall g \in L^1([-\pi, \pi]^d, r), \quad \forall E \in \mathbb{C}^{s \times s}.$$

By properties of tensor products (see Section 2.2.2),

$$\begin{aligned}
 & \Pi_{\mathbf{n},s,r}(E \otimes T_{\mathbf{n}}(g))\Pi_{\mathbf{n},s,r}^T \\
 &= \left[ \sum_{\mathbf{k}=1}^{\mathbf{n}} \mathbf{e}_{\mathbf{k}} \otimes I_s \otimes \mathbf{e}_{\mathbf{k}}^T \otimes I_r \right] (E \otimes T_{\mathbf{n}}(g)) \left[ \sum_{\mathbf{\ell}=1}^{\mathbf{n}} \mathbf{e}_{\mathbf{\ell}}^T \otimes I_s \otimes \mathbf{e}_{\mathbf{\ell}} \otimes I_r \right] \\
 &= \sum_{\mathbf{k}, \mathbf{\ell}=1}^{\mathbf{n}} (\mathbf{e}_{\mathbf{k}} \otimes I_s \otimes \mathbf{e}_{\mathbf{k}}^T \otimes I_r) (E \otimes T_{\mathbf{n}}(g)) (\mathbf{e}_{\mathbf{\ell}}^T \otimes I_s \otimes \mathbf{e}_{\mathbf{\ell}} \otimes I_r) \\
 &= \sum_{\mathbf{k}, \mathbf{\ell}=1}^{\mathbf{n}} \mathbf{e}_{\mathbf{k}} \mathbf{e}_{\mathbf{\ell}}^T \otimes E \otimes (\mathbf{e}_{\mathbf{k}}^T \otimes I_r) T_{\mathbf{n}}(g) (\mathbf{e}_{\mathbf{\ell}} \otimes I_r)
 \end{aligned}$$



$$= \sum_{k,\ell=1}^n \mathbf{e}_k \mathbf{e}_\ell^T \otimes E \otimes g_{k-\ell} = \sum_{k,\ell=1}^n \mathbf{e}_k \mathbf{e}_\ell^T \otimes (E \otimes g)_{k-\ell} = T_n(E \otimes g),$$

as required.  $\square$

**2.7. Multilevel block diagonal sampling matrices.** If  $\mathbf{n} \in \mathbb{N}^d$  and  $a : [0, 1]^d \rightarrow \mathbb{C}^{r \times r}$ , then we define the  $\mathbf{n}$ th ( $d$ -level  $r$ -block) diagonal sampling matrix generated by  $a$  as the following multilevel block diagonal matrix of size  $N(\mathbf{n})r$ :

$$D_{\mathbf{n}}(a) = \text{diag}_{i=1, \dots, n} a\left(\frac{i}{\mathbf{n}}\right),$$

where we recall that  $i$  varies from  $\mathbf{1}$  to  $\mathbf{n}$  following the lexicographic ordering. Note that  $D_{\mathbf{n}}(a)$  can also be defined through a recursive formula: if  $d = 1$ , then

$$D_{\mathbf{n}}(a) = \text{diag}_{i=1, \dots, n} a\left(\frac{i}{n}\right);$$

if  $d > 1$ , then

$$(2.23) \quad D_{\mathbf{n}}(a) = D_{n_1, \dots, n_d}(a) = \text{diag}_{i_1=1, \dots, n_1} D_{n_2, \dots, n_d}\left(a\left(\frac{i_1}{n_1}, x_2, \dots, x_d\right)\right),$$

where  $a(i_1/n_1, x_2, \dots, x_d)$  is the  $(d-1)$ -variate function defined as follows:

$$a\left(\frac{i_1}{n_1}, x_2, \dots, x_d\right) : [0, 1]^{d-1} \rightarrow \mathbb{C}^{r \times r}, \quad (x_2, \dots, x_d) \mapsto a\left(\frac{i_1}{n_1}, x_2, \dots, x_d\right).$$

If  $(\mathbb{C}^{r \times r})^{[0, 1]^d}$  denotes the space of all functions  $a : [0, 1]^d \rightarrow \mathbb{C}^{r \times r}$ , then the mapping  $D_{\mathbf{n}}(\cdot) : (\mathbb{C}^{r \times r})^{[0, 1]^d} \rightarrow \mathbb{C}^{rN(\mathbf{n}) \times rN(\mathbf{n})}$  is linear, i.e.,

$$D_{\mathbf{n}}(\alpha a + \beta b) = \alpha D_{\mathbf{n}}(a) + \beta D_{\mathbf{n}}(b), \quad \alpha, \beta \in \mathbb{C}, \quad a, b \in (\mathbb{C}^{r \times r})^{[0, 1]^d}.$$

Moreover, it is clear from the definition that  $D_{\mathbf{n}}(E) = T_{\mathbf{n}}(E)$  for all constant matrices  $E \in \mathbb{C}^{r \times r}$  and  $D_{\mathbf{n}}(a)^* = D_{\mathbf{n}}(a^*)$  for all  $a \in (\mathbb{C}^{r \times r})^{[0, 1]^d}$ . The next result, which is the version of Theorem 2.47 for multilevel block diagonal sampling matrices, highlights the connection between multilevel block diagonal sampling matrices and block matrices with multilevel block diagonal sampling blocks. It is a generalization of [39, Lemma 4].

**THEOREM 2.48.** *Let  $\mathbf{n} \in \mathbb{N}^d$ , let  $a_{ij} : [0, 1]^d \rightarrow \mathbb{C}^{r \times r}$ , for  $i, j = 1, \dots, s$ , and set  $a = [a_{ij}]_{i,j=1}^s$ . The block matrix  $D_{\mathbf{n}} = [D_{\mathbf{n}}(a_{ij})]_{i,j=1}^s$  is similar via the permutation (2.3) to the multilevel block diagonal sampling matrix  $D_{\mathbf{n}}(a)$ , that is,  $\Pi_{\mathbf{n}, s, r} D_{\mathbf{n}} \Pi_{\mathbf{n}, s, r}^T = D_{\mathbf{n}}(a)$ .*

*Proof.* With obvious adaptations, it is the same as the proof of Theorem 2.47.  $\square$

For  $n, s \in \mathbb{N}$ , we denote by  $\{x_{i,s}^{(n)}\}_{i=1}^{ns} = \{x_{1,s}^{(n)}, \dots, x_{ns,s}^{(n)}\}$  the sequence of points

$$\{x_{1,s}^{(n)}, \dots, x_{ns,s}^{(n)}\} = \left\{ \underbrace{\frac{1}{n}, \dots, \frac{1}{n}}_s, \underbrace{\frac{2}{n}, \dots, \frac{2}{n}}_s, \dots, \underbrace{\frac{n-1}{n}, \dots, \frac{n-1}{n}}_s, \underbrace{1, \dots, 1}_s \right\}.$$

In formulas,

$$x_{i,s}^{(n)} = \left( \left\lfloor \frac{i-1}{s} \right\rfloor + 1 \right) \frac{1}{n}, \quad i = 1, \dots, ns.$$

Similarly, for  $\mathbf{n}, \mathbf{s} \in \mathbb{N}^d$ , we denote by  $\{\mathbf{x}_{i,\mathbf{s}}^{(\mathbf{n})}\}_{i=1}^{n\mathbf{s}} = \{\mathbf{x}_{1,\mathbf{s}}^{(\mathbf{n})}, \dots, \mathbf{x}_{n\mathbf{s},\mathbf{s}}^{(\mathbf{n})}\}$  the sequence of points

$$\mathbf{x}_{i,\mathbf{s}}^{(\mathbf{n})} = (x_{i_1,s_1}^{(n_1)}, \dots, x_{i_d,s_d}^{(n_d)}), \quad i = 1, \dots, n\mathbf{s}.$$

For  $\mathbf{n}, \mathbf{s} \in \mathbb{N}^d$  and  $a : [0, 1]^d \rightarrow \mathbb{C}$ , we denote by  $D_{\mathbf{n},\mathbf{s}}(a)$  the  $d$ -level diagonal sampling matrix given by

$$D_{\mathbf{n},\mathbf{s}}(a) = \operatorname{diag}_{i=1,\dots,n\mathbf{s}} a(\mathbf{x}_{i,\mathbf{s}}^{(\mathbf{n})}).$$

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

$$(2.24) \quad D_{\mathbf{n},\mathbf{s}}(a) = \operatorname{diag}_{i=1,\dots,n\mathbf{s}} a(x_{i,\mathbf{s}}^{(\mathbf{n})}) = D_{\mathbf{n}}(aI_{\mathbf{s}}),$$

as defined in [8, Section 2.7]; if  $d > 1$ , then

$$(2.25) \quad D_{\mathbf{n},\mathbf{s}}(a) = D_{n_1,\dots,n_d,s_1,\dots,s_d}(a) = \operatorname{diag}_{i_1=1,\dots,n_1s_1} D_{n_2,\dots,n_d,s_2,\dots,s_d}(a(x_{i_1,s_1}^{(n_1)}, x_2, \dots, x_d)),$$

where  $a(x_{i_1,s_1}^{(n_1)}, x_2, \dots, x_d)$  is the  $(d-1)$ -variate function defined as follows:

$$a(x_{i_1,s_1}^{(n_1)}, x_2, \dots, x_d) : [0, 1]^{d-1} \rightarrow \mathbb{C}, \quad (x_2, \dots, x_d) \mapsto a(x_{i_1,s_1}^{(n_1)}, x_2, \dots, x_d).$$

If  $\mathbb{C}^{[0,1]^d}$  denotes the space of all functions  $a : [0, 1]^d \rightarrow \mathbb{C}$ , it is not difficult to see that the operator  $D_{\mathbf{n},\mathbf{s}}(\cdot) : \mathbb{C}^{[0,1]^d} \rightarrow \mathbb{C}^{N(\mathbf{n}) \times N(\mathbf{s})}$  enjoys the following properties: for every  $\mathbf{n}, \mathbf{s} \in \mathbb{N}^d$ , every  $a, b : [0, 1]^d \rightarrow \mathbb{C}$ , and every  $\alpha, \beta \in \mathbb{C}$ ,

$$(2.26) \quad D_{\mathbf{n},\mathbf{s}}(\bar{a}) = D_{\mathbf{n},\mathbf{s}}(a)^*,$$

$$(2.27) \quad D_{\mathbf{n},\mathbf{s}}(ab) = D_{\mathbf{n},\mathbf{s}}(a)D_{\mathbf{n},\mathbf{s}}(b),$$

$$(2.28) \quad D_{\mathbf{n},\mathbf{s}}(\alpha a + \beta b) = \alpha D_{\mathbf{n},\mathbf{s}}(a) + \beta D_{\mathbf{n},\mathbf{s}}(b).$$

LEMMA 2.49. *Let  $\mathbf{n}, \mathbf{s} \in \mathbb{N}^d$ , and let  $a : [0, 1]^d \rightarrow \mathbb{C}$ . Then*

$$(2.29) \quad D_{\mathbf{n},\mathbf{s}}(a) = \Gamma_{\mathbf{n},\mathbf{s}} D_{\mathbf{n}}(aI_{N(\mathbf{s})}) \Gamma_{\mathbf{n},\mathbf{s}}^T,$$

where  $\Gamma_{\mathbf{n},\mathbf{s}}$  is the permutation matrix defined in Remark 2.3.

*Proof.* Suppose first that  $a$  is a separable function, i.e.,  $a(\mathbf{x}) = a_1(x_1) \cdots a_d(x_d)$  with  $a_1, \dots, a_d : [0, 1] \rightarrow \mathbb{C}$ . Then, by definition of  $\Gamma_{\mathbf{n},\mathbf{s}}$  (see Remark 2.3),

$$\begin{aligned} D_{\mathbf{n},\mathbf{s}}(a) &= \operatorname{diag}_{i=1,\dots,n\mathbf{s}} a(\mathbf{x}_{i,\mathbf{s}}^{(\mathbf{n})}) = \operatorname{diag}_{i=1,\dots,n\mathbf{s}} a_1(x_{i_1,s_1}^{(n_1)}) \cdots a_d(x_{i_d,s_d}^{(n_d)}) \\ &= D_{n_1,s_1}(a_1) \otimes \cdots \otimes D_{n_d,s_d}(a_d) = D_{n_1}(a_1 I_{s_1}) \otimes \cdots \otimes D_{n_d}(a_d I_{s_d}) \\ &= D_{n_1}(a_1) \otimes I_{s_1} \otimes \cdots \otimes D_{n_d}(a_d) \otimes I_{s_d} \\ &= \Gamma_{\mathbf{n},\mathbf{s}} [D_{n_1}(a_1) \otimes \cdots \otimes D_{n_d}(a_d) \otimes I_{s_1} \otimes \cdots \otimes I_{s_d}] \Gamma_{\mathbf{n},\mathbf{s}}^T \\ &= \Gamma_{\mathbf{n},\mathbf{s}} [D_{\mathbf{n}}(a) \otimes I_{N(\mathbf{s})}] \Gamma_{\mathbf{n},\mathbf{s}}^T = \Gamma_{\mathbf{n},\mathbf{s}} D_{\mathbf{n}}(aI_{N(\mathbf{s})}) \Gamma_{\mathbf{n},\mathbf{s}}^T, \end{aligned}$$

which completes the proof of equation (2.29) in the case where  $a$  is separable. By the properties (2.26)–(2.28), equation (2.29) continues to hold if  $a$  belongs to the  $*$ -subalgebra of  $\mathbb{C}^{[0,1]^d}$

generated by separable functions; in particular, it continues to hold if  $a$  is a linear combinations of separable functions.

Suppose now that  $a : [0, 1]^d \rightarrow \mathbb{C}$  is an arbitrary function. We show that there exists a linear combination  $b$  of separable functions such that  $b(\mathbf{x}_{i,s}^{(n)}) = a(\mathbf{x}_{i,s}^{(n)})$  for all  $i = 1, \dots, ns$ . Once this is done, equation (2.29) is proved because, by the first part of the proof,

$$D_{n,s}(a) = D_{n,s}(b) = \Gamma_{n,s} D_n(b I_{N(s)}) \Gamma_{n,s}^T = \Gamma_{n,s} D_n(a I_{N(s)}) \Gamma_{n,s}^T.$$

To prove the existence of a function  $b$  with the required properties, it suffices to find a linear combination  $b$  of separable functions such that

$$b(i/n) = a(i/n), \quad i = 1, \dots, n,$$

because the set  $\{\mathbf{x}_{i,s}^{(n)} : i = 1, \dots, ns\}$  coincides with  $\{i/n : i = 1, \dots, n\}$  regardless of  $s$ . In the following, we provide an explicit construction of a function  $b$  satisfying the required properties. For  $i = 1, \dots, n$ , let  $b_{i,n} : [0, 1] \rightarrow \mathbb{C}$  be any function such that  $b_{i,n}(j/n) = 1$  if  $j = i$  and  $b_{i,n}(j/n) = 0$  if  $j \neq i$ . For  $i = 1, \dots, n$ , define  $b_{i,n} : [0, 1]^d \rightarrow \mathbb{C}$  by setting  $b_{i,n}(\mathbf{x}) = b_{i_1, n_1}(x_1) \cdots b_{i_d, n_d}(x_d)$ , and note that  $b_{i,n}(j/n) = 1$  if  $j = i$  and  $b_{i,n}(j/n) = 0$  if  $j \neq i$ . Finally, define

$$b(\mathbf{x}) = \sum_{i=1}^n a(i/n) b_{i,n}(\mathbf{x}),$$

and note that  $b$  satisfies all the required properties.  $\square$

Since  $\Gamma_{n,s} = I_{ns}$ , formula (2.29) is a generalization of (2.24) to the  $d$ -level case.

We conclude this section by introducing a notation that we shall use later on. If  $\mathbf{n} \in \mathbb{N}^d$  and  $a : [0, 1]^d \rightarrow \mathbb{C}$ , the  $\mathbf{n}$ th ( $d$ -level) arrow-shaped sampling matrix generated by  $a$  is the symmetric  $N(\mathbf{n}) \times N(\mathbf{n})$  matrix denoted by  $S_n(a)$  and defined as follows:

$$S_n(a) = [(D_n(a))_{i \wedge j, i \wedge j}]_{i,j=1}^n = \left[ a \left( \frac{i \wedge j}{\mathbf{n}} \right) \right]_{i,j=1}^n.$$

**3. Multilevel block locally Toeplitz sequences.** The theory of (scalar) LT sequences dates back to Tilli's pioneering paper [66]. It was then carried forward in [62, 63], and it was finally developed in a systematic way in [40, Chapter 7] and [41, Chapter 4]. The theory of block LT sequences was originally suggested in [63, Section 3.3], carried forward in [45], and developed in a systematic way in [8, Chapter 3]. In this chapter, we address the multidimensional version of the theory of block LT sequences, also known as the theory of multilevel block LT sequences. The topic is presented here on an abstract level, whereas for motivations and insights we refer the reader to the introduction of [8]; see also the introduction of Tilli's paper [66] and Section 7.1 of [40].

**3.1. Multilevel block LT operators.** Just as the theory of block LT sequences begins with the notion of block LT operators, the theory of multilevel block LT sequences begins with the notion of multilevel block LT operators.

**3.1.1. Definition of multilevel block LT operators.**

DEFINITION 3.1 (Multilevel block locally Toeplitz operator).

- Let  $m, n, s \in \mathbb{N}$ , let  $a : [0, 1] \rightarrow \mathbb{C}$ , and let  $f \in L^1([-\pi, \pi], s)$ . The (unilevel) block locally Toeplitz (LT) operator is defined as the following  $ns \times ns$  matrix:

$$\begin{aligned}
 LT_{n,s}^m(a, f) &= D_m(a) \otimes T_{\lfloor n/m \rfloor}(f) \oplus O_{s(n \bmod m)} \\
 &= \text{diag}_{i=1, \dots, m} \left[ a\left(\frac{i}{m}\right) T_{\lfloor n/m \rfloor}(f) \right] \oplus O_{s(n \bmod m)} \\
 &= \text{diag}_{i=1, \dots, m} a\left(\frac{i}{m}\right) T_{\lfloor n/m \rfloor}(f) \oplus O_{s(n \bmod m)}.
 \end{aligned}$$

It is understood that  $LT_{n,s}^m(a, f) = O_{sn}$  if  $n < m$  and that the term  $O_{s(n \bmod m)}$  is not present if  $n$  is a multiple of  $m$ . Moreover, here and in what follows, 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. Note also that in the last equality we intentionally removed the square brackets in order to illustrate a notation that will be used hereinafter to simplify the presentation (roughly speaking, we are assuming that the “diag operator” is applied before the direct sum  $\oplus$ ).

- Let  $\mathbf{m}, \mathbf{n}, \mathbf{s} \in \mathbb{N}^d$ , let  $a : [0, 1]^d \rightarrow \mathbb{C}$ , and let  $f_j \in L^1([-\pi, \pi], s_j)$  for  $j = 1, \dots, d$ . The multilevel block locally Toeplitz (LT) operator is defined as the following  $N(\mathbf{ns}) \times N(\mathbf{ns})$  matrix:

$$\begin{aligned}
 &LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(a, f_1 \otimes \dots \otimes f_d) \\
 &= LT_{n_1, \dots, n_d, s_1, \dots, s_d}^{m_1, \dots, m_d}(a(x_1, \dots, x_d), f_1 \otimes \dots \otimes f_d) \\
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{i_1=1, \dots, m_1} T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} LT_{n_2, \dots, n_d, s_2, \dots, s_d}^{m_2, \dots, m_d} \left( a\left(\frac{i_1}{m_1}, x_2, \dots, x_d\right), f_2 \otimes \dots \otimes f_d \right) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T \right. \\
 &\quad \left. \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \dots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}}.
 \end{aligned}$$

This is a recursive definition, whose base case has been given in the previous item. For example, in the case  $d = 2$  we have

$$\begin{aligned}
 &LT_{n_1, n_2, s_1, s_2}^{m_1, m_2}(a, f_1 \otimes f_2) \\
 &= \Gamma_{n_1, n_2, s_1, s_2}^T \left[ \text{diag}_{i_1=1, \dots, m_1} T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes \left[ \text{diag}_{i_2=1, \dots, m_2} a\left(\frac{i_1}{m_1}, \frac{i_2}{m_2}\right) T_{\lfloor n_2/m_2 \rfloor}(f_2) \oplus O_{s_2(n_2 \bmod m_2)} \right] \right. \\
 &\quad \left. \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2} \right] \Gamma_{n_1, n_2, s_1, s_2}.
 \end{aligned}$$

where we have used the fact that  $\Gamma_{n_2, s_2} = I_{n_2 s_2}$ ; see Remark 2.3.

In Definition 3.1, we have defined the multilevel block LT operator  $LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(a, f)$  in the case where  $f$  is a separable function of the form  $f = f_1 \otimes \dots \otimes f_d$  with  $f_j \in L^1([-\pi, \pi], s_j)$  for  $j = 1, \dots, d$ . We are going to see in Definition 3.3 that  $LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(a, f)$  is actually well-defined (in a unique way) for any  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . The crucial result in view of Definition 3.3 is Theorem 3.2. It shows that the multilevel block LT operator  $LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(a, f_1 \otimes \dots \otimes f_d)$

coincides with  $D_m(a) \otimes T_{\lfloor n/m \rfloor}(f_1 \otimes \cdots \otimes f_d) \oplus O$  up to a permutation transformation  $\Pi_{n,m}^{(s)}$  depending only on  $m, n, s$  and not on the functions  $a, f_1, \dots, f_d$ .

**THEOREM 3.2.** *For every  $m, n, s \in \mathbb{N}^d$  there exists a permutation matrix  $\Pi_{n,m}^{(s)}$  such that*

$$\begin{aligned} & LT_{n,s}^m(a, f_1 \otimes \cdots \otimes f_d) \\ &= \Pi_{n,m}^{(s)} \left[ D_m(a) \otimes T_{\lfloor n/m \rfloor}(f_1 \otimes \cdots \otimes f_d) \oplus O_{N(ns) - N(m \lfloor n/m \rfloor s)} \right] (\Pi_{n,m}^{(s)})^T \end{aligned}$$

for every  $a : [0, 1]^d \rightarrow \mathbb{C}$  and every  $f_1 \in L^1([-\pi, \pi], s_1), \dots, f_d \in L^1([-\pi, \pi], s_d)$ .

*Proof.* The proof is done by induction on  $d$ . For  $d = 1$ , the result holds with  $\Pi_{n,m}^{(s)} = I_{ns}$ . For  $d \geq 2$ , set  $\nu = (n_2, \dots, n_d)$ ,  $\mu = (m_2, \dots, m_d)$ ,  $\sigma = (s_2, \dots, s_d)$ . By definition,

$$\begin{aligned} & LT_{n,s}^m(a, f_1 \otimes \cdots \otimes f_d) = \\ & \Gamma_{n,s}^T \left[ \text{diag}_{i_1=1, \dots, m_1} T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes \Gamma_{\nu,\sigma} LT_{\nu,\sigma}^\mu \left( a \left( \frac{i_1}{m_1}, \cdot \right), f_2 \otimes \cdots \otimes f_d \right) \Gamma_{\nu,\sigma}^T \right. \\ & \left. \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \cdots s_d n_d} \right] \Gamma_{n,s}, \end{aligned} \quad (3.1)$$

where  $a(i_1/m_1, \cdot) : [0, 1]^{d-1} \rightarrow \mathbb{C}$  is the function  $(x_2, \dots, x_d) \mapsto a(i_1/m_1, x_2, \dots, x_d)$ . By induction hypothesis, setting  $N(\nu, \mu, \sigma) = N(\nu\sigma) - N(\mu \lfloor \nu/\mu \rfloor \sigma)$ , we have

$$\begin{aligned} & LT_{\nu,\sigma}^\mu \left( a \left( \frac{i_1}{m_1}, \cdot \right), f_2 \otimes \cdots \otimes f_d \right) \\ &= \Pi_{\nu,\mu}^{(\sigma)} \left[ D_\mu \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor \nu/\mu \rfloor}(f_2 \otimes \cdots \otimes f_d) \oplus O_{N(\nu,\mu,\sigma)} \right] (\Pi_{\nu,\mu}^{(\sigma)})^T. \end{aligned} \quad (3.2)$$

Let us work on the argument of the “diag operator” in (3.1). From Lemma 2.2, equation (3.2), and properties of tensor products (see Section 2.2.2), we get

$$\begin{aligned} & (3.3) \\ & T_{\lfloor n_1/m_1 \rfloor}(f_1) \otimes \Gamma_{\nu,\sigma} LT_{\nu,\sigma}^\mu \left( a \left( \frac{i_1}{m_1}, \cdot \right), f_2 \otimes \cdots \otimes f_d \right) \Gamma_{\nu,\sigma}^T \\ &= \Pi_{(N(\nu\sigma), \lfloor n_1/m_1 \rfloor s_1); [2,1]} \left\{ \Gamma_{\nu,\sigma} LT_{\nu,\sigma}^\mu \left( a \left( \frac{i_1}{m_1}, \cdot \right), f_2 \otimes \cdots \otimes f_d \right) \Gamma_{\nu,\sigma}^T \otimes T_{\lfloor n_1/m_1 \rfloor}(f_1) \right\} \\ & \quad \cdot (\Pi_{(N(\nu\sigma), \lfloor n_1/m_1 \rfloor s_1); [2,1]})^T \\ &= \Pi_{(N(\nu\sigma), \lfloor n_1/m_1 \rfloor s_1); [2,1]} \\ & \quad \cdot \left\{ \Gamma_{\nu,\sigma} \Pi_{\nu,\mu}^{(\sigma)} \left[ D_\mu \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor \nu/\mu \rfloor}(f_2 \otimes \cdots \otimes f_d) \oplus O_{N(\nu,\mu,\sigma)} \right] (\Pi_{\nu,\mu}^{(\sigma)})^T \Gamma_{\nu,\sigma}^T \right. \\ & \quad \left. \otimes T_{\lfloor n_1/m_1 \rfloor}(f_1) \right\} (\Pi_{(N(\nu\sigma), \lfloor n_1/m_1 \rfloor s_1); [2,1]})^T \\ &= \Pi_{(N(\nu\sigma), \lfloor n_1/m_1 \rfloor s_1); [2,1]} (\Gamma_{\nu,\sigma} \Pi_{\nu,\mu}^{(\sigma)} \otimes I_{s_1 \lfloor n_1/m_1 \rfloor}) \\ & \quad \cdot \left\{ \left[ D_\mu \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor \nu/\mu \rfloor}(f_2 \otimes \cdots \otimes f_d) \oplus O_{N(\nu,\mu,\sigma)} \right] \otimes T_{\lfloor n_1/m_1 \rfloor}(f_1) \right\} \\ & \quad \cdot (\Gamma_{\nu,\sigma} \Pi_{\nu,\mu}^{(\sigma)} \otimes I_{s_1 \lfloor n_1/m_1 \rfloor})^T (\Pi_{(N(\nu\sigma), \lfloor n_1/m_1 \rfloor s_1); [2,1]})^T. \end{aligned}$$

Using equation (2.10), Lemma 2.2, Lemma 2.45, and properties of tensor products and direct sums (see Section 2.2.2), we obtain

$$\begin{aligned}
 (3.4) \quad & \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor \nu/\mu \rfloor} (f_2 \otimes \cdots \otimes f_d) \oplus O_{N(\nu, \mu, \sigma)} \right] \otimes T_{\lfloor n_1/m_1 \rfloor} (f_1) \\
 &= D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor \nu/\mu \rfloor} (f_2 \otimes \cdots \otimes f_d) \otimes T_{\lfloor n_1/m_1 \rfloor} (f_1) \oplus O_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \\
 &= \Pi_{(N(\mu), \lfloor n_1/m_1 \rfloor s_1, N(\lfloor \nu/\mu \rfloor \sigma)); [1,3,2]} \\
 &\quad \cdot \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor n_1/m_1 \rfloor} (f_1) \otimes T_{\lfloor \nu/\mu \rfloor} (f_2 \otimes \cdots \otimes f_d) \right] \\
 &\quad \cdot (\Pi_{(N(\mu), \lfloor n_1/m_1 \rfloor s_1, N(\lfloor \nu/\mu \rfloor \sigma)); [1,3,2]})^T \oplus O_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \\
 &= \Pi_{(N(\mu), \lfloor n_1/m_1 \rfloor s_1, N(\lfloor \nu/\mu \rfloor \sigma)); [1,3,2]} \\
 &\quad \cdot \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes (I_{\lfloor n_1/m_1 \rfloor s_1} \otimes \Gamma_{\lfloor \nu/\mu \rfloor, \sigma})^T \Gamma_{\lfloor n/m \rfloor, s} \right. \\
 &\quad \cdot \left. T_{\lfloor n/m \rfloor} (f_1 \otimes \cdots \otimes f_d) \Gamma_{\lfloor n/m \rfloor, s}^T (I_{\lfloor n_1/m_1 \rfloor s_1} \otimes \Gamma_{\lfloor \nu/\mu \rfloor, \sigma}) \right] \\
 &\quad \cdot (\Pi_{(N(\mu), \lfloor n_1/m_1 \rfloor s_1, N(\lfloor \nu/\mu \rfloor \sigma)); [1,3,2]})^T \oplus O_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \\
 &= \left[ \Pi_{(N(\mu), \lfloor n_1/m_1 \rfloor s_1, N(\lfloor \nu/\mu \rfloor \sigma)); [1,3,2]} (I_{N(\mu)} \otimes (I_{\lfloor n_1/m_1 \rfloor s_1} \otimes \Gamma_{\lfloor \nu/\mu \rfloor, \sigma})^T \Gamma_{\lfloor n/m \rfloor, s}) \right. \\
 &\quad \left. \oplus I_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \right] \\
 &\quad \cdot \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \cdots \otimes f_d) \oplus O_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \right] \\
 &\quad \cdot \left[ \Pi_{(N(\mu), \lfloor n_1/m_1 \rfloor s_1, N(\lfloor \nu/\mu \rfloor \sigma)); [1,3,2]} (I_{N(\mu)} \otimes (I_{\lfloor n_1/m_1 \rfloor s_1} \otimes \Gamma_{\lfloor \nu/\mu \rfloor, \sigma})^T \Gamma_{\lfloor n/m \rfloor, s}) \right. \\
 &\quad \left. \oplus I_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \right]^T.
 \end{aligned}$$

Substituting (3.4) into (3.3), we arrive at

$$\begin{aligned}
 (3.5) \quad & T_{\lfloor n_1/m_1 \rfloor} (f_1) \otimes \Gamma_{\nu, \sigma} L T_{\nu, \sigma}^{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right), f_2 \otimes \cdots \otimes f_d \right) \Gamma_{\nu, \sigma}^T \\
 &= P_{\mathbf{n}, \mathbf{m}}^{(s)} \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \cdots \otimes f_d) \oplus O_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \right] (P_{\mathbf{n}, \mathbf{m}}^{(s)})^T,
 \end{aligned}$$

where

$$\begin{aligned}
 P_{\mathbf{n}, \mathbf{m}}^{(s)} &= \Pi_{(N(\nu \sigma), \lfloor n_1/m_1 \rfloor s_1); [2,1]} (\Gamma_{\nu, \sigma} \Pi_{\nu, \mu}^{(\sigma)} \otimes I_{s_1 \lfloor n_1/m_1 \rfloor}) \\
 &\quad \cdot \left[ \Pi_{(N(\mu), \lfloor n_1/m_1 \rfloor s_1, N(\lfloor \nu/\mu \rfloor \sigma)); [1,3,2]} \right. \\
 &\quad \left. \cdot (I_{N(\mu)} \otimes (I_{\lfloor n_1/m_1 \rfloor s_1} \otimes \Gamma_{\lfloor \nu/\mu \rfloor, \sigma})^T \Gamma_{\lfloor n/m \rfloor, s}) \oplus I_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \right].
 \end{aligned}$$

Combining (3.5) and (3.1), we obtain

$$\begin{aligned}
 & L T_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}} (a, f_1 \otimes \cdots \otimes f_d) \\
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \left( \bigoplus_{i_1=1}^{m_1} P_{\mathbf{n}, \mathbf{m}}^{(s)} \right)_{i_1=1, \dots, m_1} \text{diag} \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \cdots \otimes f_d) \right. \right. \\
 &\quad \left. \left. \oplus O_{N(\nu, \mu, \sigma)_{\lfloor n_1/m_1 \rfloor s_1}} \right] \left( \bigoplus_{i_1=1}^{m_1} P_{\mathbf{n}, \mathbf{m}}^{(s)} \right)^T \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \cdots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}}.
 \end{aligned}$$

From Lemma 2.4 and equations (2.10) and (2.23),

$$\begin{aligned}
 & \text{diag}_{i_1=1, \dots, m_1} \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \dots \otimes f_d) \oplus O_{N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor s_1} \right] \\
 &= \bigoplus_{i_1=1}^{m_1} \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \dots \otimes f_d) \oplus O_{N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor s_1} \right] \\
 &= V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \left[ \bigoplus_{i_1=1}^{m_1} \left[ D_{\mu} \left( a \left( \frac{i_1}{m_1}, \cdot \right) \right) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \dots \otimes f_d) \right] \right. \\
 &\quad \left. \oplus O_{N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor m_1 s_1} \right] (V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})})^T \\
 &= V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \left[ D_{\mathbf{m}}(a) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \dots \otimes f_d) \oplus O_{N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor m_1 s_1} \right] (V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})})^T,
 \end{aligned}$$

where

$$\begin{aligned}
 V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} &= V_{\mathbf{h}(\mathbf{m}, \mathbf{n}, \mathbf{s}); \sigma}, \\
 \sigma &= [1, m_1 + 1, 2, m_1 + 2, \dots, m_1, 2m_1], \\
 \mathbf{h}(\mathbf{m}, \mathbf{n}, \mathbf{s}) &= \left( \underbrace{N(\mu)N(\lfloor n/m \rfloor s), \dots, N(\mu)N(\lfloor n/m \rfloor s)}_{m_1}, \right. \\
 &\quad \left. \underbrace{N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor s_1, \dots, N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor s_1}_{m_1} \right).
 \end{aligned}$$

Thus,

$$\begin{aligned}
 & LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(a, f_1 \otimes \dots \otimes f_d) \\
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \left( \bigoplus_{i_1=1}^{m_1} P_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \right) V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \right. \\
 &\quad \cdot \left[ D_{\mathbf{m}}(a) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \dots \otimes f_d) \oplus O_{N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor m_1 s_1} \right] \\
 &\quad \left. \cdot (V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})})^T \left( \bigoplus_{i_1=1}^{m_1} P_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \right)^T \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \dots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}} \\
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \left( \bigoplus_{i_1=1}^{m_1} P_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \right) V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \oplus I_{s_1(n_1 \bmod m_1) s_2 n_2 \dots s_d n_d} \right] \\
 &\quad \cdot \left[ D_{\mathbf{m}}(a) \otimes T_{\lfloor n/m \rfloor} (f_1 \otimes \dots \otimes f_d) \oplus O_{N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor m_1 s_1 + s_1(n_1 \bmod m_1) s_2 n_2 \dots s_d n_d} \right] \\
 &\quad \cdot \left[ (V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})})^T \left( \bigoplus_{i_1=1}^{m_1} P_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \right)^T \oplus I_{s_1(n_1 \bmod m_1) s_2 n_2 \dots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}}.
 \end{aligned}$$

This concludes the proof; note that the permutation matrix  $\Pi_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})}$  is given by

$$\Pi_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} = \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \left( \bigoplus_{i_1=1}^{m_1} P_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \right) V_{\mathbf{n}, \mathbf{m}}^{(\mathbf{s})} \oplus I_{s_1(n_1 \bmod m_1) s_2 n_2 \dots s_d n_d} \right],$$

and, moreover, the number  $N(\nu, \mu, \sigma) \lfloor n_1/m_1 \rfloor m_1 s_1 + s_1(n_1 \bmod m_1) s_2 n_2 \dots s_d n_d$  is equal to  $N(\mathbf{n}\mathbf{s}) - N(\mathbf{m} \lfloor \mathbf{n}/\mathbf{m} \rfloor \mathbf{s})$ .  $\square$

**DEFINITION 3.3** (Multilevel block locally Toeplitz operator). *Let  $\mathbf{m}, \mathbf{n}, \mathbf{s} \in \mathbb{N}^d$ , let  $a : [0, 1]^d \rightarrow \mathbb{C}$ , and let  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . The multilevel block locally Toeplitz (LT)*

operator is defined as the following  $N(\mathbf{n}s) \times N(\mathbf{n}s)$  matrix:

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

where  $\Pi_{\mathbf{n},\mathbf{m}}^{(\mathbf{s})}$  is the permutation matrix appearing in Theorem 3.2.

REMARK 3.4. We note that  $LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f) = LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, g)$  whenever  $f = g$  a.e. Moreover, suppose that  $f = f_1 \otimes \cdots \otimes f_d$  a.e., with  $f_j \in L^1([-\pi, \pi], s_j)$  for  $j = 1, \dots, d$ ; then  $LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f)$  is equal to  $LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f_1 \otimes \cdots \otimes f_d)$ , as defined in Definition 3.1. This shows that Definition 3.3 is an extension of Definition 3.1.

**3.1.2. Properties of multilevel block LT operators.** For every  $\mathbf{m}, \mathbf{n}, \mathbf{s} \in \mathbb{N}^d$ , every  $a, b : [0, 1]^d \rightarrow \mathbb{C}$ , every  $f, g \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ , and every  $\alpha, \beta \in \mathbb{C}$ , we have

$$(3.6) \quad LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f)^* = LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(\bar{a}, f^*),$$

$$(3.7) \quad LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(\alpha a + \beta b, f) = \alpha LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f) + \beta LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(b, f),$$

$$(3.8) \quad LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, \alpha f + \beta g) = \alpha LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f) + \beta LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, g),$$

$$(3.9) \quad \begin{aligned} \|LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f)\|_1 &= \|D_{\mathbf{m}}(a)\|_1 \|T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}(f)\|_1 \\ &\leq \frac{1}{(2\pi)^d} \sum_{\mathbf{i}=1}^{\mathbf{m}} \left| a\left(\frac{\mathbf{i}}{\mathbf{m}}\right) \right| \|f\|_{L^1 N(\lfloor \mathbf{n}/\mathbf{m} \rfloor)}, \end{aligned}$$

where in the last inequality we invoked Theorem 2.44.

REMARK 3.5. Let  $\mathbf{s} \in \mathbb{N}^d$ , let  $a : [0, 1]^d \rightarrow \mathbb{C}$  be a bounded function, and take any sequence  $\{f_k\}_k \subset L^1([-\pi, \pi]^d, N(\mathbf{s}))$  such that  $f_k \rightarrow f$  in  $L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . By (3.8) and (3.9), for every  $k$  and every  $\mathbf{n}, \mathbf{m} \in \mathbb{N}^d$ , we have

$$\|LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f) - LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f_k)\|_1 = \|LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f - f_k)\|_1 \leq N(\mathbf{n}) \|a\|_{\infty} \|f - f_k\|_{L^1}.$$

By Theorem 2.39, this implies that  $\{LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f_k)\}_n \xrightarrow{\text{a.c.s.}} \{LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a, f)\}_n$  as  $k \rightarrow \infty$  for every  $\mathbf{m} \in \mathbb{N}^d$  and every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ .

PROPOSITION 3.6. Let  $\mathbf{s} \in \mathbb{N}^d$ , let  $a_i : [0, 1]^d \rightarrow \mathbb{C}$  be bounded, and let  $f_i \in L^{\infty}([-\pi, \pi]^d, N(\mathbf{s}))$  for  $i = 1, \dots, q$ . Then, for every  $\mathbf{n}, \mathbf{m} \in \mathbb{N}^d$ ,

$$(3.10) \quad \left\| \prod_{i=1}^q LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a_i, f_i) - LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}\left(\prod_{i=1}^q a_i, \prod_{i=1}^q f_i\right) \right\|_1 \leq \varepsilon(\lfloor \mathbf{n}/\mathbf{m} \rfloor) N(\mathbf{n}),$$

where

$$\varepsilon(\mathbf{k}) = \left\| \prod_{i=1}^q a_i \right\|_{\infty} \frac{\left\| \prod_{i=1}^q T_{\mathbf{k}}(f_i) - T_{\mathbf{k}}\left(\prod_{i=1}^q f_i\right) \right\|_1}{N(\mathbf{k})}$$

and  $\lim_{\mathbf{k} \rightarrow \infty} \varepsilon(\mathbf{k}) = 0$  by Theorem 2.46. In particular, for every  $\mathbf{m} \in \mathbb{N}^d$  and every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ ,

$$(3.11) \quad d_{\text{a.c.s.}}\left(\left\{\prod_{i=1}^q LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a_i, f_i)\right\}_n, \left\{LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}\left(\prod_{i=1}^q a_i, \prod_{i=1}^q f_i\right)\right\}_n\right) = 0.$$



*Proof.* By Definition 3.3 and the properties of tensor products and direct sums, we have

$$\begin{aligned}
 & \left\| \prod_{i=1}^q LT_{\mathbf{n},\mathbf{s}}^m(a_i, f_i) - LT_{\mathbf{n},\mathbf{s}}^m\left(\prod_{i=1}^q a_i, \prod_{i=1}^q f_i\right) \right\|_1 \\
 &= \left\| D_{\mathbf{m}}\left(\prod_{i=1}^q a_i\right) \otimes \left(\prod_{i=1}^q T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}(f_i) - T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\prod_{i=1}^q f_i\right)\right) \oplus O_{N(\mathbf{n}\mathbf{s})-N(\mathbf{m}\lfloor \mathbf{n}/\mathbf{m} \rfloor \mathbf{s})} \right\|_1 \\
 &= \left\| D_{\mathbf{m}}\left(\prod_{i=1}^q a_i\right) \right\|_1 \left\| \prod_{i=1}^q T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}(f_i) - T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\prod_{i=1}^q f_i\right) \right\|_1 \\
 &\leq N(\mathbf{n}) \left\| \prod_{i=1}^q a_i \right\|_\infty \frac{\left\| \prod_{i=1}^q T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}(f_i) - T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\prod_{i=1}^q f_i\right) \right\|_1}{N(\lfloor \mathbf{n}/\mathbf{m} \rfloor)}.
 \end{aligned}$$

This proves (3.10). Since  $\varepsilon(\mathbf{k}) \rightarrow 0$  as  $\mathbf{k} \rightarrow \infty$  by Theorem 2.46, equation (3.11) follows immediately from (3.10) and Theorems 2.18 and 2.32.  $\square$

**THEOREM 3.7.** *Let  $\mathbf{s} \in \mathbb{N}^d$ , let  $a^{(i,j)} : [0, 1]^d \rightarrow \mathbb{C}$  be Riemann-integrable and  $f^{(i,j)} \in L^\infty([-\pi, \pi]^d, N(\mathbf{s}))$ , for  $i = 1, \dots, p$  and  $j = 1, \dots, q_i$ . Then, for every  $\mathbf{m} \in \mathbb{N}^d$  and every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ ,*

$$\left\{ \sum_{i=1}^p \prod_{j=1}^{q_i} LT_{\mathbf{n},\mathbf{s}}^m(a^{(i,j)}, f^{(i,j)}) \right\}_n \sim_\sigma \sum_{i=1}^p \prod_{j=1}^{q_i} a_{\mathbf{m}}^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta}),$$

where

$$(3.12) \quad a_{\mathbf{m}}^{(i,j)}(\mathbf{x}) = \sum_{\mathbf{k}=1}^{\mathbf{m}} a^{(i,j)}\left(\frac{\mathbf{k}}{\mathbf{m}}\right) \chi_{[\frac{\mathbf{k}-1}{\mathbf{m}}, \frac{\mathbf{k}}{\mathbf{m}}]}(\mathbf{x}).$$

*Proof.* By Proposition 3.6 and Remark 2.36, it is enough to show that

$$\left\{ \sum_{i=1}^p LT_{\mathbf{n},\mathbf{s}}^m\left(\prod_{j=1}^{q_i} a^{(i,j)}, \prod_{j=1}^{q_i} f^{(i,j)}\right) \right\}_n \sim_\sigma \sum_{i=1}^p \prod_{j=1}^{q_i} a_{\mathbf{m}}^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta}).$$

Note that

$$\begin{aligned}
 & (\Pi_{\mathbf{n},\mathbf{m}}^{(\mathbf{s})})^T \left[ \sum_{i=1}^p LT_{\mathbf{n},\mathbf{s}}^m\left(\prod_{j=1}^{q_i} a^{(i,j)}, \prod_{j=1}^{q_i} f^{(i,j)}\right) \right] \Pi_{\mathbf{n},\mathbf{m}}^{(\mathbf{s})} \\
 (3.13) \quad &= \left( \sum_{i=1}^p D_{\mathbf{m}}\left(\prod_{j=1}^{q_i} a^{(i,j)}\right) \otimes T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\prod_{j=1}^{q_i} f^{(i,j)}\right) \right) \oplus O_{N(\mathbf{n}\mathbf{s})-N(\mathbf{m}\lfloor \mathbf{n}/\mathbf{m} \rfloor \mathbf{s})}.
 \end{aligned}$$

Recalling (2.21), for  $\mathbf{k} = 1, \dots, \mathbf{m}$ , the  $\mathbf{k}$ th diagonal block of size  $N(\mathbf{s}\lfloor \mathbf{n}/\mathbf{m} \rfloor)$  of the matrix (3.13) is given by

$$\sum_{i=1}^p \left( \prod_{j=1}^{q_i} a^{(i,j)}\left(\frac{\mathbf{k}}{\mathbf{m}}\right) \right) T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\prod_{j=1}^{q_i} f^{(i,j)}\right) = T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}\left(\frac{\mathbf{k}}{\mathbf{m}}\right) f^{(i,j)}\right).$$

It follows that the singular values of  $\sum_{i=1}^p LT_{\mathbf{n},\mathbf{s}}^m\left(\prod_{j=1}^{q_i} a^{(i,j)}, \prod_{j=1}^{q_i} f^{(i,j)}\right)$  are

$$\sigma_\ell \left( T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}\left(\frac{\mathbf{k}}{\mathbf{m}}\right) f^{(i,j)}\right) \right), \quad \ell = 1, \dots, N(\mathbf{s}\lfloor \mathbf{n}/\mathbf{m} \rfloor), \quad \mathbf{k} = 1, \dots, \mathbf{m},$$

plus further  $N(\mathbf{n}s) - N(\mathbf{m}\lfloor \mathbf{n}/\mathbf{m} \rfloor \mathbf{s}) = o(N(\mathbf{n}))$  singular values which are equal to 0. Therefore, by Theorem 2.43, since  $\sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}(\frac{\mathbf{k}}{\mathbf{m}}) f^{(i,j)} \in L^\infty([-\pi, \pi]^d, N(\mathbf{s}))$ , for any  $F \in C_c(\mathbb{R})$ , we have

(3.14)

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} \frac{1}{N(\mathbf{n}s)} \sum_{r=1}^{N(\mathbf{n}s)} F\left(\sigma_r\left(\sum_{i=1}^p LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}\left(\prod_{j=1}^{q_i} a^{(i,j)}, \prod_{j=1}^{q_i} f^{(i,j)}\right)\right)\right) \\
 &= \lim_{n \rightarrow \infty} \frac{N(\mathbf{m}\lfloor \mathbf{n}/\mathbf{m} \rfloor \mathbf{s})}{N(\mathbf{n}s)} \frac{1}{N(\mathbf{m})} \sum_{\mathbf{k}=1}^{\mathbf{m}} \frac{1}{N(\mathbf{s}\lfloor \mathbf{n}/\mathbf{m} \rfloor)} \\
 &\quad \times \sum_{\ell=1}^{N(\mathbf{s}\lfloor \mathbf{n}/\mathbf{m} \rfloor)} F\left(\sigma_\ell\left(T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}\left(\frac{\mathbf{k}}{\mathbf{m}}\right) f^{(i,j)}\right)\right)\right) \\
 &= \frac{1}{N(\mathbf{m})} \sum_{\mathbf{k}=1}^{\mathbf{m}} \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \frac{1}{N(\mathbf{s})} \sum_{\ell=1}^{N(\mathbf{s})} F\left(\sigma_\ell\left(\sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}\left(\frac{\mathbf{k}}{\mathbf{m}}\right) f^{(i,j)}(\boldsymbol{\theta})\right)\right) d\boldsymbol{\theta} \\
 &= \frac{1}{(2\pi)^d} \int_{[0,1]^d} \int_{[-\pi, \pi]^d} \frac{1}{N(\mathbf{s})} \sum_{\ell=1}^{N(\mathbf{s})} F\left(\sigma_\ell\left(\sum_{i=1}^p \prod_{j=1}^{q_i} a_{\mathbf{m}}^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta})\right)\right) d\boldsymbol{\theta} d\mathbf{x}.
 \end{aligned}$$

This concludes the proof.  $\square$

**THEOREM 3.8.** *Let  $\mathbf{s} \in \mathbb{N}^d$ , let  $a^{(i,j)} : [0,1]^d \rightarrow \mathbb{C}$  be Riemann-integrable and  $f^{(i,j)} \in L^\infty([-\pi, \pi]^d, N(\mathbf{s}))$ , for  $i = 1, \dots, p$  and  $j = 1, \dots, q_i$ . Then, for every  $\mathbf{m} \in \mathbb{N}^d$  and every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ ,*

$$\left\{ \Re\left(\sum_{i=1}^p \prod_{j=1}^{q_i} LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}(a^{(i,j)}, f^{(i,j)})\right) \right\}_n \sim_\lambda \Re\left(\sum_{i=1}^p \prod_{j=1}^{q_i} a_{\mathbf{m}}^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta})\right),$$

where  $a_{\mathbf{m}}^{(i,j)}$  is defined in (3.12).

*Proof.* The proof follows the same pattern as the proof of Theorem 3.7. By Proposition 3.6 and Remark 2.36, it is enough to show that

$$\left\{ \Re\left(\sum_{i=1}^p LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}\left(\prod_{j=1}^{q_i} a^{(i,j)}, \prod_{j=1}^{q_i} f^{(i,j)}\right)\right) \right\}_n \sim_\lambda \Re\left(\sum_{i=1}^p \prod_{j=1}^{q_i} a_{\mathbf{m}}^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta})\right).$$

Note that

$$\begin{aligned}
 & (\Pi_{\mathbf{n},\mathbf{m}}^{(\mathbf{s})})^T \left[ \Re\left(\sum_{i=1}^p LT_{\mathbf{n},\mathbf{s}}^{\mathbf{m}}\left(\prod_{j=1}^{q_i} a^{(i,j)}, \prod_{j=1}^{q_i} f^{(i,j)}\right)\right) \right] \Pi_{\mathbf{n},\mathbf{m}}^{(\mathbf{s})} \\
 (3.15) \quad &= \Re\left(\sum_{i=1}^p D_{\mathbf{m}}\left(\prod_{j=1}^{q_i} a^{(i,j)}\right) \otimes T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\prod_{j=1}^{q_i} f^{(i,j)}\right)\right) \oplus O_{N(\mathbf{n}s) - N(\mathbf{m}\lfloor \mathbf{n}/\mathbf{m} \rfloor \mathbf{s})}.
 \end{aligned}$$

Recalling (2.21) and (2.22), for  $\mathbf{k} = 1, \dots, \mathbf{m}$ , the  $\mathbf{k}$ th diagonal block of size  $N(\mathbf{s}\lfloor \mathbf{n}/\mathbf{m} \rfloor)$  of the matrix (3.15) is given by

$$\Re\left(\sum_{i=1}^p \left(\prod_{j=1}^{q_i} a^{(i,j)}\left(\frac{\mathbf{k}}{\mathbf{m}}\right)\right) T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\prod_{j=1}^{q_i} f^{(i,j)}\right)\right) = T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor}\left(\Re\left(\sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}\left(\frac{\mathbf{k}}{\mathbf{m}}\right) f^{(i,j)}\right)\right).$$

It follows that the eigenvalues of  $\Re(\sum_{i=1}^p LT_{\mathbf{n},\mathbf{s}}^m(\prod_{j=1}^{q_i} a^{(i,j)}, \prod_{j=1}^{q_i} f^{(i,j)}))$  are

$$\lambda_\ell \left( T_{\lfloor \mathbf{n}/\mathbf{m} \rfloor} \left( \Re \left( \sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)} \left( \frac{\mathbf{k}}{\mathbf{m}} \right) f^{(i,j)} \right) \right) \right), \quad \ell = 1, \dots, N(\mathbf{s} \lfloor \mathbf{n}/\mathbf{m} \rfloor), \quad \mathbf{k} = \mathbf{1}, \dots, \mathbf{m},$$

plus further  $N(\mathbf{n}\mathbf{s}) - N(\mathbf{m} \lfloor \mathbf{n}/\mathbf{m} \rfloor \mathbf{s}) = o(N(\mathbf{n}))$  eigenvalues which are equal to 0. Therefore, by Theorem 2.43, since  $\Re(\sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)} \left( \frac{\mathbf{k}}{\mathbf{m}} \right) f^{(i,j)}) \in L^\infty([-\pi, \pi]^d, N(\mathbf{s}))$ , following the same derivation as in (3.14) we obtain, for any  $F \in C_c(\mathbb{C})$ ,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{N(\mathbf{s}\mathbf{n})} \sum_{r=1}^{N(\mathbf{s}\mathbf{n})} F \left( \lambda_r \left( \Re \left( \sum_{i=1}^p LT_{\mathbf{n},\mathbf{s}}^m \left( \prod_{j=1}^{q_i} a^{(i,j)}, \prod_{j=1}^{q_i} f^{(i,j)} \right) \right) \right) \right) \\ &= \frac{1}{(2\pi)^d} \int_{[0,1]^d} \int_{[-\pi,\pi]^d} \frac{1}{N(\mathbf{s})} \sum_{\ell=1}^{N(\mathbf{s})} F \left( \lambda_\ell \left( \Re \left( \sum_{i=1}^p \prod_{j=1}^{q_i} a_m^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta}) \right) \right) \right) d\boldsymbol{\theta} d\mathbf{x}. \end{aligned}$$

This concludes the proof.  $\square$

**PROPOSITION 3.9.** *Let  $\mathbf{s} \in \mathbb{N}^d$ , let  $a : [0, 1]^d \rightarrow \mathbb{C}$  be a Riemann-integrable function and  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . Then, for every  $\mathbf{m} \in \mathbb{N}^d$  and every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ , we have*

$$\{LT_{\mathbf{n},\mathbf{s}}^m(a, f)\}_n \sim_\sigma a_m(\mathbf{x})f(\boldsymbol{\theta}),$$

where

$$a_m(\mathbf{x}) = \sum_{\mathbf{k}=1}^{\mathbf{m}} a \left( \frac{\mathbf{k}}{\mathbf{m}} \right) \chi_{\left[ \frac{\mathbf{k}-1}{\mathbf{m}}, \frac{\mathbf{k}}{\mathbf{m}} \right)}(\mathbf{x}).$$

*Proof.* Take any sequence  $\{f_k\}_k \subset L^\infty([-\pi, \pi]^d, N(\mathbf{s}))$  such that  $f_k \rightarrow f$  a.e. and in  $L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . We have:

- $\{LT_{\mathbf{n},\mathbf{s}}^m(a, f_k)\}_n \xrightarrow{\text{a.c.s.}} \{LT_{\mathbf{n},\mathbf{s}}^m(a, f)\}_n$  by Remark 3.5;
- $\{LT_{\mathbf{n},\mathbf{s}}^m(a, f_k)\}_n \sim_\sigma a_m(\mathbf{x})f_k(\boldsymbol{\theta})$  by Theorem 3.7;
- $a_m(\mathbf{x})f_k(\boldsymbol{\theta}) \rightarrow a_m(\mathbf{x})f(\boldsymbol{\theta})$  a.e. (and hence also in measure).

We conclude that  $\{LT_{\mathbf{n},\mathbf{s}}^m(a, f)\}_n \sim_\sigma a_m(\mathbf{x})f(\boldsymbol{\theta})$  by Theorem 2.34.  $\square$

**3.2. Definition of multilevel block LT sequences.** The notion of multilevel block LT sequences is formalized in the next definition.

**DEFINITION 3.10** (Multilevel block locally Toeplitz sequence). *Let  $\mathbf{s} \in \mathbb{N}^d$ , let  $\{A_n\}_n$  be a  $d$ -level  $N(\mathbf{s})$ -block matrix-sequence, let  $a : [0, 1]^d \rightarrow \mathbb{C}$  be Riemann-integrable and  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . We say that  $\{A_n\}_n$  is a ( $d$ -level  $\mathbf{s}$ -block) locally Toeplitz (LT) sequence with symbol  $a(\mathbf{x})f(\boldsymbol{\theta})$ , and we write  $\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ , if  $\{LT_{\mathbf{n},\mathbf{s}}^m(a, f)\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  as  $\mathbf{m} \rightarrow \infty$ .*

In what follows, unless specified otherwise, whenever we write a relation such as  $\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ , it is understood that  $\{A_n\}_n$  is a  $d$ -level  $\mathbf{s}$ -block LT sequence as in Definition 3.10, so in particular the function  $a : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable and  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ .

**3.3. Fundamental examples of multilevel block LT sequences.** In this section we provide three fundamental examples of multilevel block LT sequences: zero-distributed sequences, sequences of multilevel block diagonal sampling matrices, and multilevel block Toeplitz sequences. These may be regarded as the ‘‘building blocks’’ of the theory of multilevel block GLT sequences because from them we can construct through algebraic operations a lot of other matrix-sequences which will turn out to be multilevel block GLT sequences.

**3.3.1. Zero-distributed sequences.** We show that any zero-distributed sequence is a multilevel block LT sequence with symbol given by the zero matrix.

**THEOREM 3.11.** *Let  $\mathbf{s} \in \mathbb{N}^d$ , and let  $\{Z_n\}_n$  be a  $d$ -level  $N(\mathbf{s})$ -block matrix-sequence. The following statements are equivalent.*

1.  $\{Z_n\}_n \sim_\sigma 0$ .
2.  $\{O_{N(\mathbf{s}n)}\}_n \xrightarrow{\text{a.c.s.}} \{Z_n\}_n$ .
3.  $\{Z_n\}_n \sim_{\text{LT}} O_{N(\mathbf{s})}$ .

*Proof.* (1  $\iff$  2) By Theorem 2.32, we have  $\{O_{N(\mathbf{s}n)}\}_n \xrightarrow{\text{a.c.s.}} \{Z_n\}_n$  if and only if  $d_{\text{a.c.s.}}(\{O_{N(\mathbf{s}n)}\}_n, \{Z_n\}_n) = 0$  if and only if  $\{Z_n\}_n \sim_\sigma 0$ .

(2  $\iff$  3) This equivalence follows from Definition 3.10 and the observation that  $LT_{\mathbf{n},\mathbf{s}}^m(0, O_{N(\mathbf{s})}) = O_{N(\mathbf{s}n)}$  and  $0 O_{N(\mathbf{s})} = O_{N(\mathbf{s})}$ .  $\square$

**3.3.2. Sequences of multilevel block diagonal sampling matrices.** We are going to see in Theorem 3.14 that  $\{\Gamma_{\mathbf{n},\mathbf{s}}^T D_{\mathbf{n},\mathbf{s}}(a) \Gamma_{\mathbf{n},\mathbf{s}}\}_n = \{D_{\mathbf{n}}(a I_{N(\mathbf{s})})\}_n \sim_{\text{LT}} a(\mathbf{x}) I_{N(\mathbf{s})}$  whenever  $\mathbf{s} \in \mathbb{N}^d$ ,  $a : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable and  $\mathbf{n} = \mathbf{n}(n) \rightarrow \infty$  as  $n \rightarrow \infty$ . To prove Theorem 3.14 we shall need the following lemmas; cf. [40, Lemma 5.6] and [41, Lemma 4.1].

**LEMMA 3.12.** *Let  $C$  be an  $\ell \times \ell$  matrix and suppose that*

$$\|C\|_p^p \leq \varepsilon \ell',$$

where  $p \in [1, \infty)$ ,  $\varepsilon \geq 0$ , and  $\ell' \geq 0$ . Then we can write  $C$  in the form

$$C = R + N, \quad \text{rank}(R) \leq \varepsilon^{\frac{1}{p+1}} \ell', \quad \|N\| \leq \varepsilon^{\frac{1}{p+1}}.$$

**LEMMA 3.13.** *Let  $\mathcal{M}$  be any infinite subset of  $\mathbb{N}$ . For every  $m \in \mathcal{M}$ , let  $\{x(m, \mathbf{k})\}_{\mathbf{k} \in \mathbb{N}^d}$  be a family of numbers such that  $x(m, \mathbf{k}) \rightarrow x(m)$  as  $\mathbf{k} \rightarrow \infty$ , where  $x(m) \rightarrow 0$  as  $m \rightarrow \infty$ . Then there exists a family  $\{m(\mathbf{k})\}_{\mathbf{k} \in \mathbb{N}^d} \subseteq \mathcal{M}$  such that  $m(\mathbf{k}) \rightarrow \infty$  and  $x(m(\mathbf{k}), \mathbf{k}) \rightarrow 0$  as  $\mathbf{k} \rightarrow \infty$ .*

**THEOREM 3.14.** *If  $a : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable, then it holds that  $\{\Gamma_{\mathbf{n},\mathbf{s}}^T D_{\mathbf{n},\mathbf{s}}(a) \Gamma_{\mathbf{n},\mathbf{s}}\}_n = \{D_{\mathbf{n}}(a I_{N(\mathbf{s})})\}_n \sim_{\text{LT}} a(\mathbf{x}) I_{N(\mathbf{s})}$  for every  $\mathbf{s} \in \mathbb{N}^d$  and every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ .*

*Proof.* The proof consists of two steps. We first show that the thesis holds if  $a$  is continuous. Then, by using an approximation argument, we show that it holds for any Riemann-integrable function  $a$ .

*Step 1.* We prove by induction on  $d$  that if  $a \in C([0, 1]^d)$  and  $\omega_a(\cdot)$  is the modulus of continuity of  $a$ , then

$$(3.16) \quad \begin{aligned} \Gamma_{\mathbf{n},\mathbf{s}}^T D_{\mathbf{n},\mathbf{s}}(a) \Gamma_{\mathbf{n},\mathbf{s}} &= LT_{\mathbf{n},\mathbf{s}}^m(a, I_{N(\mathbf{s})}) + R_{\mathbf{n},\mathbf{m}} + N_{\mathbf{n},\mathbf{m}}, \\ \text{rank}(R_{\mathbf{n},\mathbf{m}}) &\leq N(\mathbf{s}n) \sum_{i=1}^d \frac{m_i}{n_i}, \quad \|N_{\mathbf{n},\mathbf{m}}\| \leq \sum_{i=1}^d \omega_a\left(\frac{1}{m_i} + \frac{m_i}{n_i}\right). \end{aligned}$$

Since  $\omega_a(\delta) \rightarrow 0$  as  $\delta \rightarrow 0$ , the convergence  $\{LT_{\mathbf{n},\mathbf{s}}^m(a, I_{N(\mathbf{s})})\}_n \xrightarrow{\text{a.c.s.}} \{\Gamma_{\mathbf{n},\mathbf{s}}^T D_{\mathbf{n},\mathbf{s}}(a) \Gamma_{\mathbf{n},\mathbf{s}}\}_n$  (and hence the relation  $\{\Gamma_{\mathbf{n},\mathbf{s}}^T D_{\mathbf{n},\mathbf{s}}(a) \Gamma_{\mathbf{n},\mathbf{s}}\}_n \sim_{\text{LT}} a(\mathbf{x}) I_{N(\mathbf{s})}$ ) follows immediately from Definition 2.40 (take  $n_{\mathbf{m}}$  such that  $\mathbf{n} \geq \mathbf{m}^2$  for  $n \geq n_{\mathbf{m}}$ , and take  $c(\mathbf{m}) = \sum_{i=1}^d 1/m_i$  and  $\omega(\mathbf{m}) = \sum_{i=1}^d \omega_a(2/m_i)$ ).

In the case  $d = 1$ , we have  $\mathbf{s} = (s)$ ,  $\mathbf{m} = (m)$ , and  $\mathbf{n} = \mathbf{n}(n) = (d_n)$  for some sequence of numbers  $\{d_n\}_n$  such that  $d_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Considering that  $\Gamma_{d_n,s} = I_{d_n,s}$ , equation (3.16) reduces to

$$\begin{aligned} D_{d_n,s}(a) &= LT_{d_n,s}^m(a, I_s) + R_{d_n,m} + N_{d_n,m}, \\ \text{rank}(R_{d_n,m}) &\leq sm, \quad \|N_{d_n,m}\| \leq \omega_a\left(\frac{1}{m} + \frac{m}{d_n}\right). \end{aligned}$$

This is nothing else than equation (3.15) from [8] with  $d_n$  in place of  $n$ , and it was already proved in [8]. In the case  $d > 1$ , we have

(3.17)

$$\begin{aligned}
 & LT_{\mathbf{n}, \mathbf{s}}^m(a, I_N(\mathbf{s})) = LT_{n_1, \dots, n_d, s_1, \dots, s_d}^{m_1, \dots, m_d}(a, I_{s_1} \otimes \cdots \otimes I_{s_d}) \\
 & = \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{j_1=1, \dots, m_1} \left[ I_{s_1 \lfloor n_1/m_1 \rfloor} \right. \right. \\
 & \quad \left. \left. \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} LT_{n_2, \dots, n_d, s_2, \dots, s_d}^{m_2, \dots, m_d} \left( a \left( \frac{j_1}{m_1}, \cdot \right), I_{s_2} \otimes \cdots \otimes I_{s_d} \right) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T \right] \right. \\
 & \quad \left. \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \cdots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}}, \\
 & = \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{j_1=1, \dots, m_1} \left[ \text{diag}_{i_1=(j_1-1)s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, j_1 s_1 \lfloor n_1/m_1 \rfloor} \right. \right. \\
 & \quad \left. \left. \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} LT_{n_2, \dots, n_d, s_2, \dots, s_d}^{m_2, \dots, m_d} \left( a \left( \frac{j_1}{m_1}, \cdot \right), I_{s_2 \cdots s_d} \right) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T \right] \right. \\
 & \quad \left. \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \cdots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}},
 \end{aligned}$$

where, for any  $\hat{x}_1 \in [0, 1]$ , the function  $a(\hat{x}_1, \cdot)$  is defined as follows:

$$a(\hat{x}_1, \cdot) : [0, 1]^{d-1} \rightarrow \mathbb{C}, \quad (x_2, \dots, x_d) \mapsto a(\hat{x}_1, x_2, \dots, x_d).$$

Moreover, by (2.25),

(3.18)

$$\begin{aligned}
 & \Gamma_{\mathbf{n}, \mathbf{s}}^T D_{\mathbf{n}, \mathbf{s}}(a) \Gamma_{\mathbf{n}, \mathbf{s}} = \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{i_1=1, \dots, n_1 s_1} D_{n_2, \dots, n_d, s_2, \dots, s_d} \left( a(x_{i_1, s_1}^{(n_1)}, \cdot) \right) \right] \Gamma_{\mathbf{n}, \mathbf{s}} \\
 & = \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{j_1=1, \dots, m_1} \left[ \text{diag}_{i_1=(j_1-1)s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, j_1 s_1 \lfloor n_1/m_1 \rfloor} D_{n_2, \dots, n_d, s_2, \dots, s_d} \left( a(x_{i_1, s_1}^{(n_1)}, \cdot) \right) \right] \right. \\
 & \quad \left. \oplus_{i_1=m_1 s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, n_1 s_1} \text{diag}_{i_1=m_1 s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, n_1 s_1} D_{n_2, \dots, n_d, s_2, \dots, s_d} \left( a(x_{i_1, s_1}^{(n_1)}, \cdot) \right) \right] \Gamma_{\mathbf{n}, \mathbf{s}}.
 \end{aligned}$$

For  $j_1 = 1, \dots, m_1$  and  $i_1 = (j_1 - 1)s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, j_1 s_1 \lfloor n_1/m_1 \rfloor$ , by the induction hypothesis we have

$$\begin{aligned}
 (3.19) \quad & \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T D_{n_2, \dots, n_d, s_2, \dots, s_d} \left( a(x_{i_1, s_1}^{(n_1)}, \cdot) \right) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} \\
 & \quad - LT_{n_2, \dots, n_d, s_2, \dots, s_d}^{m_2, \dots, m_d} \left( a \left( \frac{j_1}{m_1}, \cdot \right), I_{s_2 \cdots s_d} \right) \\
 & = \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T D_{n_2, \dots, n_d, s_2, \dots, s_d} \left( a(x_{i_1, s_1}^{(n_1)}, \cdot) \right) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} \\
 & \quad - \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T D_{n_2, \dots, n_d, s_2, \dots, s_d} \left( a \left( \frac{j_1}{m_1}, \cdot \right) \right) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} \\
 & \quad + R_{n_2, \dots, n_d, m_2, \dots, m_d}^{[j_1/m_1]} + N_{n_2, \dots, n_d, m_2, \dots, m_d}^{[j_1/m_1]},
 \end{aligned}$$

where

$$\begin{aligned} \text{rank}(R_{n_2, \dots, n_d, m_2, \dots, m_d}^{[j_1/m_1]}) &\leq s_2 n_2 \cdots s_d n_d \sum_{k=2}^d \frac{m_k}{n_k}, \\ \|N_{n_2, \dots, n_d, m_2, \dots, m_d}^{[j_1/m_1]}\| &\leq \sum_{k=2}^d \omega_a(j_1/m_1, \cdot) \left( \frac{1}{m_k} + \frac{m_k}{n_k} \right) \leq \sum_{k=2}^d \omega_a \left( \frac{1}{m_k} + \frac{m_k}{n_k} \right). \end{aligned}$$

Moreover, since for  $j_1 = 1, \dots, m_1$  and  $i_1 = (j_1 - 1)s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, j_1 s_1 \lfloor n_1/m_1 \rfloor$  we have

$$(j_1 - 1) \left\lfloor \frac{n_1}{m_1} \right\rfloor + 1 \leq \left\lfloor \frac{i_1 - 1}{s_1} \right\rfloor + 1 \leq j_1 \left\lfloor \frac{n_1}{m_1} \right\rfloor$$

and

$$\begin{aligned} \left| x_{i_1, s_1}^{(n_1)} - \frac{j_1}{m_1} \right| &= \left| \frac{1}{n_1} \left( \left\lfloor \frac{i_1 - 1}{s_1} \right\rfloor + 1 \right) - \frac{j_1}{m_1} \right| \leq \frac{j_1}{m_1} - \frac{(j_1 - 1) \lfloor n_1/m_1 \rfloor}{n_1} \\ &\leq \frac{j_1}{m_1} - \frac{(j_1 - 1)(n_1/m_1 - 1)}{n_1} \leq \frac{1}{m_1} + \frac{m_1}{n_1}, \end{aligned}$$

we infer that

$$\left\| D_{n_2, \dots, n_d, s_2, \dots, s_d} (a(x_{i_1, s_1}^{(n_1)}, \cdot)) - D_{n_2, \dots, n_d, s_2, \dots, s_d} \left( a \left( \frac{j_1}{m_1}, \cdot \right) \right) \right\| \leq \omega_a \left( \frac{1}{m_1} + \frac{m_1}{n_1} \right).$$

Thus, recalling (3.19),

$$\begin{aligned} &D_{n_2, \dots, n_d, s_2, \dots, s_d} (a(x_{i_1, s_1}^{(n_1)}, \cdot)) \\ &\quad - \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} LT_{n_2, \dots, n_d, s_2, \dots, s_d}^{m_2, \dots, m_d} \left( a \left( \frac{j_1}{m_1}, \cdot \right), I_{s_2 \cdots s_d} \right) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T \\ &= \hat{R}_{n_2, \dots, n_d, m_2, \dots, m_d}^{[j_1/m_1]} + \hat{N}_{\mathbf{n}, \mathbf{m}}^{[j_1/m_1, i_1/n_1]}, \\ (3.20) \quad \text{rank}(\hat{R}_{n_2, \dots, n_d, m_2, \dots, m_d}^{[j_1/m_1]}) &\leq s_2 n_2 \cdots s_d n_d \sum_{k=2}^d \frac{m_k}{n_k}, \\ \|\hat{N}_{\mathbf{n}, \mathbf{m}}^{[j_1/m_1, i_1/n_1]}\| &\leq \sum_{k=1}^d \omega_a \left( \frac{1}{m_k} + \frac{m_k}{n_k} \right). \end{aligned}$$

Hence, by (3.17) and (3.18),

$$\begin{aligned} &\Gamma_{\mathbf{n}, \mathbf{s}}^T D_{\mathbf{n}, \mathbf{s}}(a) \Gamma_{\mathbf{n}, \mathbf{s}} - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(a, I_{N(\mathbf{s})}) \\ &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{j_1=1, \dots, m_1} \left[ \text{diag}_{i_1=(j_1-1)s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, j_1 s_1 \lfloor n_1/m_1 \rfloor} \right. \right. \\ &\quad \left. \left. \begin{aligned} &D_{n_2, \dots, n_d, s_2, \dots, s_d} (a(x_{i_1, s_1}^{(n_1)}, \cdot)) \\ &- \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} LT_{n_2, \dots, n_d, s_2, \dots, s_d}^{m_2, \dots, m_d} \left( a \left( \frac{j_1}{m_1}, \cdot \right), I_{s_2 \cdots s_d} \right) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T \end{aligned} \right] \right] \\ &\oplus \text{diag}_{i_1=m_1 s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, n_1 s_1} D_{n_2, \dots, n_d, s_2, \dots, s_d} (a(x_{i_1, s_1}^{(n_1)}, \cdot)) \Gamma_{\mathbf{n}, \mathbf{s}} \end{aligned}$$

$$\begin{aligned}
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{j_1=1, \dots, m_1} \left[ \text{diag}_{i_1=(j_1-1)s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, j_1 s_1 \lfloor n_1/m_1 \rfloor} \right. \right. \\
 &\quad \left. \left. \left( \hat{R}_{n_2, \dots, n_d, m_2, \dots, m_d}^{[j_1/m_1]} + \hat{N}_{\mathbf{n}, \mathbf{m}}^{[j_1/m_1, i_1/n_1]} \right) \right] \right. \\
 &\quad \left. \oplus \text{diag}_{i_1=m_1 s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, n_1 s_1} D_{n_2, \dots, n_d, s_2, \dots, s_d} (a(x_{i_1, s_1}^{(n_1)}, \cdot)) \right] \Gamma_{\mathbf{n}, \mathbf{s}} \\
 &= R_{\mathbf{n}, \mathbf{m}} + N_{\mathbf{n}, \mathbf{m}},
 \end{aligned}$$

where

$$\begin{aligned}
 R_{\mathbf{n}, \mathbf{m}} &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{j_1=1, \dots, m_1} \left[ \text{diag}_{i_1=(j_1-1)s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, j_1 s_1 \lfloor n_1/m_1 \rfloor} \right. \right. \\
 &\quad \left. \left. \hat{R}_{n_2, \dots, n_d, m_2, \dots, m_d}^{[j_1/m_1]} \right] \right. \\
 &\quad \left. \oplus \text{diag}_{i_1=m_1 s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, n_1 s_1} D_{n_2, \dots, n_d, s_2, \dots, s_d} (a(x_{i_1, s_1}^{(n_1)}, \cdot)) \right] \Gamma_{\mathbf{n}, \mathbf{s}}, \\
 N_{\mathbf{n}, \mathbf{m}} &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{j_1=1, \dots, m_1} \left[ \text{diag}_{i_1=(j_1-1)s_1 \lfloor n_1/m_1 \rfloor + 1, \dots, j_1 s_1 \lfloor n_1/m_1 \rfloor} \right. \right. \\
 &\quad \left. \left. \hat{N}_{\mathbf{n}, \mathbf{m}}^{[j_1/m_1, i_1/n_1]} \right] \right. \\
 &\quad \left. \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \cdots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}}.
 \end{aligned}$$

By (3.20), (2.7), and (2.8), we have

$$\begin{aligned}
 \text{rank}(R_{\mathbf{n}, \mathbf{m}}) &\leq m_1 s_1 \left\lfloor \frac{n_1}{m_1} \right\rfloor s_2 n_2 \cdots s_d n_d \sum_{k=2}^d \frac{m_k}{n_k} + s_1 (n_1 \bmod m_1) s_2 n_2 \cdots s_d n_d \\
 &\leq s_1 n_1 s_2 n_2 \cdots s_d n_d \sum_{k=2}^d \frac{m_k}{n_k} + s_1 m_1 s_2 n_2 \cdots s_d n_d = N(\mathbf{s}, \mathbf{n}) \sum_{k=1}^d \frac{m_k}{n_k}, \\
 \|N_{\mathbf{n}, \mathbf{m}}\| &\leq \sum_{k=1}^d \omega_a \left( \frac{1}{m_k} + \frac{m_k}{n_k} \right),
 \end{aligned}$$

and (3.16) is proved.

*Step 2.* Let  $a : [0, 1]^d \rightarrow \mathbb{C}$  be any Riemann-integrable function. Take any sequence of continuous functions  $a_m : [0, 1]^d \rightarrow \mathbb{C}$  such that  $a_m \rightarrow a$  in  $L^1([0, 1]^d)$ . By Step 1, we have  $\{D_{\mathbf{n}}(a_m I_{N(\mathbf{s})})\}_n \sim_{\text{LT}} a_m(\mathbf{x}) I_{N(\mathbf{s})}$ . Hence,  $\{LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a_m, I_{N(\mathbf{s})})\}_n \xrightarrow{\text{a.c.s.}} \{D_{\mathbf{n}}(a_m I_{N(\mathbf{s})})\}_n$  as  $\mathbf{k} \rightarrow \infty$ , i.e., for every  $m$  and every  $\mathbf{k} \in \mathbb{N}^d$  there is  $n_{m, \mathbf{k}}$  such that, for  $n \geq n_{m, \mathbf{k}}$ ,

$$\begin{aligned}
 D_{\mathbf{n}}(a_m I_{N(\mathbf{s})}) &= LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a_m, I_{N(\mathbf{s})}) + R_{\mathbf{n}, \mathbf{m}, \mathbf{k}} + N_{\mathbf{n}, \mathbf{m}, \mathbf{k}}, \\
 \text{rank}(R_{\mathbf{n}, \mathbf{m}, \mathbf{k}}) &\leq c(m, \mathbf{k}) N(\mathbf{n}), \quad \|N_{\mathbf{n}, \mathbf{m}, \mathbf{k}}\| \leq \omega(m, \mathbf{k}),
 \end{aligned}$$

where  $\lim_{\mathbf{k} \rightarrow \infty} c(m, \mathbf{k}) = \lim_{\mathbf{k} \rightarrow \infty} \omega(m, \mathbf{k}) = 0$ . Moreover, we have  $\{D_{\mathbf{n}}(a_m I_{N(\mathbf{s})})\}_n \xrightarrow{\text{a.c.s.}} \{D_{\mathbf{n}}(a I_{N(\mathbf{s})})\}_n$ . Indeed,

$$\|D_{\mathbf{n}}(a I_{N(\mathbf{s})}) - D_{\mathbf{n}}(a_m I_{N(\mathbf{s})})\|_1 = N(\mathbf{s}) \sum_{j=1}^n \left| a\left(\frac{\mathbf{j}}{\mathbf{n}}\right) - a_m\left(\frac{\mathbf{j}}{\mathbf{n}}\right) \right| = \varepsilon(m, \mathbf{n}) N(\mathbf{n}),$$

where

$$(3.21) \quad \varepsilon(m, \mathbf{n}) = \frac{N(\mathbf{s})}{N(\mathbf{n})} \sum_{j=1}^n \left| a\left(\frac{\mathbf{j}}{\mathbf{n}}\right) - a_m\left(\frac{\mathbf{j}}{\mathbf{n}}\right) \right|.$$

By the Riemann-integrability of  $|a - a_m|$  and the fact that  $a_m \rightarrow a$  in  $L^1([0, 1]^d)$ , the quantity  $\varepsilon(m, \mathbf{n})$  satisfies

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

By Theorem 2.39, this implies that  $\{D_{\mathbf{n}}(a_m I_{N(\mathbf{s})})\}_n \xrightarrow{\text{a.c.s.}} \{D_{\mathbf{n}}(a I_{N(\mathbf{s})})\}_n$ . Thus, for every  $m$  there exists  $n_m$  such that, for  $n \geq n_m$ ,

$$\begin{aligned} D_{\mathbf{n}}(a I_{N(\mathbf{s})}) &= D_{\mathbf{n}}(a_m I_{N(\mathbf{s})}) + R_{\mathbf{n}, m} + N_{\mathbf{n}, m}, \\ \text{rank}(R_{\mathbf{n}, m}) &\leq c(m)N(\mathbf{n}), \quad \|N_{\mathbf{n}, m}\| \leq \omega(m), \end{aligned}$$

where  $\lim_{m \rightarrow \infty} c(m) = \lim_{m \rightarrow \infty} \omega(m) = 0$ . It follows that, for every  $m$ , every  $\mathbf{k} \in \mathbb{N}^d$ , and every  $n \geq \max(n_m, n_{m, \mathbf{k}})$ ,

$$\begin{aligned} D_{\mathbf{n}}(a I_{N(\mathbf{s})}) &= LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a, I_{N(\mathbf{s})}) + [LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a_m, I_{N(\mathbf{s})}) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a, I_{N(\mathbf{s})})] \\ &\quad + (R_{\mathbf{n}, m} + R_{\mathbf{n}, m, \mathbf{k}}) + (N_{\mathbf{n}, m} + N_{\mathbf{n}, m, \mathbf{k}}), \end{aligned}$$

$$\text{rank}(R_{\mathbf{n}, m} + R_{\mathbf{n}, m, \mathbf{k}}) \leq (c(m) + c(m, \mathbf{k}))N(\mathbf{n}),$$

$$\|N_{\mathbf{n}, m} + N_{\mathbf{n}, m, \mathbf{k}}\| \leq \omega(m) + \omega(m, \mathbf{k}),$$

$$\|LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a_m, I_{N(\mathbf{s})}) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a, I_{N(\mathbf{s})})\|_1 \leq \frac{N(\mathbf{s}\mathbf{n})}{N(\mathbf{k})} \sum_{j=1}^{\mathbf{k}} \left| a\left(\frac{j}{\mathbf{k}}\right) - a_m\left(\frac{j}{\mathbf{k}}\right) \right| = \varepsilon(m, \mathbf{k})N(\mathbf{n}),$$

where the last inequality follows from (3.7)–(3.9) and  $\varepsilon(m, \mathbf{k})$  is defined as in (3.21) with “ $\mathbf{n}$ ” replaced by “ $\mathbf{k}$ ”. Let  $\{m(\mathbf{k})\}_{\mathbf{k} \in \mathbb{N}^d}$  be a family of indices such that  $m(\mathbf{k}) \rightarrow \infty$  as  $\mathbf{k} \rightarrow \infty$  and

$$\lim_{\mathbf{k} \rightarrow \infty} \varepsilon(m(\mathbf{k}), \mathbf{k}) = \lim_{\mathbf{k} \rightarrow \infty} c(m(\mathbf{k}), \mathbf{k}) = \lim_{\mathbf{k} \rightarrow \infty} \omega(m(\mathbf{k}), \mathbf{k}) = 0.$$

Such a family exists by Lemma 3.13 (apply the lemma with  $x(m, \mathbf{k}) = \varepsilon(m, \mathbf{k}) + c(m, \mathbf{k}) + \omega(m, \mathbf{k})$ ). Then, for every  $\mathbf{k} \in \mathbb{N}^d$  and every  $n \geq \max(n_{m(\mathbf{k})}, n_{m(\mathbf{k}), \mathbf{k}})$ ,

$$\begin{aligned} D_{\mathbf{n}}(a I_{N(\mathbf{s})}) &= LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a, I_{N(\mathbf{s})}) + [LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a_{m(\mathbf{k})}, I_{N(\mathbf{s})}) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a, I_{N(\mathbf{s})})] \\ &\quad + (R_{\mathbf{n}, m(\mathbf{k})} + R_{\mathbf{n}, m(\mathbf{k}), \mathbf{k}}) + (N_{\mathbf{n}, m(\mathbf{k})} + N_{\mathbf{n}, m(\mathbf{k}), \mathbf{k}}), \end{aligned}$$

$$\text{rank}(R_{\mathbf{n}, m(\mathbf{k})} + R_{\mathbf{n}, m(\mathbf{k}), \mathbf{k}}) \leq (c(m(\mathbf{k})) + c(m(\mathbf{k}), \mathbf{k}))N(\mathbf{n}),$$

$$\|N_{\mathbf{n}, m(\mathbf{k})} + N_{\mathbf{n}, m(\mathbf{k}), \mathbf{k}}\| \leq \omega(m(\mathbf{k})) + \omega(m(\mathbf{k}), \mathbf{k}),$$

$$\|LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a_{m(\mathbf{k})}, I_{N(\mathbf{s})}) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a, I_{N(\mathbf{s})})\|_1 \leq \varepsilon(m(\mathbf{k}), \mathbf{k})N(\mathbf{n}).$$

By Lemma 3.12, we can decompose  $LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a_{m(\mathbf{k})}, I_{N(\mathbf{s})}) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a, I_{N(\mathbf{s})})$  as the sum of a small-rank term  $\hat{R}_{\mathbf{n}, \mathbf{k}}$ , with rank bounded by  $\sqrt{\varepsilon(m(\mathbf{k}), \mathbf{k})}N(\mathbf{n})$ , plus a small-norm term  $\hat{N}_{\mathbf{n}, \mathbf{k}}$ , with norm bounded by  $\sqrt{\varepsilon(m(\mathbf{k}), \mathbf{k})}$ . This shows that  $\{LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(a, I_{N(\mathbf{s})})\}_n \xrightarrow{\text{a.c.s.}} \{D_{\mathbf{n}}(a I_{N(\mathbf{s})})\}_n$  as  $\mathbf{k} \rightarrow \infty$ , hence  $\{D_{\mathbf{n}}(a I_{N(\mathbf{s})})\}_n \sim_{\text{LT}} a(\mathbf{x})I_{N(\mathbf{s})}$ .  $\square$

### 3.3.3. Multilevel block Toeplitz sequences.

**THEOREM 3.15.** *If  $\mathbf{s} \in \mathbb{N}^d$  and  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ , then  $\{T_{\mathbf{n}}(f)\}_n \sim_{\text{LT}} f(\boldsymbol{\theta})$  for every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ .*

*Proof.* The proof consists of three steps. We first show that the thesis holds if  $f$  is a separable  $d$ -variate matrix-valued trigonometric polynomial. Then, by linearity, we show



that it holds if  $f$  is an arbitrary  $d$ -variate matrix-valued trigonometric polynomial. Finally, by using an approximation argument, we prove the theorem under the sole assumption that  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ .

*Step 1.* We show by induction on  $d$  that if  $f$  is a separable  $d$ -variate matrix-valued trigonometric polynomial of the form  $f = f_1 \otimes \cdots \otimes f_d$ , with  $f_j$  a univariate  $s_j \times s_j$  matrix-valued trigonometric polynomial of degree  $q_j$ , then

$$(3.22) \quad T_{\mathbf{n}}(f) = LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(1, f) + R_{\mathbf{n}, \mathbf{m}}, \quad \text{rank}(R_{\mathbf{n}, \mathbf{m}}) \leq N(\mathbf{s}\mathbf{n}) \sum_{i=1}^d \frac{(2q_i + 1)m_i}{n_i}.$$

Once this is done, the convergence  $\{LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(1, f)\}_n \xrightarrow{\text{a.c.s.}} \{T_{\mathbf{n}}(f)\}_n$  (and hence the relation  $\{T_{\mathbf{n}}(f)\}_n \sim_{\text{LT}} f(\boldsymbol{\theta})$ ) follows immediately from Definition 2.40 (take  $n_{\mathbf{m}}$  such that  $\mathbf{n} \geq \mathbf{m}^2$ , for  $n \geq n_{\mathbf{m}}$ , and take  $c(\mathbf{m}) = \sum_{i=1}^d (2q_i + 1)/m_i$  and  $\omega(\mathbf{m}) = 0$ ).

In the case  $d = 1$ , we have  $\mathbf{s} = (s)$ ,  $\mathbf{m} = (m)$ , and  $\mathbf{n} = \mathbf{n}(n) = (d_n)$  for some sequence of numbers  $\{d_n\}_n$  such that  $d_n \rightarrow \infty$  as  $n \rightarrow \infty$ . In this case, equation (3.22) reduces to

$$(3.23) \quad T_{d_n}(f) = LT_{d_n, s}^m(1, f) + R_{d_n, m}, \quad \text{rank}(R_{d_n, m}) \leq s(2q + 1)m,$$

where  $q$  is the degree of  $f$ . This is nothing else than equation (3.17) from [8] with  $d_n$  in place of  $n$ , and it was already proved in [8]. In the case  $d > 1$ , let  $f = f_1 \otimes \cdots \otimes f_d$  with  $f_j$  a univariate  $s_j \times s_j$  matrix-valued trigonometric polynomial of degree  $q_j$ . By induction hypothesis,

$$\begin{aligned} LT_{n_2, \dots, n_d, s_2, \dots, s_d}^{m_2, \dots, m_d}(1, f_2 \otimes \cdots \otimes f_d) &= T_{n_2, \dots, n_d}(f_2 \otimes \cdots \otimes f_d) - R_{n_2, \dots, n_d, m_2, \dots, m_d}, \\ \text{rank}(R_{n_2, \dots, n_d, m_2, \dots, m_d}) &\leq s_2 n_2 \cdots s_d n_d \sum_{i=2}^d \frac{(2q_i + 1)m_i}{n_i}. \end{aligned}$$

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

$$\begin{aligned} &LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{m}}(1, f) \\ &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \text{diag}_{j_1=1, \dots, m_1} \left[ T_{\lfloor n_1/m_1 \rfloor}(f_1) \right. \right. \\ &\quad \left. \left. \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} LT_{n_2, \dots, n_d, s_2, \dots, s_d}^{m_2, \dots, m_d}(1, f_2 \otimes \cdots \otimes f_d) \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T \right. \right. \\ &\quad \left. \left. \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \cdots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}} \\ &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \left[ \text{diag}_{j_1=1, \dots, m_1} T_{\lfloor n_1/m_1 \rfloor}(f_1) \right] \right. \\ &\quad \left. \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} \left[ T_{n_2, \dots, n_d}(f_2 \otimes \cdots \otimes f_d) - R_{n_2, \dots, n_d, m_2, \dots, m_d} \right] \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T \right. \\ &\quad \left. \oplus O_{s_1(n_1 \bmod m_1) s_2 n_2 \cdots s_d n_d} \right] \Gamma_{\mathbf{n}, \mathbf{s}} \\ &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ \left[ \text{diag}_{j_1=1, \dots, m_1} T_{\lfloor n_1/m_1 \rfloor}(f_1) \oplus O_{s_1(n_1 \bmod m_1)} \right] \right. \\ &\quad \left. \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d} \left[ T_{n_2, \dots, n_d}(f_2 \otimes \cdots \otimes f_d) - R_{n_2, \dots, n_d, m_2, \dots, m_d} \right] \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}^T \right] \Gamma_{\mathbf{n}, \mathbf{s}} \end{aligned}$$

$$\begin{aligned}
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}) \\
 &\quad \cdot \left[ LT_{n_1, s_1}^{m_1}(1, f_1) \otimes [T_{n_2, \dots, n_d}(f_2 \otimes \dots \otimes f_d) - R_{n_2, \dots, n_d, m_2, \dots, m_d}] \right] \\
 &\quad \cdot (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d})^T \Gamma_{\mathbf{n}, \mathbf{s}} \\
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}) \left[ LT_{n_1, s_1}^{m_1}(1, f_1) \otimes T_{n_2, \dots, n_d}(f_2 \otimes \dots \otimes f_d) - \tilde{R}_{\mathbf{n}, \mathbf{m}} \right] \\
 &\quad \cdot (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d})^T \Gamma_{\mathbf{n}, \mathbf{s}},
 \end{aligned}$$

where  $\tilde{R}_{\mathbf{n}, \mathbf{m}} = LT_{n_1, s_1}^{m_1}(1, f_1) \otimes R_{n_2, \dots, n_d, m_2, \dots, m_d}$  satisfies

$$\text{rank}(\tilde{R}_{\mathbf{n}, \mathbf{m}}) \leq N(\mathbf{n}, \mathbf{s}) \sum_{i=2}^d \frac{(2q_i + 1)m_i}{n_i}.$$

Using (3.23) with  $d_n = n_1(n) = n_1$ , we can decompose  $LT_{n_1, s_1}^{m_1}(1, f_1)$  into the sum of  $T_{n_1}(f_1)$  plus a small-rank matrix  $-R_{n_1, m_1}$ , whose rank is bounded by  $s_1(2q_1 + 1)m_1$ . Invoking Lemma 2.45, we obtain

$$\begin{aligned}
 &LT_{\mathbf{n}, \mathbf{s}}^m(1, f) \\
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}) \left[ (T_{n_1}(f_1) - R_{n_1, m_1}) \otimes T_{n_2, \dots, n_d}(f_2 \otimes \dots \otimes f_d) - \tilde{R}_{\mathbf{n}, \mathbf{m}} \right] \\
 &\quad \cdot (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d})^T \Gamma_{\mathbf{n}, \mathbf{s}} \\
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}) \left[ T_{n_1}(f_1) \otimes T_{n_2, \dots, n_d}(f_2 \otimes \dots \otimes f_d) - \hat{R}_{\mathbf{n}, \mathbf{m}} \right] \\
 &\quad \cdot (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d})^T \Gamma_{\mathbf{n}, \mathbf{s}} \\
 &= \Gamma_{\mathbf{n}, \mathbf{s}}^T \left[ T_{n_1}(f_1) \otimes T_{n_2}(f_2) \otimes \dots \otimes T_{n_d}(f_d) - \hat{R}_{\mathbf{n}, \mathbf{m}} \right] \Gamma_{\mathbf{n}, \mathbf{s}} \\
 &= T_{\mathbf{n}}(f_1 \otimes f_2 \otimes \dots \otimes f_d) + R_{\mathbf{n}, \mathbf{m}},
 \end{aligned}$$

where

$$\begin{aligned}
 R_{\mathbf{n}, \mathbf{m}} &= \Gamma_{\mathbf{n}, \mathbf{s}}^T (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d}) \left[ -R_{n_1, m_1} \otimes T_{n_2, \dots, n_d}(f_2 \otimes \dots \otimes f_d) \right. \\
 &\quad \left. - \tilde{R}_{\mathbf{n}, \mathbf{m}} \right] (I_{n_1 s_1} \otimes \Gamma_{n_2, \dots, n_d, s_2, \dots, s_d})^T \Gamma_{\mathbf{n}, \mathbf{s}}
 \end{aligned}$$

satisfies

$$\begin{aligned}
 \text{rank}(R_{\mathbf{n}, \mathbf{m}}) &\leq s_1(2q_1 + 1)m_1 s_2 n_2 \dots s_d n_d + N(\mathbf{n}, \mathbf{s}) \sum_{i=2}^d \frac{(2q_i + 1)m_i}{n_i} \\
 &= N(\mathbf{n}, \mathbf{s}) \sum_{i=1}^d \frac{(2q_i + 1)m_i}{n_i}.
 \end{aligned}$$

This completes the proof of (3.22).

*Step 2.* Let  $f$  be any  $d$ -variate  $N(\mathbf{s}) \times N(\mathbf{s})$  matrix-valued trigonometric polynomial. By definition,  $f$  is a finite linear combination of the  $d$ -variate Fourier frequencies  $e^{j \cdot \theta}$ ,  $j \in \mathbb{Z}^d$ ,

and so we can write  $f(\boldsymbol{\theta}) = \sum_{j=-q}^q f_j e^{ij \cdot \boldsymbol{\theta}}$  for some  $f_j \in \mathbb{C}^{N(\mathbf{s}) \times N(\mathbf{s})}$ . Hence,

$$\begin{aligned} f(\boldsymbol{\theta}) &= \sum_{j=-q}^q f_j e^{ij \cdot \boldsymbol{\theta}} = \sum_{j=-q}^q \sum_{\ell, \mathbf{k}=1}^s e^{ij \cdot \boldsymbol{\theta}} (f_j)_{\ell \mathbf{k}} E_{\ell \mathbf{k}}^{(s)} \\ &= \sum_{j=-q}^q \sum_{\ell, \mathbf{k}=1}^s (f_j)_{\ell \mathbf{k}} (e^{ij_1 \theta_1} E_{\ell_1 k_1}^{(s_1)}) \otimes \cdots \otimes (e^{ij_d \theta_d} E_{\ell_d k_d}^{(s_d)}). \end{aligned}$$

By linearity,

$$\begin{aligned} T_{\mathbf{n}}(f) &= \sum_{j=-q}^q \sum_{\ell, \mathbf{k}=1}^s (f_j)_{\ell \mathbf{k}} T_{\mathbf{n}}((e^{ij_1 \theta_1} E_{\ell_1 k_1}^{(s_1)}) \otimes \cdots \otimes (e^{ij_d \theta_d} E_{\ell_d k_d}^{(s_d)})), \\ LT_{\mathbf{n}, \mathbf{s}}^m(1, f) &= \sum_{j=-q}^q \sum_{\ell, \mathbf{k}=1}^s (f_j)_{\ell \mathbf{k}} LT_{\mathbf{n}, \mathbf{s}}^m(1, (e^{ij_1 \theta_1} E_{\ell_1 k_1}^{(s_1)}) \otimes \cdots \otimes (e^{ij_d \theta_d} E_{\ell_d k_d}^{(s_d)})). \end{aligned}$$

The thesis now follows from Step 1 and Remark 2.42.

*Step 3.* Let  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . Since the set of  $d$ -variate  $N(\mathbf{s}) \times N(\mathbf{s})$  matrix-valued trigonometric polynomials is dense in  $L^1([-\pi, \pi]^d, N(\mathbf{s}))$  (see, e.g., [40, Lemma 2.2]), there is a sequence of  $d$ -variate  $N(\mathbf{s}) \times N(\mathbf{s})$  matrix-valued trigonometric polynomials  $f_m : [-\pi, \pi]^d \rightarrow \mathbb{C}^{N(\mathbf{s}) \times N(\mathbf{s})}$  such that  $f_m \rightarrow f$  in  $L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . By Step 2,  $\{T_{\mathbf{n}}(f_m)\}_n \sim_{\text{LT}} f_m(\boldsymbol{\theta})$ . Hence,  $\{LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f_m)\}_n \xrightarrow{\text{a.c.s.}} \{T_{\mathbf{n}}(f_m)\}_n$  as  $\mathbf{k} \rightarrow \infty$ , i.e., for every  $m$  and every  $\mathbf{k} \in \mathbb{N}^d$  there is  $n_{m, \mathbf{k}}$  such that, for  $n \geq n_{m, \mathbf{k}}$ ,

$$\begin{aligned} T_{\mathbf{n}}(f_m) &= LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f_m) + R_{\mathbf{n}, m, \mathbf{k}} + N_{\mathbf{n}, m, \mathbf{k}}, \\ \text{rank}(R_{\mathbf{n}, m, \mathbf{k}}) &\leq c(m, \mathbf{k})N(\mathbf{n}), \quad \|N_{\mathbf{n}, m, \mathbf{k}}\| \leq \omega(m, \mathbf{k}), \end{aligned}$$

where  $\lim_{\mathbf{k} \rightarrow \infty} c(m, \mathbf{k}) = \lim_{\mathbf{k} \rightarrow \infty} \omega(m, \mathbf{k}) = 0$ . Moreover, by Theorem 2.44,

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

and so  $\{T_{\mathbf{n}}(f_m)\}_n \xrightarrow{\text{a.c.s.}} \{T_{\mathbf{n}}(f)\}_n$  by Theorem 2.39. Thus, for every  $m$  there exists  $n_m$  such that, for  $n \geq n_m$ ,

$$\begin{aligned} T_{\mathbf{n}}(f) &= T_{\mathbf{n}}(f_m) + R_{\mathbf{n}, m} + N_{\mathbf{n}, m}, \\ \text{rank}(R_{\mathbf{n}, m}) &\leq c(m)N(\mathbf{n}), \quad \|N_{\mathbf{n}, m}\| \leq \omega(m), \end{aligned}$$

where  $\lim_{m \rightarrow \infty} c(m) = \lim_{m \rightarrow \infty} \omega(m) = 0$ . It follows that, for every  $m$ , every  $\mathbf{k} \in \mathbb{N}^d$ , and every  $n \geq \max(n_m, n_{m, \mathbf{k}})$ ,

$$\begin{aligned} T_{\mathbf{n}}(f) &= LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f) + [LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f_m) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f)] \\ &\quad + (R_{\mathbf{n}, m} + R_{\mathbf{n}, m, \mathbf{k}}) + (N_{\mathbf{n}, m} + N_{\mathbf{n}, m, \mathbf{k}}), \\ \text{rank}(R_{\mathbf{n}, m} + R_{\mathbf{n}, m, \mathbf{k}}) &\leq (c(m) + c(m, \mathbf{k}))N(\mathbf{n}), \\ \|N_{\mathbf{n}, m} + N_{\mathbf{n}, m, \mathbf{k}}\| &\leq \omega(m) + \omega(m, \mathbf{k}), \\ \|LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f_m) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f)\|_1 &= \|LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f_m - f)\|_1 \leq N(\mathbf{n})\|f_m - f\|_{L^1}, \end{aligned}$$

where the last inequality follows from (3.9). Let  $\{m(\mathbf{k})\}_{\mathbf{k} \in \mathbb{N}^d}$  be a family of indices such that  $m(\mathbf{k}) \rightarrow \infty$  as  $\mathbf{k} \rightarrow \infty$  and

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

Such a family exists by Lemma 3.13 (apply the lemma with  $x(m, \mathbf{k}) = c(m, \mathbf{k}) + \omega(m, \mathbf{k})$ ). Then, for every  $\mathbf{k} \in \mathbb{N}^d$  and every  $n \geq \max(n_{m(\mathbf{k})}, n_{m(\mathbf{k}), \mathbf{k}})$ ,

$$\begin{aligned} T_{\mathbf{n}}(f) &= LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f) + [LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f_{m(\mathbf{k})}) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f)] \\ &\quad + (R_{\mathbf{n}, m(\mathbf{k})} + R_{\mathbf{n}, m(\mathbf{k}), \mathbf{k}}) + (N_{\mathbf{n}, m(\mathbf{k})} + N_{\mathbf{n}, m(\mathbf{k}), \mathbf{k}}), \\ \text{rank}(R_{\mathbf{n}, m(\mathbf{k})} + R_{\mathbf{n}, m(\mathbf{k}), \mathbf{k}}) &\leq (c(m(\mathbf{k})) + c(m(\mathbf{k}), \mathbf{k}))N(\mathbf{n}), \\ \|N_{\mathbf{n}, m(\mathbf{k})} + N_{\mathbf{n}, m(\mathbf{k}), \mathbf{k}}\| &\leq \omega(m(\mathbf{k})) + \omega(m(\mathbf{k}), \mathbf{k}), \\ \|LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f_{m(\mathbf{k})}) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f)\|_1 &\leq N(\mathbf{n})\|f_{m(\mathbf{k})} - f\|_{L^1}. \end{aligned}$$

By Lemma 3.12, we can decompose  $LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f_{m(\mathbf{k})}) - LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f)$  as the sum of a small-rank term  $\hat{R}_{\mathbf{n}, \mathbf{k}}$ , with rank bounded by  $\sqrt{\|f_{m(\mathbf{k})} - f\|_{L^1}} N(\mathbf{n})$ , plus a small-norm term  $\hat{N}_{\mathbf{n}, \mathbf{k}}$ , with norm bounded by  $\sqrt{\|f_{m(\mathbf{k})} - f\|_{L^1}}$ . This shows that  $\{LT_{\mathbf{n}, \mathbf{s}}^{\mathbf{k}}(1, f)\}_n \xrightarrow{\text{a.c.s.}} \{T_{\mathbf{n}}(f)\}_n$  as  $\mathbf{k} \rightarrow \infty$ , hence  $\{T_{\mathbf{n}}(f)\}_n \sim_{\text{LT}} f(\boldsymbol{\theta})$ .  $\square$

**3.4. Singular value and spectral distribution of sums of products of multilevel block LT sequences.** The main results of this section are Theorems 3.17 and 3.18. In order to prove them, we shall need the following lemma.

LEMMA 3.16. *If  $\{A_{\mathbf{n}}\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ , then  $\{A_{\mathbf{n}}\}_n \sim_{\sigma} a(\mathbf{x})f(\boldsymbol{\theta})$  and  $\{A_{\mathbf{n}}\}_n$  is s.u.*

*Proof.* We have:

- $\{LT_{\mathbf{n}, \mathbf{s}}^m(a, f)\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}\}_n$  by definition of multilevel block LT sequences;
- $\{LT_{\mathbf{n}, \mathbf{s}}^m(a, f)\}_n \sim_{\sigma} a_m(\mathbf{x})f(\boldsymbol{\theta})$  with  $a_m(\mathbf{x}) = \sum_{\mathbf{k}=1}^m a(\frac{\mathbf{k}}{m})\chi_{[\frac{\mathbf{k}-1}{m}, \frac{\mathbf{k}}{m}]}(\mathbf{x})$  by Proposition 3.9;
- $a_m(\mathbf{x})f(\boldsymbol{\theta}) \rightarrow a(\mathbf{x})f(\boldsymbol{\theta})$  a.e. (and hence also in measure) as  $m \rightarrow \infty$  by Lemma 2.10 because  $a(\mathbf{x})$  is Riemann-integrable.

We conclude that  $\{A_{\mathbf{n}}\}_n \sim_{\sigma} a(\mathbf{x})f(\boldsymbol{\theta})$  by Theorem 2.34 and Remark 2.41, and so  $\{A_{\mathbf{n}}\}_n$  is s.u. by Proposition 2.22.  $\square$

THEOREM 3.17. *If  $\{A_{\mathbf{n}}^{(i,j)}\}_n \sim_{\text{LT}} a^{(i,j)}(\mathbf{x})f^{(i,j)}(\boldsymbol{\theta})$ , for  $i = 1, \dots, p$  and  $j = 1, \dots, q_i$ , then*

$$\left\{ \sum_{i=1}^p \prod_{j=1}^{q_i} A_{\mathbf{n}}^{(i,j)} \right\}_n \sim_{\sigma} \sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}(\mathbf{x})f^{(i,j)}(\boldsymbol{\theta}).$$

*Proof.* Let

$$\begin{aligned} A_{\mathbf{n}} &= \sum_{i=1}^p \prod_{j=1}^{q_i} A_{\mathbf{n}}^{(i,j)}, & A_{\mathbf{n}, m} &= \sum_{i=1}^p \prod_{j=1}^{q_i} LT_{\mathbf{n}, \mathbf{s}}^m(a^{(i,j)}, f^{(i,j)}), \\ \kappa(\mathbf{x}, \boldsymbol{\theta}) &= \sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}(\mathbf{x})f^{(i,j)}(\boldsymbol{\theta}), & \kappa_m(\mathbf{x}, \boldsymbol{\theta}) &= \sum_{i=1}^p \prod_{j=1}^{q_i} a_m^{(i,j)}(\mathbf{x})f^{(i,j)}(\boldsymbol{\theta}), \end{aligned}$$

where

$$a_m^{(i,j)}(\mathbf{x}) = \sum_{\mathbf{k}=1}^m a^{(i,j)}\left(\frac{\mathbf{k}}{m}\right)\chi_{[\frac{\mathbf{k}-1}{m}, \frac{\mathbf{k}}{m}]}(\mathbf{x}).$$

Since  $\{LT_{\mathbf{n}, \mathbf{s}}^m(a^{(i,j)}, f^{(i,j)})\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}^{(i,j)}\}_n$  by definition of multilevel block LT sequences, we have:

- $\{A_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  by Lemma 3.16, Proposition 2.21, and Remark 2.42;
- $\{A_{n,m}\}_n \sim_\sigma \kappa_m(\mathbf{x}, \boldsymbol{\theta})$  by Theorem 3.7;
- $\kappa_m(\mathbf{x}, \boldsymbol{\theta}) \rightarrow \kappa(\mathbf{x}, \boldsymbol{\theta})$  a.e. (and hence also in measure) as  $m \rightarrow \infty$  by Lemma 2.10 because each  $a^{(i,j)}(\mathbf{x})$  is Riemann-integrable.

We conclude that  $\{A_n\}_n \sim_\sigma \kappa(\mathbf{x}, \boldsymbol{\theta})$  by Theorem 2.34 and Remark 2.41.  $\square$

**THEOREM 3.18.** *If  $\{A_n^{(i,j)}\}_n \sim_{\text{LT}} a^{(i,j)}(\mathbf{x})f^{(i,j)}(\boldsymbol{\theta})$ , for  $i = 1, \dots, p$  and  $j = 1, \dots, q_i$ , then*

$$\left\{ \Re \left( \sum_{i=1}^p \prod_{j=1}^{q_i} A_n^{(i,j)} \right) \right\}_n \sim_\lambda \Re \left( \sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta}) \right).$$

*Proof.* The proof is essentially the same as the proof of Theorem 3.17. Define the matrices  $A_n$ ,  $A_{n,m}$  and the functions  $\kappa(\mathbf{x}, \boldsymbol{\theta})$ ,  $\kappa_m(\mathbf{x}, \boldsymbol{\theta})$  as in the proof of Theorem 3.17. Since  $\{LT_{n,s}^m(a^{(i,j)}, f^{(i,j)})\}_n \xrightarrow{\text{a.c.s.}} \{A_n^{(i,j)}\}_n$  by definition of multilevel block LT sequences, we have:

- $\{\Re(A_{n,m})\}_n \xrightarrow{\text{a.c.s.}} \{\Re(A_n)\}_n$  by Lemma 3.16, Proposition 2.21, and Remark 2.42;
- $\{\Re(A_{n,m})\}_n \sim_\lambda \Re(\kappa_m(\mathbf{x}, \boldsymbol{\theta}))$  by Theorem 3.8;
- $\Re(\kappa_m(\mathbf{x}, \boldsymbol{\theta})) \rightarrow \Re(\kappa(\mathbf{x}, \boldsymbol{\theta}))$  a.e. (and hence also in measure) as  $m \rightarrow \infty$  by Lemma 2.10 because each  $a^{(i,j)}(\mathbf{x})$  is Riemann-integrable.

We conclude that  $\{\Re(A_n)\}_n \sim_\lambda \Re(\kappa(\mathbf{x}, \boldsymbol{\theta}))$  by Theorem 2.35 and Remark 2.41.  $\square$

**3.5. Algebraic properties of multilevel block LT sequences.** Proposition 3.19 collects the most elementary algebraic properties of multilevel block LT sequences, which follow from Definition 3.10, the properties of the multilevel block LT operator (see (3.6)–(3.8)), and Remark 2.42.

**PROPOSITION 3.19.** *The following properties hold.*

- If  $\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ , then  $\{A_n^*\}_n \sim_{\text{LT}} \bar{a}(\mathbf{x})\bar{f}(\boldsymbol{\theta}) = (a(\mathbf{x})f(\boldsymbol{\theta}))^*$ .
- If  $\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ , then  $\{\alpha A_n\}_n \sim_{\text{LT}} \alpha a(\mathbf{x})f(\boldsymbol{\theta})$  for all  $\alpha \in \mathbb{C}$ .
- If  $\{A_n^{(i)}\}_n \sim_{\text{LT}} a(\mathbf{x})f_i(\boldsymbol{\theta})$ ,  $i = 1, \dots, r$ , then  $\{\sum_{i=1}^r A_n^{(i)}\}_n \sim_{\text{LT}} a(\mathbf{x})\sum_{i=1}^r f_i(\boldsymbol{\theta})$ .
- If  $\{A_n^{(i)}\}_n \sim_{\text{LT}} a_i(\mathbf{x})f(\boldsymbol{\theta})$ ,  $i = 1, \dots, r$ , then  $\{\sum_{i=1}^r A_n^{(i)}\}_n \sim_{\text{LT}} \sum_{i=1}^r a_i(\mathbf{x})f(\boldsymbol{\theta})$ .

In Theorem 3.20, we show, under mild assumptions, that the product of multilevel block LT sequences is again a multilevel block LT sequence with symbol given by the product of the symbols.

**THEOREM 3.20.** *Suppose that  $\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$  and  $\{\tilde{A}_n\}_n \sim_{\text{LT}} \tilde{a}(\mathbf{x})\tilde{f}(\boldsymbol{\theta})$ , where  $f \in L^p([-\pi, \pi]^d, N(\mathbf{s}))$ ,  $\tilde{f} \in L^q([-\pi, \pi]^d, N(\mathbf{s}))$ , and  $1 \leq p, q \leq \infty$  are conjugate exponents. Then*

$$\{A_n \tilde{A}_n\}_n \sim_{\text{LT}} a(\mathbf{x})\tilde{a}(\mathbf{x})f(\boldsymbol{\theta})\tilde{f}(\boldsymbol{\theta}).$$

*Proof.* By Lemma 3.16, every multilevel block LT sequence is s.u., so in particular  $\{A_n\}_n$  and  $\{\tilde{A}_n\}_n$  are s.u. Since, by definition of multilevel block LT sequences,

$$\begin{aligned} \{LT_{n,s}^m(a, f)\}_n &\xrightarrow{\text{a.c.s.}} \{A_n\}_n \text{ as } m \rightarrow \infty, \\ \{LT_{n,s}^m(\tilde{a}, \tilde{f})\}_n &\xrightarrow{\text{a.c.s.}} \{\tilde{A}_n\}_n \text{ as } m \rightarrow \infty, \end{aligned}$$

Remark 2.42 yields

$$\{LT_{n,s}^m(a, f)LT_{n,s}^m(\tilde{a}, \tilde{f})\}_n \xrightarrow{\text{a.c.s.}} \{A_n \tilde{A}_n\}_n \text{ as } m \rightarrow \infty.$$

Using Proposition 3.6, especially (3.11), we obtain

$$\{LT_{n,s}^m(a\tilde{a}, f\tilde{f})\}_n \xrightarrow{\text{a.c.s.}} \{A_n \tilde{A}_n\}_n \text{ as } m \rightarrow \infty,$$

hence  $\{A_n \tilde{A}_n\}_n \sim_{\text{LT}} a(\mathbf{x}) \tilde{a}(\mathbf{x}) f(\boldsymbol{\theta}) \tilde{f}(\boldsymbol{\theta})$ .  $\square$

As a consequence of Theorems 3.14, 3.15, and 3.20, we immediately obtain the following result.

**THEOREM 3.21.** *If  $\mathbf{s} \in \mathbb{N}^d$ ,  $a : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable, and  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ , then we have  $\{\Gamma_{\mathbf{n}, \mathbf{s}}^T D_{\mathbf{n}, \mathbf{s}}(a) \Gamma_{\mathbf{n}, \mathbf{s}} T_{\mathbf{n}}(f)\}_n = \{D_{\mathbf{n}}(a I_{N(\mathbf{s})}) T_{\mathbf{n}}(f)\}_n \sim_{\text{LT}} a(\mathbf{x}) f(\boldsymbol{\theta})$  for every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ .*

### 3.6. Characterizations and simplified definition of multilevel block LT sequences.

Theorem 3.21 shows that, for any  $a, f$  as in Definition 3.10, there always exists a  $d$ -level  $s$ -block LT sequence  $\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x}) f(\boldsymbol{\theta})$ . Indeed, it suffices to take  $A_n = D_{\mathbf{n}}(a I_{N(\mathbf{s})}) T_{\mathbf{n}}(f)$ . Theorem 3.22 shows that the sequences of the form  $\{D_{\mathbf{n}}(a I_{N(\mathbf{s})}) T_{\mathbf{n}}(f)\}_n$  play a central role in the world of multilevel block LT sequences. Indeed,

$$\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x}) f(\boldsymbol{\theta}) \iff d_{\text{a.c.s.}}(\{A_n\}_n, \{D_{\mathbf{n}}(a I_{N(\mathbf{s})}) T_{\mathbf{n}}(f)\}_n) = 0.$$

**THEOREM 3.22.** *Let  $\{A_n\}_n$  be a  $d$ -level  $N(\mathbf{s})$ -block matrix-sequence, let  $a : [0, 1]^d \rightarrow \mathbb{C}$  be a Riemann-integrable function, and let  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . The following conditions are equivalent.*

1.  $\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x}) f(\boldsymbol{\theta})$ .
2. For all sequences  $\{a_m\}_m, \{f_m\}_m, \{\{A_n^{(m)}\}_n\}_m$  such that
  - $a_m : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable and  $f_m \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ ,
  - $a_m(\mathbf{x}) f_m(\boldsymbol{\theta}) \rightarrow a(\mathbf{x}) f(\boldsymbol{\theta})$  in measure,
  - $\{A_n^{(m)}\}_n \sim_{\text{LT}} a_m(\mathbf{x}) f_m(\boldsymbol{\theta})$ ,
 we have  $\{A_n^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .
3. There exist sequences  $\{a_m\}_m, \{f_m\}_m$  such that
  - $a_m : [0, 1]^d \rightarrow \mathbb{C}$  is continuous,  $\|a_m\|_{\infty} \leq \|a\|_{L^{\infty}}$  for all  $m$ , and  $a_m \rightarrow a$  a.e.,
  - $f_m : [-\pi, \pi]^d \rightarrow \mathbb{C}^{N(\mathbf{s}) \times N(\mathbf{s})}$  is a matrix-valued trigonometric polynomial with  $\|(f_m)_{\alpha\beta}\|_{\infty} \leq \text{ess sup}_{[-\pi, \pi]^d} |f_{\alpha\beta}|$  for all  $m$  and all  $\alpha, \beta = 1, \dots, s$ , and  $f_m \rightarrow f$  a.e. and in  $L^1([-\pi, \pi]^d, N(\mathbf{s}))$ ,
  - $\{D_{\mathbf{n}}(a_m I_{N(\mathbf{s})}) T_{\mathbf{n}}(f_m)\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .
4. There exist sequences  $\{a_m\}_m, \{f_m\}_m, \{\{A_n^{(m)}\}_n\}_m$  such that
  - $a_m : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable and  $f_m \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ ,
  - $a_m(\mathbf{x}) f_m(\boldsymbol{\theta}) \rightarrow a(\mathbf{x}) f(\boldsymbol{\theta})$  in measure,
  - $\{A_n^{(m)}\}_n \sim_{\text{LT}} a_m(\mathbf{x}) f_m(\boldsymbol{\theta})$  and  $\{A_n^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .
5.  $\{D_{\mathbf{n}}(a I_{N(\mathbf{s})}) T_{\mathbf{n}}(f)\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .
6.  $A_n = D_{\mathbf{n}}(a I_{N(\mathbf{s})}) T_{\mathbf{n}}(f) + Z_n$  for every  $n$ , where  $\{Z_n\}_n$  is zero-distributed.

*Proof.* (1  $\implies$  2) Suppose that  $\{A_n\}_n \sim_{\text{LT}} a(\mathbf{x}) f(\boldsymbol{\theta})$ , and let  $\{a_m\}_m, \{f_m\}_m, \{\{A_n^{(m)}\}_n\}_m$  be sequences with the properties specified in item 2. By Theorems 2.33 and 3.17, we have

$$\begin{aligned} d_{\text{a.c.s.}}(\{A_n^{(m)}\}_n, \{A_n\}_n) &= p_{\text{a.c.s.}}(\{A_n^{(m)} - A_n\}_n) \\ &= p_{\text{measure}}(a_m(\mathbf{x}) f_m(\boldsymbol{\theta}) - a(\mathbf{x}) f(\boldsymbol{\theta})) \\ &= d_{\text{measure}}(a_m(\mathbf{x}) f_m(\boldsymbol{\theta}), a(\mathbf{x}) f(\boldsymbol{\theta})), \end{aligned}$$

which tends to 0 as  $m \rightarrow \infty$  because  $a_m(\mathbf{x}) f_m(\boldsymbol{\theta}) \rightarrow a(\mathbf{x}) f(\boldsymbol{\theta})$  in measure.

(2  $\implies$  3) Since any Riemann-integrable function is bounded by definition, we have  $a \in L^{\infty}([0, 1]^d)$ . Hence, by [40, Theorem 2.2], there exists a sequence of continuous functions  $a_m : [0, 1]^d \rightarrow \mathbb{C}$  such that  $\|a_m\|_{\infty} \leq \|a\|_{L^{\infty}}$  for all  $m$  and  $a_m \rightarrow a$  a.e. The

sequence  $\{a_m\}_m$  satisfies the properties in item 3. Note also that  $a_m \rightarrow a$  in  $L^1([0, 1]^d)$  by the dominated convergence theorem.

Since  $f \in L^1([-\pi, \pi]^d, N(\mathbf{s}))$ , by [41, Lemma 2.3] there exists a sequence of matrix-valued trigonometric polynomials  $f_m : [-\pi, \pi]^d \rightarrow \mathbb{C}^{N(\mathbf{s}) \times N(\mathbf{s})}$  which satisfy the estimate  $\|(f_m)_{\alpha\beta}\|_\infty \leq \text{ess sup}_{[-\pi, \pi]^d} |f_{\alpha\beta}|$  for all  $m$  and all  $\alpha, \beta = 1, \dots, N(\mathbf{s})$ , and  $f_m \rightarrow f$  a.e. and in  $L^1([-\pi, \pi]^d, N(\mathbf{s}))$ . The sequence  $\{f_m\}_m$  satisfies the properties in item 3.

Since  $a_m(\mathbf{x})f_m(\boldsymbol{\theta}) \rightarrow a(\mathbf{x})f(\boldsymbol{\theta})$  in  $L^1([0, 1]^d \times [-\pi, \pi]^d, N(\mathbf{s}))$  (and hence also in measure), item 2 and Theorem 3.21 imply that  $\{D_{\mathbf{n}}(a_m I_{N(\mathbf{s})})T_{\mathbf{n}}(f_m)\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}\}_n$ , and the proof is complete.

(3  $\implies$  4) Simply note that, under the assumptions in item 3,  $a_m(\mathbf{x})f_m(\boldsymbol{\theta}) \rightarrow a(\mathbf{x})f(\boldsymbol{\theta})$  in measure, and  $\{D_{\mathbf{n}}(a_m I_{N(\mathbf{s})})T_{\mathbf{n}}(f_m)\}_n \sim_{\text{LT}} a_m(\mathbf{x})f_m(\boldsymbol{\theta})$  by Theorem 3.21.

(4  $\implies$  1) Let  $\{B_{\mathbf{n}}\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ . We can take, e.g.,  $B_{\mathbf{n}} = D_{\mathbf{n}}(a I_{N(\mathbf{s})})T_{\mathbf{n}}(f)$  thanks to Theorem 3.21. By Theorems 2.33 and 3.17,

$$\begin{aligned} d_{\text{a.c.s.}}(\{A_{\mathbf{n}}^{(m)}\}_n, \{B_{\mathbf{n}}\}_n) &= p_{\text{a.c.s.}}(\{A_{\mathbf{n}}^{(m)} - B_{\mathbf{n}}\}_n) \\ &= p_{\text{measure}}(a_m(\mathbf{x})f_m(\boldsymbol{\theta}) - a(\mathbf{x})f(\boldsymbol{\theta})) \\ &= d_{\text{measure}}(a_m(\mathbf{x})f_m(\boldsymbol{\theta}), a(\mathbf{x})f(\boldsymbol{\theta})), \end{aligned}$$

which tends to 0 as  $m \rightarrow \infty$  because  $a_m(\mathbf{x})f_m(\boldsymbol{\theta}) \rightarrow a(\mathbf{x})f(\boldsymbol{\theta})$  in measure. Thus,  $\{A_{\mathbf{n}}^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{B_{\mathbf{n}}\}_n$ , and since  $\{A_{\mathbf{n}}^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}\}_n$  by assumption, we conclude that  $d_{\text{a.c.s.}}(\{A_{\mathbf{n}}\}_n, \{B_{\mathbf{n}}\}_n) = 0$ . Considering that  $\{LT_{\mathbf{n}, \mathbf{s}}^m(a, f)\}_n \xrightarrow{\text{a.c.s.}} \{B_{\mathbf{n}}\}_n$ , we obtain  $\{LT_{\mathbf{n}, \mathbf{s}}^m(a, f)\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}\}_n$ , i.e.,  $\{A_{\mathbf{n}}\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ .

(5  $\iff$  6) Item 5 is equivalent to  $d_{\text{a.c.s.}}(\{A_{\mathbf{n}}\}_n, \{D_{\mathbf{n}}(a I_{N(\mathbf{s})})T_{\mathbf{n}}(f)\}_n) = 0$ , which, by Theorem 2.32, is equivalent to  $\{A_{\mathbf{n}} - D_{\mathbf{n}}(a I_{N(\mathbf{s})})T_{\mathbf{n}}(f)\}_n \sim_{\sigma} 0$ .

(2  $\implies$  5) Obvious (take  $a_m = a$ ,  $f_m = f$  and  $A_{\mathbf{n}}^{(m)} = D_{\mathbf{n}}(a I_{N(\mathbf{s})})T_{\mathbf{n}}(f)$ ).

(5  $\implies$  4) Obvious (take  $a_m = a$ ,  $f_m = f$  and  $A_{\mathbf{n}}^{(m)} = D_{\mathbf{n}}(a I_{N(\mathbf{s})})T_{\mathbf{n}}(f)$ ).  $\square$

REMARK 3.23. Suppose that  $\mathbf{r}, \mathbf{s} \in \mathbb{N}^d$  are such that  $N(\mathbf{r}) = N(\mathbf{s})$ , and let  $\{A_{\mathbf{n}}\}_n$  be a  $d$ -level  $\mathbf{s}$ -block LT sequence with symbol  $a(\mathbf{x})f(\boldsymbol{\theta})$ . Then,  $\{A_{\mathbf{n}}\}_n$  is also a  $d$ -level  $\mathbf{r}$ -block LT sequence with symbol  $a(\mathbf{x})f(\boldsymbol{\theta})$ . Indeed, by Theorem 3.22, there exists a zero-distributed sequence  $\{Z_{\mathbf{n}}\}_n$  such that

$$A_{\mathbf{n}} = D_{\mathbf{n}}(a I_{N(\mathbf{s})})T_{\mathbf{n}}(f) + Z_{\mathbf{n}} = D_{\mathbf{n}}(a I_{N(\mathbf{r})})T_{\mathbf{n}}(f) + Z_{\mathbf{n}},$$

and so  $\{A_{\mathbf{n}}\}_n$  is a  $d$ -level  $\mathbf{r}$ -block LT sequence with symbol  $a(\mathbf{x})f(\boldsymbol{\theta})$ . This remark shows that the notion of  $d$ -level  $\mathbf{s}$ -block LT sequences depends on  $N(\mathbf{s})$  but not on  $\mathbf{s}$ . In other words, the set of  $d$ -level  $\mathbf{s}$ -block LT sequences coincides with the set of  $d$ -level  $\mathbf{r}$ -block LT sequences whenever  $N(\mathbf{s}) = N(\mathbf{r})$ . We can therefore give a simplified definition of multilevel block LT sequences.

DEFINITION 3.24 (Multilevel block locally Toeplitz sequence). *Let  $\{A_{\mathbf{n}}\}_n$  be a  $d$ -level  $\mathbf{s}$ -block matrix-sequence, let  $a : [0, 1]^d \rightarrow \mathbb{C}$  be Riemann-integrable, and  $f \in L^1([-\pi, \pi]^d, \mathbf{s})$ . We say that  $\{A_{\mathbf{n}}\}_n$  is a  $d$ -level  $\mathbf{s}$ -block locally Toeplitz (LT) sequence with symbol  $a(\mathbf{x})f(\boldsymbol{\theta})$ , and we write  $\{A_{\mathbf{n}}\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ , if  $\{A_{\mathbf{n}} - D_{\mathbf{n}}(a I_{\mathbf{s}})T_{\mathbf{n}}(f)\}_n$  is zero-distributed.*

A  $d$ -level  $\mathbf{s}$ -block LT sequence in the sense of Definition 3.10 is a  $d$ -level  $N(\mathbf{s})$ -block LT sequence in the sense of Definition 3.24. Vice versa, a  $d$ -level  $\mathbf{s}$ -block LT sequence in the sense of Definition 3.24 is a  $d$ -level  $\mathbf{s}$ -block LT sequence in the sense of Definition 3.10 for all  $\mathbf{s} \in \mathbb{N}^d$  such that  $N(\mathbf{s}) = \mathbf{s}$ . From now on, unless specified otherwise, whenever we write a relation such as  $\{A_{\mathbf{n}}\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})$ , it is understood that  $\{A_{\mathbf{n}}\}_n$  is a  $d$ -level  $\mathbf{s}$ -block LT sequence as in Definition 3.24, so in particular  $a : [0, 1]^d \rightarrow \mathbb{C}$  is Riemann-integrable and  $f \in L^1([-\pi, \pi]^d, \mathbf{s})$ .

**4. Multilevel block generalized locally Toeplitz sequences.** In this chapter we develop the theory of multilevel block GLT sequences, by correcting/extending the results in [8, 46] and [63, Section 3.3].

**4.1. Equivalent definitions of multilevel block GLT sequences.** Multilevel block GLT sequences can be defined in several different ways. We begin with what we may call the “classical definition”.

**DEFINITION 4.1** (Multilevel block generalized locally Toeplitz sequences). *Let  $\{A_n\}_n$  be a  $d$ -level  $s$ -block matrix-sequence, and let  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  be measurable. We say that  $\{A_n\}_n$  is a ( $d$ -level  $s$ -block) generalized locally Toeplitz (GLT) sequence with symbol  $\kappa$ , and we write  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ , if the following condition is met.*

*For every  $m \in \mathbb{N}$  there exists a finite number of  $d$ -level  $s$ -block LT sequences*

$$\{A_{n,m}^{(i,j)}\}_n \sim_{\text{LT}} a_m^{(i,j)}(\mathbf{x}) f_m^{(i,j)}(\boldsymbol{\theta}), \quad i = 1, \dots, N_m, \quad j = 1, \dots, M_{m,i},$$

*such that:*

- $\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} a_m^{(i,j)}(\mathbf{x}) f_m^{(i,j)}(\boldsymbol{\theta}) \rightarrow \kappa(\mathbf{x}, \boldsymbol{\theta})$  in measure;

- $\{\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} A_{n,m}^{(i,j)}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

In what follows, unless specified otherwise, whenever we write a relation such as  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ , it is understood that  $\{A_n\}_n$  is a  $d$ -level  $s$ -block GLT sequence as in Definition 4.1, so in particular  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  is measurable.

**REMARK 4.2.** It is clear that any sum of products of  $d$ -level  $s$ -block LT sequences is a  $d$ -level  $s$ -block GLT sequence. More precisely, if  $\{A_n^{(i,j)}\}_n \sim_{\text{LT}} a^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta})$ , for  $i = 1, \dots, p$  and  $j = 1, \dots, q_i$ , then

$$\left\{ \sum_{i=1}^p \prod_{j=1}^{q_i} A_n^{(i,j)} \right\}_n \sim_{\text{GLT}} \sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}(\mathbf{x}) f^{(i,j)}(\boldsymbol{\theta}).$$

**REMARK 4.3.** If  $a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$ , then

$$D_n(a) = \sum_{i,j=1}^s D_n(a_{ij} E_{ij}^{(s)}) = \sum_{i,j=1}^s D_n(a_{ij} I_s) T_n(E_{ij}^{(s)}).$$

Hence, by Remark 4.2 and Theorem 3.21 (applied with any  $s \in \mathbb{N}^d$  such that  $N(\mathbf{s}) = s$ ), if  $a$  is Riemann-integrable, then

$$\{D_n(a)\}_n \sim_{\text{GLT}} \sum_{i,j=1}^s a_{ij}(\mathbf{x}) E_{ij}^{(s)} = a(\mathbf{x}).$$

**REMARK 4.4.** Let  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{B_n\}_n \sim_{\text{GLT}} \xi$ . Then,  $\{A_n^*\}_n \sim_{\text{GLT}} \kappa^*$  and  $\{\alpha A_n + \beta B_n\}_n \sim_{\text{GLT}} \alpha \kappa + \beta \xi$  for all  $\alpha, \beta \in \mathbb{C}$ . This follows immediately from Definition 4.1, Proposition 3.19, and Theorem 2.37.

In the remainder of this section, we present another equivalent definition of multilevel block GLT sequences, which is illuminating for many purposes. Fix a sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ , and set

$$\mathcal{E}^{(s)} = \{\{A_n\}_n : A_n \in \mathbb{C}^{N(\mathbf{n})s \times N(\mathbf{n})s}\},$$

$$\mathfrak{M}_d^{(s)} = \{\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s} : \kappa \text{ is measurable}\},$$

$$\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)} = \{(\{A_n\}_n, \kappa) : \{A_n\}_n \in \mathcal{E}^{(s)}, \kappa \in \mathfrak{M}_d^{(s)}\}.$$

We make the following observations.



- $\mathcal{E}^{(s)}$  is a  $*$ -algebra with respect to the natural pointwise operations (i.e.,  $\{A_n\}_n^* = \{A_n^*\}_n$ ,  $\alpha\{A_n\}_n + \beta\{B_n\}_n = \{\alpha A_n + \beta B_n\}_n$ ,  $\{A_n\}_n\{B_n\}_n = \{A_n B_n\}_n$ ), and it is also a pseudometric space with respect to the pseudometric  $d_{\text{a.c.s.}}$  inducing the a.c.s. topology  $\tau_{\text{a.c.s.}}$ .
- $\mathfrak{M}_d^{(s)}$  is a  $*$ -algebra with respect to the natural pointwise operations, and it is also a pseudometric space with respect to the pseudometric  $d_{\text{measure}}$  inducing the topology  $\tau_{\text{measure}}$  of convergence in measure.
- $\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}$  is a  $*$ -algebra with respect to the natural pointwise operations (that is,  $(\{A_n\}_n, \kappa)^* = (\{A_n^*\}_n, \kappa^*)$ ,  $\alpha(\{A_n\}_n, \kappa) + \beta(\{B_n\}_n, \xi) = (\{\alpha A_n + \beta B_n\}_n, \alpha\kappa + \beta\xi)$ ,  $(\{A_n\}_n, \kappa)(\{B_n\}_n, \xi) = (\{A_n B_n\}_n, \kappa\xi)$ ), and it is also a pseudometric space with respect to the product pseudometric

$$(d_{\text{a.c.s.}} \times d_{\text{measure}})((\{A_n\}_n, \kappa), (\{B_n\}_n, \xi)) = d_{\text{a.c.s.}}(\{A_n\}_n, \{B_n\}_n) + d_{\text{measure}}(\kappa, \xi)$$

inducing the product topology  $\tau_{\text{a.c.s.}} \times \tau_{\text{measure}}$ .

Let  $\mathcal{A}_d^{(s)}$  be the  $*$ -subalgebra of  $\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}$  generated by the set of “ $d$ -level  $s$ -block LT pairs”

$$\mathcal{L}_d^{(s)} = \{(\{A_n\}_n, a(\mathbf{x})f(\boldsymbol{\theta})) \in \mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)} : \{A_n\}_n \sim_{\text{LT}} a(\mathbf{x})f(\boldsymbol{\theta})\}.$$

Using Proposition 3.19, it is not difficult to see that

$$\mathcal{A}_d^{(s)} = \left\{ \left( \sum_{i=1}^p \prod_{j=1}^{q_i} A_n^{(i,j)}, \sum_{i=1}^p \prod_{j=1}^{q_i} a^{(i,j)}(\mathbf{x})f^{(i,j)}(\boldsymbol{\theta}) \right) : \right. \\ \left. p, q_1, \dots, q_p \in \mathbb{N}, \quad \{A_n^{(i,j)}\}_n \sim_{\text{LT}} a^{(i,j)}(\mathbf{x})f^{(i,j)}(\boldsymbol{\theta}) \text{ for all } i, j \right\}.$$

We can now reformulate Definition 4.1 as follows.

**DEFINITION 4.5** (Multilevel block generalized locally Toeplitz sequence). *Let  $\{A_n\}_n$  be a  $d$ -level  $s$ -block matrix-sequence, and let  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  be measurable. We say that  $\{A_n\}_n$  is a ( $d$ -level  $s$ -block) generalized locally Toeplitz (GLT) sequence with symbol  $\kappa$ , and we write  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  if the pair  $(\{A_n\}_n, \kappa)$  belongs to the closure of  $\mathcal{A}_d^{(s)}$  in  $(\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}, d_{\text{a.c.s.}} \times d_{\text{measure}})$ . In other words, the set of “ $d$ -level  $s$ -block GLT pairs”*

$$(4.1) \quad \mathcal{G}_d^{(s)} = \{(\{A_n\}_n, \kappa) \in \mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)} : \{A_n\}_n \sim_{\text{GLT}} \kappa\}$$

is defined as the closure of  $\mathcal{A}_d^{(s)}$  in  $(\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}, d_{\text{a.c.s.}} \times d_{\text{measure}})$ .

In the light of this algebraic-topological definition of multilevel block GLT sequences, the following theorem is obvious.

**THEOREM 4.6.** *Let  $\{A_n\}_n$  be a  $d$ -level  $s$ -block matrix-sequence, and let  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  be measurable. Suppose that:*

1.  $\{B_{n,m}\}_n \sim_{\text{GLT}} \kappa_m$  for every  $m$ ;
2.  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ ;
3.  $\kappa_m \rightarrow \kappa$  in measure.

Then  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ .

#### 4.2. Singular value and spectral distribution of multilevel block GLT sequences.

In this section we prove the main singular value and eigenvalue distribution results for multilevel block GLT sequences.

**THEOREM 4.7.** *If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ , then  $\{A_n\}_n \sim_{\sigma} \kappa$ .*

*Proof.* By definition, for every  $m \in \mathbb{N}$  there exist multilevel block LT sequences  $\{A_{n,m}^{(i,j)}\}_n \sim_{\text{LT}} a_m^{(i,j)}(\mathbf{x})f_m^{(i,j)}(\boldsymbol{\theta})$ ,  $i = 1, \dots, N_m$ ,  $j = 1, \dots, M_{m,i}$ , such that

- $\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} a_m^{(i,j)}(\mathbf{x}) f_m^{(i,j)}(\boldsymbol{\theta}) \rightarrow \kappa(\mathbf{x}, \boldsymbol{\theta})$  in measure,
- $\{\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} A_{n,m}^{(i,j)}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

Moreover, by Theorem 3.17,

- $\{\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} A_{n,m}^{(i,j)}\}_n \sim_{\sigma} \sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} a_m^{(i,j)}(\mathbf{x}) f_m^{(i,j)}(\boldsymbol{\theta})$ .

We conclude that  $\{A_n\}_n \sim_{\sigma} \kappa$  by Theorem 2.34.  $\square$

REMARK 4.8. Any multilevel block GLT sequence  $\{A_n\}_n$  is s.u. This follows from Theorem 4.7 and Proposition 2.22.

Using Theorem 4.7 we now show that the symbol of a multilevel block GLT sequence is essentially unique and that the symbol of a multilevel block GLT sequence formed by Hermitian matrices is Hermitian a.e.

PROPOSITION 4.9. *If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{A_n\}_n \sim_{\text{GLT}} \xi$ , then  $\kappa = \xi$  a.e.*

*Proof.* By Remark 4.4 we have  $\{O_{N(n)s}\}_n = \{A_n - A_n\}_n \sim_{\text{GLT}} \kappa - \xi$ . Hence, by Theorem 4.7, we also have  $\{O_{N(n)s}\}_n \sim_{\sigma} \kappa - \xi$ , i.e.,

$$F(0) = \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \int_{[0, 1]^d} \frac{\sum_{j=1}^{N(s)} F(\sigma_j(\kappa(\mathbf{x}, \boldsymbol{\theta}) - \xi(\mathbf{x}, \boldsymbol{\theta})))}{N(s)} dx d\boldsymbol{\theta}, \quad \forall F \in C_c(\mathbb{R}).$$

We conclude that  $\kappa - \xi = O_s$  a.e. by Remark 2.9.  $\square$

PROPOSITION 4.10. *If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and the  $A_n$  are Hermitian, then  $\kappa$  is Hermitian a.e.*

*Proof.* Since the  $A_n$  are Hermitian, by Remark 4.4 we have  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{A_n\}_n \sim_{\text{GLT}} \kappa^*$ . Thus, by Proposition 4.9,  $\kappa = \kappa^*$  a.e.  $\square$

THEOREM 4.11. *If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and the  $A_n$  are Hermitian, then  $\{A_n\}_n \sim_{\lambda} \kappa$ .*

*Proof.* By definition, for every  $m \in \mathbb{N}$  there exist multilevel block LT sequences  $\{A_{n,m}^{(i,j)}\}_n \sim_{\text{LT}} a_m^{(i,j)}(\mathbf{x}) f_m^{(i,j)}(\boldsymbol{\theta})$ ,  $i = 1, \dots, N_m$ ,  $j = 1, \dots, M_{m,i}$ , such that

- $\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} a_m^{(i,j)}(\mathbf{x}) f_m^{(i,j)}(\boldsymbol{\theta}) \rightarrow \kappa(\mathbf{x}, \boldsymbol{\theta})$  in measure,
- $\{\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} A_{n,m}^{(i,j)}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

Thus:

- $\{\Re(\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} A_{n,m}^{(i,j)})\}_n \xrightarrow{\text{a.c.s.}} \{\Re(A_n)\}_n$  by Theorem 2.37;
- $\{\Re(\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} A_{n,m}^{(i,j)})\}_n \sim_{\lambda} \Re(\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} a_m^{(i,j)}(\mathbf{x}) f_m^{(i,j)}(\boldsymbol{\theta}))$  by Theorem 3.18;
- $\Re(\sum_{i=1}^{N_m} \prod_{j=1}^{M_{m,i}} a_m^{(i,j)}(\mathbf{x}) f_m^{(i,j)}(\boldsymbol{\theta})) \rightarrow \Re(\kappa(\mathbf{x}, \boldsymbol{\theta}))$  in measure.

We conclude that  $\{\Re(A_n)\}_n \sim_{\lambda} \Re(\kappa)$  by Theorem 2.35. Since the matrices  $A_n$  are Hermitian, we have  $\Re(A_n) = A_n$  and  $\Re(\kappa) = \kappa$  a.e. by Proposition 4.10. Hence, the spectral distribution  $\{\Re(A_n)\}_n \sim_{\lambda} \Re(\kappa)$  yields  $\{A_n\}_n \sim_{\lambda} \kappa$ .  $\square$

We end this section with a spectral distribution result for (compressions of) multilevel block GLT sequences formed by perturbed Hermitian matrices.

THEOREM 4.12. *Let  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $A_n = X_n + Y_n$ . Assume that*

1. every  $X_n$  is Hermitian,
2.  $\|Y_n\|_2 = o(\sqrt{N(n)})$ .

*Then  $\{P_n^* A_n P_n\}_n \sim_{\sigma, \lambda} \kappa$  for every sequence  $\{P_n\}_n$  such that  $P_n \in \mathbb{C}^{N(n)s \times \delta_n}$ ,  $P_n^* P_n = I_{\delta_n}$ ,  $\delta_n \leq N(n)s$ , and  $\delta_n / (N(n)s) \rightarrow 1$ . In particular,  $\{A_n\}_n \sim_{\sigma, \lambda} \kappa$  (take  $P_n = I_{N(n)s}$ ).*

*Proof.*  $\{Y_n\}_n$  is zero-distributed by Theorem 2.18, so  $\{Y_n\}_n \sim_{\text{GLT}} O_s$  by Theorem 3.11 (applied with any  $s \in \mathbb{N}^d$  such that  $N(s) = s$ ). Since  $X_n = A_n - Y_n$  and the matrices  $X_n$  are Hermitian, we have  $\{X_n\}_n \sim_{\text{GLT}} \kappa$  by Remark 4.4 and  $\{X_n\}_n \sim_{\sigma, \lambda} \kappa$  by Theorems 4.7 and 4.11. Consider the decomposition

$$P_n^* A_n P_n = P_n^* X_n P_n + P_n^* Y_n P_n.$$

Let  $\hat{P}_n = [P_n \mid O]$  be the  $N(n)s \times N(n)s$  matrix obtained from  $P_n$  by adding zeros, and observe that  $\hat{P}_n^* B \hat{P}_n = P_n^* B P_n \oplus O_{N(n)s - \delta_n}$  for all  $B \in \mathbb{C}^{N(n)s \times N(n)s}$ . We have:

- $\{P_n^* X_n P_n\}_n \sim_{\sigma, \lambda} \kappa$  by Theorem 2.30;
- $\|P_n^* Y_n P_n\|_2 = \|\hat{P}_n^* Y_n \hat{P}_n\|_2 \leq \|\hat{P}_n^*\| \|Y_n\|_2 \|\hat{P}_n\| = \|Y_n\|_2 = o(\sqrt{N(n)}) = o(\sqrt{\delta_n})$ , where the inequality follows from (2.6), while the second equality follows from the fact that the nonzero singular values of  $\hat{P}_n$  and  $\hat{P}_n^*$  are all equal to 1 due to the relation  $\hat{P}_n^* \hat{P}_n = P_n^* P_n \oplus O_{N(n)s - \delta_n} = I_{\delta_n} \oplus O_{N(n)s - \delta_n}$ . In particular,  $\{P_n^* Y_n P_n\}_n$  is zero-distributed by Theorem 2.18.

We conclude that  $d_{\text{a.c.s.}}(\{P_n^* X_n P_n\}_n, \{P_n^* A_n P_n\}_n) = 0$ , the singular value distribution  $\{P_n^* A_n P_n\}_n \sim_{\sigma} \kappa$  follows from Remark 2.36, and the spectral distribution  $\{P_n^* A_n P_n\}_n \sim_{\lambda} \kappa$  follows from Theorem 2.28.  $\square$

As shown in this section, the theory of multilevel block GLT sequences allows one to compute the spectral distribution of a given multilevel block GLT sequence  $\{A_n\}_n$  in the case where  $A_n$  is either a Hermitian matrix or a (small) perturbation of a Hermitian matrix. Similar results can be formulated in the case where  $A_n$  is either a normal matrix or a (small) perturbation of a normal matrix. For the case of purely non-normal matrices, the main spectral distribution results obtained so far are [30, Theorems 9 and 10], which are based on the previous works [29, 32] and, especially, on Tilli's pioneering paper [68].

**4.3. Multilevel block GLT sequences and matrix-valued measurable functions.** We prove in this section that every  $s \times s$  matrix-valued measurable function  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  is the symbol of a suitable  $d$ -level  $s$ -block GLT sequence  $\{A_n\}_n$ . In combination with results already proved before, this will allow us to show that the map associating with each  $d$ -level  $s$ -block GLT sequence  $\{A_n\}_n$  its symbol  $\kappa$  is an isometry with respect to  $d_{\text{a.c.s.}}$  on the space  $\mathcal{G}^{(s)}$  of  $d$ -level  $s$ -block GLT sequences and  $d_{\text{measure}}$  on the space  $\mathfrak{M}_d^{(s)}$  of  $s \times s$  matrix-valued measurable functions defined on  $[0, 1]^d \times [-\pi, \pi]^d$ .

LEMMA 4.13. *Let  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  be a sequence of  $d$ -indices such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ , and let  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  be any measurable function. Then there exists a sequence of  $d$ -level  $s$ -block GLT pairs  $(\{A_{\mathbf{n}, m}\}_n, \kappa_m)$  such that  $\kappa_m \rightarrow \kappa$  in measure.*

*Proof.* By [41, Lemma 2.4], for every  $\alpha, \beta = 1, \dots, s$  there exists a sequence of measurable functions  $\kappa_{m, \alpha\beta} : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}$  such that  $\kappa_{m, \alpha\beta}$  is of the form

$$\kappa_{m, \alpha\beta}(\mathbf{x}, \boldsymbol{\theta}) = \sum_{j=-N_m}^{N_m} a_{j, m, \alpha\beta}(\mathbf{x}) e^{ij \cdot \boldsymbol{\theta}},$$

with  $N_m \in \mathbb{N}^d$  and  $a_{j, m, \alpha\beta} : [0, 1]^d \rightarrow \mathbb{C}$  belonging to  $C^\infty([0, 1]^d)$ , and  $\kappa_{m, \alpha\beta} \rightarrow \kappa_{\alpha\beta}$  a.e. Take

$$\begin{aligned} \kappa_m(\mathbf{x}, \boldsymbol{\theta}) &= [\kappa_{m, \alpha\beta}(\mathbf{x}, \boldsymbol{\theta})]_{\alpha, \beta=1}^s = \sum_{\alpha, \beta=1}^s \kappa_{m, \alpha\beta}(\mathbf{x}, \boldsymbol{\theta}) E_{\alpha\beta}^{(s)} \\ &= \sum_{\alpha, \beta=1}^s \sum_{j=-N_m}^{N_m} a_{j, m, \alpha\beta}(\mathbf{x}) e^{ij \cdot \boldsymbol{\theta}} E_{\alpha\beta}^{(s)}, \\ A_{\mathbf{n}, m} &= \sum_{\alpha, \beta=1}^s \sum_{j=-N_m}^{N_m} D_{\mathbf{n}}(a_{j, m, \alpha\beta} I_s) T_{\mathbf{n}}(e^{ij \cdot \boldsymbol{\theta}} E_{\alpha\beta}^{(s)}). \end{aligned}$$

Clearly,  $\kappa_m \rightarrow \kappa$  a.e. and hence also in measure. Moreover,  $\{A_{n,m}\}_n \sim_{\text{GLT}} \kappa_m$  by Theorem 3.21 and Remark 4.4.  $\square$

**THEOREM 4.14.** *Let  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  be a sequence of  $d$ -indices such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ , and let  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  be any measurable function. Then there exists a  $d$ -level  $s$ -block GLT sequence  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ .*

*Proof.* By Lemma 4.13, there exist  $d$ -level  $s$ -block GLT sequences  $\{A_{n,m}\}_n \sim_{\text{GLT}} \kappa_m$  such that  $\kappa_m \rightarrow \kappa$  in measure. Since  $\{\kappa_m\}_m$  converges in measure, it is a Cauchy sequence with respect to the pseudometric  $d_{\text{measure}}$  inducing the convergence in measure. It follows that  $\{\{A_{n,m}\}_n\}_m$  is a Cauchy sequence with respect to the pseudometric  $d_{\text{a.c.s.}}$  inducing the a.c.s. convergence because for each pair of indices  $m, m'$ ,  $\{A_{n,m} - A_{n,m'}\}_n \sim_{\text{GLT}} \kappa_m - \kappa_{m'}$  by Remark 4.4,  $\{A_{n,m} - A_{n,m'}\}_n \sim_{\sigma} \kappa_m - \kappa_{m'}$  by Theorem 4.7, and

$$\begin{aligned} d_{\text{a.c.s.}}(\{A_{n,m}\}_n, \{A_{n,m'}\}_n) &= p_{\text{a.c.s.}}(\{A_{n,m} - A_{n,m'}\}_n) \\ &= p_{\text{measure}}(\kappa_m - \kappa_{m'}) \\ &= d_{\text{measure}}(\kappa_m, \kappa_{m'}) \end{aligned}$$

by Theorem 2.33. Since  $d_{\text{a.c.s.}}$  is complete on the space  $\mathcal{E}^{(s)}$  of  $d$ -level  $s$ -block matrix-sequences corresponding to the sequence of  $d$ -indices  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  (by Theorem 2.32), we infer that  $\{A_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  for some  $d$ -level  $s$ -block matrix-sequence  $\{A_n\}_n$ . We conclude that  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  by Theorem 4.6.  $\square$

**REMARK 4.15** (Isometry between  $d$ -level  $s$ -block GLT sequences and  $s \times s$  matrix-valued measurable functions). With the notation used in Definition 4.5, suppose we identify two  $d$ -level  $s$ -block matrix-sequences  $\{A_n\}_n, \{B_n\}_n \in \mathcal{E}^{(s)}$  whenever their difference  $\{A_n - B_n\}_n$  is zero-distributed and two measurable functions  $\kappa, \xi \in \mathfrak{M}_d^{(s)}$  whenever their difference  $\kappa - \xi$  equals  $O_s$  a.e. Let  $\mathcal{G}^{(s)} \subset \mathcal{E}^{(s)}$  be the subspace of  $\mathcal{E}^{(s)}$  consisting of  $d$ -level  $s$ -block GLT sequences and consider the application  $\mathbb{J}$  from  $\mathcal{G}^{(s)}$  to  $\mathfrak{M}_d^{(s)}$  that associates with each  $\{A_n\}_n \in \mathcal{G}^{(s)}$  its symbol  $\kappa \in \mathfrak{M}_d^{(s)}$ . This application is well-defined by Proposition 4.9. Moreover, if  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ , then  $\{B_n\}_n \sim_{\text{GLT}} \kappa$  is equivalent to  $\{A_n - B_n\}_n \sim_{\sigma} 0$  by Remark 4.4 and Theorems 3.11 and 4.7. This means that, after identification of two  $d$ -level  $s$ -block matrix-sequences  $\{A_n\}_n, \{B_n\}_n$  whenever  $\{A_n - B_n\}_n \sim_{\sigma} 0$ , the application  $\mathbb{J}$  is still well-defined and, moreover, it is injective. By Theorem 4.14,  $\mathbb{J}$  is also surjective. Finally, by Theorems 2.33 and 4.7, given any gauge function  $\varphi$ , if  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{B_n\}_n \sim_{\text{GLT}} \xi$ , then

$$d_{\text{a.c.s.}}^{\varphi}(\{A_n\}_n, \{B_n\}_n) = p_{\text{a.c.s.}}^{\varphi}(\{A_n - B_n\}_n) = p_{\text{measure}}^{\varphi}(\kappa - \xi) = d_{\text{measure}}^{\varphi}(\kappa, \xi),$$

which means that  $\mathbb{J}$  is an isometry with respect to  $d_{\text{a.c.s.}}^{\varphi}$  on  $\mathcal{G}^{(s)}$  and  $d_{\text{measure}}^{\varphi}$  on  $\mathfrak{M}_d^{(s)}$ .

**4.4. The multilevel block GLT algebra.** The next theorems are of fundamental importance. In particular, the first one shows that the set of  $d$ -level  $s$ -block GLT pairs  $\mathcal{G}_d^{(s)}$  defined in (4.1) is a  $*$ -subalgebra of  $\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}$ .

**THEOREM 4.16.** *Let  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{B_n\}_n \sim_{\text{GLT}} \xi$ . Then*

1.  $\{A_n^*\}_n \sim_{\text{GLT}} \kappa^*$ ;
2.  $\{\alpha A_n + \beta B_n\}_n \sim_{\text{GLT}} \alpha \kappa + \beta \xi$  for all  $\alpha, \beta \in \mathbb{C}$ ;
3.  $\{A_n B_n\}_n \sim_{\text{GLT}} \kappa \xi$ .

*Proof.* The first two statements have already been settled before (see Remark 4.4). We prove the third one. By Definition 4.5, there exist  $(\{A_{n,m}\}_n, \kappa_m), (\{B_{n,m}\}_n, \xi_m) \in \mathcal{A}_d^{(s)}$  such that  $(\{A_{n,m}\}_n, \kappa_m) \rightarrow (\{A_n\}_n, \kappa)$  and  $(\{B_{n,m}\}_n, \xi_m) \rightarrow (\{B_n\}_n, \xi)$  in the space  $(\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}, \tau_{\text{a.c.s.}} \times \tau_{\text{measure}})$ , i.e.,

- $\{A_{\mathbf{n},m}\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}\}_n$  and  $\{B_{\mathbf{n},m}\}_n \xrightarrow{\text{a.c.s.}} \{B_{\mathbf{n}}\}_n$ ,
- $\kappa_m \rightarrow \kappa$  in measure and  $\xi_m \rightarrow \xi$  in measure.

Considering that every multilevel block GLT sequence is s.u. (see Remark 4.8), from Theorem 2.37 and Lemma 2.7 we obtain

- $\{A_{\mathbf{n},m}B_{\mathbf{n},m}\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}B_{\mathbf{n}}\}_n$ ,
- $\kappa_m\xi_m \rightarrow \kappa\xi$  in measure.

Since  $(\{A_{\mathbf{n},m}B_{\mathbf{n},m}\}_n, \kappa_m\xi_m) \in \mathcal{A}_d^{(s)}$ , by definition we have  $\{A_{\mathbf{n}}B_{\mathbf{n}}\}_n \sim_{\text{GLT}} \kappa\xi$ .  $\square$

**THEOREM 4.17.** *If  $\{A_{\mathbf{n}}\}_n \sim_{\text{GLT}} \kappa$  and  $\kappa$  is invertible a.e., then  $\{A_{\mathbf{n}}^\dagger\}_n \sim_{\text{GLT}} \kappa^{-1}$ .*

*Proof.* Since  $\kappa$  is measurable and invertible a.e., its inverse  $\kappa^{-1}$  is a well-defined measurable function. Thus, by Theorem 4.14, there exists a  $d$ -level  $s$ -block GLT sequence  $\{B_{\mathbf{n}}\}_n \sim_{\text{GLT}} \kappa^{-1}$ . By Theorem 4.16 we have  $\{B_{\mathbf{n}}A_{\mathbf{n}} - I_{N(\mathbf{n})s}\}_n \sim_{\text{GLT}} \kappa^{-1}\kappa - I_s$ , which implies that  $\{B_{\mathbf{n}}A_{\mathbf{n}} - I_{N(\mathbf{n})s}\}_n \sim_\sigma 0$  by Theorem 4.7, since  $\kappa^{-1}\kappa - I_s = O_s$  a.e. Hence,  $\{B_{\mathbf{n}}A_{\mathbf{n}}\}_n \xrightarrow{\text{a.c.s.}} \{I_{N(\mathbf{n})s}\}_n$ . Since  $\kappa$  is invertible a.e.,  $\{A_{\mathbf{n}}\}_n$  is s.v. by Theorem 4.7 and Proposition 2.27. It follows that  $A_{\mathbf{n}}^\dagger$  is s.u. (see Remark 2.25), and so, by Theorem 2.37,  $\{B_{\mathbf{n}}A_{\mathbf{n}}A_{\mathbf{n}}^\dagger\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}^\dagger\}_n$ , i.e.,

$$(4.2) \quad \{B_{\mathbf{n}}A_{\mathbf{n}}A_{\mathbf{n}}^\dagger - A_{\mathbf{n}}^\dagger\}_n \sim_\sigma 0.$$

Now we observe that, by definition of  $A_{\mathbf{n}}^\dagger$ ,

$$A_{\mathbf{n}}A_{\mathbf{n}}^\dagger = I_{N(\mathbf{n})s} + S_{\mathbf{n}}, \quad \text{rank}(S_{\mathbf{n}}) = \#\{i \in \{1, \dots, N(\mathbf{n})s\} : \sigma_i(A_{\mathbf{n}}) = 0\}.$$

Considering that  $\{A_{\mathbf{n}}\}_n$  is s.v., we have

$$\lim_{n \rightarrow \infty} \frac{\text{rank}(S_{\mathbf{n}})}{N(\mathbf{n})} = 0.$$

Hence, from (4.2) we obtain

$$\{B_{\mathbf{n}} + Z_{\mathbf{n}} - A_{\mathbf{n}}^\dagger\}_n \sim_\sigma 0,$$

where  $Z_{\mathbf{n}} = B_{\mathbf{n}}S_{\mathbf{n}}$  is zero-distributed by Theorem 2.18. Thus,  $A_{\mathbf{n}}^\dagger = B_{\mathbf{n}} + Z_{\mathbf{n}}$ , and it follows from Theorems 3.11 and 4.16 that  $\{A_{\mathbf{n}}^\dagger\}_n \sim_{\text{GLT}} \kappa^{-1}$ .  $\square$

**THEOREM 4.18.** *If  $\{A_{\mathbf{n}}\}_n \sim_{\text{GLT}} \kappa$  and each  $A_{\mathbf{n}}$  is Hermitian, then  $\{f(A_{\mathbf{n}})\}_n \sim_{\text{GLT}} f(\kappa)$  for any continuous function  $f : \mathbb{C} \rightarrow \mathbb{C}$ .*

*Proof.* Since every  $A_{\mathbf{n}}$  is Hermitian by assumption and  $\kappa$  is Hermitian a.e. by Proposition 4.10, it suffices to prove the theorem for real continuous functions  $f : \mathbb{R} \rightarrow \mathbb{R}$ . Indeed, suppose that we have proved the theorem for this kind of functions, and let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be any continuous complex function. Denote by  $\alpha, \beta : \mathbb{R} \rightarrow \mathbb{R}$  the real and imaginary parts of the restriction of  $f$  to  $\mathbb{R}$ . Then,  $\alpha, \beta$  are continuous functions such that  $f(x) = \alpha(x) + i\beta(x)$  for all  $x \in \mathbb{R}$ , and since the eigenvalues of  $A_{\mathbf{n}}$  are real, we have  $f(A_{\mathbf{n}}) = \alpha(A_{\mathbf{n}}) + i\beta(A_{\mathbf{n}})$ . In view of the relations  $\{\alpha(A_{\mathbf{n}})\}_n \sim_{\text{GLT}} \alpha(\kappa)$  and  $\{\beta(A_{\mathbf{n}})\}_n \sim_{\text{GLT}} \beta(\kappa)$ , Theorem 4.16 yields  $\{f(A_{\mathbf{n}})\}_n \sim_{\text{GLT}} \alpha(\kappa) + i\beta(\kappa)$ , and so  $\{A_{\mathbf{n}}\}_n \sim_{\text{GLT}} f(\kappa)$  because  $\alpha(\kappa) + i\beta(\kappa) = f(\kappa)$  a.e. as  $\kappa$  is Hermitian a.e.

Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a real continuous function. For each  $M > 0$ , let  $\{p_{m,M}\}_m$  be a sequence of polynomials that converges uniformly to  $f$  over  $[-M, M]$ :

$$\lim_{m \rightarrow \infty} \|f - p_{m,M}\|_{\infty, [-M, M]} = 0.$$

Note that such a sequence exists by the Weierstrass theorem; see, e.g., [58, Theorem 7.26]. By replacing  $p_{m,M}$  with  $p_{m,M} + f(0) - p_{m,M}(0)$  if necessary, we may assume, without loss

of generality, that  $p_{m,M}(0) = f(0)$ . Since any multilevel block GLT sequence is s.u. (by Remark 4.8), the sequence  $\{A_n\}_n$  is s.u. Hence, by Remark 2.23, for all  $M > 0$  there exists  $n_M$  such that, for  $n \geq n_M$ ,

$$(4.3) \quad A_n = \hat{A}_{n,M} + \tilde{A}_{n,M}, \quad \text{rank}(\hat{A}_{n,M}) \leq r(M)N(\mathbf{n}), \quad \|\tilde{A}_{n,M}\| \leq M,$$

where  $r(M) \rightarrow 0$  as  $M \rightarrow \infty$ , the matrices  $\hat{A}_{n,M}$  and  $\tilde{A}_{n,M}$  are Hermitian, and for all functions  $g : \mathbb{R} \rightarrow \mathbb{R}$  satisfying  $g(0) = 0$  we have

$$g(\hat{A}_{n,M} + \tilde{A}_{n,M}) = g(\hat{A}_{n,M}) + g(\tilde{A}_{n,M}).$$

Taking into account that  $(f - p_{m,M})(0) = 0$ , for every  $M > 0$ , every  $m$  and every  $n \geq n_M$ , we can write

$$\begin{aligned} f(A_n) &= p_{m,M}(A_n) + f(A_n) - p_{m,M}(A_n) \\ &= p_{m,M}(A_n) + (f - p_{m,M})(\hat{A}_{n,M}) + (f - p_{m,M})(\tilde{A}_{n,M}) \\ &= p_{m,M}(A_n) + R_{n,m,M} + N_{n,m,M}, \end{aligned}$$

where, in view of (4.3),  $R_{n,m,M} = (f - p_{m,M})(\hat{A}_{n,M})$  and  $N_{n,m,M} = (f - p_{m,M})(\tilde{A}_{n,M})$  satisfy

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

Choose a sequence  $\{M_m\}_m$  such that

$$(4.4) \quad M_m \rightarrow \infty, \quad \|f - p_{m,M_m}\|_{\infty,[-M_m,M_m]} \rightarrow 0.$$

Then, for every  $m$  and every  $n \geq n_{M_m}$ ,

$$\begin{aligned} f(A_n) &= p_{m,M_m}(A_n) + R_{n,m,M_m} + N_{n,m,M_m}, \\ \text{rank}(R_{n,m,M_m}) &\leq r(M_m)N(\mathbf{n}), \\ \|N_{n,m,M_m}\| &\leq \|f - p_{m,M_m}\|_{\infty,[-M_m,M_m]}, \end{aligned}$$

which implies that

$$\{p_{m,M_m}(A_n)\}_n \xrightarrow{\text{a.c.s.}} \{f(A_n)\}_n.$$

Moreover, by Theorem 4.16,

$$\{p_{m,M_m}(A_n)\}_n \sim_{\text{GLT}} p_{m,M_m}(\kappa).$$

Finally, by (4.4),

$$\begin{aligned} \|f(\kappa) - p_{m,M_m}(\kappa)\| &= \max_{i=1,\dots,s} |(f - p_{m,M_m})(\lambda_i(\kappa))| \\ &\leq \|f - p_{m,M_m}\|_{\infty,[-\|\kappa\|,\|\kappa\|]} \rightarrow 0 \quad \text{a.e.}, \end{aligned}$$

which implies that

$$p_{m,M_m}(\kappa) \rightarrow f(\kappa) \quad \text{a.e.}$$

All the hypotheses of Theorem 4.6 are then satisfied and  $\{f(A_n)\}_n \sim_{\text{GLT}} f(\kappa)$ .  $\square$

**4.5. Topological density results for multilevel block GLT sequences.** Theorem 4.6 admits the following simple but important converse, which will allow us to prove the main result of this section (Theorem 4.20).

**THEOREM 4.19.** *Let  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{B_{n,m}\}_n \sim_{\text{GLT}} \kappa_m$  for every  $m$ . Then*

$$\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n \iff \kappa_m \rightarrow \kappa \text{ in measure.}$$

*Proof.* By Remark 4.15,

$$d_{\text{a.c.s.}}(\{A_n\}_n, \{B_{n,m}\}_n) = d_{\text{measure}}(\kappa, \kappa_m),$$

and the proof is complete.  $\square$

**THEOREM 4.20.** *Let  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ . Then, for all functions  $a_{i,m}, f_{i,m}, i = 1, \dots, N_m$ , such that*

- $a_{i,m} : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$  is Riemann-integrable and  $f_{i,m} \in L^1([-\pi, \pi]^d, s)$ ,
- $\sum_{i=1}^{N_m} a_{i,m}(\mathbf{x}) f_{i,m}(\boldsymbol{\theta}) \rightarrow \kappa(\mathbf{x}, \boldsymbol{\theta})$  in measure,

*we have  $\{\sum_{i=1}^{N_m} D_n(a_{i,m}) T_n(f_{i,m})\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ . In particular,  $\{A_n\}_n$  admits an a.c.s. of the form*

$$(4.5) \quad \left\{ \left\{ \sum_{\alpha, \beta=1}^s \sum_{j=-N_m}^{N_m} D_n(a_{\alpha\beta, j}^{(m)}) T_n(e^{ij \cdot \boldsymbol{\theta}} E_{\alpha\beta}^{(s)}) \right\}_n \right\}_m \\ = \left\{ \left\{ \sum_{j=-N_m}^{N_m} D_n(a_j^{(m)}) T_n(e^{ij \cdot \boldsymbol{\theta}} I_s) \right\}_n \right\}_m,$$

where  $N_m \in \mathbb{N}^d$ ,  $a_{\alpha\beta, j}^{(m)} : [0, 1]^d \rightarrow \mathbb{C}$  belongs to  $C^\infty([0, 1]^d)$ ,

$$a_j^{(m)} = \sum_{\alpha, \beta=1}^s a_{\alpha\beta, j}^{(m)} E_{\alpha\beta}^{(s)} = [a_{\alpha\beta, j}^{(m)}]_{\alpha, \beta=1}^s,$$

and

$$\sum_{\alpha, \beta=1}^s \sum_{j=-N_m}^{N_m} a_{\alpha\beta, j}^{(m)}(\mathbf{x}) e^{ij \cdot \boldsymbol{\theta}} E_{\alpha\beta}^{(s)} = \sum_{j=-N_m}^{N_m} a_j^{(m)}(\mathbf{x}) e^{ij \cdot \boldsymbol{\theta}} \rightarrow \kappa(\mathbf{x}, \boldsymbol{\theta}) \text{ a.e.}$$

*Proof.* Let  $a_{i,m}, f_{i,m}, i = 1, \dots, N_m$ , be functions with the properties specified in the statement of the theorem. Then

$$\left\{ \sum_{i=1}^{N_m} D_n(a_{i,m}) T_n(f_{i,m}) \right\}_n \sim_{\text{GLT}} \sum_{i=1}^{N_m} a_{i,m}(\mathbf{x}) f_{i,m}(\boldsymbol{\theta})$$

by Remark 4.3, Theorem 3.15 (applied with any  $\mathbf{s} \in \mathbb{N}^d$  such that  $N(\mathbf{s}) = s$ ), and Theorem 4.16. Therefore, the convergence

$$\left\{ \sum_{i=1}^{N_m} D_n(a_{i,m}) T_n(f_{i,m}) \right\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$$

follows from Theorem 4.19 applied with

$$B_{n,m} = \sum_{i=1}^{N_m} D_n(a_{i,m}) T_n(f_{i,m}), \quad \kappa_m(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{N_m} a_{i,m}(\mathbf{x}) f_{i,m}(\boldsymbol{\theta}).$$

To obtain for  $\{A_n\}_n$  an a.c.s. of the form (4.5), we use the result of this theorem in combination with [41, Lemma 2.4]. The details are as follows. By [41, Lemma 2.4], for every  $\alpha, \beta = 1, \dots, s$  there exists a sequence of measurable functions  $\kappa_{m,\alpha\beta} : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}$  such that  $\kappa_{m,\alpha\beta}$  is of the form

$$\kappa_{m,\alpha\beta}(\mathbf{x}, \boldsymbol{\theta}) = \sum_{j=-N_m}^{N_m} a_{\alpha\beta,j}^{(m)}(\mathbf{x}) e^{j \cdot \boldsymbol{\theta}},$$

with  $N_m \in \mathbb{N}^d$  and  $a_{\alpha\beta,j}^{(m)} : [0, 1]^d \rightarrow \mathbb{C}$  belonging to  $C^\infty([0, 1]^d)$ , and  $\kappa_{m,\alpha\beta} \rightarrow \kappa_{\alpha\beta}$  a.e. Then,  $\sum_{\alpha,\beta=1}^s \kappa_{m,\alpha\beta} E_{\alpha\beta}^{(s)} \rightarrow \sum_{\alpha,\beta=1}^s \kappa_{\alpha\beta} E_{\alpha\beta}^{(s)} = \kappa$  a.e., and so, by the result of this theorem,

$$\left\{ \sum_{\alpha,\beta=1}^s \sum_{j=-N_m}^{N_m} D_n(a_{\alpha\beta,j}^{(m)} I_s) T_n(e^{j \cdot \boldsymbol{\theta}} E_{\alpha\beta}^{(s)}) \right\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n. \quad \square$$

REMARK 4.21 (Topological density in the space of multilevel block GLT sequences). With the notation introduced in Section 4.1, we recall that the set of  $d$ -level  $s$ -block GLT pairs

$$\mathcal{G}_d^{(s)} = \{(\{A_n\}_n, \kappa) \in \mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)} : \{A_n\}_n \sim_{\text{GLT}} \kappa\}$$

is closed in  $(\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}, \tau_{\text{a.c.s.}} \times \tau_{\text{measure}})$  by Definition 4.5. Consider the subset of  $\mathcal{G}_d^{(s)}$  consisting of the  $d$ -level  $s$ -block GLT pairs of the form

$$\left( \sum_{i=1}^N D_n(a_i I_s) T_n(f_i), \sum_{i=1}^N a_i(\mathbf{x}) f_i(\boldsymbol{\theta}) \right),$$

where  $a_i : [0, 1]^d \rightarrow \mathbb{C}$  belongs to  $C^\infty([0, 1]^d)$ ,  $f_i$  is a trigonometric monomial in  $\{e^{j \cdot \boldsymbol{\theta}} E_{\alpha\beta}^{(s)} : \mathbf{j} \in \mathbb{Z}^d, 1 \leq \alpha, \beta \leq s\}$  for all  $i = 1, \dots, N$ , and  $N \in \mathbb{N}$ . Then, according to Theorem 4.20, this subset is dense in  $\mathcal{G}_d^{(s)}$ , i.e., its closure in  $(\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}, \tau_{\text{a.c.s.}} \times \tau_{\text{measure}})$  coincides precisely with  $\mathcal{G}_d^{(s)}$ .

**4.6. Characterizations of multilevel block GLT sequences.** The next result is a characterization theorem for multilevel block GLT sequences. All the provided characterizations have already been proved before, but it is anyway useful to collect them in a single statement.

THEOREM 4.22. *Let  $\{A_n\}_n$  be a  $d$ -level  $s$ -block matrix-sequence, and let  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  be a measurable function. The following conditions are equivalent:*

1.  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ .
2. For all sequences  $\{\kappa_m\}_m, \{\{B_{n,m}\}_n\}_m$  such that
  - $\{B_{n,m}\}_n \sim_{\text{GLT}} \kappa_m$  for every  $m$ ,
  - $\kappa_m \rightarrow \kappa$  in measure,
 we have  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .
3. There exist functions  $a_{i,m}, f_{i,m}$ ,  $i = 1, \dots, N_m$ , such that
  - $a_{i,m} : [0, 1]^d \rightarrow \mathbb{C}$  belongs to  $C^\infty([0, 1]^d)$  and  $f_{i,m}$  is a trigonometric monomial belonging to  $\{e^{j \cdot \boldsymbol{\theta}} E_{\alpha\beta}^{(s)} : \mathbf{j} \in \mathbb{Z}^d, 1 \leq \alpha, \beta \leq s\}$ ,
  - $\sum_{i=1}^{N_m} a_{i,m}(\mathbf{x}) f_{i,m}(\boldsymbol{\theta}) \rightarrow \kappa(\mathbf{x}, \boldsymbol{\theta})$  a.e.,
  - $\{\sum_{i=1}^{N_m} D_n(a_{i,m} I_s) T_n(f_{i,m})\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .
4. There exist sequences  $\{\kappa_m\}_m, \{\{B_{n,m}\}_n\}_m$  such that
  - $\{B_{n,m}\}_n \sim_{\text{GLT}} \kappa_m$  for every  $m$ ,



- $\kappa_m \rightarrow \kappa$  in measure,
- $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

*Proof.* The implication  $1 \implies 2$  follows from Theorem 4.19. The implication  $2 \implies 3$  follows from the observation that we can find functions  $a_{i,m}, f_{i,m}, i = 1, \dots, N_m$ , with the first two properties specified in item 3 (by [41, Lemma 2.4], as we have already seen in the proof of Theorem 4.20), and, moreover,  $\{\sum_{i=1}^{N_m} D_n(a_{i,m} I_s) T_n(f_{i,m})\}_n \sim_{\text{GLT}} \sum_{i=1}^{N_m} a_{i,m}(\mathbf{x}) f_{i,m}(\boldsymbol{\theta})$  (by Theorems 3.21 and 4.16). The implication  $3 \implies 4$  is obvious (it suffices to take  $B_{n,m} = \sum_{i=1}^{N_m} D_n(a_{i,m} I_s) T_n(f_{i,m})$  and  $\kappa_m(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{N_m} a_{i,m}(\mathbf{x}) f_{i,m}(\boldsymbol{\theta})$ ). Finally, the implication  $4 \implies 1$  is Theorem 4.6.  $\square$

**4.7. Sequences of multilevel block diagonal sampling matrices.** We have encountered in Section 3.3 and Remark 4.3 the three most important examples of multilevel block GLT sequences, i.e., zero-distributed sequences, multilevel block Toeplitz sequences, and sequences of multilevel block diagonal sampling matrices. Concerning the latter kind of sequences, we have proved that  $\{D_n(a)\}_n \sim_{\text{GLT}} a(\mathbf{x})$  whenever  $a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$  is Riemann-integrable. From a mathematical point of view, however, the GLT relation  $\{D_n(a)\}_n \sim_{\text{GLT}} a(\mathbf{x})$  makes sense for all measurable functions  $a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$ , and it is therefore natural to ask whether we can drop the Riemann-integrability assumption. In Theorem 4.23 we show that the relation  $\{D_n(a)\}_n \sim_{\text{GLT}} a(\mathbf{x})$  holds for all functions  $a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$  that are continuous a.e. in  $[0, 1]^d$ . Since a function  $a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$  is Riemann-integrable if and only if  $a$  is bounded and continuous a.e. (see Section 2.3.5), Theorem 4.23 is an extension of both Theorem 3.14 and Remark 4.3. More precisely, in Theorem 4.23 we are dropping the boundedness assumption.

**THEOREM 4.23.** *If  $a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$  is continuous a.e., then  $\{D_n(a)\}_n \sim_{\text{GLT}} a(\mathbf{x})$ .*

*Proof.* By looking at the decomposition of  $D_n(a)$  considered in Remark 4.3, it is immediately clear that, in order to prove the theorem, it is enough to show that

$$\{D_n(a I_s)\}_n \sim_{\text{GLT}} a(\mathbf{x}) I_s$$

whenever  $a : [0, 1]^d \rightarrow \mathbb{C}$  is a scalar a.e. continuous function. Moreover, for an arbitrary a.e. continuous function  $a : [0, 1]^d \rightarrow \mathbb{C}$ , we can write  $a = \alpha_+ - \alpha_- + i\beta_+ - i\beta_-$ , where  $\alpha_{\pm}, \beta_{\pm} : [0, 1]^d \rightarrow \mathbb{R}$  are nonnegative a.e. continuous functions; simply take

$$\begin{aligned} \alpha_+ &= \max(\Re(a), 0), & \alpha_- &= -\min(\Re(a), 0), \\ \beta_+ &= \max(\Im(a), 0), & \beta_- &= -\min(\Im(a), 0). \end{aligned}$$

Hence, by Theorem 4.16 and the linearity of  $D_n(a I_s)$  with respect to its argument  $a$ , it suffices to prove the relation  $\{D_n(a I_s)\}_n \sim_{\text{GLT}} a(\mathbf{x})$  in the case where  $a : [0, 1]^d \rightarrow \mathbb{R}$  is a nonnegative a.e. continuous function.

Let  $a : [0, 1]^d \rightarrow [0, \infty)$  be a nonnegative a.e. continuous function. Denote by  $a_m$  the truncation of  $a$  at level  $m$ , i.e.,

$$a_m(\mathbf{x}) = \begin{cases} a(\mathbf{x}), & \text{if } a(\mathbf{x}) \leq m, \\ m, & \text{if } a(\mathbf{x}) > m. \end{cases}$$

Since  $a_m$  is bounded and continuous a.e.,  $a_m$  is Riemann-integrable, hence

$$\{D_n(a_m I_s)\}_n \sim_{\text{GLT}} a_m(\mathbf{x}) I_s$$

by Theorem 3.14. Moreover, it is clear that  $a_m \rightarrow a$  pointwise, so

$$a_m \rightarrow a \text{ in measure.}$$

We show that

$$\{D_{\mathbf{n}}(a_m I_s)\}_n \xrightarrow{\text{a.c.s.}} \{D_{\mathbf{n}}(a I_s)\}_n,$$

after which the application of Theorem 4.6 concludes the proof. In order to show that  $\{D_{\mathbf{n}}(a_m I_s)\}_n \xrightarrow{\text{a.c.s.}} \{D_{\mathbf{n}}(a I_s)\}_n$ , we observe that  $D_{\mathbf{n}}(a_m I_s) = D_{\mathbf{n}}(a_m) \otimes I_s$  and that  $D_{\mathbf{n}}(a I_s) = D_{\mathbf{n}}(a) \otimes I_s$ . Therefore, since the convergence  $\{D_{\mathbf{n}}(a_m)\}_n \xrightarrow{\text{a.c.s.}} \{D_{\mathbf{n}}(a)\}_n$  was proved in [41, pp. 101–102], it is immediately clear from the definition of a.c.s. and the properties of tensor products that  $\{D_{\mathbf{n}}(a_m I_s)\}_n \xrightarrow{\text{a.c.s.}} \{D_{\mathbf{n}}(a I_s)\}_n$ .  $\square$

**4.8. Sequences of block matrices with multilevel block GLT blocks.** This section is devoted to proving the following theorem, which says that, up to a suitable permutation transformation, a sequence of block matrices with multilevel block GLT blocks is a multilevel block GLT sequence whose symbol is obtained by “putting together” the symbols of each multilevel block GLT block. Recall that  $\Pi_{\mathbf{n},s,r}$  denotes the special permutation matrix (2.3).

**THEOREM 4.24.** *For  $i, j = 1, \dots, s$ , let  $\{A_{\mathbf{n},ij}\}_n$  be a  $d$ -level  $r$ -block GLT sequence with symbol  $\kappa_{ij} : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{r \times r}$ . Set  $A_{\mathbf{n}} = [A_{\mathbf{n},ij}]_{i,j=1}^s$  and  $\kappa = [\kappa_{ij}]_{i,j=1}^s$ . Then  $\{\Pi_{\mathbf{n},s,r} A_{\mathbf{n}} \Pi_{\mathbf{n},s,r}^T\}_n$  is a  $d$ -level  $rs$ -block GLT sequence with symbol  $\kappa$ .*

*Proof.* The proof consists of the following two steps.

*Step 1.* We first prove the theorem under the additional assumption that  $A_{\mathbf{n},ij}$  is of the form

$$(4.6) \quad A_{\mathbf{n},ij} = \sum_{\ell=1}^{L_{ij}} D_{\mathbf{n}}(a_{\ell,ij}) T_{\mathbf{n}}(f_{\ell,ij}),$$

where  $L_{ij} \in \mathbb{N}$ ,  $a_{\ell,ij} : [0, 1]^d \rightarrow \mathbb{C}^{r \times r}$  is continuous a.e., and  $f_{\ell,ij} : [-\pi, \pi]^d \rightarrow \mathbb{C}^{r \times r}$  belongs to  $L^1([-\pi, \pi]^d, r)$ . Note that the symbol of  $\{A_{\mathbf{n},ij}\}_n$  is

$$\kappa_{ij}(\mathbf{x}, \boldsymbol{\theta}) = \sum_{\ell=1}^{L_{ij}} a_{\ell,ij}(\mathbf{x}) f_{\ell,ij}(\boldsymbol{\theta}).$$

By setting  $L = \max_{i,j=1,\dots,s} L_{ij}$  and adding zero matrices of the form  $D_{\mathbf{n}}(O_r) T_{\mathbf{n}}(O_r)$  in the summation (4.6) whenever  $L_{ij} < L$ , we can assume, without loss of generality, that

$$A_{\mathbf{n},ij} = \sum_{\ell=1}^L D_{\mathbf{n}}(a_{\ell,ij}) T_{\mathbf{n}}(f_{\ell,ij}), \quad \kappa_{ij}(\mathbf{x}, \boldsymbol{\theta}) = \sum_{\ell=1}^L a_{\ell,ij}(\mathbf{x}) f_{\ell,ij}(\boldsymbol{\theta}),$$

with  $L$  independent of  $i, j$ . Then,

$$\begin{aligned} \Pi_{\mathbf{n},s,r} A_{\mathbf{n}} \Pi_{\mathbf{n},s,r}^T &= \sum_{\ell=1}^L \Pi_{\mathbf{n},s,r} [D_{\mathbf{n}}(a_{\ell,ij}) T_{\mathbf{n}}(f_{\ell,ij})]_{i,j=1}^s \Pi_{\mathbf{n},s,r}^T \\ &= \sum_{\ell=1}^L \Pi_{\mathbf{n},s,r} \left[ \sum_{i,j=1}^s E_{ij}^{(s)} \otimes D_{\mathbf{n}}(a_{\ell,ij}) T_{\mathbf{n}}(f_{\ell,ij}) \right] \Pi_{\mathbf{n},s,r}^T \\ &= \sum_{\ell=1}^L \Pi_{\mathbf{n},s,r} \left[ \sum_{i,j=1}^s (E_{ij}^{(s)} \otimes D_{\mathbf{n}}(a_{\ell,ij})) (I_s \otimes T_{\mathbf{n}}(f_{\ell,ij})) \right] \Pi_{\mathbf{n},s,r}^T \\ &= \sum_{\ell=1}^L \sum_{i,j=1}^s \Pi_{\mathbf{n},s,r} (E_{ij}^{(s)} \otimes D_{\mathbf{n}}(a_{\ell,ij})) \Pi_{\mathbf{n},s,r}^T \Pi_{\mathbf{n},s,r} (I_s \otimes T_{\mathbf{n}}(f_{\ell,ij})) \Pi_{\mathbf{n},s,r}^T. \end{aligned}$$

By Theorems 2.47 and 2.48,

$$\begin{aligned}\Pi_{\mathbf{n},s,r}(E_{ij}^{(s)} \otimes D_{\mathbf{n}}(a_{\ell,ij}))\Pi_{\mathbf{n},s,r}^T &= D_{\mathbf{n}}(E_{ij}^{(s)} \otimes a_{\ell,ij}), \\ \Pi_{\mathbf{n},s,r}(I_s \otimes T_{\mathbf{n}}(f_{\ell,ij}))\Pi_{\mathbf{n},s,r}^T &= T_{\mathbf{n}}(I_s \otimes f_{\ell,ij}).\end{aligned}$$

It follows that

$$\Pi_{\mathbf{n},s,r}A_{\mathbf{n}}\Pi_{\mathbf{n},s,r}^T = \sum_{\ell=1}^L \sum_{i,j=1}^s D_{\mathbf{n}}(E_{ij}^{(s)} \otimes a_{\ell,ij})T_{\mathbf{n}}(I_s \otimes f_{\ell,ij}).$$

Considering that  $E_{ij}^{(s)} \otimes a_{\ell,ij}$  and  $I_s \otimes f_{\ell,ij}$  are  $rs \times rs$  matrix-valued functions, Theorems 3.15, 4.16, and 4.23 imply that  $\{\Pi_{\mathbf{n},s,r}A_{\mathbf{n}}\Pi_{\mathbf{n},s,r}^T\}_n$  is a  $d$ -level  $rs$ -block GLT sequence with symbol

$$\kappa(\mathbf{x}, \boldsymbol{\theta}) = \sum_{\ell=1}^L \sum_{i,j=1}^s E_{ij}^{(s)} \otimes a_{\ell,ij}(\mathbf{x})f_{\ell,ij}(\boldsymbol{\theta}) = [\kappa_{ij}(\mathbf{x}, \boldsymbol{\theta})]_{i,j=1}^s.$$

*Step 2.* We now prove the theorem in its full generality. Since  $\{A_{\mathbf{n},ij}\}_n$  is a  $d$ -level  $r$ -block GLT sequence with symbol  $\kappa_{ij}$ , by Theorem 4.20, there exist functions  $a_{\ell,ij}^{(m)}, f_{\ell,ij}^{(m)}$ ,  $\ell = 1, \dots, L_{ij}^{(m)}$ , such that

- $a_{\ell,ij}^{(m)} : [0, 1]^d \rightarrow \mathbb{C}^{r \times r}$  belongs to  $C^\infty([0, 1]^d)$  and  $f_{\ell,ij}^{(m)} : [-\pi, \pi]^d \rightarrow \mathbb{C}^{r \times r}$  is a trigonometric polynomial,
  - $\kappa_{ij}^{(m)}(\mathbf{x}, \boldsymbol{\theta}) = \sum_{\ell=1}^{L_{ij}^{(m)}} a_{\ell,ij}^{(m)}(\mathbf{x})f_{\ell,ij}^{(m)}(\boldsymbol{\theta}) \rightarrow \kappa_{ij}(\mathbf{x}, \boldsymbol{\theta})$  a.e.,
  - $\{A_{\mathbf{n},ij}^{(m)} = \sum_{\ell=1}^{L_{ij}^{(m)}} D_{\mathbf{n}}(a_{\ell,ij}^{(m)})T_{\mathbf{n}}(f_{\ell,ij}^{(m)})\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n},ij}\}_n$ .
- Set  $A_{\mathbf{n}}^{(m)} = [A_{\mathbf{n},ij}^{(m)}]_{i,j=1}^s$  and  $\kappa^{(m)} = [\kappa_{ij}^{(m)}]_{i,j=1}^s$ . We have
- $\{\Pi_{\mathbf{n},s,r}A_{\mathbf{n}}^{(m)}\Pi_{\mathbf{n},s,r}^T\}_n \sim_{\text{GLT}} \kappa^{(m)}$  by Step 1;
  - $\kappa^{(m)} \rightarrow \kappa$  a.e. (and hence also in measure);
  - $\{\Pi_{\mathbf{n},s,r}A_{\mathbf{n}}^{(m)}\Pi_{\mathbf{n},s,r}^T\}_n \xrightarrow{\text{a.c.s.}} \{\Pi_{\mathbf{n},s,r}A_{\mathbf{n}}\Pi_{\mathbf{n},s,r}^T\}_n$  because  $\{A_{\mathbf{n}}^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{A_{\mathbf{n}}\}_n$  by Theorem 2.38.

We conclude that  $\{\Pi_{\mathbf{n},s,r}A_{\mathbf{n}}\Pi_{\mathbf{n},s,r}^T\}_n \sim_{\text{GLT}} \kappa$  by Theorem 4.6.  $\square$

**4.9. Further possible definitions of multilevel block GLT sequences.** In this section, we discuss a couple of possible alternative definitions of multilevel block GLT sequences. We will use the same notation as in Definition 4.5.

Fix a sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ . By Definition 4.5 and Theorem 4.16, the set of  $d$ -level  $s$ -block GLT pairs

$$\mathcal{G}_d^{(s)} = \{(\{A_{\mathbf{n}}\}_n, \kappa) : \{A_{\mathbf{n}}\}_n \sim_{\text{GLT}} \kappa\} \subseteq \mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}$$

is a closed  $*$ -subalgebra of  $\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}$ . By Theorems 3.11, 3.15, and 4.23,  $\mathcal{G}_d^{(s)}$  contains the set

$$\begin{aligned}\mathcal{B}_d^{(s)} &= \{(\{T_{\mathbf{n}}(f)\}_n, \kappa(\mathbf{x}, \boldsymbol{\theta}) = f(\boldsymbol{\theta})) : f \in L^1([-\pi, \pi]^d, s)\} \\ &\cup \{(\{D_{\mathbf{n}}(a)\}_n, \kappa(\mathbf{x}, \boldsymbol{\theta}) = a(\mathbf{x})) : a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s} \text{ is continuous a.e.}\} \\ &\cup \{(\{Z_{\mathbf{n}}\}_n, \kappa(\mathbf{x}, \boldsymbol{\theta}) = O_s) : \{Z_{\mathbf{n}}\}_n \sim_\sigma 0\}.\end{aligned}$$

By the results in Section 4.5, the algebra generated by  $\mathcal{B}_d^{(s)}$  is dense in  $\mathcal{G}_d^{(s)}$ . In conclusion,

the set of  $d$ -level  $s$ -block GLT pairs  $\mathcal{G}_d^{(s)}$  is the closed  $*$ -subalgebra of  $\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}$  generated by  $\mathcal{B}_d^{(s)}$ , i.e., the smallest closed  $*$ -subalgebra of  $\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}$  containing  $\mathcal{B}_d^{(s)}$ .

Looking more carefully at the results in Section 4.5, we also note that if we let

$$\begin{aligned} \mathcal{C}_d^{(s)} = & \{(\{D_{\mathbf{n}}(aI_s)\}_n, \kappa(\mathbf{x}, \boldsymbol{\theta}) = a(\mathbf{x})I_s) : a : [0, 1]^d \rightarrow \mathbb{C} \text{ belongs to } C^\infty([0, 1]^d)\} \\ & \cup \{(\{T_{\mathbf{n}}(e^{i\mathbf{j}\cdot\boldsymbol{\theta}} E_{\alpha\beta}^{(s)})\}_n, \kappa(\mathbf{x}, \boldsymbol{\theta}) = e^{i\mathbf{j}\cdot\boldsymbol{\theta}} E_{\alpha\beta}^{(s)}) : \mathbf{j} \in \mathbb{Z}^d, 1 \leq \alpha, \beta \leq s\}, \end{aligned}$$

then

the set of  $d$ -level  $s$ -block GLT pairs  $\mathcal{G}_d^{(s)}$  is the closure of the subalgebra of  $\mathcal{E}^{(s)} \times \mathfrak{M}_d^{(s)}$  generated by  $\mathcal{C}_d^{(s)}$ .

**5. Summary of the theory.** We conclude the theory of multilevel block GLT sequences by providing a *self-contained* summary, which contains *everything one needs to know* in order to understand the applications presented in the next chapter. It is assumed that anyone who reads this summary is aware of the notation and terminology used throughout this work, which will be only partially repeated here for the sake of brevity. The reader can find most of the notation and terminology in Section 2.1.

**Multi-index notation.** A multi-index  $\mathbf{i}$  of size  $d$ , also called a  $d$ -index, is a (row) vector in  $\mathbb{Z}^d$ ; its components are denoted by  $i_1, \dots, i_d$ .  $\mathbf{0}, \mathbf{1}, \mathbf{2}, \dots$  are the vectors of all zeros, all ones, all twos,  $\dots$  (their size will be clear from the context). For any  $d$ -index  $\mathbf{m}$ , we set  $N(\mathbf{m}) = \prod_{j=1}^d m_j$ , and we write  $\mathbf{m} \rightarrow \infty$  to indicate that  $\min(\mathbf{m}) \rightarrow \infty$ . The notation  $N(\boldsymbol{\alpha}) = \prod_{j=1}^d \alpha_j$  will be used for any vector  $\boldsymbol{\alpha}$  with  $d$  components and not only for  $d$ -indices. If  $\mathbf{h}, \mathbf{k}$  are  $d$ -indices, then  $\mathbf{h} \leq \mathbf{k}$  means that  $h_r \leq k_r$  for all  $r = 1, \dots, d$ . If  $\mathbf{h}, \mathbf{k}$  are  $d$ -indices such that  $\mathbf{h} \leq \mathbf{k}$ , then the multi-index (or  $d$ -index) range  $\mathbf{h}, \dots, \mathbf{k}$  is the set  $\{\mathbf{j} \in \mathbb{Z}^d : \mathbf{h} \leq \mathbf{j} \leq \mathbf{k}\}$ . We assume for this set the standard lexicographic ordering:

$$\left[ \dots \left[ \left[ (j_1, \dots, j_d) \right]_{j_d=h_d, \dots, k_d} \right]_{j_{d-1}=h_{d-1}, \dots, k_{d-1}} \dots \right]_{j_1=h_1, \dots, k_1}.$$

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

$$(h_1, h_2), (h_1, h_2 + 1), \dots, (h_1, k_2), (h_1 + 1, h_2), (h_1 + 1, h_2 + 1), \dots, (h_1 + 1, k_2), \\ \dots \dots \dots, (k_1, h_2), (k_1, h_2 + 1), \dots, (k_1, k_2).$$

When a multi-index  $\mathbf{j}$  varies over a multi-index range  $\mathbf{h}, \dots, \mathbf{k}$  (this is often written as  $\mathbf{j} = \mathbf{h}, \dots, \mathbf{k}$ ), it is understood that  $\mathbf{j}$  varies from  $\mathbf{h}$  to  $\mathbf{k}$  following the lexicographic ordering. For instance, if  $\mathbf{m} \in \mathbb{N}^d$  and we write  $\mathbf{x} = [x_i]_{i=1}^{\mathbf{m}}$ , then  $\mathbf{x}$  is a vector of size  $N(\mathbf{m})$  whose components  $x_i$ ,  $i = 1, \dots, \mathbf{m}$ , are ordered in accordance with the lexicographic ordering: the first component is  $x_1 = x_{(1, \dots, 1, 1)}$ , the second component is  $x_{(1, \dots, 1, 2)}$ , and so on until the last component, which is  $x_{\mathbf{m}} = x_{(m_1, \dots, m_d)}$ . Similarly, if  $X = [x_{ij}]_{i, j=1}^{\mathbf{m}}$ , then  $X$  is an  $N(\mathbf{m}) \times N(\mathbf{m})$  matrix whose components are indexed by a pair of  $d$ -indices  $\mathbf{i}, \mathbf{j}$ , both varying from  $\mathbf{1}$  to  $\mathbf{m}$  according to the lexicographic ordering. If  $\mathbf{h}, \mathbf{k}$  are  $d$ -indices such that  $\mathbf{h} \leq \mathbf{k}$ , then the notation  $\sum_{\mathbf{j}=\mathbf{h}}^{\mathbf{k}}$  indicates the summation over all  $\mathbf{j}$  in  $\mathbf{h}, \dots, \mathbf{k}$ . If  $\mathbf{i}, \mathbf{j}$  are  $d$ -indices, then  $\mathbf{i} \preceq \mathbf{j}$  means that  $\mathbf{i}$  precedes (or equals)  $\mathbf{j}$  in the lexicographic ordering (which is a total ordering on  $\mathbb{Z}^d$ ). Moreover, we define

$$\mathbf{i} \wedge \mathbf{j} = \begin{cases} \mathbf{i}, & \text{if } \mathbf{i} \preceq \mathbf{j}, \\ \mathbf{j}, & \text{if } \mathbf{i} \succ \mathbf{j}. \end{cases}$$

Note that  $i \wedge j$  is the minimum among  $i$  and  $j$  with respect to the lexicographic ordering. Operations involving  $d$ -indices that have no meaning in the vector space  $\mathbb{Z}^d$  must always be interpreted in the componentwise sense. For instance,  $ij = (i_1j_1, \dots, i_dj_d)$ ,  $\alpha i/j = (\alpha i_1/j_1, \dots, \alpha i_d/j_d)$  for all  $\alpha \in \mathbb{C}$ , etc.

**Matrix norms.** Here is a list of important inequalities involving  $p$ -norms and Schatten  $p$ -norms of matrices.

- N1.**  $\|X\| \leq \sqrt{|X|_1 |X|_\infty} \leq \max(|X|_1, |X|_\infty)$  for all  $X \in \mathbb{C}^{m \times m}$ .
- N2.**  $\|X\|_1 \leq \text{rank}(X) \|X\| \leq m \|X\|$  for all  $X \in \mathbb{C}^{m \times m}$ .
- N3.**  $\|X\|_1 \leq \sum_{i,j=1}^m |x_{ij}|$  for all  $X \in \mathbb{C}^{m \times m}$ .
- N4.**  $\|X\|_2 \leq \sqrt{\|X\| \|X\|_1}$  for all  $X \in \mathbb{C}^{m \times m}$ .
- N5.**  $\|AXB\|_p \leq \|A\| \|X\|_p \|B\|$  for all  $A, X, B \in \mathbb{C}^{m \times m}$ .

**Tensor products.** If  $X \in \mathbb{C}^{m_1 \times m_2}$  and  $Y \in \mathbb{C}^{\ell_1 \times \ell_2}$ , the tensor (Kronecker) product of  $X$  and  $Y$  is the  $m_1 \ell_1 \times m_2 \ell_2$  matrix defined by

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

Here is a list of important properties satisfied by tensor products.

- P1.** Associativity:  $(X \otimes Y) \otimes Z = X \otimes (Y \otimes Z)$  for all matrices  $X, Y, Z$ .
- P2.** Bilinearity:  $(\alpha X + \beta Y) \otimes (\gamma W + \eta Z) = \alpha\gamma(X \otimes W) + \alpha\eta(X \otimes Z) + \beta\gamma(Y \otimes W) + \beta\eta(Y \otimes Z)$  for all  $\alpha, \beta, \gamma, \eta \in \mathbb{C}$  and for all matrices  $X, Y, W, Z$  such that  $X, Y$  are summable and  $W, Z$  are summable.
- P3.**  $(X \otimes Y)^* = X^* \otimes Y^*$  and  $(X \otimes Y)^T = X^T \otimes Y^T$  for all matrices  $X, Y$ .
- P4.**  $(X \otimes Y)(W \otimes Z) = (XW) \otimes (YZ)$  for all matrices  $X, Y, W, Z$  such that  $X, W$  are multipliable and  $Y, Z$  are multipliable.
- P5.**  $\|X \otimes Y\|_p = \|X\|_p \|Y\|_p$  for all square matrices  $X, Y$  and all  $p \in [1, \infty]$ .
- P6.**  $\text{rank}(X \otimes Y) = \text{rank}(X) \text{rank}(Y)$  for all matrices  $X, Y$ .
- P7.** If  $X \in \mathbb{C}^{m \times m}$  and  $Y \in \mathbb{C}^{\ell \times \ell}$ , the eigenvalues (resp., singular values) of  $X \otimes Y$  are given by  $\{\lambda_i(X) \lambda_j(Y) : i = 1, \dots, m, j = 1, \dots, \ell\}$  (resp.,  $\{\sigma_i(X) \sigma_j(Y) : i = 1, \dots, m, j = 1, \dots, \ell\}$ ).
- P8.** If  $X_i \in \mathbb{C}^{m_i \times \ell_i}$  for  $i = 1, \dots, d$  and we set  $\mathbf{m} = (m_1, \dots, m_d)$  and  $\ell = (\ell_1, \dots, \ell_d)$ , then

$$(X_1 \otimes \cdots \otimes X_d)_{ij} = (X_1)_{i_1 j_1} \cdots (X_d)_{i_d j_d}, \quad \mathbf{i} = \mathbf{1}, \dots, \mathbf{m}, \quad \mathbf{j} = \mathbf{1}, \dots, \ell.$$

- P9.** If  $X_i, Y_i \in \mathbb{C}^{m_i \times m_i}$  for  $i = 1, \dots, d$  and  $\mathbf{m} = (m_1, \dots, m_d)$ , then

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

- P10.** Given  $\mathbf{m} \in \mathbb{N}^d$  and a permutation  $\sigma$  of  $\{1, \dots, d\}$ , there exists a unique permutation matrix  $\Pi_{\mathbf{m}; \sigma}$  such that

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

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

**Special permutation matrices.** For every  $s \in \mathbb{N}$  and  $\mathbf{n} \in \mathbb{N}^d$ , we denote by  $\Pi_{\mathbf{n},s}$  the  $N(\mathbf{n})_s \times N(\mathbf{n})_s$  permutation matrix given by

$$\Pi_{\mathbf{n},s} = \begin{bmatrix} I_s \otimes \mathbf{e}_1^T \\ I_s \otimes \mathbf{e}_2^T \\ \vdots \\ I_s \otimes \mathbf{e}_n^T \end{bmatrix} = \sum_{k=1}^n \mathbf{e}_k \otimes I_s \otimes \mathbf{e}_k^T,$$

where  $\mathbf{e}_i, i = 1, \dots, n$ , are the vectors of the canonical basis of  $\mathbb{C}^{N(\mathbf{n})}$ , which, for convenience, are indexed by a  $d$ -index  $i = 1, \dots, n$ . For every  $s, r \in \mathbb{N}$  and  $\mathbf{n} \in \mathbb{N}^d$ , we define the permutation matrix

$$\Pi_{\mathbf{n},s,r} = \Pi_{\mathbf{n},s} \otimes I_r.$$

With reference to **P 10**, if  $\mathbf{m}, \mathbf{s} \in \mathbb{N}^d$ , and  $\sigma$  is the permutation of  $\{1, \dots, 2d\}$  given by  $\sigma = [1, d+1, 2, d+2, \dots, d, 2d]$ , we define  $\Gamma_{\mathbf{m},\mathbf{s}} = \Pi_{(\mathbf{m},\mathbf{s});\sigma}$ . In other words,  $\Gamma_{\mathbf{m},\mathbf{s}}$  is the unique permutation matrix such that

$$X_1 \otimes X_{d+1} \otimes X_2 \otimes X_{d+2} \otimes \dots \otimes X_d \otimes X_{2d} = \Gamma_{\mathbf{m},\mathbf{s}}(X_1 \otimes \dots \otimes X_{2d})\Gamma_{\mathbf{m},\mathbf{s}}^T$$

for all  $X_1 \in \mathbb{C}^{m_1 \times m_1}, \dots, X_d \in \mathbb{C}^{m_d \times m_d}, X_{d+1} \in \mathbb{C}^{s_1 \times s_1}, \dots, X_{2d} \in \mathbb{C}^{s_d \times s_d}$ .

**Sequences of matrices and multilevel block matrix-sequences.** A *sequence of matrices* is a sequence of the form  $\{A_n\}_n$ , where  $A_n$  is a square matrix of size  $d_n$  such that  $d_n \rightarrow \infty$  as  $n \rightarrow \infty$ . If  $\{A_n\}_n$  is a sequence of matrices with  $A_n$  of size  $d_n$ , we say that  $\{A_n\}_n$  is *sparsely unbounded (s.u.)* if

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\#\{i \in \{1, \dots, d_n\} : \sigma_i(A_n) > M\}}{d_n} = 0,$$

and we say that  $\{A_n\}_n$  is *sparsely vanishing (s.v.)* if

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\#\{i \in \{1, \dots, d_n\} : \sigma_i(A_n) < 1/M\}}{d_n} = 0.$$

A *d-level r-block matrix-sequence* is a sequence of matrices  $\{A_n\}_n$ , where

- $n$  varies in some infinite subset of  $\mathbb{N}$ ;
- $\mathbf{n} = \mathbf{n}(n) \in \mathbb{N}^d$  and  $\mathbf{n} \rightarrow \infty$  (i.e.,  $\min(\mathbf{n}) \rightarrow \infty$ ) as  $n \rightarrow \infty$ ;
- $A_n$  is a square matrix of size  $N(\mathbf{n})r$ .

**Singular value and eigenvalue distribution of a sequence of matrices.** Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ , and let  $f : D \subset \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  be a measurable function defined on a set  $D$  with  $0 < \mu_k(D) < \infty$ .

- We say that  $\{A_n\}_n$  has a singular value distribution described by  $f$ , and we write  $\{A_n\}_n \sim_\sigma f$ , if

$$\lim_{n \rightarrow \infty} \frac{1}{d_n} \sum_{i=1}^{d_n} F(\sigma_i(A_n)) = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^r F(\sigma_i(f(\mathbf{x})))}{r} d\mathbf{x}, \quad \forall F \in C_c(\mathbb{R}).$$

In this case,  $f$  is called the *singular value symbol* of  $\{A_n\}_n$ .

- We say that  $\{A_n\}_n$  has a spectral (or eigenvalue) distribution described by  $f$ , and we write  $\{A_n\}_n \sim_\lambda f$ , if

$$\lim_{n \rightarrow \infty} \frac{1}{d_n} \sum_{i=1}^{d_n} F(\lambda_i(A_n)) = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^r F(\lambda_i(f(\mathbf{x})))}{r} d\mathbf{x}, \quad \forall F \in C_c(\mathbb{C}).$$

In this case,  $f$  is called the *spectral (or eigenvalue) symbol* of  $\{A_n\}_n$ .

When we write a relation such as  $\{A_n\}_n \sim_\sigma f$  or  $\{A_n\}_n \sim_\lambda f$ , it is understood that  $\{A_n\}_n$  is a sequence of matrices and  $f$  is a measurable function defined on a subset  $D$  of some  $\mathbb{R}^k$  with  $0 < \mu_k(D) < \infty$  and taking values in  $\mathbb{C}^{r \times r}$  for some  $r \geq 1$ . In what follows, “iff” is an abbreviation of “if and only if”.

**S1.** If  $\{A_n\}_n \sim_\sigma f$ , then  $\{A_n\}_n$  is s.u.

**S2.** If  $\{A_n\}_n \sim_\sigma f$ , then  $\{A_n\}_n$  is s.v. iff  $f$  is invertible a.e.

**S3.** If  $\{A_n\}_n \sim_\lambda f$  and  $\Lambda(A_n) \subseteq S$  for all  $n$ , then  $\Lambda(f) \subseteq \bar{S}$  a.e.

**S4.** If  $A_n = X_n + Y_n \in \mathbb{C}^{d_n \times d_n}$  and

- each  $X_n$  is Hermitian and  $\{X_n\}_n \sim_\lambda f$ ,

- $\lim_{n \rightarrow \infty} (d_n)^{-1/2} \|Y_n\|_2 = 0$ ,

then  $\{A_n\}_n \sim_\lambda f$ .

**S5.** Let  $X_n \in \mathbb{C}^{d_n \times d_n}$  and  $P_n \in \mathbb{C}^{d_n \times \delta_n}$ , where  $P_n^* P_n = I_{\delta_n}$ ,  $\delta_n \leq d_n$ ,  $\delta_n/d_n \rightarrow 1$ .

- $\{X_n\}_n \sim_\sigma f$  iff  $\{P_n^* X_n P_n\}_n \sim_\sigma f$ .

- If the matrices  $X_n$  are Hermitian, then  $\{X_n\}_n \sim_\lambda f$  iff  $\{P_n^* X_n P_n\}_n \sim_\lambda f$ .

**Informal meaning.** Assuming that  $f : D \subset \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  possesses  $r$  a.e. continuous eigenvalue functions  $\lambda_i(f(\mathbf{x}))$ ,  $i = 1, \dots, r$ , the spectral distribution  $\{A_n\}_n \sim_\lambda f$  has the following informal meaning: all the eigenvalues of  $A_n$ , except possibly for  $o(d_n)$  outliers (with  $d_n$  being the size of  $A_n$ ), can be subdivided into  $r$  different subsets of approximately the same cardinality and the eigenvalues belonging to the  $i$ th subset are approximately equal to the samples of the  $i$ th eigenvalue function  $\lambda_i(f(\mathbf{x}))$  over a uniform grid in the domain  $D$  (for  $n$  large enough). For instance, if  $k = 1$ ,  $d_n = nr$ , and  $D = [a, b]$ , then, assuming we have no outliers, the eigenvalues of  $A_n$  are approximately equal to

$$\lambda_i \left( f \left( a + j \frac{b-a}{n} \right) \right), \quad j = 1, \dots, n, \quad i = 1, \dots, r,$$

for  $n$  large enough. Similarly, if  $k = 2$ ,  $d_n = n^2 r$ , and  $D = [a_1, b_1] \times [a_2, b_2]$ , then, assuming we have no outliers, the eigenvalues of  $A_n$  are approximately equal to

$$\lambda_i \left( f \left( a_1 + j_1 \frac{b_1 - a_1}{n}, a_2 + j_2 \frac{b_2 - a_2}{n} \right) \right), \quad j_1, j_2 = 1, \dots, n, \quad i = 1, \dots, r,$$

for  $n$  large enough. A completely analogous meaning can also be given for the singular value distribution  $\{A_n\}_n \sim_\sigma f$ .

**Rearrangement.** Assuming that  $D = [a_1, b_1] \times \dots \times [a_k, b_k]$  is a hyperrectangle in  $\mathbb{R}^k$  and  $f : D \rightarrow \mathbb{C}^{r \times r}$  is a measurable function possessing  $r$  real-valued a.e. continuous eigenvalue functions  $\lambda_i(f(\mathbf{x}))$ ,  $i = 1, \dots, r$ , compute, for each  $\rho \in \mathbb{N}$ , the uniform samples

$$\lambda_i \left( f \left( a_1 + j_1 \frac{b_1 - a_1}{\rho}, \dots, a_k + j_k \frac{b_k - a_k}{\rho} \right) \right), \quad j_1, \dots, j_k = 1, \dots, \rho, \quad i = 1, \dots, r,$$

sort them in non-decreasing order, and put them into a vector  $(\varsigma_1, \varsigma_2, \dots, \varsigma_{r\rho^k})$ . Let  $\phi_\rho : [0, 1] \rightarrow \mathbb{R}$  be the piecewise linear non-decreasing function that interpolates the samples

$(s_0 = s_1, s_1, s_2, \dots, s_{r\rho^k})$  over the nodes  $(0, \frac{1}{r\rho^k}, \frac{2}{r\rho^k}, \dots, 1)$ , i.e.,

$$\begin{cases} \phi_\rho\left(\frac{i}{r\rho^k}\right) = s_i, & i = 0, \dots, r\rho^k, \\ \phi_\rho \text{ linear on } \left[\frac{i}{r\rho^k}, \frac{i+1}{r\rho^k}\right], & i = 0, \dots, r\rho^k - 1. \end{cases}$$

When  $\rho \rightarrow \infty$ , the function  $\phi_\rho$  converges a.e. to a function  $\phi$ , which is non-decreasing on  $(0, 1)$  and satisfies

$$\int_0^1 F(\phi(t))dt = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^r F(\lambda_i(f(\mathbf{x})))}{r} d\mathbf{x}, \quad \forall F \in C_c(\mathbb{C}).$$

The function  $\phi$  is referred to as the *canonical rearranged version* of  $f$ . What is interesting about  $\phi$  is that if  $\{A_n\}_n \sim_\lambda f$ , then  $\{A_n\}_n \sim_\lambda \phi$ . In particular, if we have  $\{A_n\}_n \sim_\lambda f$  (and hence also  $\{A_n\}_n \sim_\lambda \phi$ ), then, for  $n$  large enough, the eigenvalues of  $A_n$ , with the possible exception of  $o(d_n)$  outliers ( $d_n = \text{size}(A_n)$ ), are approximately equal to the samples of  $\phi$  over a uniform grid in  $[0, 1]$ .

**Clustering and attraction.** In what follows, if  $f : D \subseteq \mathbb{R}^k \rightarrow \mathbb{C}^{r \times r}$  is a measurable matrix-valued function, its essential range is denoted by  $\mathcal{ER}(f)$  and is defined as

$$\mathcal{ER}(f) = \{z \in \mathbb{C} : \mu_k\{\exists j \in \{1, \dots, r\} : \lambda_j(f) \in D(z, \varepsilon)\} > 0 \text{ for all } \varepsilon > 0\}.$$

- Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ , and let  $S$  be a nonempty subset of  $\mathbb{C}$ . We say that  $\{A_n\}_n$  is weakly clustered at  $S$  if

$$\lim_{n \rightarrow \infty} \frac{\#\{j \in \{1, \dots, d_n\} : \lambda_j(A_n) \notin D(S, \varepsilon)\}}{d_n} = 0, \quad \forall \varepsilon > 0.$$

- Let  $\{A_n\}_n$  be a sequence of matrices with  $A_n$  of size  $d_n$ , and let  $z \in \mathbb{C}$ . We say that  $z$  strongly attracts the spectrum  $\Lambda(A_n)$  with infinite order if, once we have ordered the eigenvalues of  $A_n$  according to their distance from  $z$ ,

$$|\lambda_1(A_n) - z| \leq |\lambda_2(A_n) - z| \leq \dots \leq |\lambda_{d_n}(A_n) - z|,$$

the following limit relation holds for each fixed  $j \geq 1$  :

$$\lim_{n \rightarrow \infty} |\lambda_j(A_n) - z| = 0.$$

**CA 1.** If  $\{A_n\}_n \sim_\lambda f$ , then  $\{A_n\}_n$  is weakly clustered at  $\mathcal{ER}(f)$  and each  $z \in \mathcal{ER}(f)$  strongly attracts  $\Lambda(A_n)$  with infinite order.

**Zero-distributed sequences.** A sequence of matrices  $\{Z_n\}_n$  such that  $\{Z_n\}_n \sim_\sigma 0$  is referred to as a zero-distributed sequence. In other words,  $\{Z_n\}_n$  is zero-distributed iff

$$\lim_{n \rightarrow \infty} \frac{1}{d_n} \sum_{i=1}^{d_n} F(\sigma_i(Z_n)) = F(0), \quad \forall F \in C_c(\mathbb{R}),$$

where  $d_n$  is the size of  $Z_n$ . Given a sequence of matrices  $\{Z_n\}_n$  with  $Z_n$  of size  $d_n$ , the following properties hold. In what follows, we use the natural convention  $C/\infty = 0$  for all numbers  $C$ .



**Z1.**  $\{Z_n\}_n \sim_\sigma 0$  iff  $Z_n = R_n + N_n$  with  $\lim_{n \rightarrow \infty} (d_n)^{-1} \text{rank}(R_n) = \lim_{n \rightarrow \infty} \|N_n\| = 0$ .

**Z2.**  $\{Z_n\}_n \sim_\sigma 0$  if there exists a  $p \in [1, \infty]$  such that  $\lim_{n \rightarrow \infty} (d_n)^{-1/p} \|Z_n\|_p = 0$ .

**Sequences of multilevel block diagonal sampling matrices.** If  $\mathbf{n} \in \mathbb{N}^d$  and  $a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$ , then the  $n$ th multilevel block (or  $d$ -level  $s$ -block) diagonal sampling matrix generated by  $a$  is the  $N(\mathbf{n})s \times N(\mathbf{n})s$  block diagonal matrix given by

$$D_{\mathbf{n}}(a) = \text{diag}_{i=1, \dots, n} a\left(\frac{\mathbf{i}}{\mathbf{n}}\right).$$

Each  $d$ -level  $s$ -block matrix-sequence of the form  $\{D_{\mathbf{n}}(a)\}_n$  with  $\mathbf{n} = \mathbf{n}(n) \rightarrow \infty$  as  $n \rightarrow \infty$  is referred to as a sequence of multilevel block (or  $d$ -level  $s$ -block) diagonal sampling matrices generated by  $a$ . If  $n, s \in \mathbb{N}$ , we denote by  $\{x_{i,s}^{(n)}\}_{i=1}^{ns} = \{x_{1,s}^{(n)}, \dots, x_{ns,s}^{(n)}\}$  the sequence of points

$$\{x_{1,s}^{(n)}, \dots, x_{ns,s}^{(n)}\} = \left\{ \underbrace{\frac{1}{n}, \dots, \frac{1}{n}}_s, \underbrace{\frac{2}{n}, \dots, \frac{2}{n}}_s, \dots, \underbrace{\frac{n-1}{n}, \dots, \frac{n-1}{n}}_s, \underbrace{1, \dots, 1}_s \right\},$$

i.e.,

$$x_{i,s}^{(n)} = \left( \left\lfloor \frac{i-1}{s} \right\rfloor + 1 \right) \frac{1}{n}, \quad i = 1, \dots, ns.$$

If  $\mathbf{n}, \mathbf{s} \in \mathbb{N}^d$  and  $a : [0, 1]^d \rightarrow \mathbb{C}$ , we denote by  $D_{\mathbf{n}, \mathbf{s}}(a)$  the  $d$ -level diagonal sampling matrix given by

$$D_{\mathbf{n}, \mathbf{s}}(a) = \text{diag}_{i=1, \dots, n\mathbf{s}} a(\mathbf{x}_{i, \mathbf{s}}^{(\mathbf{n})}),$$

where  $\{\mathbf{x}_{i, \mathbf{s}}^{(\mathbf{n})}\}_{i=1}^{n\mathbf{s}} = \{\mathbf{x}_{1, \mathbf{s}}^{(\mathbf{n})}, \dots, \mathbf{x}_{n\mathbf{s}, \mathbf{s}}^{(\mathbf{n})}\}$  is the sequence of points

$$\mathbf{x}_{i, \mathbf{s}}^{(\mathbf{n})} = (x_{i_1, s_1}^{(n_1)}, \dots, x_{i_d, s_d}^{(n_d)}), \quad \mathbf{i} = 1, \dots, n\mathbf{s}.$$

In what follows, we denote by  $E^{[0,1]^d}$  the space of all functions from  $[0, 1]^d$  to the set  $E$ .

**D1.** For every  $\mathbf{n} \in \mathbb{N}^d$  the map  $D_{\mathbf{n}}(\cdot) : (\mathbb{C}^{s \times s})^{[0,1]^d} \rightarrow \mathbb{C}^{sN(\mathbf{n}) \times sN(\mathbf{n})}$

- is linear:  $D_{\mathbf{n}}(\alpha a + \beta b) = \alpha D_{\mathbf{n}}(a) + \beta D_{\mathbf{n}}(b)$ ,

- satisfies  $D_{\mathbf{n}}(a)^* = D_{\mathbf{n}}(a^*)$ .

**D2.** If  $\mathbf{n} \in \mathbb{N}^d$ ,  $a_{ij} : [0, 1]^d \rightarrow \mathbb{C}^{r \times r}$ , for  $i, j = 1, \dots, s$ ,  $a = [a_{ij}]_{i,j=1}^s$ , and  $D_{\mathbf{n}} = [D_{\mathbf{n}}(a_{ij})]_{i,j=1}^s$ , then we have  $\Pi_{\mathbf{n}, s, r} D_{\mathbf{n}} \Pi_{\mathbf{n}, s, r}^T = D_{\mathbf{n}}(a)$ .

**D3.** If  $\mathbf{n}, \mathbf{s} \in \mathbb{N}^d$  and  $a : [0, 1]^d \rightarrow \mathbb{C}$ , then  $D_{\mathbf{n}, \mathbf{s}}(a) = \Gamma_{\mathbf{n}, \mathbf{s}} D_{\mathbf{n}}(a I_{N(\mathbf{s})}) \Gamma_{\mathbf{n}, \mathbf{s}}^T$ .

**Multilevel block Toeplitz sequences.** If  $\mathbf{n} \in \mathbb{N}^d$  and  $f : [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  is a function in  $L^1([-\pi, \pi]^d, s)$ , then the  $n$ th ( $d$ -level  $s$ -block) Toeplitz matrix generated by  $f$  is the  $sN(\mathbf{n}) \times sN(\mathbf{n})$  matrix given by

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

where the  $s \times s$  blocks  $f_{\mathbf{k}}$  are the Fourier coefficients of  $f$ ,

$$f_{\mathbf{k}} = \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} f(\boldsymbol{\theta}) e^{-i\mathbf{k} \cdot \boldsymbol{\theta}} d\boldsymbol{\theta} \in \mathbb{C}^{s \times s}, \quad \mathbf{k} \in \mathbb{Z}^d,$$

and the integrals in the previous formula are computed componentwise.  $\{T_n(f)\}_n$  is the ( $d$ -level  $s$ -block) Toeplitz sequence generated by  $f$ . In what follows, for any function  $f$  in  $L^p([-\pi, \pi]^d, s)$ , we define

$$\|f\|_{L^p} = \begin{cases} (\int_D \|f(\mathbf{x})\|_p^p d\mathbf{x})^{1/p}, & \text{if } 1 \leq p < \infty, \\ \text{ess sup}_{\mathbf{x} \in D} \|f(\mathbf{x})\|, & \text{if } p = \infty. \end{cases}$$

**T1.** For every  $\mathbf{n} \in \mathbb{N}^d$  the map  $T_n(\cdot) : L^1([-\pi, \pi]^d, s) \rightarrow \mathbb{C}^{sN(\mathbf{n}) \times sN(\mathbf{n})}$

- is linear:  $T_n(\alpha f + \beta g) = \alpha T_n(f) + \beta T_n(g)$ ,
- satisfies  $T_n(f)^* = T_n(f^*)$ .

**T2.** If  $f$  is Hermitian a.e., then  $T_n(f)$  is Hermitian for all  $\mathbf{n} \in \mathbb{N}^d$ .

**T3.** If  $1 \leq p \leq \infty$  and  $f \in L^p([-\pi, \pi]^d, s)$ , then  $\|T_n(f)\|_p \leq \frac{N(\mathbf{n})^{1/p}}{(2\pi)^{d/p}} \|f\|_{L^p}$ .

**T4.** If  $f \in L^1([-\pi, \pi]^d, s)$  and  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  is such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ , then  $\{T_n(f)\}_n \sim_\sigma f$ . If in addition  $f$  is Hermitian a.e., then  $\{T_n(f)\}_n \sim_\lambda f$ .

**T5.** If  $f_1, \dots, f_q \in L^\infty([-\pi, \pi]^d, s)$ , then  $N(\mathbf{n})^{-1} \|\prod_{i=1}^q T_n(f_i) - T_n(\prod_{i=1}^q f_i)\|_1 \rightarrow 0$  as  $\mathbf{n} \rightarrow \infty$ .

**T6.** If  $\mathbf{n} \in \mathbb{N}^d$ ,  $f_{ij} \in L^1([-\pi, \pi]^d, r)$ , for  $i, j = 1, \dots, s$ ,  $f = [f_{ij}]_{i,j=1}^s$ , and  $T_n = [T_n(f_{ij})]_{i,j=1}^s$ , then we have  $\Pi_{\mathbf{n},s,r} T_n \Pi_{\mathbf{n},s,r}^T = T_n(f)$ .

**T7.** If  $\mathbf{n}, \mathbf{s} \in \mathbb{N}^d$ , then

$$T_n(f_1 \otimes \dots \otimes f_d) = \Gamma_{\mathbf{n},\mathbf{s}}^T (T_{n_1}(f_1) \otimes \dots \otimes T_{n_d}(f_d)) \Gamma_{\mathbf{n},\mathbf{s}}$$

for all  $f_1 \in L^1([-\pi, \pi], s_1), \dots, f_d \in L^1([-\pi, \pi], s_d)$ .

**Approximating classes of sequences.** Let  $\{A_n\}_n$  a sequence of matrices and  $\{\{B_{n,m}\}_m\}_n$  a sequence of sequences of matrices with  $A_n$  and  $B_{n,m}$  of size  $d_n$ . We say that  $\{\{B_{n,m}\}_m\}_n$  is an approximating class of sequences (a.c.s.) for  $\{A_n\}_n$  if the following condition is met: for every  $m$  there exists  $n_m$  such that, for  $n \geq n_m$ ,

$$A_n = B_{n,m} + R_{n,m} + N_{n,m}, \quad \text{rank}(R_{n,m}) \leq c(m)d_n, \quad \|N_{n,m}\| \leq \omega(m),$$

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

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

We use the abbreviation ‘‘a.c.s.’’ for both the singular ‘‘approximating class of sequences’’ and the plural ‘‘approximating classes of sequences’’. It turns out that, for each fixed sequence of positive integers  $d_n$  such that  $d_n \rightarrow \infty$ , the notion of a.c.s. is a notion of convergence in the space  $\mathcal{E} = \{\{A_n\}_n : A_n \in \mathbb{C}^{d_n \times d_n} \text{ for every } n\}$ . More precisely, for every gauge function  $\varphi$  and every  $A \in \mathbb{C}^{\ell \times \ell}$ , let

$$p^\varphi(A) = \frac{1}{\ell} \sum_{i=1}^{\ell} \varphi(\sigma_i(A))$$

and define

$$p_{\text{a.c.s.}}^\varphi(\{A_n\}_n) = \limsup_{n \rightarrow \infty} p(A_n), \quad \{A_n\}_n \in \mathcal{E},$$

$$d_{\text{a.c.s.}}^\varphi(\{A_n\}_n, \{B_n\}_n) = p_{\text{a.c.s.}}^\varphi(\{A_n - B_n\}_n), \quad \{A_n\}_n, \{B_n\}_n \in \mathcal{E}.$$

Then,  $d_{\text{a.c.s.}}^\varphi$  is a distance on  $\mathcal{E}$  such that  $d_{\text{a.c.s.}}^\varphi(\{A_n\}_n, \{B_n\}_n) = 0$  iff  $\{A_n - B_n\}_n$  is zero-distributed. Moreover,  $d_{\text{a.c.s.}}^\varphi$  turns  $\mathcal{E}$  into a complete pseudometric space  $(\mathcal{E}, d_{\text{a.c.s.}}^\varphi)$ , where

the statement “ $\{\{B_{n,m}\}_n\}_m$  converges to  $\{A_n\}_n$ ” is equivalent to “ $\{\{B_{n,m}\}_n\}_m$  is an a.c.s. for  $\{A_n\}_n$ ”. In particular, we can reformulate the definition of a.c.s. in the following way: *a sequence of sequences of matrices  $\{\{B_{n,m}\}_n\}_m$  is said to be an a.c.s. for  $\{A_n\}_n$  if  $\{B_{n,m}\}_n$  converges to  $\{A_n\}_n$  in  $(\mathcal{E}, d_{\text{a.c.s.}}^\varphi)$  as  $m \rightarrow \infty$ , i.e., if  $d_{\text{a.c.s.}}^\varphi(\{B_{n,m}\}_n, \{A_n\}_n) \rightarrow 0$  as  $m \rightarrow \infty$ . The theory of a.c.s. may then be interpreted as an approximation theory for sequences of matrices, and for this reason we will use the convergence notation  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  to indicate that  $\{\{B_{n,m}\}_n\}_m$  is an a.c.s. for  $\{A_n\}_n$ . In view of what follows, let  $D \subset \mathbb{R}^k$  be a measurable set such that  $0 < \mu_k(D) < \infty$  and, for every gauge function  $\varphi$ , define*

$$p_{\text{measure}}^\varphi(f) = \frac{1}{\mu_k(D)} \int_D \frac{\sum_{i=1}^r \varphi(\sigma_i(f(\mathbf{x})))}{r} d\mathbf{x}, \quad f \in \mathfrak{M}_D^{(r)},$$

$$d_{\text{measure}}^\varphi(f, g) = p_{\text{measure}}^\varphi(f - g), \quad f, g \in \mathfrak{M}_D^{(r)}.$$

Then,  $d_{\text{measure}}^\varphi$  is a distance on  $\mathfrak{M}_D^{(r)}$  such that  $d_{\text{measure}}^\varphi(f, g) = 0$  iff  $f = g$  a.e. Moreover,  $d_{\text{measure}}^\varphi$  turns  $\mathfrak{M}_D^{(r)}$  into a complete pseudometric space  $(\mathfrak{M}_D^{(r)}, d_{\text{measure}}^\varphi)$  where the statement “ $f_m$  converges to  $f$ ” is equivalent to “ $f_m$  converges to  $f$  in measure”.

**ACS 1.** If  $\{A_n\}_n \sim_\sigma f$ , then  $p_{\text{a.c.s.}}^\varphi(\{A_n\}_n) = p_{\text{measure}}^\varphi(f)$  for all gauge functions  $\varphi$ .

**ACS 2.**  $\{A_n\}_n \sim_\sigma f$  iff there exist sequences of matrices  $\{B_{n,m}\}_n \sim_\sigma f_m$  such that  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  and  $f_m \rightarrow f$  in measure.

**ACS 3.** Suppose each  $A_n$  is Hermitian. Then,  $\{A_n\}_n \sim_\lambda f$  iff there exist sequences of Hermitian matrices  $\{B_{n,m}\}_n \sim_\lambda f_m$  such that  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  and  $f_m \rightarrow f$  in measure.

**ACS 4.** If  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  and  $\{B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A'_n\}_n$  with  $A_n$  and  $A'_n$  of the same size  $d_n$ , then

- $\{B_{n,m}^*\}_n \xrightarrow{\text{a.c.s.}} \{A_n^*\}_n$ ,
- $\{\alpha B_{n,m} + \beta B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{\alpha A_n + \beta A'_n\}_n$  for all  $\alpha, \beta \in \mathbb{C}$ ,
- $\{B_{n,m} B'_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n A'_n\}_n$  whenever  $\{A_n\}_n, \{A'_n\}_n$  are s.u.,
- $\{B_{n,m} C_n\}_n \xrightarrow{\text{a.c.s.}} \{A_n C_n\}_n$  whenever  $\{C_n\}_n$  is s.u.

**ACS 5.** If  $A_n = [A_{n,ij}]_{i,j=1}^s$ ,  $B_n^{(m)} = [B_{n,ij}^{(m)}]_{i,j=1}^s$  and  $\{B_{n,ij}^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{A_{n,ij}^{(m)}\}_n$  for  $i, j = 1, \dots, s$ , then  $\{B_n^{(m)}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

**ACS 6.** Let  $p \in [1, \infty]$  and assume that for each  $m$  there is  $n_m$  such that, for  $n \geq n_m$ ,  $\|A_n - B_{n,m}\|_p \leq \varepsilon(m, n)(d_n)^{1/p}$ , where  $\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \varepsilon(m, n) = 0$  and  $d_n$  is the size of both  $A_n$  and  $B_{n,m}$ . Then  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

**Multilevel block generalized locally Toeplitz sequences.** A  $d$ -level  $s$ -block generalized locally Toeplitz (GLT) sequence  $\{A_n\}_n$  is a special  $d$ -level  $s$ -block matrix-sequence equipped with a measurable function  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$ , the so-called *symbol* (or *kernel*). Unless specified otherwise, the notation  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  means that  $\{A_n\}_n$  is a  $d$ -level  $s$ -block GLT sequence with symbol  $\kappa$ . The symbol of a  $d$ -level  $s$ -block GLT sequence is unique in the sense that if  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{A_n\}_n \sim_{\text{GLT}} \xi$ , then  $\kappa = \xi$  a.e. in  $[0, 1]^d \times [-\pi, \pi]^d$ . Conversely, if  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\kappa = \xi$  a.e. in  $[0, 1]^d \times [-\pi, \pi]^d$ , then  $\{A_n\}_n \sim_{\text{GLT}} \xi$ . In addition, any measurable function  $\kappa : [0, 1]^d \times [-\pi, \pi]^d \rightarrow \mathbb{C}^{s \times s}$  is the symbol of some  $d$ -level  $s$ -block GLT sequence  $\{A_n\}_n$ .

**GLT 1.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ , then  $\{A_n\}_n \sim_\sigma \kappa$ . If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and the matrices  $A_n$  are Hermitian, then  $\kappa$  is Hermitian a.e. and  $\{A_n\}_n \sim_\lambda \kappa$ .

**GLT 2.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  with  $A_n = X_n + Y_n$ , and

- every  $X_n$  is Hermitian,
- $N(\mathbf{n})^{-1/2} \|Y_n\|_2 \rightarrow 0$ ,

then  $\{P_n^* A_n P_n\}_n \sim_{\sigma, \lambda} \kappa$  for every sequence  $\{P_n\}_n$  such that  $P_n \in \mathbb{C}^{N(\mathbf{n}) \times \delta_n}$ ,  $P_n^* P_n = I_{\delta_n}$ ,  $\delta_n \leq N(\mathbf{n})s$ , and  $\delta_n / (N(\mathbf{n})s) \rightarrow 1$ .

**GLT 3.** We have

- $\{T_n(f)\}_n \sim_{\text{GLT}} \kappa(\mathbf{x}, \boldsymbol{\theta}) = f(\boldsymbol{\theta})$  if  $f \in L^1([-\pi, \pi]^d, s)$ ,
- $\{D_n(a)\}_n \sim_{\text{GLT}} \kappa(\mathbf{x}, \boldsymbol{\theta}) = a(\mathbf{x})$  if  $a : [0, 1]^d \rightarrow \mathbb{C}^{s \times s}$  is continuous a.e.,
- $\{Z_n\}_n \sim_{\text{GLT}} \kappa(\mathbf{x}, \boldsymbol{\theta}) = O_s$  iff  $\{Z_n\}_n \sim_{\sigma} 0$ .

**GLT 4.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{B_n\}_n \sim_{\text{GLT}} \xi$ , then

- $\{A_n^*\}_n \sim_{\text{GLT}} \kappa^*$ ,
- $\{\alpha A_n + \beta B_n\}_n \sim_{\text{GLT}} \alpha \kappa + \beta \xi$  for all  $\alpha, \beta \in \mathbb{C}$ ,
- $\{A_n B_n\}_n \sim_{\text{GLT}} \kappa \xi$ ,
- $\{A_n^\dagger\}_n \sim_{\text{GLT}} \kappa^{-1}$  if  $\kappa$  is invertible a.e.

**GLT 5.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and each  $A_n$  is Hermitian, then  $\{f(A_n)\}_n \sim_{\text{GLT}} f(\kappa)$  for every continuous function  $f : \mathbb{C} \rightarrow \mathbb{C}$ .

**GLT 6.** If  $\{A_{n,ij}\}_n$  is a  $d$ -level  $r$ -block GLT sequence with symbol  $\kappa_{ij}$ , for  $i, j = 1, \dots, s$ , and  $A_n = [A_{n,ij}]_{i,j=1}^s$ , then  $\{\Pi_{n,s,r} A_n \Pi_{n,s,r}^T\}_n$  is a  $d$ -level  $rs$ -block GLT sequence with symbol  $\kappa = [\kappa_{ij}]_{i,j=1}^s$ .

**GLT 7.**  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  iff there exist  $d$ -level  $s$ -block GLT sequences  $\{B_{n,m}\}_n \sim_{\text{GLT}} \kappa_m$  such that  $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$  and  $\kappa_m \rightarrow \kappa$  in measure.

**GLT 8.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$  and  $\{B_n\}_n \sim_{\text{GLT}} \xi$ , then  $d_{\text{a.c.s.}}^\varphi(\{A_n\}_n, \{B_n\}_n) = d_{\text{measure}}^\varphi(\kappa, \xi)$  for any gauge function  $\varphi$ .

**GLT 9.** If  $\{A_n\}_n \sim_{\text{GLT}} \kappa$ , then there exist functions  $a_{i,m}, f_{i,m}$ ,  $i = 1, \dots, N_m$ , such that

- $a_{i,m} : [0, 1]^d \rightarrow \mathbb{C}$  belongs to  $C^\infty([0, 1]^d)$  and  $f_{i,m}$  is a trigonometric monomial in  $\{e^{i\mathbf{j} \cdot \boldsymbol{\theta}} E_{\alpha\beta}^{(s)} : \mathbf{j} \in \mathbb{Z}^d, 1 \leq \alpha, \beta \leq s\}$ ,
- $\sum_{i=1}^{N_m} a_{i,m}(\mathbf{x}) f_{i,m}(\boldsymbol{\theta}) \rightarrow \kappa(\mathbf{x}, \boldsymbol{\theta})$  a.e.,
- $\{\sum_{i=1}^{N_m} D_n(a_{i,m} I_s) T_n(f_{i,m})\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ .

**6. Applications.** In this chapter we present several emblematic applications of the theory of multilevel block GLT sequences for the computation of the singular value and eigenvalue distribution of sequences of matrices arising from the numerical discretization of PDEs. In order to understand the content of this chapter, it is enough that the reader knows the summary of Chapter 5 and possesses the necessary prerequisites, most of which have been addressed in Chapter 2. Indeed, our derivations here will never refer to Chapters 1–4, i.e., they will only rely on the summary of Chapter 5.

**6.1. FD discretization of systems of PDEs.** Consider the following system of PDEs:

$$(6.1) \quad \begin{cases} -\nabla \cdot A \nabla u + \mathbf{b} \cdot \nabla v = f, & \text{in } (0, 1)^d, \\ \mathbf{c} \cdot \nabla u + \rho v = g, & \text{in } (0, 1)^d, \\ u = v = 0, & \text{on } \partial((0, 1)^d), \end{cases} \\ \Leftrightarrow \begin{cases} -\sum_{\ell, k=1}^d \frac{\partial}{\partial x_\ell} \left( a_{\ell k} \frac{\partial u}{\partial x_k} \right) + \sum_{k=1}^d b_k \frac{\partial v}{\partial x_k} = f, & \text{in } (0, 1)^d, \\ \sum_{k=1}^d c_k \frac{\partial u}{\partial x_k} + \rho v = g, & \text{in } (0, 1)^d, \\ u = v = 0, & \text{on } \partial((0, 1)^d), \end{cases}$$

where  $a_{\ell k}, b_k, c_k, \rho, f, g$  are given functions,  $A = [a_{\ell k}]_{\ell, k=1}^d$ ,  $\mathbf{b} = [b_k]_{k=1}^d$  and  $\mathbf{c} = [c_k]_{k=1}^d$ . In this section we consider the classical central FD discretization of (6.1). Through the theory

of multilevel block GLT sequences we show that, under suitable assumptions on the PDE coefficients, the corresponding sequence of (normalized) FD discretization matrices enjoys a spectral distribution described by a  $2 \times 2$  matrix-valued function. We remark that the number 2, which identifies the matrix space  $\mathbb{C}^{2 \times 2}$  where the spectral symbol takes values, coincides with the number of equations that compose the system (6.1).

**FD discretization.** Problem (6.1) can be reformulated as follows:

$$\left\{ \begin{array}{l} - \sum_{\ell, k=1}^d a_{\ell k} \frac{\partial^2 u}{\partial x_\ell \partial x_k} + \sum_{k=1}^d s_k \frac{\partial u}{\partial x_k} + \sum_{k=1}^d b_k \frac{\partial v}{\partial x_k} = f, \quad \text{in } (0, 1)^d, \\ \sum_{k=1}^d c_k \frac{\partial u}{\partial x_k} + \rho v = g, \quad \text{in } (0, 1)^d, \\ u = v = 0, \quad \text{on } \partial((0, 1)^d), \end{array} \right.$$

where  $\mathbf{s} = [s_k]_{k=1}^d$  is given by

$$s_k = - \sum_{\ell=1}^d \frac{\partial a_{\ell k}}{\partial x_\ell}, \quad k = 1, \dots, d.$$

Let  $\mathbf{n} \in \mathbb{N}^d$ , set  $\mathbf{h} = \frac{1}{\mathbf{n}+1}$  and  $\mathbf{x}_j = \mathbf{j}\mathbf{h}$ , for  $\mathbf{j} = \mathbf{0}, \dots, \mathbf{n} + \mathbf{1}$ .<sup>1</sup> Let  $\mathbf{e}_k$  be the  $k$ th vector of the canonical basis of  $\mathbb{R}^d$ . For  $\mathbf{j} = \mathbf{1}, \dots, \mathbf{n}$ , we have

$$\begin{aligned} a_{kk} \frac{\partial^2 u}{\partial x_k^2} \Big|_{\mathbf{x}=\mathbf{x}_j} &\approx a_{kk}(\mathbf{x}_j) \frac{u(\mathbf{x}_j + h_k \mathbf{e}_k) - 2u(\mathbf{x}_j) + u(\mathbf{x}_j - h_k \mathbf{e}_k)}{h_k^2} \\ &= a_{kk}(\mathbf{x}_j) \frac{u(\mathbf{x}_j + \mathbf{e}_k) - 2u(\mathbf{x}_j) + u(\mathbf{x}_j - \mathbf{e}_k)}{h_k^2}, \end{aligned}$$

for  $k = 1, \dots, d$ ,

$$\begin{aligned} a_{\ell k} \frac{\partial^2 u}{\partial x_\ell \partial x_k} \Big|_{\mathbf{x}=\mathbf{x}_j} &\approx a_{\ell k}(\mathbf{x}_j) \frac{\frac{\partial u}{\partial x_\ell}(\mathbf{x}_j + h_k \mathbf{e}_k) - \frac{\partial u}{\partial x_\ell}(\mathbf{x}_j - h_k \mathbf{e}_k)}{2h_k} \\ &\approx a_{\ell k}(\mathbf{x}_j) \frac{1}{2h_k} \left[ \frac{u(\mathbf{x}_j + h_k \mathbf{e}_k + h_\ell \mathbf{e}_\ell) - u(\mathbf{x}_j + h_k \mathbf{e}_k - h_\ell \mathbf{e}_\ell)}{2h_\ell} \right. \\ &\quad \left. - \frac{u(\mathbf{x}_j - h_k \mathbf{e}_k + h_\ell \mathbf{e}_\ell) - u(\mathbf{x}_j - h_k \mathbf{e}_k - h_\ell \mathbf{e}_\ell)}{2h_\ell} \right] \\ &= a_{\ell k}(\mathbf{x}_j) \frac{u(\mathbf{x}_j + \mathbf{e}_k + \mathbf{e}_\ell) - u(\mathbf{x}_j + \mathbf{e}_k - \mathbf{e}_\ell) - u(\mathbf{x}_j - \mathbf{e}_k + \mathbf{e}_\ell) + u(\mathbf{x}_j - \mathbf{e}_k - \mathbf{e}_\ell)}{4h_\ell h_k}, \end{aligned}$$

for  $\ell, k = 1, \dots, d$  with  $\ell \neq k$ ,

$$s_k \frac{\partial u}{\partial x_k} \Big|_{\mathbf{x}=\mathbf{x}_j} \approx s_k(\mathbf{x}_j) \frac{u(\mathbf{x}_j + h_k \mathbf{e}_k) - u(\mathbf{x}_j - h_k \mathbf{e}_k)}{2h_k} = s_k(\mathbf{x}_j) \frac{u(\mathbf{x}_j + \mathbf{e}_k) - u(\mathbf{x}_j - \mathbf{e}_k)}{2h_k},$$

<sup>1</sup>Recall that operations involving  $d$ -indices that have no meaning in  $\mathbb{Z}^d$  must be interpreted in the component-wise sense. In the present case, given  $\mathbf{n} = (n_1, \dots, n_d)$  and  $\mathbf{j} = (j_1, \dots, j_d)$ , the vector of discretization steps  $\mathbf{h} = \frac{1}{\mathbf{n}+1}$  and the grid point  $\mathbf{x}_j = \mathbf{j}\mathbf{h}$  are given by  $\mathbf{h} = (\frac{1}{n_1+1}, \dots, \frac{1}{n_d+1}) = (h_1, \dots, h_d)$  and  $\mathbf{x}_j = (j_1 h_1, \dots, j_d h_d)$ .

for  $k = 1, \dots, d$ ,

$$b_k \frac{\partial v}{\partial x_k} \Big|_{\mathbf{x}=\mathbf{x}_j} \approx b_k(\mathbf{x}_j) \frac{v(\mathbf{x}_j+\mathbf{e}_k) - v(\mathbf{x}_j-\mathbf{e}_k)}{2h_k},$$

for  $k = 1, \dots, d$ ,

$$c_k \frac{\partial u}{\partial x_k} \Big|_{\mathbf{x}=\mathbf{x}_j} \approx c_k(\mathbf{x}_j) \frac{u(\mathbf{x}_j+\mathbf{e}_k) - u(\mathbf{x}_j-\mathbf{e}_k)}{2h_k},$$

for  $k = 1, \dots, d$ ,

$$(6.2) \quad \rho v|_{\mathbf{x}=\mathbf{x}_j} = \rho(\mathbf{x}_j)v(\mathbf{x}_j).$$

Thus, for every  $\mathbf{j} = \mathbf{0}, \dots, \mathbf{n} + \mathbf{1}$ , we approximate the evaluations  $u(\mathbf{x}_j)$  and  $v(\mathbf{x}_j)$  by the values  $u_j$  and  $v_j$ , where  $u_j = v_j = 0$  if  $\mathbf{j} \notin \{\mathbf{1}, \dots, \mathbf{n}\}$  and the vectors  $\mathbf{u} = (u_1, \dots, u_n)^T$  and  $\mathbf{v} = (v_1, \dots, v_n)^T$  solve the linear system

$$\begin{aligned} & - \sum_{k=1}^d a_{kk}(\mathbf{x}_j) \frac{u_{j+\mathbf{e}_k} - 2u_j + u_{j-\mathbf{e}_k}}{h_k^2} \\ & - \sum_{\substack{\ell, k=1 \\ \ell \neq k}}^d a_{\ell k}(\mathbf{x}_j) \frac{u_{j+\mathbf{e}_k+\mathbf{e}_\ell} - u_{j+\mathbf{e}_k-\mathbf{e}_\ell} - u_{j-\mathbf{e}_k+\mathbf{e}_\ell} + u_{j-\mathbf{e}_k-\mathbf{e}_\ell}}{4h_\ell h_k} \\ & + \sum_{k=1}^d s_k(\mathbf{x}_j) \frac{u_{j+\mathbf{e}_k} - u_{j-\mathbf{e}_k}}{2h_k} + \sum_{k=1}^d b_k(\mathbf{x}_j) \frac{v_{j+\mathbf{e}_k} - v_{j-\mathbf{e}_k}}{2h_k} = f(\mathbf{x}_j), \\ & \sum_{k=1}^d c_k(\mathbf{x}_j) \frac{u_{j+\mathbf{e}_k} - u_{j-\mathbf{e}_k}}{2h_k} + \rho(\mathbf{x}_j)v_j = g(\mathbf{x}_j), \end{aligned}$$

for  $\mathbf{j} = \mathbf{1}, \dots, \mathbf{n}$ . This linear system can be written in matrix form as follows:

$$(6.3) \quad A_{\mathbf{n}} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix},$$

where  $\mathbf{f} = [f(\mathbf{x}_j)]_{j=1}^n$ ,  $\mathbf{g} = [g(\mathbf{x}_j)]_{j=1}^n$ ,

$$A_{\mathbf{n}} = \begin{bmatrix} B_{\mathbf{n}} & C_{\mathbf{n}} \\ D_{\mathbf{n}} & E_{\mathbf{n}} \end{bmatrix},$$

$$B_{\mathbf{n}} = \sum_{\ell, k=1}^d K_{\mathbf{n}, \ell k}(a_{\ell k}) + \sum_{k=1}^d H_{\mathbf{n}, k}(s_k),$$

$$C_{\mathbf{n}} = \sum_{k=1}^d H_{\mathbf{n}, k}(b_k),$$

$$D_{\mathbf{n}} = \sum_{k=1}^d H_{\mathbf{n}, k}(c_k),$$

$$E_{\mathbf{n}} = I_{\mathbf{n}}(\rho),$$

and the matrices  $K_{\mathbf{n},\ell k}(\alpha)$ ,  $H_{\mathbf{n},k}(\alpha)$ ,  $I_{\mathbf{n}}(\alpha)$  are defined for all functions  $\alpha : [0, 1]^d \rightarrow \mathbb{C}$  by

$$\begin{aligned} K_{\mathbf{n},\ell k}(\alpha) &= \frac{1}{h_\ell h_k} \left( \text{diag}_{j=1,\dots,\mathbf{n}} \alpha(\mathbf{x}_j) \right) K_{\mathbf{n},\ell k}, \quad \ell, k = 1, \dots, d, \\ H_{\mathbf{n},k}(\alpha) &= \frac{1}{h_k} \left( \text{diag}_{j=1,\dots,\mathbf{n}} \alpha(\mathbf{x}_j) \right) H_{\mathbf{n},k}, \quad k = 1, \dots, d, \\ I_{\mathbf{n}}(\alpha) &= \left( \text{diag}_{j=1,\dots,\mathbf{n}} \alpha(\mathbf{x}_j) \right) I_{\mathbf{n}}, \end{aligned}$$

with  $K_{\mathbf{n},\ell k}$ ,  $H_{\mathbf{n},k}$ ,  $I_{\mathbf{n}}$  being defined by their actions on a generic vector  $\mathbf{w} \in \mathbb{R}^{N(\mathbf{n})}$  as follows:

$$(6.4) \quad (K_{\mathbf{n},kk} \mathbf{w})_j = -w_{j-\mathbf{e}_k} + 2w_j - w_{j+\mathbf{e}_k}, \quad \mathbf{j} = \mathbf{1}, \dots, \mathbf{n},$$

for  $k = 1, \dots, d$ ,

$$(6.5) \quad (K_{\mathbf{n},\ell k} \mathbf{w})_j = -\frac{1}{4}(w_{j-\mathbf{e}_\ell-\mathbf{e}_k} - w_{j-\mathbf{e}_\ell+\mathbf{e}_k} - w_{j+\mathbf{e}_\ell-\mathbf{e}_k} + w_{j+\mathbf{e}_\ell+\mathbf{e}_k}), \quad \mathbf{j} = \mathbf{1}, \dots, \mathbf{n},$$

for  $\ell, k = 1, \dots, d$  with  $\ell \neq k$ ,

$$(6.6) \quad (H_{\mathbf{n},k} \mathbf{w})_j = \frac{1}{2}(-w_{j-\mathbf{e}_k} + w_{j+\mathbf{e}_k}), \quad \mathbf{j} = \mathbf{1}, \dots, \mathbf{n},$$

for  $k = 1, \dots, d$ ,

$$(6.7) \quad (I_{\mathbf{n}} \mathbf{w})_j = w_j, \quad \mathbf{j} = \mathbf{1}, \dots, \mathbf{n}.$$

In (6.4)–(6.7), it is assumed that  $w_i = 0$  whenever  $i \notin \{1, \dots, \mathbf{n}\}$ . It is clear that  $I_{\mathbf{n}} = I_{N(\mathbf{n})}$ . For the matrices  $K_{\mathbf{n},\ell k}$ ,  $H_{\mathbf{n},k}$  we have the following result [41, Remark 7.4].

LEMMA 6.1. *For every  $\mathbf{n} \in \mathbb{N}^d$ , we have*

$$\begin{aligned} K_{\mathbf{n},kk} &= T_{\mathbf{n}}(2 - 2 \cos \theta_k), & k &= 1, \dots, d, \\ K_{\mathbf{n},\ell k} &= T_{\mathbf{n}}(\sin \theta_\ell \sin \theta_k), & \ell, k &= 1, \dots, d, \quad \ell \neq k, \\ H_{\mathbf{n},k} &= -i T_{\mathbf{n}}(\sin \theta_k), & k &= 1, \dots, d. \end{aligned}$$

In particular,

$$K_{\mathbf{n},\ell k} = T_{\mathbf{n}}(H_{\ell k}), \quad \ell, k = 1, \dots, d,$$

where  $H(\boldsymbol{\theta})$  is the  $d \times d$  symmetric matrix defined as follows:

$$H_{\ell k}(\boldsymbol{\theta}) = \begin{cases} 2 - 2 \cos \theta_k, & \text{if } \ell = k, \\ \sin \theta_\ell \sin \theta_k, & \text{if } \ell \neq k. \end{cases}$$

**GLT analysis of the FD discretization matrices.** In what follows, we assume that  $\mathbf{n} + \mathbf{1} = \gamma \mathbf{n}$ , where  $\gamma \in \mathbb{Q}^d$  is a fixed vector with positive components and  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} + \mathbf{1} = \gamma \mathbf{n} \in \mathbb{N}^d$ . This assumption essentially says that each stepsize  $h_i = \frac{1}{n_i+1}$  tends to 0 with the same asymptotic speed as the others. The linear system (6.3) is equivalent to

$$\begin{bmatrix} n^{-1} I_{N(\mathbf{n})} & O \\ O & I_{N(\mathbf{n})} \end{bmatrix} A_{\mathbf{n}} \begin{bmatrix} n^{-1} I_{N(\mathbf{n})} & O \\ O & I_{N(\mathbf{n})} \end{bmatrix} \begin{bmatrix} n\mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} n^{-1}\mathbf{f} \\ \mathbf{g} \end{bmatrix} \iff \hat{A}_{\mathbf{n}} \begin{bmatrix} n\mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} n^{-1}\mathbf{f} \\ \mathbf{g} \end{bmatrix}$$

where

$$(6.8) \quad \hat{A}_{\mathbf{n}} = \begin{bmatrix} n^{-2}B_{\mathbf{n}} & n^{-1}C_{\mathbf{n}} \\ n^{-1}D_{\mathbf{n}} & E_{\mathbf{n}} \end{bmatrix}.$$

In the main result of this section (Theorem 6.3), we show that  $\{\Pi_{\mathbf{n},2}\hat{A}_{\mathbf{n}}\Pi_{\mathbf{n},2}^T\}_n$  is a  $d$ -level 2-block GLT sequence whose symbol  $\kappa(\mathbf{x}, \boldsymbol{\theta})$  is a  $2 \times 2$  matrix-valued function obtained by replacing the  $d$ -level (1-block) GLT sequences  $\{n^{-2}B_{\mathbf{n}}\}_n$ ,  $\{n^{-1}C_{\mathbf{n}}\}_n$ ,  $\{n^{-1}D_{\mathbf{n}}\}_n$ ,  $\{E_{\mathbf{n}}\}_n$  appearing in (6.8) with the corresponding symbols.<sup>2</sup> To prove Theorem 6.3, we need the following lemma; see [41, statement and proof of Theorem 7.2]. Recall that  $S_{\mathbf{n}}(a)$  denotes the  $\mathbf{n}$ th ( $d$ -level) arrow-shaped sampling matrix generated by  $a$ , as defined at the end of Section 2.7.

LEMMA 6.2. *Let  $a : [0, 1]^d \rightarrow \mathbb{C}$  be continuous, and let  $f(\boldsymbol{\theta}) = \sum_{j=-r}^r f_j e^{ij \cdot \boldsymbol{\theta}}$  be a  $d$ -variate trigonometric polynomial. Then,*

$$\begin{aligned} & |S_{\mathbf{n}}(a) \circ T_{\mathbf{n}}(f) - D_{\mathbf{n}}(a)T_{\mathbf{n}}(f)|_1, |S_{\mathbf{n}}(a) \circ T_{\mathbf{n}}(f) - D_{\mathbf{n}}(a)T_{\mathbf{n}}(f)|_{\infty} \\ & \leq (2|\mathbf{r}|_{\infty} + 1)^d \|f\|_{\infty} \omega_a \left( \frac{|\mathbf{r}|_{\infty}}{\min(\mathbf{n})} \right) \end{aligned}$$

for every  $\mathbf{n} \in \mathbb{N}^d$ , and

$$\{S_{\mathbf{n}}(a) \circ T_{\mathbf{n}}(f)\}_n \sim_{\text{GLT}} a(\mathbf{x})f(\boldsymbol{\theta})$$

for every sequence  $\{\mathbf{n} = \mathbf{n}(n)\}_n \subseteq \mathbb{N}^d$  such that  $\mathbf{n} \rightarrow \infty$  as  $n \rightarrow \infty$ .

THEOREM 6.3. *Suppose that the following conditions on the PDE coefficients are satisfied:*

- for every  $\ell, k = 1, \dots, d$ , the function  $a_{\ell k} : [0, 1]^d \rightarrow \mathbb{R}$  belongs to  $C([0, 1]^d)$  and its partial derivatives  $\partial a_{\ell k} / \partial x_1, \dots, \partial a_{\ell k} / \partial x_d : [0, 1]^d \rightarrow \mathbb{R}$  are bounded,
- for every  $k = 1, \dots, d$ , the functions  $b_k, c_k, \rho : [0, 1]^d \rightarrow \mathbb{R}$  belong to  $C([0, 1]^d)$ .

Let  $\boldsymbol{\gamma} \in \mathbb{Q}^d$  be a vector with positive components and assume that  $\mathbf{n} + \mathbf{1} = \boldsymbol{\gamma}n$  (it is understood that  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} + \mathbf{1} = \boldsymbol{\gamma}n \in \mathbb{N}^d$ ). Then,

$$(6.9) \quad \{\Pi_{\mathbf{n},2}\hat{A}_{\mathbf{n}}\Pi_{\mathbf{n},2}^T\}_n \sim_{\text{GLT}} \kappa^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}) = \begin{bmatrix} \kappa_{11}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}) & \kappa_{12}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}) \\ \kappa_{21}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}) & \rho(\mathbf{x}) \end{bmatrix},$$

where

$$\begin{aligned} \kappa_{11}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}) &= \sum_{\ell, k=1}^d \gamma_{\ell} \gamma_k a_{\ell k}(\mathbf{x}) H_{\ell k}(\boldsymbol{\theta}), \\ \kappa_{12}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}) &= -i \sum_{k=1}^d \gamma_k b_k(\mathbf{x}) \sin \theta_k, \\ \kappa_{21}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}) &= -i \sum_{k=1}^d \gamma_k c_k(\mathbf{x}) \sin \theta_k, \end{aligned}$$

and  $H(\boldsymbol{\theta})$  is defined in Lemma 6.1. Moreover, we have

$$(6.10) \quad \{\hat{A}_{\mathbf{n}}\}_n \sim_{\sigma} \kappa^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}).$$

<sup>2</sup>We shall see in the proof of Theorem 6.3 that  $\{n^{-2}B_{\mathbf{n}}\}_n$ ,  $\{n^{-1}C_{\mathbf{n}}\}_n$ ,  $\{n^{-1}D_{\mathbf{n}}\}_n$ ,  $\{E_{\mathbf{n}}\}_n$  are indeed  $d$ -level GLT sequences.



If in addition  $c_k = -b_k$  for every  $k = 1, \dots, d$ , then we also have

$$(6.11) \quad \{\hat{A}_n\}_n \sim_\lambda \kappa^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}).$$

*Proof.* The proof consists of the following steps.

*Step 1.* Consider the sequences  $\{n^{-2}B_n\}_n$ ,  $\{n^{-1}C_n\}_n$ ,  $\{n^{-1}D_n\}_n$ ,  $\{E_n\}_n$ , which “compose” the sequence  $\{\hat{A}_n\}_n$ . In Step 2, we show that these sequences are  $d$ -level GLT sequences, and precisely that

$$(6.12) \quad \{n^{-2}B_n\}_n \sim_{\text{GLT}} \kappa_{11}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}),$$

$$(6.13) \quad \{n^{-1}C_n\}_n \sim_{\text{GLT}} \kappa_{12}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}),$$

$$(6.14) \quad \{n^{-1}D_n\}_n \sim_{\text{GLT}} \kappa_{21}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}),$$

$$(6.15) \quad \{E_n\}_n \sim_{\text{GLT}} \rho(\mathbf{x}).$$

Once this is done, the GLT relation (6.9) follows immediately from **GLT 6**, and the singular value distribution (6.10) follows from (6.9) and **GLT 1**. We then prove in Step 3 the eigenvalue distribution (6.11) under the additional assumption that  $b_k = -c_k$ , for all  $k = 1, \dots, d$ .

*Step 2.* To prove the GLT relations (6.12)–(6.15) it suffices to prove the following:

$$(6.16) \quad \{n^{-2}K_{n,\ell k}(\alpha)\}_n \sim_{\text{GLT}} \gamma_\ell \gamma_k \alpha(\mathbf{x}) H_{\ell k}(\boldsymbol{\theta}), \quad \ell, k = 1, \dots, d, \quad \alpha \in C([0, 1]^d),$$

$$(6.17) \quad \{n^{-1}H_{n,k}(\alpha)\}_n \sim_{\text{GLT}} -i\gamma_k \alpha(\mathbf{x}) \sin \theta_k, \quad k = 1, \dots, d, \quad \alpha \in C([0, 1]^d),$$

$$(6.18) \quad \{I_n(\alpha)\}_n \sim_{\text{GLT}} \alpha(\mathbf{x}), \quad \alpha \in C([0, 1]^d).$$

Actually, we only prove (6.17) because the proofs of (6.16) and (6.18) are completely analogous. By **T 3**, the definition of  $H_{n,k}(\alpha)$ , and Lemma 6.1,

$$\begin{aligned} \|n^{-1}H_{n,k}(\alpha) + i\gamma_k D_n(\alpha) T_n(\sin \theta_k)\| &\leq \gamma_k \left\| \text{diag}_{j=1, \dots, n} \alpha(\mathbf{x}_j) - D_n(\alpha) \right\| \|T_n(\sin \theta_k)\| \\ &\leq \gamma_k \max_{j=1, \dots, n} \left| \alpha(\mathbf{x}_j) - \alpha\left(\frac{j}{n}\right) \right| \\ &\leq \gamma_k \omega_\alpha\left(\frac{1}{\min(\mathbf{n})}\right). \end{aligned}$$

As  $\omega_\alpha(1/\min(\mathbf{n})) \rightarrow 0$  for  $n \rightarrow \infty$ , we have  $\{n^{-1}H_{n,k}(\alpha) + i\gamma_k D_n(\alpha) T_n(\sin \theta_k)\}_n \sim_\sigma 0$  by **Z 1**, and so **GLT 3** and **GLT 4** immediately yield (6.17).

*Step 3.* We prove the eigenvalue distribution (6.11) in the case where  $c_k = -b_k$  for all  $k = 1, \dots, d$ . In this case, we have  $D_n = -C_n$  and

$$\hat{A}_n = \begin{bmatrix} n^{-2}B_n & n^{-1}C_n \\ -n^{-1}C_n & E_n \end{bmatrix}.$$

Consider the symmetric approximation of  $\hat{A}_n$  given by

$$\tilde{A}_n = \begin{bmatrix} n^{-2}\tilde{B}_n & n^{-1}\tilde{C}_n \\ -n^{-1}\tilde{C}_n & E_n \end{bmatrix},$$

where

$$\tilde{B}_n = \sum_{\ell, k=1}^d \tilde{K}_{n,\ell k}(a_{\ell k}), \quad \tilde{C}_n = \sum_{k=1}^d \tilde{H}_{n,k}(b_k),$$

and the matrices  $\tilde{K}_{\mathbf{n},\ell k}(\alpha)$ ,  $\tilde{H}_{\mathbf{n},k}(\alpha)$  are defined for all  $\alpha : [0, 1]^d \rightarrow \mathbb{C}$  as follows:

$$\begin{aligned}\tilde{K}_{\mathbf{n},\ell k}(\alpha) &= \frac{1}{h_\ell h_k} S_{\mathbf{n}}(\alpha) \circ K_{\mathbf{n},\ell k} = \frac{1}{h_\ell h_k} S_{\mathbf{n}}(\alpha) \circ T_{\mathbf{n}}(H_{\ell k}), & \ell, k = 1, \dots, d, \\ \tilde{H}_{\mathbf{n},k}(\alpha) &= \frac{1}{h_k} S_{\mathbf{n}}(\alpha) \circ H_{\mathbf{n},k} = -i \frac{1}{h_k} S_{\mathbf{n}}(\alpha) \circ T_{\mathbf{n}}(\sin \theta_k), & k = 1, \dots, d.\end{aligned}$$

Using **T3**, Lemmas **6.1**, **6.2**, and the assumptions on the PDE coefficients, it can be shown that both  $|\hat{A}_{\mathbf{n}} - \tilde{A}_{\mathbf{n}}|_1$  and  $|\hat{A}_{\mathbf{n}} - \tilde{A}_{\mathbf{n}}|_\infty$  tend to 0 as  $n \rightarrow \infty$ , and, consequently,  $\|\hat{A}_{\mathbf{n}} - \tilde{A}_{\mathbf{n}}\| \rightarrow 0$  as  $n \rightarrow \infty$  by **N1**. Thus, setting  $\hat{A}'_{\mathbf{n}} = \Pi_{\mathbf{n},2} \hat{A}_{\mathbf{n}} \Pi_{\mathbf{n},2}^T$  and  $\tilde{A}'_{\mathbf{n}} = \Pi_{\mathbf{n},2} \tilde{A}_{\mathbf{n}} \Pi_{\mathbf{n},2}^T$ , we have  $\|\hat{A}'_{\mathbf{n}} - \tilde{A}'_{\mathbf{n}}\| = \|\hat{A}_{\mathbf{n}} - \tilde{A}_{\mathbf{n}}\| \rightarrow 0$  as  $n \rightarrow \infty$ . Therefore, we finally obtain the estimate  $\|\hat{A}'_{\mathbf{n}} - \tilde{A}'_{\mathbf{n}}\|_2 \leq \sqrt{2N(\mathbf{n})} \|\hat{A}'_{\mathbf{n}} - \tilde{A}'_{\mathbf{n}}\| = o(\sqrt{N(\mathbf{n})})$  as  $n \rightarrow \infty$ . By **GLT2** applied to the decomposition  $\hat{A}'_{\mathbf{n}} = \tilde{A}'_{\mathbf{n}} + (\hat{A}'_{\mathbf{n}} - \tilde{A}'_{\mathbf{n}})$ , taking into account the symmetry of  $\tilde{A}'_{\mathbf{n}}$ , and the fact that  $\{\hat{A}'_{\mathbf{n}}\}_n \sim_{\text{GLT}} \kappa^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta})$  by (6.9), we infer that  $\{\hat{A}'_{\mathbf{n}}\}_n \sim_\lambda \kappa^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta})$ , which immediately implies (6.11).  $\square$

**6.2. Higher-order FE discretization of diffusion equations.** Consider the diffusion problem

$$(6.19) \quad \begin{cases} -\nabla \cdot A \nabla u = f, & \text{in } (0, 1)^d, \\ u = 0, & \text{on } \partial((0, 1)^d), \end{cases} \\ \iff \begin{cases} -\sum_{\ell, q=1}^d \frac{\partial}{\partial x_\ell} \left( a_{\ell q} \frac{\partial u}{\partial x_q} \right) = f, & \text{in } (0, 1)^d, \\ u = 0, & \text{on } \partial((0, 1)^d), \end{cases}$$

where  $a_{\ell q}$ ,  $f$  are given functions and  $A = [a_{\ell q}]_{\ell, q=1}^d$ . In this section we consider the higher-order FE discretization of (6.19). Through the theory of multilevel block GLT sequences we show that the corresponding sequence of (normalized) FE discretization matrices enjoys a spectral distribution described by a  $N(\mathbf{p}-\mathbf{k}) \times N(\mathbf{p}-\mathbf{k})$  matrix-valued function, where  $p_i$  and  $k_i$  represent, respectively, the degree and the smoothness in the  $i$ th direction of the piecewise polynomial functions involved in the FE approximation. Note that this result essentially proves [44, Conjecture 2].

**FE discretization.** The weak form of (6.19) reads as follows [21, Chapter 9]: find  $u \in H_0^1((0, 1)^d)$  such that

$$a(u, w) = f(w), \quad \forall w \in H_0^1((0, 1)^d),$$

where

$$a(u, w) = \int_{(0,1)^d} (\nabla w)^T A \nabla u, \quad f(w) = \int_{(0,1)^d} f w.$$

In the FE method [54], we fix a set of basis functions  $\{\varphi_1, \dots, \varphi_N\} \subset H_0^1((0, 1)^d)$  and we look for an approximation of the exact solution in the space  $\mathcal{W} = \text{span}(\varphi_1, \dots, \varphi_N)$  by solving the following discrete problem: find  $u_{\mathcal{W}} \in \mathcal{W}$  such that

$$a(u_{\mathcal{W}}, w) = f(w), \quad \forall w \in \mathcal{W}.$$

Since  $\{\varphi_1, \dots, \varphi_N\}$  is a basis of  $\mathcal{W}$ , we can write  $u_{\mathcal{W}} = \sum_{j=1}^N u_j \varphi_j$  for a unique vector  $\mathbf{u} = (u_1, \dots, u_N)^T$ . By linearity, the computation of  $u_{\mathcal{W}}$  (i.e., of  $\mathbf{u}$ ) reduces to solving the

linear system

$$A^{\text{stiff}} \mathbf{u} = \mathbf{f},$$

where  $\mathbf{f} = (f(\varphi_1), \dots, f(\varphi_N))^T$  and  $A^{\text{stiff}}$  is the stiffness matrix,

$$(6.20) \quad A^{\text{stiff}} = [a(\varphi_j, \varphi_i)]_{i,j=1}^N = \left[ \int_{(0,1)^d} (\nabla \varphi_i)^T A \nabla \varphi_j \right]_{i,j=1}^N.$$

**Tensor-product  $p$ -degree  $C^k$  B-spline basis functions.** Following the higher-order FE approach, the basis functions  $\varphi_1, \dots, \varphi_N$  will be chosen as piecewise polynomials of degree  $p_i \geq 1$  in the  $i$ th direction,  $i = 1, \dots, d$ . More precisely, for  $p, n \geq 1$  and  $0 \leq k \leq p - 1$ , let  $B_{1,[p,k]}, \dots, B_{n(p-k)+k+1,[p,k]} : \mathbb{R} \rightarrow \mathbb{R}$  be the B-splines of degree  $p$  and smoothness  $C^k$  defined on the knot sequence

$$(6.21) \quad \begin{aligned} & \{\tau_1, \dots, \tau_{n(p-k)+p+k+2}\} \\ & = \left\{ \underbrace{0, \dots, 0}_{p+1}, \underbrace{\frac{1}{n}, \dots, \frac{1}{n}}_{p-k}, \underbrace{\frac{2}{n}, \dots, \frac{2}{n}}_{p-k}, \dots, \underbrace{\frac{n-1}{n}, \dots, \frac{n-1}{n}}_{p-k}, \underbrace{1, \dots, 1}_{p+1} \right\}. \end{aligned}$$

We collect here a few properties of  $B_{1,[p,k]}, \dots, B_{n(p-k)+k+1,[p,k]}$  that we shall need later on. For the formal definition of B-splines as well as for the proof of the properties listed below, see [24, 60].

- The support of the  $i$ th B-spline is given by

$$(6.22) \quad \text{supp}(B_{i,[p,k]}) = [\tau_i, \tau_{i+p+1}], \quad i = 1, \dots, n(p-k) + k + 1.$$

In particular, for the measure of the support we have

$$(6.23) \quad \mu_1(\text{supp}(B_{i,[p,k]})) \leq \frac{p+1}{n}, \quad i = 1, \dots, n(p-k) + k + 1.$$

- Except for the first and the last ones, all the other B-splines vanish on the boundary of  $[0, 1]$ , i.e.,

$$(6.24) \quad B_{i,[p,k]}(0) = B_{i,[p,k]}(1) = 0, \quad i = 2, \dots, n(p-k) + k.$$

- $\{B_{1,[p,k]}, \dots, B_{n(p-k)+k+1,[p,k]}\}$  is a basis for the space of piecewise polynomial functions of degree  $p$  and smoothness  $C^k$ , that is,

$$\mathcal{V}_{n,[p,k]} = \left\{ v \in C^k([0, 1]) : v|_{\left[\frac{i}{n}, \frac{i+1}{n}\right]} \in \mathbb{P}_p \text{ for } i = 0, \dots, n-1 \right\},$$

where  $\mathbb{P}_p$  is the space of polynomials of degree less than or equal to  $p$ . Moreover,  $\{B_{2,[p,k]}, \dots, B_{n(p-k)+k,[p,k]}\}$  is a basis for the space

$$\mathcal{W}_{n,[p,k]} = \{w \in \mathcal{V}_{n,[p,k]} : w(0) = w(1) = 0\}.$$

- The B-splines form a non-negative partition of unity over  $[0, 1]$ :

$$(6.25) \quad B_{i,[p,k]} \geq 0 \text{ over } \mathbb{R}, \quad i = 1, \dots, n(p-k) + k + 1,$$

$$(6.26) \quad \sum_{i=1}^{n(p-k)+k+1} B_{i,[p,k]} = 1 \text{ over } [0, 1].$$

- The derivatives of the B-splines satisfy

$$(6.27) \quad \sum_{i=1}^{n(p-k)+k+1} |B'_{i,[p,k]}| \leq C_p n \text{ over } \mathbb{R},$$

where  $C_p$  is a constant depending only on  $p$ . Note that the derivatives  $B'_{i,[p,k]}$  may not be defined at some of the grid points  $0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1$  in the case of  $C^0$  smoothness ( $k = 0$ ). In (6.27) it is assumed that the undefined values are excluded from the summation.

- For every  $\mathbf{y} = (y_1, \dots, y_{n(p-k)+k+1}) \in \mathbb{R}^{n(p-k)+k+1}$ , we have

$$(6.28) \quad \left\| \sum_{i=1}^{n(p-k)+k+1} y_i B_{i,[p,k]} \right\|_{L^2([0,1])}^2 = \int_0^1 \left( \sum_{i=1}^{n(p-k)+k+1} y_i B_{i,[p,k]} \right)^2 \geq \frac{c_p}{n} \|\mathbf{y}\|^2,$$

where  $c_p$  is a constant depending only on  $p$ .

- All the B-splines, except for the first  $k + 1$  and the last  $k + 1$ , are uniformly shifted-scaled versions of  $p - k$  fixed reference functions  $\beta_{1,[p,k]}, \dots, \beta_{p-k,[p,k]}$ , namely the first  $p - k$  B-splines defined on the reference knot sequence

$$\underbrace{0, \dots, 0}_{p-k}, \underbrace{1, \dots, 1}_{p-k}, \dots, \underbrace{\eta, \dots, \eta}_{p-k}, \quad \eta = \left\lceil \frac{p+1}{p-k} \right\rceil.$$

The precise formula we shall need later on is the following: setting

$$(6.29) \quad \nu = \left\lceil \frac{k+1}{p-k} \right\rceil,$$

then, for the B-splines  $B_{k+2,[p,k]}, \dots, B_{k+1+(n-\nu)(p-k),[p,k]}$ , we have

$$B_{k+1+(p-k)(r-1)+q,[p,k]}(x) = \beta_{q,[p,k]}(nx - r + 1), \\ r = 1, \dots, n - \nu, \quad q = 1, \dots, p - k.$$

We point out that the supports of the reference B-splines  $\beta_{q,[p,k]}$  satisfy

$$(6.30) \quad \text{supp}(\beta_{1,[p,k]}) \subseteq \text{supp}(\beta_{2,[p,k]}) \subseteq \dots \subseteq \text{supp}(\beta_{p-k,[p,k]}) = [0, \eta].$$

Figures 6.1 and 6.2 display the graphs of the B-splines  $B_{1,[p,k]}, \dots, B_{n(p-k)+k+1,[p,k]}$  for the degree  $p = 3$  and the smoothness  $k = 1$  and the graphs of the associated reference B-splines  $\beta_{1,[p,k]}, \beta_{2,[p,k]}$ .

Now, for  $\mathbf{p}, \mathbf{n} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ , define the tensor-product  $\mathbf{p}$ -degree  $C^{\mathbf{k}}$  B-splines

$$B_{\mathbf{i},[\mathbf{p},\mathbf{k}]} = B_{i_1,[p_1,k_1]} \otimes \dots \otimes B_{i_d,[p_d,k_d]}, \quad \mathbf{i} = \mathbf{1}, \dots, \mathbf{n}(\mathbf{p} - \mathbf{k}) + \mathbf{k} + \mathbf{1},$$

where  $B_{i_j,[p_j,k_j]}$ ,  $i_j = 1, \dots, n_j(p_j - k_j) + k_j + 1$ , are the B-splines  $B_{i,[p,k]}$ ,  $i = 1, \dots, n(p - k) + k + 1$ , corresponding to  $n = n_j$ ,  $p = p_j$ ,  $k = k_j$ . Let

$$\boldsymbol{\tau}_{\mathbf{i}} = (\tau_{i_1}, \dots, \tau_{i_d}), \quad \mathbf{i} = \mathbf{1}, \dots, \mathbf{n}(\mathbf{p} - \mathbf{k}) + \mathbf{p} + \mathbf{k} + \mathbf{2},$$

where  $\{\tau_{i_j} : i_j = 1, \dots, n_j(p_j - k_j) + p_j + k_j + 2\}$  is the knot sequence  $\{\tau_i : i = 1, \dots, n(p - k) + p + k + 2\}$  in (6.21) corresponding to  $n = n_j$ ,  $p = p_j$ ,  $k = k_j$ . As a consequence

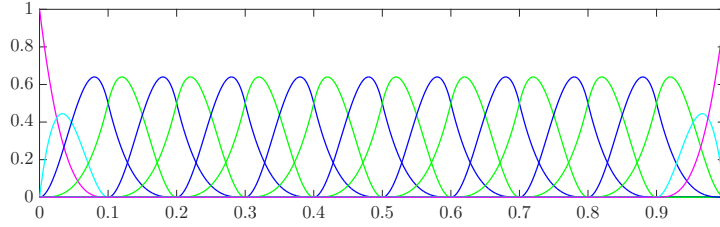


FIG. 6.1.  $B$ -splines  $B_{1,[p,k]}, \dots, B_{n(p-k)+k+1,[p,k]}$  for  $p = 3$  and  $k = 1$ , with  $n = 10$ .

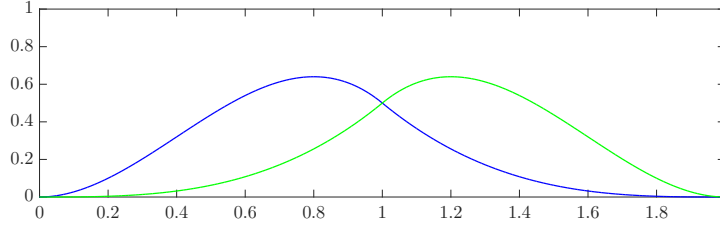


FIG. 6.2. Reference  $B$ -splines  $\beta_{1,[p,k]}, \beta_{2,[p,k]}$  for  $p = 3$  and  $k = 1$ .

of (6.22)–(6.27), we have

$$(6.31) \quad \text{supp}(B_{i,[p,k]}) = [\tau_i, \tau_{i+p+1}], \quad i = 1, \dots, n(p-k) + k + 1,$$

$$(6.32) \quad \mu_d(\text{supp}(B_{i,[p,k]})) \leq N \left( \frac{p+1}{n} \right), \quad i = 1, \dots, n(p-k) + k + 1,$$

$$(6.33) \quad B_{i,[p,k]} = 0 \text{ over } \partial([0, 1]^d), \quad i = 2, \dots, n(p-k) + k,$$

$$(6.34) \quad B_{i,[p,k]} \geq 0 \text{ over } \mathbb{R}^d, \quad i = 1, \dots, n(p-k) + k + 1,$$

$$(6.35) \quad \sum_{i=1}^{n(p-k)+k+1} B_{i,[p,k]} = 1 \text{ over } [0, 1]^d,$$

$$(6.36) \quad \sum_{i=1}^{n(p-k)+k+1} \left| \frac{\partial B_{i,[p,k]}}{\partial x_\ell} \right| \leq C_{p_\ell} n_\ell \text{ over } \mathbb{R}^d,$$

where  $C_{p_\ell}$  is a constant depending only on  $p_\ell$ . Note that the derivatives  $\partial B_{i,[p,k]}/\partial x_\ell$  may not be defined at some of the grid points  $\mathbf{j}/n$ ,  $\mathbf{j} = \mathbf{0}, \dots, \mathbf{n}$ , in the case of  $C^0$  smoothness in the  $\ell$ th direction ( $k_\ell = 0$ ). In (6.36) it is assumed that the undefined values are excluded from the summation. The basis functions  $\{\varphi_1, \dots, \varphi_N\} = \{\varphi_1, \dots, \varphi_{n(p-k)+k-1}\}$  are defined as follows:

$$(6.37) \quad \varphi_i = B_{i+1,[p,k]}, \quad i = 1, \dots, n(p-k) + k - 1.$$

In particular, we have  $N = N(n(p-k) + k - 1)$ .

**Higher-order FE discretization matrices.** The stiffness matrix (6.20) resulting from the choice of the basis functions as in (6.37) will be denoted by  $A_{n,[p,k]}(A)$ :

$$(6.38) \quad A_{n,[p,k]}(A) = \left[ \int_{(0,1)^d} (\nabla B_{i+1,[p,k]})^T A \nabla B_{j+1,[p,k]} \right]_{i,j=1}^{n(p-k)+k-1} = \sum_{\ell,q=1}^d A_{n,[p,k],\ell q}(a_{\ell q}),$$

where the matrix  $A_{n,[p,k],\ell q}(\alpha)$  is defined for all functions  $\alpha \in L^1((0,1)^d)$  as follows:

$$(6.39) \quad A_{n,[p,k],\ell q}(\alpha) = \left[ \int_{(0,1)^d} \alpha \frac{\partial B_{i+1,[p,k]}}{\partial x_\ell} \frac{\partial B_{j+1,[p,k]}}{\partial x_q} \right]_{i,j=1}^{n(p-k)+k-1}.$$

In the next lemmas, we investigate the structure of the matrix  $A_{n,[p,k],\ell q}(\alpha)$  for  $\alpha = 1$  identically. This is necessary for the GLT analysis that we tackle below. In view of what follows, for  $p, n \geq 1$  and  $0 \leq k \leq p-1$ , we define the  $(p-k) \times (p-k)$  blocks

$$\begin{aligned} K_{[p,k]}^{[\ell]} &= \left[ \int_{\mathbb{R}} \beta'_{j,[p,k]}(t) \beta'_{i,[p,k]}(t-\ell) dt \right]_{i,j=1}^{p-k}, & \ell \in \mathbb{Z}, \\ H_{[p,k]}^{[\ell]} &= \left[ \int_{\mathbb{R}} \beta'_{j,[p,k]}(t) \beta_{i,[p,k]}(t-\ell) dt \right]_{i,j=1}^{p-k}, & \ell \in \mathbb{Z}, \\ M_{[p,k]}^{[\ell]} &= \left[ \int_{\mathbb{R}} \beta_{j,[p,k]}(t) \beta_{i,[p,k]}(t-\ell) dt \right]_{i,j=1}^{p-k}, & \ell \in \mathbb{Z}, \end{aligned}$$

and the matrix-valued functions  $\kappa_{[p,k]}, \xi_{[p,k]}, \mu_{[p,k]} : [-\pi, \pi] \rightarrow \mathbb{C}^{(p-k) \times (p-k)}$ ,

$$(6.40) \quad \kappa_{[p,k]}(\theta) = \sum_{\ell \in \mathbb{Z}} K_{[p,k]}^{[\ell]} e^{i\ell\theta} = K_{[p,k]}^{[0]} + \sum_{\ell > 0} \left( K_{[p,k]}^{[\ell]} e^{i\ell\theta} + (K_{[p,k]}^{[\ell]})^T e^{-i\ell\theta} \right),$$

$$(6.41) \quad \xi_{[p,k]}(\theta) = \sum_{\ell \in \mathbb{Z}} H_{[p,k]}^{[\ell]} e^{i\ell\theta} = H_{[p,k]}^{[0]} + \sum_{\ell > 0} \left( H_{[p,k]}^{[\ell]} e^{i\ell\theta} - (H_{[p,k]}^{[\ell]})^T e^{-i\ell\theta} \right),$$

$$(6.42) \quad \mu_{[p,k]}(\theta) = \sum_{\ell \in \mathbb{Z}} M_{[p,k]}^{[\ell]} e^{i\ell\theta} = M_{[p,k]}^{[0]} + \sum_{\ell > 0} \left( M_{[p,k]}^{[\ell]} e^{i\ell\theta} + (M_{[p,k]}^{[\ell]})^T e^{-i\ell\theta} \right).$$

Due to the compact support of the reference B-splines  $\beta_{1,[p,k]}, \dots, \beta_{p-k,[p,k]}$  (see (6.30)), there are only a finite number of nonzero blocks  $K_{[p,k]}^{[\ell]}, H_{[p,k]}^{[\ell]}, M_{[p,k]}^{[\ell]}$ . Consequently, the series in (6.40)–(6.42) are actually finite sums.

NOTATION 6.4. *From now on, we will use the following notation:*

- If  $p, n \geq 1$ ,  $0 \leq k \leq p-1$ , and  $X$  is a matrix of size  $n(p-k) + k - 1$ , we denote by  $\tilde{X}$  the principal submatrix of  $X$  corresponding to the row and column indices  $i, j = k+1, \dots, k+(n-\nu)(p-k)$ , where  $\nu = \lceil (k+1)/(p-k) \rceil$  as in (6.29).
- If  $p, n \geq 1$ ,  $0 \leq k \leq p-1$ , and  $X$  is a matrix of size  $n(p-k) + k - 1$ , we denote by  $\hat{X}$  any block diagonal matrix of the form

$$\hat{X} = \begin{bmatrix} D_{k(p-k)-k} & & \\ & X & \\ & & \delta \end{bmatrix} = D_{k(p-k)-k} \oplus X \oplus [\delta],$$

where  $D_{k(p-k)-k}$  is any real diagonal matrix,  $\delta \in \mathbb{R}$ , and it is understood that the block  $D_{k(p-k)-k}$  is not present if  $k(p-k) - k = 0$  (i.e., if  $k = 0$  or  $k = p-1$ ). Note that  $\hat{X}$  has the following key properties:

- Its size  $(n+k)(p-k)$  is a multiple of  $p-k$  and is such that the difference  $(n+k)(p-k) - (n(p-k) + k - 1) = k(p-k) - k + 1 > 0$  is independent of  $n$ ;
- It contains  $X$  as a principal submatrix in such a way that  $\tilde{X}$  is the principal submatrix of  $\hat{X}$  corresponding to the row and column indices  $i, j = k(p-k) + 1, \dots, k(p-k) + (n-\nu)(p-k)$ ;

– It satisfies the matrix identity  $X = P_{n,[p,k]}^* \hat{X} P_{n,[p,k]}$ , where  $P_{n,[p,k]}$  is the  $(n+k)(p-k) \times (n(p-k)+k-1)$  matrix given by

$$P_{n,[p,k]} = \begin{bmatrix} O \\ \frac{I_{n(p-k)+k-1}}{\mathbf{0}^T} \end{bmatrix};$$

– Its eigenvalues (resp., singular values) are given by the eigenvalues (resp., singular values) of  $X$  plus further  $k(p-k)-k+1$  eigenvalues (resp., singular values) that are equal to  $\delta_1, \dots, \delta_{k(p-k)-k}, \delta$  (resp.,  $|\delta_1|, \dots, |\delta_{k(p-k)-k}|, |\delta|$ ), where  $\delta_1, \dots, \delta_{k(p-k)-k}$  are the diagonal entries of  $D_{k(p-k)-k}$ .

Throughout this work, whenever we formulate a statement regarding  $\hat{X}$  without further specifications, it is understood that the statement holds for any  $\hat{X}$ , whatever  $D_{k(p-k)-k}$  and  $\delta$ .

- If  $p, n \geq 1, 0 \leq k \leq p-1$ , and  $X$  is a matrix of size  $n(p-k)+k-1$ , we denote by  $\hat{X}^0$  (resp.,  $\hat{X}^1$ ) the matrix  $\hat{X}$  corresponding to the choice  $D_{k(p-k)-k} = O_{k(p-k)-k}$  and  $\delta = 0$  (resp.,  $D_{k(p-k)-k} = I_{k(p-k)-k}$  and  $\delta = 1$ ). The matrix  $\hat{X}^0$  satisfies not only the matrix identity  $X = P_{n,[p,k]}^* \hat{X}^0 P_{n,[p,k]}$ , but also the inverse matrix identity  $\hat{X}^0 = P_{n,[p,k]} X P_{n,[p,k]}^*$ .
- If  $p, n \geq 1, \mathbf{0} \leq k \leq p-1$ , and  $X$  is a matrix of size  $N(n(p-k)+k-1)$ , we define the matrix  $\hat{X}^0$  of size  $N((n+k)(p-k))$  as follows:

$$\hat{X}^0 = P_{n,[p,k]} X P_{n,[p,k]}^*,$$

where  $P_{n,[p,k]} = P_{n_1,[p_1,k_1]} \otimes \dots \otimes P_{n_d,[p_d,k_d]}$  and  $P_{n_i,[p_i,k_i]}$  is the matrix  $P_{n,[p,k]}$  with  $n, p, k$  replaced by  $n_i, p_i, k_i$ . Using **P8**, it can be shown that  $P_{n,[p,k]}$  is a matrix of 0 and 1, in which every row possesses at most one 1 and every column possesses exactly one 1. Moreover, by **P4**,

$$(6.43) \quad P_{n,[p,k]}^* P_{n,[p,k]} = I_{N(n(p-k)+k-1)}.$$

The matrix  $\hat{X}^0$  has the following key properties:

- $X$  is the principal submatrix of  $\hat{X}^0$  corresponding to the indices  $k(p-k)-k+1, \dots, k(p-k)+n(p-k)-1$ . Indeed, for all  $i, j = 1, \dots, n(p-k)+k-1$ , by **P8** we have

$$\begin{aligned} (\hat{X}^0)_{k(p-k)-k+i, k(p-k)-k+j} &= (P_{n,[p,k]} X P_{n,[p,k]}^*)_{k(p-k)-k+i, k(p-k)-k+j} \\ &= \sum_{a,b=1}^{n(p-k)+k-1} X_{ab} (P_{n,[p,k]})_{k(p-k)-k+i, a} (P_{n,[p,k]}^*)_{b, k(p-k)-k+j} \\ &= \sum_{a,b=1}^{n(p-k)+k-1} X_{ab} (P_{n,[p,k]})_{k(p-k)-k+i, a} (P_{n,[p,k]})_{k(p-k)-k+j, b} \\ &= \sum_{a,b=1}^{n(p-k)+k-1} X_{ab} \prod_{r=1}^d (P_{n_r,[p_r,k_r]})_{k_r(p_r-k_r)-k_r+i, a_r} (P_{n_r,[p_r,k_r]})_{k_r(p_r-k_r)-k_r+j, b_r} \\ &= X_{ij}. \end{aligned}$$

– If  $X = X_1 \otimes \cdots \otimes X_d$  with  $X_i$  of size  $n_i(p_i - k_i) + k_i - 1$ , then, by **P 4**,

$$\begin{aligned}\hat{X}^0 &= P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* = P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} (X_1 \otimes \cdots \otimes X_d) P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* = \hat{X}_1^0 \otimes \cdots \otimes \hat{X}_d^0, \\ X &= X_1 \otimes \cdots \otimes X_d = P_{n_1,[p_1,k_1]}^* \hat{X}_1^0 P_{n_1,[p_1,k_1]} \otimes \cdots \otimes P_{n_d,[p_d,k_d]}^* \hat{X}_d^0 P_{n_d,[p_d,k_d]} \\ &= P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* \hat{X}^0 P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}.\end{aligned}$$

Since any matrix  $X$  of size  $N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)$  can be written as a linear combination of tensor products of the form  $X_1 \otimes \cdots \otimes X_d$  with  $X_i$  of size  $n_i(p_i - k_i) + k_i - 1$ , we infer that the equalities

$$\begin{aligned}\hat{X}^0 &= P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^*, \\ X &= P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* \hat{X}^0 P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}.\end{aligned}$$

actually hold for all matrices  $X$  of size  $N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)$ .

- If  $i \in \{1, \dots, (\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k})\}$  is an index such that the  $i$ th row of  $P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$  is zero, then the  $i$ th row and column of  $\hat{X}^0$  are zero. This can be verified by direct computation.
- If  $\mathbf{p}, \mathbf{n} \geq 1$ ,  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p}-1$ , and  $X$  is a matrix of size  $N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)$ , we define

$$\hat{X}^1 = \hat{X}^0 + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0,$$

where  $R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0$  is the  $N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k})) \times N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k}))$  diagonal matrix of 0 and 1 such that  $(R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0)_{ii} = 1$  if and only if the  $i$ th row of  $P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$  is zero. Note that

$$(6.44) \quad \text{rank}(R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0) = N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k})) - N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1),$$

$$(6.45) \quad R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} = O = P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0,$$

$$(6.46) \quad (R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0)^2 = R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0,$$

$$(6.46) \quad P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 = I_{N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k}))}.$$

The matrix  $\hat{X}^1$  has the following properties:

- The eigenvalues (resp., singular values) of  $\hat{X}^1$  coincide with the eigenvalues (resp., singular values) of  $X$  plus further  $N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k})) - N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)$  eigenvalues (resp., singular values) that are equal to 1. Indeed, if  $\tilde{P}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$  is the square matrix of size  $N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)$  obtained from  $P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$  by deleting the zero rows, then  $\tilde{P}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$  is a permutation matrix and, by (6.46), and the definition of  $R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0$ , we have

$$\begin{aligned}\det(\hat{X}^1 - \lambda I) &= \det(P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 - \lambda I) \\ &= \det(P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 - \lambda P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* - \lambda R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0) \\ &= \det(P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} (X - \lambda I) P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + (1 - \lambda) R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0) \\ &= (1 - \lambda)^{N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k})) - N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)} \det(\tilde{P}_{\mathbf{n},[\mathbf{p},\mathbf{k}]} (X - \lambda I) \tilde{P}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^*) \\ &= (1 - \lambda)^{N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k})) - N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)} \det(X - \lambda I),\end{aligned}$$

and, similarly,

$$\det((\hat{X}^1)^* \hat{X}^1 - \lambda I) = (1 - \lambda)^{N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k})) - N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)} \det(X^* X - \lambda I).$$



– If  $X$  is invertible then  $\hat{X}^1$  is invertible as well and

$$(\hat{X}^1)^{-1} = P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X^{-1} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 = (\widehat{X^{-1}})^1.$$

Indeed, by (6.43), (6.44), and (6.45), we have

$$\begin{aligned} \hat{X}^1 (\widehat{X^{-1}})^1 &= (\hat{X}^0 + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0) (\widehat{X^{-1}})^0 + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 \\ &= (P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0) (P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X^{-1} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0) \\ &= P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X^{-1} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 \\ &\quad + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} X^{-1} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + (R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0)^2 \\ &= P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* + R_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0 = I. \end{aligned}$$

LEMMA 6.5. Let  $\mathbf{p}, \mathbf{n} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ . Then,

$$(6.47) \quad A_{\mathbf{n},[\mathbf{p},\mathbf{k}],q\ell}(1) = \left( \bigotimes_{r=1}^{q-1} M_{n_r, [p_r, k_r]} \right) \otimes K_{n_q, [p_q, k_q]} \otimes \left( \bigotimes_{r=q+1}^d M_{n_r, [p_r, k_r]} \right),$$

for  $q = 1, \dots, d$ , and

$$(6.48) \quad \begin{aligned} A_{\mathbf{n},[\mathbf{p},\mathbf{k}],q\ell}(1) &= A_{\mathbf{n},[\mathbf{p},\mathbf{k}],\ell q}(1) \\ &= - \left( \bigotimes_{r=1}^{\ell-1} M_{n_r, [p_r, k_r]} \right) \otimes H_{n_\ell, [p_\ell, k_\ell]} \otimes \left( \bigotimes_{r=\ell+1}^{q-1} M_{n_r, [p_r, k_r]} \right) \\ &\quad \otimes H_{n_q, [p_q, k_q]} \otimes \left( \bigotimes_{r=q+1}^d M_{n_r, [p_r, k_r]} \right), \end{aligned}$$

for  $1 \leq \ell < q \leq d$ , where the matrices  $K_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$ ,  $H_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$ ,  $M_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$  are defined in terms of the  $B$ -splines  $B_{2,[\mathbf{p},\mathbf{k}]}, \dots, B_{n(\mathbf{p}-\mathbf{k})+\mathbf{k},[\mathbf{p},\mathbf{k}]}$  as follows:

$$(6.49) \quad K_{\mathbf{n},[\mathbf{p},\mathbf{k}]} = \left[ \int_0^1 B'_{j+1,[\mathbf{p},\mathbf{k}]}(x) B'_{i+1,[\mathbf{p},\mathbf{k}]}(x) dx \right]_{i,j=1}^{n(\mathbf{p}-\mathbf{k})+\mathbf{k}-1},$$

$$(6.50) \quad \begin{aligned} H_{\mathbf{n},[\mathbf{p},\mathbf{k}]} &= \left[ \int_0^1 B'_{j+1,[\mathbf{p},\mathbf{k}]}(x) B_{i+1,[\mathbf{p},\mathbf{k}]}(x) dx \right]_{i,j=1}^{n(\mathbf{p}-\mathbf{k})+\mathbf{k}-1} \\ &= \left[ - \int_0^1 B_{j+1,[\mathbf{p},\mathbf{k}]}(x) B'_{i+1,[\mathbf{p},\mathbf{k}]}(x) dx \right]_{i,j=1}^{n(\mathbf{p}-\mathbf{k})+\mathbf{k}-1}, \end{aligned}$$

$$(6.51) \quad M_{\mathbf{n},[\mathbf{p},\mathbf{k}]} = \left[ \int_0^1 B_{j+1,[\mathbf{p},\mathbf{k}]}(x) B_{i+1,[\mathbf{p},\mathbf{k}]}(x) dx \right]_{i,j=1}^{n(\mathbf{p}-\mathbf{k})+\mathbf{k}-1}.$$

*Proof.* We only prove (6.48) because (6.47) is proved in the same way. For convenience, throughout this proof we write  $B_{i+1}$  instead of  $B_{i+1,[\mathbf{p},\mathbf{k}]}$  and  $B_{i_r+1}$  instead of  $B_{i_r+1,[p_r, k_r]}$ . In view of (6.39), for  $1 \leq \ell < q \leq d$  and  $i, j = 1, \dots, n(\mathbf{p} - \mathbf{k}) + \mathbf{k} - 1$ , we have

$$(A_{\mathbf{n},[\mathbf{p},\mathbf{k}],q\ell}(1))_{ij} = \int_{(0,1)^d} \frac{\partial B_{i+1}}{\partial x_\ell}(\mathbf{x}) \frac{\partial B_{j+1}}{\partial x_q}(\mathbf{x}) dx$$

$$\begin{aligned}
 &= \int_{(0,1)^d} B'_{i_\ell+1}(x_\ell) B_{j_\ell+1}(x_\ell) B_{i_q+1}(x_q) B'_{j_q+1}(x_q) \prod_{\substack{r=1 \\ r \neq \ell, q}}^d B_{i_r+1}(x_r) B_{j_r+1}(x_r) dx \\
 &= \int_0^1 B'_{i_\ell+1}(x_\ell) B_{j_\ell+1}(x_\ell) dx_\ell \int_0^1 B_{i_q+1}(x_q) B'_{j_q+1}(x_q) dx_q \\
 &\quad \times \prod_{\substack{r=1 \\ r \neq \ell, q}}^d \int_0^1 B_{i_r+1}(x_r) B_{j_r+1}(x_r) dx_r \\
 &= -(H_{n_\ell, [p_\ell, k_\ell]})_{i_\ell j_\ell} (H_{n_q, [p_q, k_q]})_{i_q j_q} \prod_{\substack{r=1 \\ r \neq \ell, q}}^d (M_{n_r, [p_r, k_r]})_{i_r j_r} \\
 &= - \left( \left( \bigotimes_{r=1}^{\ell-1} M_{n_r, [p_r, k_r]} \right) \otimes H_{n_\ell, [p_\ell, k_\ell]} \otimes \left( \bigotimes_{r=\ell+1}^{q-1} M_{n_r, [p_r, k_r]} \right) \right. \\
 &\quad \left. \otimes H_{n_q, [p_q, k_q]} \otimes \left( \bigotimes_{r=q+1}^d M_{n_r, [p_r, k_r]} \right) \right)_{ij},
 \end{aligned}$$

where the last equality follows from **P8**.  $\square$

LEMMA 6.6. Let  $p, n \geq 1$  and  $0 \leq k \leq p-1$ . Then, for the matrices in (6.49)–(6.51) we have

$$(6.52) \quad \tilde{K}_{n, [p, k]} = nT_{n-\nu}(\kappa_{[p, k]}),$$

$$(6.53) \quad \tilde{H}_{n, [p, k]} = T_{n-\nu}(\xi_{[p, k]}),$$

$$(6.54) \quad \tilde{M}_{n, [p, k]} = n^{-1}T_{n-\nu}(\mu_{[p, k]}),$$

and

$$(6.55) \quad \hat{K}_{n, [p, k]} = nT_{n+k}(\kappa_{[p, k]}) + R_{n, [p, k]}, \quad \text{rank}(R_{n, [p, k]}) \leq 2(p-k)(k+\nu),$$

$$(6.56) \quad \hat{H}_{n, [p, k]} = T_{n+k}(\xi_{[p, k]}) + S_{n, [p, k]}, \quad \text{rank}(S_{n, [p, k]}) \leq 2(p-k)(k+\nu),$$

$$(6.57) \quad \hat{M}_{n, [p, k]} = n^{-1}T_{n+k}(\mu_{[p, k]}) + Q_{n, [p, k]}, \quad \text{rank}(Q_{n, [p, k]}) \leq 2(p-k)(k+\nu).$$

*Proof.* For the proof of (6.52)–(6.54), see [8, Lemma 6.10]. To prove (6.55), simply use (6.52) and note that, in view of our Notation 6.4,  $\hat{K}_{n, [p, k]}$  is the principal submatrix of  $\tilde{K}_{n, [p, k]}$  corresponding to the row and column indices  $i, j = k(p-k)+1, \dots, k(p-k)+(n-\nu)(p-k)$ . The proof of (6.56) and (6.57) is the same as the proof of (6.55).  $\square$

LEMMA 6.7. Let  $p, n \geq 1$  and  $0 \leq k \leq p-1$ . In view of Lemma 6.5, we define

$$\hat{A}_{n, [p, k], qq}(1) = \left( \bigotimes_{r=1}^{q-1} \hat{M}_{n_r, [p_r, k_r]} \right) \otimes \hat{K}_{n_q, [p_q, k_q]} \otimes \left( \bigotimes_{r=q+1}^d \hat{M}_{n_r, [p_r, k_r]} \right),$$

for  $q = 1, \dots, d$ , and

$$\begin{aligned}
 \hat{A}_{n, [p, k], q\ell}(1) &= \hat{A}_{n, [p, k], \ell q}(1) \\
 &= - \left( \bigotimes_{r=1}^{\ell-1} \hat{M}_{n_r, [p_r, k_r]} \right) \otimes \hat{H}_{n_\ell, [p_\ell, k_\ell]} \otimes \left( \bigotimes_{r=\ell+1}^{q-1} \hat{M}_{n_r, [p_r, k_r]} \right) \\
 &\quad \otimes \hat{H}_{n_q, [p_q, k_q]} \otimes \left( \bigotimes_{r=q+1}^d \hat{M}_{n_r, [p_r, k_r]} \right),
 \end{aligned}$$

for  $1 \leq \ell < q \leq d$ . Then, for every  $\ell, q = 1, \dots, d$ ,

$$\hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}],\ell q}(1) = \frac{n_\ell n_q}{N(\mathbf{n})} \Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}} T_{\mathbf{n}+\mathbf{k}} ((H_{[\mathbf{p},\mathbf{k}]})_{\ell q}) \Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}}^T + E_{\mathbf{n},[\mathbf{p},\mathbf{k}]},$$

$$\text{rank}(E_{\mathbf{n},[\mathbf{p},\mathbf{k}]}) \leq N((\mathbf{n} + \mathbf{k})(\mathbf{p} - \mathbf{k})) \sum_{i=1}^d \frac{2(k_i + \nu_i)}{n_i + k_i},$$

where  $\boldsymbol{\nu} = (\nu_1, \dots, \nu_d)$ , each  $\nu_i$  is defined as in (6.29) with  $p_i, k_i$  in place of  $p, k$ , the matrix  $\Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}}$  is defined as in Remark 2.3 with  $\mathbf{n} + \mathbf{k}, \mathbf{p} - \mathbf{k}$  in place of  $\mathbf{m}, \mathbf{s}$ , and

$$(6.58) \quad (H_{[\mathbf{p},\mathbf{k}]})_{\ell q} = \begin{cases} (\otimes_{r=1}^{\ell-1} \mu_{[p_r, k_r]}) \otimes \kappa_{[p_\ell, k_\ell]} \otimes (\otimes_{r=\ell+1}^d \mu_{[p_r, k_r]}), & \ell = q, \\ -(\otimes_{r=1}^{\ell-1} \mu_{[p_r, k_r]}) \otimes \xi_{[p_\ell, k_\ell]} \otimes (\otimes_{r=\ell+1}^{q-1} \mu_{[p_r, k_r]}) \otimes \xi_{[p_q, k_q]} \otimes (\otimes_{r=q+1}^d \mu_{[p_r, k_r]}), & \ell < q, \\ -(\otimes_{r=1}^{q-1} \mu_{[p_r, k_r]}) \otimes \xi_{[p_q, k_q]} \otimes (\otimes_{r=q+1}^{\ell-1} \mu_{[p_r, k_r]}) \otimes \xi_{[p_\ell, k_\ell]} \otimes (\otimes_{r=\ell+1}^d \mu_{[p_r, k_r]}), & \ell > q. \end{cases}$$

*Proof.* The result follows from (6.55)–(6.57) and the properties **P9** and **T7**.  $\square$

**GLT analysis of the higher-order FE discretization matrices.** In what follows, we assume that  $\mathbf{n} = \gamma n$ , where  $\gamma \in \mathbb{Q}^d$  is a fixed vector with positive components and  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} = \gamma n \in \mathbb{N}^d$ . This assumption essentially says that each stepsize  $h_i = \frac{1}{n_i}$  tends to 0 with the same asymptotic speed as the others. The main result of this section is Theorem 6.8, which gives the spectral distribution of the normalized sequence  $\{n^{d-2} A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A)\}_n$ .

**THEOREM 6.8.** *Suppose that  $a_{\ell q} \in L^1((0, 1)^d)$ , for every  $\ell, q = 1, \dots, d$ , and that the matrix  $A(\mathbf{x}) = [a_{\ell q}(\mathbf{x})]_{\ell, q=1}^d$  is symmetric for every  $\mathbf{x} \in (0, 1)^d$ . Let  $\mathbf{p} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ , let  $\gamma \in \mathbb{Q}^d$  be a vector with positive components, and assume that  $\mathbf{n} = \gamma n$  (it is understood that  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} = \gamma n \in \mathbb{N}^d$ ). Then*

$$\{n^{d-2} A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A)\}_n \sim_{\sigma, \lambda} \kappa_{[\mathbf{p},\mathbf{k}]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) = \frac{1}{N(\gamma)} \sum_{\ell, q=1}^d \gamma_\ell \gamma_q a_{\ell q}(\mathbf{x}) (H_{[\mathbf{p},\mathbf{k}]})_{\ell q}(\boldsymbol{\theta}),$$

where  $H_{[\mathbf{p},\mathbf{k}]}$  is defined in (6.58).

*Proof.* The proof consists of the following steps. Throughout this proof, the letter  $C$  denotes a generic constant independent of  $n$ .

*Step 1.* Let  $L^1((0, 1)^d, \mathbb{R}^{d \times d})$  be the space of functions  $L : (0, 1)^d \rightarrow \mathbb{R}^{d \times d}$  such that  $L_{ij} \in L^1((0, 1)^d)$  for all  $i, j = 1, \dots, d$ . Consider the linear operator

$$A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(\cdot) : L^1((0, 1)^d, \mathbb{R}^{d \times d}) \rightarrow \mathbb{R}^{N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1) \times N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1)},$$

$$A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(L) = \left[ \int_{(0,1)^d} (\nabla B_{i+1, [\mathbf{p}, \mathbf{k}]})^T L \nabla B_{j+1, [\mathbf{p}, \mathbf{k}]} \right]_{i, j=1}^{\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1}.$$

The next steps are devoted to showing that

(6.59)

$$\{n^{d-2} \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(L) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{1}{N(\gamma)} \sum_{\ell, q=1}^d \gamma_\ell \gamma_q L_{\ell q}(\mathbf{x}) (H_{[\mathbf{p}, \mathbf{k}]})_{\ell q}(\boldsymbol{\theta}),$$

where, according to our Notation 6.4,  $\hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(L) = P_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]} A_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(L) P_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^*$ . Once this is done, the theorem is proved. Indeed, since  $A(\mathbf{x}) = [a_{\ell q}(\mathbf{x})]_{\ell, q=1}^d$  is symmetric for all

$\mathbf{x} \in (0, 1)^d$  by assumption, the matrix  $\hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(A)$  is symmetric as well. Hence, it follows from (6.59) and **GLT 1** that

$$\begin{aligned} \{n^{d-2} \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(A) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n &\sim_{\sigma, \lambda} \frac{1}{N(\boldsymbol{\gamma})} \sum_{\ell, q=1}^d \gamma_\ell \gamma_q a_{\ell q}(\mathbf{x}) (H_{[\mathbf{p}, \mathbf{k}]})_{\ell q}(\boldsymbol{\theta}) \\ &= \kappa_{[\mathbf{p}, \mathbf{k}]}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}), \end{aligned}$$

which implies

$$\{n^{d-2} \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(A)\}_n \sim_{\sigma, \lambda} \kappa_{[\mathbf{p}, \mathbf{k}]}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}),$$

which implies

$$\{n^{d-2} A_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(A)\}_n \sim_{\sigma, \lambda} \kappa_{[\mathbf{p}, \mathbf{k}]}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta})$$

by **S 5** (taking into account that  $A_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(A) = P_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^* \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(A) P_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}$ ; see Notation 6.4).

*Step 2.* We first prove (6.59) in the constant-coefficient case where  $L(\mathbf{x}) = E_{\ell q}^{(d)}$  identically.

In this case, we have  $A_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(E_{\ell q}^{(d)}) = A_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], \ell q}(1)$ . By Lemma 6.7,

$$\begin{aligned} n^{d-2} \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(E_{\ell q}^{(d)}) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}} &= n^{d-2} \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], \ell q}^0(1) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}} \\ &= \frac{\gamma_\ell \gamma_q}{N(\boldsymbol{\gamma})} T_{\mathbf{n}+\mathbf{k}}((H_{[\mathbf{p}, \mathbf{k}]})_{\ell q}) + F_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}, \end{aligned}$$

where  $\text{rank}(F_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}) \leq Cn^{d-1}$ . It follows from **GLT 3** and **GLT 4** that

$$\{n^{d-2} \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(E_{\ell q}^{(d)}) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{\gamma_\ell \gamma_q}{N(\boldsymbol{\gamma})} (H_{[\mathbf{p}, \mathbf{k}]})_{\ell q}(\boldsymbol{\theta}),$$

which is (6.59) for  $L(\mathbf{x}) = E_{\ell q}^{(d)}$ .

*Step 3.* Now we prove (6.59) in the case where  $L(\mathbf{x}) = a(\mathbf{x}) E_{\ell q}^{(d)}$  with  $a \in C([0, 1]^d)$ . Let

$$Z_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]} = n^{d-2} \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(a E_{\ell q}^{(d)}) - n^{d-2} D_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}(a) \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(E_{\ell q}^{(d)}).$$

Let  $\boldsymbol{\nu} = (\nu_1, \dots, \nu_d)$ , with  $\nu_i$  defined as in (6.29) with  $p_i, k_i$  replaced by  $p, k$ . Taking into account Notation 6.4 and the properties (6.31)–(6.36), for  $\mathbf{r}, \mathbf{R} = \mathbf{1}, \dots, \mathbf{n} - \boldsymbol{\nu}$  and  $\mathbf{s}, \mathbf{S} = \mathbf{1}, \dots, \mathbf{p} - \mathbf{k}$ , we have

$$\begin{aligned} & |(n^{2-d} Z_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]})_{\mathbf{k}(\mathbf{p}-\mathbf{k})+(\mathbf{p}-\mathbf{k})(\mathbf{r}-\mathbf{1})+\mathbf{s}, \mathbf{k}(\mathbf{p}-\mathbf{k})+(\mathbf{p}-\mathbf{k})(\mathbf{R}-\mathbf{1})+\mathbf{S}}| \\ &= \left| (A_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(a E_{\ell q}^{(d)}))_{\mathbf{k}+(\mathbf{p}-\mathbf{k})(\mathbf{r}-\mathbf{1})+\mathbf{s}, \mathbf{k}+(\mathbf{p}-\mathbf{k})(\mathbf{R}-\mathbf{1})+\mathbf{S}} \right. \\ &\quad \left. - (D_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}(a))_{\mathbf{k}(\mathbf{p}-\mathbf{k})+(\mathbf{p}-\mathbf{k})(\mathbf{r}-\mathbf{1})+\mathbf{s}, \mathbf{k}(\mathbf{p}-\mathbf{k})+(\mathbf{p}-\mathbf{k})(\mathbf{R}-\mathbf{1})+\mathbf{s}} \right. \\ &\quad \left. \times (A_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(E_{\ell q}^{(d)}))_{\mathbf{k}+(\mathbf{p}-\mathbf{k})(\mathbf{r}-\mathbf{1})+\mathbf{s}, \mathbf{k}+(\mathbf{p}-\mathbf{k})(\mathbf{R}-\mathbf{1})+\mathbf{S}} \right| \\ &= \left| \int_{(0,1)^d} \left[ a(\mathbf{x}) - a\left(\frac{\mathbf{k} + \mathbf{r}}{\mathbf{n} + \mathbf{k}}\right) \right] \frac{\partial B_{\mathbf{k}+\mathbf{1}+(\mathbf{p}-\mathbf{k})(\mathbf{r}-\mathbf{1})+\mathbf{s}, [\mathbf{p}, \mathbf{k}]}(\mathbf{x})}{\partial x_\ell} \right. \\ &\quad \left. \times \frac{\partial B_{\mathbf{k}+\mathbf{1}+(\mathbf{p}-\mathbf{k})(\mathbf{R}-\mathbf{1})+\mathbf{S}, [\mathbf{p}, \mathbf{k}]}(\mathbf{x})}{\partial x_q} \right| \mathrm{d}\mathbf{x} \\ &\leq C_{p_\ell} n_\ell C_{p_q} n_q \int_{[\boldsymbol{\tau}_{\mathbf{k}+\mathbf{1}+(\mathbf{p}-\mathbf{k})(\mathbf{r}-\mathbf{1})+\mathbf{s}}, \boldsymbol{\tau}_{\mathbf{k}+\mathbf{1}+(\mathbf{p}-\mathbf{k})(\mathbf{r}-\mathbf{1})+\mathbf{s}+\mathbf{p}+\mathbf{1}}]} \left| a(\mathbf{x}) - a\left(\frac{\mathbf{k} + \mathbf{r}}{\mathbf{n} + \mathbf{k}}\right) \right| \mathrm{d}\mathbf{x} \end{aligned}$$

$$\begin{aligned}
 &\leq C_{p_\ell} C_{p_q} n_\ell n_q \int_{[(r-1)/n, (r+p)/n]} \left| a(\mathbf{x}) - a\left(\frac{\mathbf{k} + \mathbf{r}}{\mathbf{n} + \mathbf{k}}\right) \right| d\mathbf{x} \\
 &\leq C_{p_\ell} C_{p_q} n_\ell n_q \omega_a\left(\frac{2 \max(\mathbf{p})}{\min(\mathbf{n})}\right) \prod_{i=1}^d \frac{p_i + 1}{n_i},
 \end{aligned}$$

where  $\omega_a(\cdot)$  is the modulus of continuity of  $a$  and the last inequality is justified by the fact that the maximum distance (in the  $\infty$ -norm) of a point  $\mathbf{x}$  in the hyperrectangle  $[(r-1)/n, (r+p)/n]$  from the point  $(\mathbf{k} + \mathbf{r})/(\mathbf{n} + \mathbf{k})$  is not larger than  $\max(2\mathbf{p}/\mathbf{n}) \leq 2 \max(\mathbf{p})/\min(\mathbf{n})$ . It follows that, if we denote by  $Z_n$  the principal submatrix of  $Z_{n, [\mathbf{p}, \mathbf{k}]}$  corresponding to the indices  $\mathbf{k}(\mathbf{p} - \mathbf{k}) + \mathbf{1}, \dots, \mathbf{k}(\mathbf{p} - \mathbf{k}) + (\mathbf{n} - \nu)(\mathbf{p} - \mathbf{k})$ , each entry of  $Z_n$  is bounded in modulus by  $C\omega_a(1/\min(\mathbf{n}))$ . Moreover, the number of nonzero entries in each row and column of  $Z_n$  is bounded by a constant  $C$  independent of  $n$ . This follows from the following more general property:

for every  $L \in L^1((0, 1)^d, \mathbb{R}^{d \times d})$ , the number of nonzero entries in each row and column of  $A_{n, [\mathbf{p}, \mathbf{k}]}(L)$  is bounded by  $(2|\mathbf{p}|_\infty + 1)^d$  because  $(A_{n, [\mathbf{p}, \mathbf{k}]}(L))_{ij} = 0$  whenever  $|i - j|_\infty > |\mathbf{p}|_\infty$ ; this is due to the fact that, for  $|i - j|_\infty > |\mathbf{p}|_\infty$ , the intersection of the supports of  $B_{i+1, [\mathbf{p}, \mathbf{k}]}$  and  $B_{j+1, [\mathbf{p}, \mathbf{k}]}$  has zero measure by the local support property (6.31).

Thus, by **N1**,  $\|Z_n\| \leq C\omega_a(1/\min(\mathbf{n})) \rightarrow 0$  as  $n \rightarrow \infty$ . Recalling that  $Z_n$  is the principal submatrix of  $Z_{n, [\mathbf{p}, \mathbf{k}]}$  corresponding to the indices  $\mathbf{k}(\mathbf{p} - \mathbf{k}) + \mathbf{1}, \dots, \mathbf{k}(\mathbf{p} - \mathbf{k}) + (\mathbf{n} - \nu)(\mathbf{p} - \mathbf{k})$ , we arrive at

$$Z_{n, [\mathbf{p}, \mathbf{k}]} = N_{n, [\mathbf{p}, \mathbf{k}]} + R_{n, [\mathbf{p}, \mathbf{k}]},$$

where  $\|N_{n, [\mathbf{p}, \mathbf{k}]}\| = \|Z_n\| \rightarrow 0$  as  $n \rightarrow \infty$  and  $\text{rank}(R_{n, [\mathbf{p}, \mathbf{k}]}) \leq Cn^{d-1}$ . It follows from **Z1** that  $\{Z_{n, [\mathbf{p}, \mathbf{k}]}\}_n$  is zero-distributed. Since

$$n^{d-2} \hat{A}_{n, [\mathbf{p}, \mathbf{k}]}^0(aE_{\ell_q}^{(d)}) = n^{d-2} D_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}}(a) \hat{A}_{n, [\mathbf{p}, \mathbf{k}]}^0(E_{\ell_q}^{(d)}) + Z_{n, [\mathbf{p}, \mathbf{k}]},$$

we infer from **D3** that

$$\begin{aligned}
 &n^{d-2} \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{n, [\mathbf{p}, \mathbf{k}]}^0(aE_{\ell_q}^{(d)}) \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}} \\
 &= n^{d-2} D_{n+\mathbf{k}}(aI_{N(\mathbf{p}-\mathbf{k})}) \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{n, [\mathbf{p}, \mathbf{k}]}^0(E_{\ell_q}^{(d)}) \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}} \\
 &\quad + \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T Z_{n, [\mathbf{p}, \mathbf{k}]} \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}},
 \end{aligned}$$

and we conclude that

$$\{n^{d-2} \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{n, [\mathbf{p}, \mathbf{k}]}^0(aE_{\ell_q}^{(d)}) \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{\gamma_\ell \gamma_q}{N(\gamma)} a(\mathbf{x})(H_{[\mathbf{p}, \mathbf{k}]})_{\ell_q}(\boldsymbol{\theta})$$

by Step 2, **GLT3**, and **GLT4**. The previous relation is (6.59) for  $L(\mathbf{x}) = a(\mathbf{x})E_{\ell_q}^{(d)}$ .

*Step 4.* Now we prove (6.59) in the case where  $L(\mathbf{x}) = a(\mathbf{x})E_{\ell_q}^{(d)}$  with  $a \in L^1((0, 1)^d)$ . By the density of  $C([0, 1]^d)$  in  $L^1((0, 1)^d)$ , there exist functions  $a_m \in C([0, 1]^d)$  such that  $a_m \rightarrow a$  in  $L^1((0, 1)^d)$ . By Step 3,

$$\{n^{d-2} \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{n, [\mathbf{p}, \mathbf{k}]}^0(a_m E_{\ell_q}^{(d)}) \Gamma_{n+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{\gamma_\ell \gamma_q}{N(\gamma)} a_m(\mathbf{x})(H_{[\mathbf{p}, \mathbf{k}]})_{\ell_q}(\boldsymbol{\theta}).$$

Moreover,

$$\frac{\gamma_\ell \gamma_q}{N(\gamma)} a_m(\mathbf{x})(H_{[\mathbf{p}, \mathbf{k}]})_{\ell_q}(\boldsymbol{\theta}) \rightarrow \frac{\gamma_\ell \gamma_q}{N(\gamma)} a(\mathbf{x})(H_{[\mathbf{p}, \mathbf{k}]})_{\ell_q}(\boldsymbol{\theta}) \text{ in measure.}$$

We show that

$$(6.60) \quad \begin{aligned} & \{n^{d-2}\Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(a_m E_{\ell_q}^{(d)}) \Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}}\}_n \\ & \xrightarrow{\text{a.c.s.}} \{n^{d-2}\Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(a E_{\ell_q}^{(d)}) \Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}}\}_n. \end{aligned}$$

Once this is done, the thesis (6.59) follows immediately from **GLT 7**. To prove (6.60), let  $\hat{P}_{\mathbf{n},[\mathbf{p},\mathbf{k}]} = [P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}|O]$  be the square matrix of size  $N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k}))$  obtained from  $P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}$  by padding with zeros. Taking into account that  $\|\hat{P}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}\| = 1$ , by **N 3**, **N 5**, and (6.36), we have

$$\begin{aligned} & \|\hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(a E_{\ell_q}^{(d)}) - \hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(a_m E_{\ell_q}^{(d)})\|_1 = \|\hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0((a - a_m) E_{\ell_q}^{(d)})\|_1 \\ & = \|P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}((a - a_m) E_{\ell_q}^{(d)}) P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^*\|_1 \\ & = \|\hat{P}_{\mathbf{n},[\mathbf{p},\mathbf{k}]} [A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}((a - a_m) E_{\ell_q}^{(d)}) \oplus O] \hat{P}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^*\|_1 \\ & \leq \|A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}((a - a_m) E_{\ell_q}^{(d)})\|_1 \\ & \leq \sum_{i,j=1}^{n(\mathbf{p}-\mathbf{k})+\mathbf{k}-1} \left| \int_{(0,1)^d} [a(\mathbf{x}) - a_m(\mathbf{x})] \frac{\partial B_{i+1,[\mathbf{p},\mathbf{k}]}(\mathbf{x})}{\partial x_\ell} \frac{\partial B_{j+1,[\mathbf{p},\mathbf{k}]}(\mathbf{x})}{\partial x_q} d\mathbf{x} \right| \\ & \leq \int_{(0,1)^d} |a(\mathbf{x}) - a_m(\mathbf{x})| \sum_{i,j=1}^{n(\mathbf{p}-\mathbf{k})+\mathbf{k}-1} \left| \frac{\partial B_{i+1,[\mathbf{p},\mathbf{k}]}(\mathbf{x})}{\partial x_\ell} \right| \left| \frac{\partial B_{j+1,[\mathbf{p},\mathbf{k}]}(\mathbf{x})}{\partial x_q} \right| d\mathbf{x} \\ & \leq C_{p_\ell} C_{p_q} n_\ell n_q \|a - a_m\|_{L^1}. \end{aligned}$$

Thus, the a.c.s. convergence (6.60) follows from **ACS 6**.

*Step 5.* Finally, we prove (6.59) for an arbitrary  $L \in L^1((0,1)^d, \mathbb{R}^{d \times d})$ . Write

$$L(\mathbf{x}) = \sum_{\ell,q=1}^d L_{\ell q}(\mathbf{x}) E_{\ell q}^{(d)},$$

and note that, by linearity,

$$\hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(L) = \sum_{\ell,q=1}^d \hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(L_{\ell q} E_{\ell q}^{(d)}).$$

Hence, by Step 4 and **GLT 4**,

$$\{n^{d-2}\Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(L) \Gamma_{\mathbf{n}+\mathbf{k},\mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{1}{N(\boldsymbol{\gamma})} \sum_{\ell,q=1}^d \gamma_\ell \gamma_q L_{\ell q}(\mathbf{x}) (H_{[\mathbf{p},\mathbf{k}]}(\boldsymbol{\theta}))_{\ell q}(\boldsymbol{\theta}),$$

which concludes the proof.  $\square$

**REMARK 6.9** (Space-time higher-order FE-DG discretization of time-dependent diffusion equations). Consider the time-dependent diffusion equation

$$(6.61) \quad \begin{cases} \partial_t u(t, \mathbf{x}) - \nabla \cdot A(\mathbf{x}) \nabla u(t, \mathbf{x}) = f(t, \mathbf{x}), & (t, \mathbf{x}) \in (0, T) \times (0, 1)^d, \\ u(t, \mathbf{x}) = 0, & (t, \mathbf{x}) \in (0, T) \times \partial((0, 1)^d), \\ u(t, \mathbf{x}) = 0, & (t, \mathbf{x}) \in \{0\} \times (0, 1)^d. \end{cases}$$

If we discretize (6.61) by the space-time higher-order FE-DG approximation technique considered in [12], the resulting (normalized) FE-DG discretization matrices enjoy an asymptotic

spectral distribution described by a  $(q + 1)N(\mathbf{p} - \mathbf{k}) \times (q + 1)N(\mathbf{p} - \mathbf{k})$  matrix-valued function. This result was (only partially) proved in [12, Theorem 5.2] by a direct (complicated and cumbersome) approach. By following step by step the proof of Theorem 6.8, we can give an alternative (much more lucid and simpler) proof of [12, Theorem 5.2] based on the theory of multilevel block GLT sequences.

REMARK 6.10 (Formal structure of the symbol). From a formal point of view (i.e., disregarding the regularity of  $A$  and  $u$ ), problem (6.19) can be rewritten in the form

$$\begin{cases} - \sum_{\ell, q=1}^d a_{\ell q} \frac{\partial^2 u}{\partial x_\ell \partial x_q} - \sum_{\ell, q=1}^d \frac{\partial a_{\ell q}}{\partial x_\ell} \frac{\partial u}{\partial x_q} = f, & \text{in } (0, 1)^d, \\ u = 0, & \text{on } \partial((0, 1)^d). \end{cases}$$

The formal structure of the (singular value and spectral) symbol

$$\kappa_{[\mathbf{p}, \mathbf{k}]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) = \sum_{\ell, q=1}^d \frac{\gamma_\ell \gamma_q}{N(\gamma)} a_{\ell q}(\mathbf{x}) (H_{[\mathbf{p}, \mathbf{k}]})_{\ell q}(\boldsymbol{\theta})$$

is deeply connected with the structure of the higher-order differential operator

$$- \sum_{\ell, q=1}^d a_{\ell q}(\mathbf{x}) \frac{\partial^2 u}{\partial x_\ell \partial x_q}(\mathbf{x})$$

associated with problem (6.19), whereas the lower-order differential operator

$$- \sum_{\ell, q=1}^d \frac{\partial a_{\ell q}}{\partial x_\ell}(\mathbf{x}) \frac{\partial u}{\partial x_q}(\mathbf{x})$$

does not enter the expression of the symbol. For more insights into this topic, we refer the reader to [41, Section 7.2]; see also [41, Remarks 7.4, 7.6, 7.9, 7.12].

**6.3. Higher-order FE discretization of convection-diffusion-reaction equations.** Suppose we add to the diffusion equation (6.19) a convection and a reaction term. In this way, we obtain the following convection-diffusion-reaction problem:

$$(6.62) \quad \begin{cases} -\nabla \cdot A \nabla u + \mathbf{b} \cdot \nabla u + cu = f, & \text{in } (0, 1)^d, \\ u = 0, & \text{on } \partial((0, 1)^d), \end{cases} \\ \iff \begin{cases} - \sum_{\ell, q=1}^d \frac{\partial}{\partial x_\ell} \left( a_{\ell q} \frac{\partial u}{\partial x_q} \right) + \sum_{q=1}^d b_q \frac{\partial u}{\partial x_q} + cu = f, & \text{in } (0, 1)^d, \\ u = 0, & \text{on } \partial((0, 1)^d), \end{cases}$$

Based on Remark 6.10 and the discussion in [41, Section 7.2], we expect the term  $\mathbf{b} \cdot \nabla u + cu$ , which only involves lower-order derivatives of  $u$ , not to enter the expression of the symbol. In other words, if we consider for problem (6.62) the same higher-order FE discretization as in Section 6.2, the symbol of the resulting sequence of (normalized) FE discretization matrices should be again  $\kappa_{[\mathbf{p}, \mathbf{k}]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta})$  as per Theorem 6.8. We are going to show that this is in fact the case.

**FE discretization.** The weak form of (6.62) reads as follows [21, Chapter 9]: find  $u \in H_0^1((0, 1)^d)$  such that

$$a(u, w) = f(w), \quad \forall w \in H_0^1((0, 1)^d),$$

where

$$\begin{aligned} a(u, w) &= \int_{(0,1)^d} ((\nabla w)^T A \nabla u + (\nabla u)^T \mathbf{b} w + c u w), \\ f(w) &= \int_{(0,1)^d} f w. \end{aligned}$$

In the FE method, we fix a set of basis functions  $\{\varphi_1, \dots, \varphi_N\} \subset H_0^1((0, 1)^d)$ , and we look for an approximation of the exact solution in the space  $\mathcal{W} = \text{span}(\varphi_1, \dots, \varphi_N)$  by solving the following discrete problem: find  $u_{\mathcal{W}} \in \mathcal{W}$  such that

$$a(u_{\mathcal{W}}, w) = f(w), \quad \forall w \in \mathcal{W}.$$

Since  $\{\varphi_1, \dots, \varphi_N\}$  is a basis for  $\mathcal{W}$ , we can write  $u_{\mathcal{W}} = \sum_{j=1}^N u_j \varphi_j$  for a unique vector  $\mathbf{u} = (u_1, \dots, u_N)^T$ . By linearity, the computation of  $u_{\mathcal{W}}$  (i.e., of  $\mathbf{u}$ ) reduces to solving the linear system

$$S^{\text{stiff}} \mathbf{u} = \mathbf{f},$$

where  $\mathbf{f} = (f(\varphi_1), \dots, f(\varphi_N))^T$  and  $S$  is the stiffness matrix,

$$S^{\text{stiff}} = [a(\varphi_j, \varphi_i)]_{i,j=1}^N.$$

Note that  $S$  admits the following decomposition:

$$(6.63) \quad S^{\text{stiff}} = A^{\text{stiff}} + Z^{\text{stiff}},$$

where

$$A^{\text{stiff}} = \left[ \int_{(0,1)^d} (\nabla \varphi_i)^T A \nabla \varphi_j \right]_{i,j=1}^N$$

is the diffusion matrix and

$$Z^{\text{stiff}} = \left[ \int_{(0,1)^d} (\nabla \varphi_j)^T \mathbf{b} \varphi_i + \int_{(0,1)^d} c \varphi_j \varphi_i \right]_{i,j=1}^N$$

is the sum of the convection and reaction matrices.

**GLT analysis of the higher-order FE discretization matrices.** Following the higher-order FE approach as in Section 6.2, the basis functions  $\varphi_1, \dots, \varphi_N$  are chosen as in (6.37). The stiffness matrix resulting from this choice will be denoted by  $S_{\mathbf{n}, [p, k]}(A, \mathbf{b}, c)$ . According to (6.63), it can be decomposed as follows:

$$S_{\mathbf{n}, [p, k]}(A, \mathbf{b}, c) = A_{\mathbf{n}, [p, k]}(A) + Z_{\mathbf{n}, [p, k]}(\mathbf{b}, c),$$



where

$$\begin{aligned}
 A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A) &= \left[ \int_{(0,1)^d} (\nabla B_{\mathbf{i}+1,[\mathbf{p},\mathbf{k}]})^T A \nabla B_{\mathbf{j}+1,[\mathbf{p},\mathbf{k}]} \right]_{\mathbf{i},\mathbf{j}=1}^{\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1}, \\
 Z_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(\mathbf{b}, c) &= \left[ \int_{(0,1)^d} (\nabla B_{\mathbf{j}+1,[\mathbf{p},\mathbf{k}]})^T \mathbf{b} B_{\mathbf{i}+1,[\mathbf{p},\mathbf{k}]} \right. \\
 &\quad \left. + \int_{(0,1)^d} c B_{\mathbf{j}+1,[\mathbf{p},\mathbf{k}]} B_{\mathbf{i}+1,[\mathbf{p},\mathbf{k}]} \right]_{\mathbf{i},\mathbf{j}=1}^{\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-1}.
 \end{aligned}$$

Note that  $A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A)$  is the same as in (6.38) and Theorem 6.8. The main result of this section is Theorem 6.11, which shows that Theorem 6.8 holds unchanged with  $S_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A, \mathbf{b}, c)$  in place of  $A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A)$ . This highlights a general aspect: *lower-order terms such as  $\mathbf{b} \cdot \nabla u + cu$  do not enter the expression of the symbol and do not affect in any way the asymptotic singular value and eigenvalue distribution of PDE discretization matrices.*

**THEOREM 6.11.** *Suppose that  $a_{\ell q} \in L^1((0,1)^d)$ , for  $\ell, q = 1, \dots, d$ , that  $b_q, c \in L^\infty((0,1)^d)$ , for  $q = 1, \dots, d$ , and that the matrix  $A(\mathbf{x}) = [a_{\ell q}(\mathbf{x})]_{\ell, q=1}^d$  is symmetric for every  $\mathbf{x} \in (0,1)^d$ . Let  $\mathbf{p} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ , let  $\boldsymbol{\gamma} \in \mathbb{Q}^d$  be a vector with positive components, and assume that  $\mathbf{n} = \boldsymbol{\gamma} n$  (it is understood that  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} = \boldsymbol{\gamma} n \in \mathbb{N}^d$ ). Then*

(6.64)

$$\{n^{d-2} S_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A, \mathbf{b}, c)\}_n \sim_{\sigma, \lambda} \kappa_{[\mathbf{p},\mathbf{k}]}^{(\boldsymbol{\gamma})}(\mathbf{x}, \boldsymbol{\theta}) = \frac{1}{N(\boldsymbol{\gamma})} \sum_{\ell, q=1}^d \gamma_\ell \gamma_q a_{\ell q}(\mathbf{x}) (H_{[\mathbf{p},\mathbf{k}]})_{\ell q}(\boldsymbol{\theta}),$$

where  $H_{[\mathbf{p},\mathbf{k}]}$  is defined in (6.58).

*Proof.* Throughout this proof, we make use of Notation 6.4, and we use the letter  $C$  to denote a generic constant independent of  $n$ . We are going to show that

$$(6.65) \quad \|Z_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(\mathbf{b}, c)\| \leq C n^{1-d}.$$

Once this is done, the thesis is proved. Indeed, it follows from (6.65) that

$$(6.66) \quad \begin{aligned}
 \|n^{d-2} Z_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(\mathbf{b}, c)\| &\leq C n^{-1}, \\
 \|n^{d-2} Z_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(\mathbf{b}, c)\|_2 &\leq C n^{d/2-1} = o(n^{d/2}).
 \end{aligned}$$

Thus:

- the singular value distribution in (6.64) follows from Theorem 6.8 and ACS 2, taking into account that, in view of (6.66) and the decomposition

$$(6.67) \quad n^{d-2} S_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A, \mathbf{b}, c) = n^{d-2} A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A) + n^{d-2} Z_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(\mathbf{b}, c),$$

we have  $\{n^{d-2} A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A)\}_n \xrightarrow{\text{a.c.s.}} \{n^{d-2} S_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A, \mathbf{b}, c)\}_n$ ,

- the eigenvalue distribution in (6.64) follows from S 4 applied to the decomposition (6.67), taking into account Theorem 6.8.

It only remains to prove (6.65). We first note that the number of nonzero entries in each row and column of  $Z_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(\mathbf{b}, c)$  is bounded by  $(2|\mathbf{p}|_\infty + 1)^d$  because  $(Z_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(\mathbf{b}, c))_{\mathbf{i}\mathbf{j}} = 0$  whenever  $|\mathbf{i} - \mathbf{j}|_\infty > |\mathbf{p}|_\infty$ . Indeed, for  $|\mathbf{i} - \mathbf{j}|_\infty > |\mathbf{p}|_\infty$ , the intersection of the supports of

$B_{i+1, [\mathbf{p}, \mathbf{k}]}$  and  $B_{j+1, [\mathbf{p}, \mathbf{k}]}$  has zero measure by the local support property (6.31). Moreover, by (6.31)–(6.36), for every  $i, j = 1, \dots, \mathbf{n}(\mathbf{p} - \mathbf{k}) + \mathbf{k} - 1$ , we have

$$\begin{aligned}
 |(Z_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(\mathbf{b}, c))_{ij}| &= \left| \int_{(0,1)^d} (\nabla B_{j+1, [\mathbf{p}, \mathbf{k}]})^T \mathbf{b} B_{i+1, [\mathbf{p}, \mathbf{k}]} + \int_{(0,1)^d} c B_{j+1, [\mathbf{p}, \mathbf{k}]} B_{i+1, [\mathbf{p}, \mathbf{k}]} \right| \\
 &\leq \int_{\text{supp}(B_{i+1, [\mathbf{p}, \mathbf{k}]})} \sum_{q=1}^d |b_q| \left| \frac{\partial B_{j+1, [\mathbf{p}, \mathbf{k}]} }{\partial x_q} \right| |B_{i+1, [\mathbf{p}, \mathbf{k}]}| \\
 &\quad + \int_{\text{supp}(B_{i+1, [\mathbf{p}, \mathbf{k}]})} |c| |B_{j+1, [\mathbf{p}, \mathbf{k}]}| |B_{i+1, [\mathbf{p}, \mathbf{k}]}| \\
 &\leq \int_{\text{supp}(B_{i+1, [\mathbf{p}, \mathbf{k}]})} \max_{\ell=1, \dots, d} \|b_\ell\|_{L^\infty} \sum_{q=1}^d C_{p_q} n_q + \int_{\text{supp}(B_{i+1, [\mathbf{p}, \mathbf{k}]})} \|c\|_{L^\infty} \\
 &\leq \max_{\ell=1, \dots, d} \|b_\ell\|_{L^\infty} \sum_{q=1}^d C_{p_q} n_q N\left(\frac{\mathbf{p} + \mathbf{1}}{\mathbf{n}}\right) + \|c\|_{L^\infty} N\left(\frac{\mathbf{p} + \mathbf{1}}{\mathbf{n}}\right) \\
 &\leq C n^{1-d},
 \end{aligned}$$

and (6.65) is proved.  $\square$

**6.4. Higher-order FE discretization of systems of PDEs.** Consider again the same system of PDEs as in Section 6.1, i.e.,

$$(6.68) \quad \begin{cases} -\nabla \cdot A \nabla u + \mathbf{b} \cdot \nabla v = f, & \text{in } (0, 1)^d, \\ \mathbf{c} \cdot \nabla u + \rho v = g, & \text{in } (0, 1)^d, \\ u = v = 0, & \text{on } \partial((0, 1)^d), \end{cases} \\
 \Leftrightarrow \begin{cases} -\sum_{\ell, q=1}^d \frac{\partial}{\partial x_\ell} \left( a_{\ell q} \frac{\partial u}{\partial x_q} \right) + \sum_{q=1}^d b_q \frac{\partial v}{\partial x_q} = f, & \text{in } (0, 1)^d, \\ \sum_{q=1}^d c_q \frac{\partial u}{\partial x_q} + \rho v = g, & \text{in } (0, 1)^d, \\ u = v = 0, & \text{on } \partial((0, 1)^d), \end{cases}$$

where  $a_{\ell q}, b_q, c_q, \rho, f, g$  are given functions,  $A = [a_{\ell q}]_{\ell, q=1}^d$ ,  $\mathbf{b} = [b_q]_{q=1}^d$ , and  $\mathbf{c} = [c_q]_{q=1}^d$ . In this section we consider the higher-order FE discretization of (6.68). Through the theory of multilevel block GLT sequences we show that, under suitable assumptions on the PDE coefficients, the corresponding sequence of (normalized) FE discretization matrices enjoys a spectral distribution described by a  $2N(\mathbf{p} - \mathbf{k}) \times 2N(\mathbf{p} - \mathbf{k})$  matrix-valued function, where  $p_i$  and  $k_i$  are, respectively, the degree and the smoothness in the  $i$ th direction of the piecewise polynomial functions involved in the FE approximation, while the number 2 in front of  $N(\mathbf{p} - \mathbf{k})$  coincides with the number of equations that compose the system (6.68).

**FE discretization.** The weak form of (6.68) reads as follows: find  $u, v \in H_0^1((0, 1)^d)$  such that, for all  $w \in H_0^1((0, 1)^d)$ ,

$$\begin{cases} \int_{(0,1)^d} (\nabla w)^T A \nabla u + \int_{(0,1)^d} (\nabla v)^T \mathbf{b} w = \int_{(0,1)^d} f w, \\ \int_{(0,1)^d} (\nabla u)^T \mathbf{c} w + \int_{(0,1)^d} \rho v w = \int_{(0,1)^d} g w. \end{cases}$$

In the FE method, we fix a set of basis functions  $\{\varphi_1, \dots, \varphi_N\} \subset H_0^1((0, 1)^d)$ , and we look for approximations  $u_{\mathcal{W}}, v_{\mathcal{W}}$  of the exact solutions  $u, v$  in the space  $\mathcal{W} = \text{span}(\varphi_1, \dots, \varphi_N)$  by solving the following discrete problem: find  $u_{\mathcal{W}}, v_{\mathcal{W}} \in \mathcal{W}$  such that, for all  $w \in \mathcal{W}$ ,

$$\begin{cases} \int_{(0,1)^d} (\nabla w)^T A \nabla u_{\mathcal{W}} + \int_{(0,1)^d} (\nabla v_{\mathcal{W}})^T \mathbf{b} w = \int_{(0,1)^d} f w, \\ \int_{(0,1)^d} (\nabla u_{\mathcal{W}})^T \mathbf{c} w + \int_{(0,1)^d} \rho v_{\mathcal{W}} w = \int_{(0,1)^d} g w. \end{cases}$$

Since  $\{\varphi_1, \dots, \varphi_N\}$  is a basis of  $\mathcal{W}$ , we can write  $u_{\mathcal{W}} = \sum_{j=1}^N u_j \varphi_j$  and  $v_{\mathcal{W}} = \sum_{j=1}^N v_j \varphi_j$  for unique vectors  $\mathbf{u} = (u_1, \dots, u_N)^T$  and  $\mathbf{v} = (v_1, \dots, v_N)^T$ . By linearity, the computation of  $u_{\mathcal{W}}, v_{\mathcal{W}}$  (i.e., of  $\mathbf{u}, \mathbf{v}$ ) reduces to solving the linear system

$$(6.69) \quad S^{\text{stiff}} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix},$$

where  $\mathbf{f} = [\int_{(0,1)^d} f \varphi_i]_{i=1}^N$ ,  $\mathbf{g} = [\int_{(0,1)^d} g \varphi_i]_{i=1}^N$ , and  $S^{\text{stiff}}$  is the stiffness matrix, which has the following block structure:

$$(6.70) \quad S^{\text{stiff}} = \begin{bmatrix} A^{\text{stiff}}(A) & H^{\text{stiff}}(\mathbf{b}) \\ H^{\text{stiff}}(\mathbf{c}) & M^{\text{stiff}}(\rho) \end{bmatrix},$$

where, for any  $W = [w_{\ell q}]_{\ell, q=1}^d$ ,  $\mathbf{w} = [w_q]_{q=1}^d$ ,  $w$  with  $w_{\ell q}, w_q, w \in L^1((0, 1)^d)$ ,

$$(6.71) \quad A^{\text{stiff}}(W) = \left[ \int_{(0,1)^d} (\nabla \varphi_i)^T W \nabla \varphi_j \right]_{i,j=1}^N,$$

$$(6.72) \quad H^{\text{stiff}}(\mathbf{w}) = \left[ \int_{(0,1)^d} (\nabla \varphi_j)^T \mathbf{w} \varphi_i \right]_{i,j=1}^N,$$

$$(6.73) \quad M^{\text{stiff}}(w) = \left[ \int_{(0,1)^d} w \varphi_j \varphi_i \right]_{i,j=1}^N.$$

Note that for any  $r, s \neq 0$ , the system (6.69) is equivalent to

$$B^{(r,s)} \begin{bmatrix} \mathbf{u} \\ r^{-1} \mathbf{v} \end{bmatrix} = s \begin{bmatrix} r^{-1} \mathbf{f} \\ \mathbf{g} \end{bmatrix},$$

where

$$\begin{aligned} B^{(r,s)} &= s(r^{-1} I_N \oplus I_N) S^{\text{stiff}} (I_N \oplus r I_N) \\ &= s \begin{bmatrix} r^{-1} I_N & O_N \\ O_N & I_N \end{bmatrix} \begin{bmatrix} A^{\text{stiff}}(A) & H^{\text{stiff}}(\mathbf{b}) \\ H^{\text{stiff}}(\mathbf{c}) & M^{\text{stiff}}(\rho) \end{bmatrix} \begin{bmatrix} I_N & O_N \\ O_N & r I_N \end{bmatrix} \\ (6.74) \quad &= \begin{bmatrix} r^{-1} s A^{\text{stiff}}(A) & s H^{\text{stiff}}(\mathbf{b}) \\ s H^{\text{stiff}}(\mathbf{c}) & r s M^{\text{stiff}}(\rho) \end{bmatrix}. \end{aligned}$$

**Higher-order FE discretization matrices.** Following the higher-order FE approach as in Sections 6.2 and 6.3, the basis functions  $\varphi_1, \dots, \varphi_N$  are chosen as in (6.37). The stiffness matrix  $S^{\text{stiff}}$  resulting from this choice and its normalized version  $B^{(r,s)}$  will be denoted

by  $S_{n,[p,k]}(A, \mathbf{b}, \mathbf{c}, \rho)$  and  $B_{n,[p,k]}^{(r,s)}(A, \mathbf{b}, \mathbf{c}, \rho)$ , respectively. According to (6.70)–(6.73) and (6.74), we have

$$(6.75) \quad \begin{aligned} S_{n,[p,k]}(A, \mathbf{b}, \mathbf{c}, \rho) &= \begin{bmatrix} A_{n,[p,k]}(A) & H_{n,[p,k]}(\mathbf{b}) \\ H_{n,[p,k]}(\mathbf{c}) & M_{n,[p,k]}(\rho) \end{bmatrix}, \\ B_{n,[p,k]}^{(r,s)}(A, \mathbf{b}, \mathbf{c}, \rho) &= \begin{bmatrix} r^{-1} s A_{n,[p,k]}(A) & s H_{n,[p,k]}(\mathbf{b}) \\ s H_{n,[p,k]}(\mathbf{c}) & r s M_{n,[p,k]}(\rho) \end{bmatrix}, \end{aligned}$$

where, for any  $W = [w_{\ell q}]_{\ell,q=1}^d$ ,  $\mathbf{w} = [w_q]_{q=1}^d$ ,  $w$  with  $w_{\ell q}, w_q, w \in L^1((0, 1)^d)$ ,

$$\begin{aligned} A_{n,[p,k]}(W) &= \left[ \int_{(0,1)^d} (\nabla B_{i+1,[p,k]})^T W \nabla B_{j+1,[p,k]} \right]_{i,j=1}^{n(p-k)+k-1}, \\ H_{n,[p,k]}(\mathbf{w}) &= \left[ \int_{(0,1)^d} (\nabla B_{j+1,[p,k]})^T \mathbf{w} B_{i+1,[p,k]} \right]_{i,j=1}^{n(p-k)+k-1}, \\ M_{n,[p,k]}(w) &= \left[ \int_{(0,1)^d} w B_{j+1,[p,k]} B_{i+1,[p,k]} \right]_{i,j=1}^{n(p-k)+k-1}. \end{aligned}$$

Note that  $A_{n,[p,k]}(W)$  is the same as in (6.38) and Theorem 6.8 with the only difference that  $A$  is replaced by  $W$ . For the matrix  $H_{n,[p,k]}(\mathbf{w})$ , we have the following decomposition:

$$H_{n,[p,k]}(\mathbf{w}) = \sum_{q=1}^d H_{n,[p,k],q}(w_i),$$

where  $H_{n,[p,k],q}(w)$  is defined for all functions  $w \in L^1((0, 1)^d)$  as follows:

$$(6.76) \quad H_{n,[p,k],q}(w) = \left[ \int_{(0,1)^d} w \frac{\partial B_{j+1,[p,k]}}{\partial x_q} B_{i+1,[p,k]} \right]_{i,j=1}^{n(p-k)+k-1}.$$

In the next lemmas, we investigate the structure of the matrices  $H_{n,[p,k],q}(w)$  and  $M_{n,[p,k]}(w)$  for  $w = 1$ . This is necessary for the GLT analysis that we tackle below. In what follows, we use Notation 6.4.

LEMMA 6.12. *Let  $p, n \geq 1$  and  $0 \leq k \leq p - 1$ . Then,*

$$(6.77) \quad H_{n,[p,k],q}(1) = \left( \bigotimes_{r=1}^{q-1} M_{n_r,[p_r,k_r]} \right) \otimes H_{n_q,[p_q,k_q]} \otimes \left( \bigotimes_{r=q+1}^d M_{n_r,[p_r,k_r]} \right),$$

for  $q = 1, \dots, d$ , and

$$(6.78) \quad M_{n,[p,k]}(1) = \bigotimes_{r=1}^d M_{n_r,[p_r,k_r]},$$

where the matrices  $H_{n,[p,k]}$  and  $M_{n,[p,k]}$  are defined in (6.50) and (6.51).

*Proof.* We only prove (6.77) because (6.78) is proved in the same way. For convenience, throughout this proof we write  $B_{i+1}$  instead of  $B_{i+1,[p,k]}$  and  $B_{i_r+1}$  instead of  $B_{i_r+1,[p_r,k_r]}$ . For  $q = 1, \dots, d$  and  $i, j = 1, \dots, n(p-k) + k - 1$ ,

$$(H_{n,[p,k],q}(1))_{ij} = \int_{(0,1)^d} \frac{\partial B_{j+1}}{\partial x_q}(\mathbf{x}) B_{i+1}(\mathbf{x}) d\mathbf{x}$$

$$\begin{aligned}
 &= \int_{(0,1)^d} B'_{j_q+1}(x_q) B_{i_q+1}(x_q) \prod_{\substack{r=1 \\ r \neq q}}^d B_{j_r+1}(x_r) B_{i_r+1}(x_r) d\mathbf{x} \\
 &= \int_0^1 B'_{j_q+1}(x_q) B_{i_q+1}(x_q) dx_q \prod_{\substack{r=1 \\ r \neq q}}^d \int_0^1 B_{j_r+1}(x_r) B_{i_r+1}(x_r) dx_r \\
 &= (H_{n_q, [p_q, k_q]})_{i_q j_q} \prod_{\substack{r=1 \\ r \neq q}}^d (M_{n_r, [p_r, k_r]})_{i_r j_r} \\
 &= \left( \left( \bigotimes_{r=1}^{q-1} M_{n_r, [p_r, k_r]} \right) \otimes H_{n_q, [p_q, k_q]} \otimes \left( \bigotimes_{r=q+1}^d M_{n_r, [p_r, k_r]} \right) \right)_{ij},
 \end{aligned}$$

where the last equality follows from **P8**.  $\square$

LEMMA 6.13. Let  $\mathbf{p}, \mathbf{n} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ . In view of Lemma 6.12, we define

$$\begin{aligned}
 \hat{H}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(1) &= \left( \bigotimes_{r=1}^{q-1} \hat{M}_{n_r, [p_r, k_r]} \right) \otimes \hat{H}_{n_q, [p_q, k_q]} \otimes \left( \bigotimes_{r=q+1}^d \hat{M}_{n_r, [p_r, k_r]} \right), \\
 \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(1) &= \bigotimes_{r=1}^d \hat{M}_{n_r, [p_r, k_r]}.
 \end{aligned}$$

Then, for every  $q = 1, \dots, d$ ,

$$\begin{aligned}
 \hat{H}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(1) &= \frac{n_q}{N(\mathbf{n})} \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}} T_{\mathbf{n}+\mathbf{k}}(\xi_{[\mathbf{p}, \mathbf{k}], q}) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T + F_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}, \\
 \text{rank}(F_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}) &\leq N((\mathbf{n} + \mathbf{k})(\mathbf{p} - \mathbf{k})) \sum_{i=1}^d \frac{2(k_i + \nu_i)}{n_i + k_i},
 \end{aligned}$$

and

$$\begin{aligned}
 \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(1) &= \frac{1}{N(\mathbf{n})} \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}} T_{\mathbf{n}+\mathbf{k}}(\mu_{[\mathbf{p}, \mathbf{k}]}) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T + G_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}, \\
 \text{rank}(G_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}) &\leq N((\mathbf{n} + \mathbf{k})(\mathbf{p} - \mathbf{k})) \sum_{i=1}^d \frac{2(k_i + \nu_i)}{n_i + k_i},
 \end{aligned}$$

where  $\boldsymbol{\nu} = (\nu_1, \dots, \nu_d)$ , each  $\nu_i$  is defined as in (6.29) with  $p_i, k_i$  in place of  $p, k$ , the matrix  $\Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}$  is defined as in Remark 2.3 with  $\mathbf{n} + \mathbf{k}, \mathbf{p} - \mathbf{k}$  in place of  $\mathbf{m}, \mathbf{s}$ , and

$$(6.79) \quad \xi_{[\mathbf{p}, \mathbf{k}], q} = \left( \bigotimes_{r=1}^{q-1} \mu_{[p_r, k_r]} \right) \otimes \xi_{[p_q, k_q]} \otimes \left( \bigotimes_{r=q+1}^d \mu_{[p_r, k_r]} \right),$$

$$(6.80) \quad \mu_{[\mathbf{p}, \mathbf{k}]} = \bigotimes_{r=1}^d \mu_{[p_r, k_r]},$$

with  $\xi_{[p, k]}$  and  $\mu_{[p, k]}$  being defined in (6.41) and (6.42).

*Proof.* The result follows from (6.56)–(6.57) and properties **P9** and **T7**.  $\square$

LEMMA 6.14. Let  $\mathbf{p} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ . Suppose that  $\mathbf{n} = \gamma \mathbf{n}$ , where  $\gamma \in \mathbb{Q}^d$  is a fixed vector with positive components and  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} =$

$\gamma\mathbf{n} \in \mathbb{N}^d$ . Then, for any  $W = [w_{\ell q}]_{\ell, q=1}^d$ ,  $\mathbf{w} = [w_q]_{q=1}^d$ ,  $w$  with  $w_{\ell q}, w_q, w \in L^1((0, 1)^d)$ ,

(6.81)

$$\{n^{d-2}\Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(W)\Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{1}{N(\boldsymbol{\gamma})} \sum_{\ell, q=1}^d \gamma_{\ell} \gamma_q w_{\ell q}(\mathbf{x})(H_{[\mathbf{p}, \mathbf{k}]})_{\ell q}(\boldsymbol{\theta}),$$

(6.82)

$$\{n^{d-1}\Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{H}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(\mathbf{w})\Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{1}{N(\boldsymbol{\gamma})} \sum_{q=1}^d \gamma_q w_q(\mathbf{x}) \xi_{[\mathbf{p}, \mathbf{k}], q}(\boldsymbol{\theta}),$$

(6.83)

$$\{n^d \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(w)\Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{1}{N(\boldsymbol{\gamma})} w(\mathbf{x}) \mu_{[\mathbf{p}, \mathbf{k}]}(\boldsymbol{\theta}).$$

*Proof.* Except for the fact that  $W$  is replaced by  $A$ , relation (6.81) is nothing else than (6.59), which has been proved in the proof of Theorem 6.8. The proofs of (6.82) and (6.83) are analogous to the proof of (6.81); they are left to the reader.  $\square$

The last lemma shows that  $H_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(w)$  is “almost” skew-symmetric as long as the function  $w$  is continuous. In this regard, we note that  $H_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(1)$  is skew-symmetric, as it is clear from Lemma 6.12 and **P 3**.

**LEMMA 6.15.** *Let  $\mathbf{p}, \mathbf{n} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ . For  $i = 1, \dots, \mathbf{n}(\mathbf{p} - \mathbf{k}) + \mathbf{k} - \mathbf{1}$ , let  $\mathbf{x}_i$  be any point in the support of the B-spline  $B_{i+1, [\mathbf{p}, \mathbf{k}]}$ . Then, for all functions  $w \in C([0, 1]^d)$  and for  $q = 1, \dots, d$ ,*

$$(6.84) \quad \|H_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(w) - \Delta_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(w)H_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(1)\| \leq C \frac{n_q}{N(\mathbf{n})} \omega_w\left(\frac{1}{\min(\mathbf{n})}\right),$$

$$(6.85) \quad \|H_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(w) - H_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(1)\Delta_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(w)\| \leq C \frac{n_q}{N(\mathbf{n})} \omega_w\left(\frac{1}{\min(\mathbf{n})}\right),$$

where  $C$  is a constant independent of  $\mathbf{n}$  and  $\Delta_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(w) = \text{diag}_{i=1, \dots, \mathbf{n}(\mathbf{p} - \mathbf{k}) + \mathbf{k} - \mathbf{1}} w(\mathbf{x}_i)$ .

*Proof.* Throughout this proof, the letter  $C$  denotes a generic constant independent of  $\mathbf{n}$ . Let  $Z = H_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(w) - \Delta_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(w)H_{\mathbf{n}, [\mathbf{p}, \mathbf{k}], q}(1)$ . By (6.31)–(6.36), for  $i, j = 1, \dots, \mathbf{n}(\mathbf{p} - \mathbf{k}) + \mathbf{k} - \mathbf{1}$ , we have

$$\begin{aligned} |Z_{ij}| &= \left| \int_{\text{supp}(B_{i+1, [\mathbf{p}, \mathbf{k}]})} (w(\mathbf{x}) - w(\mathbf{x}_i)) \frac{\partial B_{j+1, [\mathbf{p}, \mathbf{k}]}(\mathbf{x})}{\partial x_q} B_{i+1, [\mathbf{p}, \mathbf{k}]}(\mathbf{x}) d\mathbf{x} \right| \\ &= \int_{\text{supp}(B_{i+1, [\mathbf{p}, \mathbf{k}]})} |w(\mathbf{x}) - w(\mathbf{x}_i)| \left| \frac{\partial B_{j+1, [\mathbf{p}, \mathbf{k}]}(\mathbf{x})}{\partial x_q} \right| |B_{i+1, [\mathbf{p}, \mathbf{k}]}(\mathbf{x})| d\mathbf{x} \\ &\leq \max_{\mathbf{x} \in \text{supp}(B_{i+1, [\mathbf{p}, \mathbf{k}]})} |w(\mathbf{x}) - w(\mathbf{x}_i)| C_{p_q} n_q \mu_d(\text{supp}(B_{i+1, [\mathbf{p}, \mathbf{k}]})) \\ &\leq C_{p_q} n_q \omega_w\left(\frac{\max(\mathbf{p} + \mathbf{1})}{\min(\mathbf{n})}\right) N\left(\frac{\mathbf{p} + \mathbf{1}}{\mathbf{n}}\right) \leq C \frac{n_q}{N(\mathbf{n})} \omega_w\left(\frac{1}{\min(\mathbf{n})}\right). \end{aligned}$$

The number of nonzero entries in each row and column of  $Z$  is bounded by  $(2|\mathbf{p}|_{\infty} + 1)^d$  because  $Z_{ij} = 0$  whenever  $|\mathbf{i} - \mathbf{j}|_{\infty} > |\mathbf{p}|_{\infty}$ . Indeed, for  $|\mathbf{i} - \mathbf{j}|_{\infty} > |\mathbf{p}|_{\infty}$ , the intersection of the supports of  $B_{i+1, [\mathbf{p}, \mathbf{k}]}$  and  $B_{j+1, [\mathbf{p}, \mathbf{k}]}$  has zero measure by the local support property (6.31). Thus, by **N 1**,

$$\|Z\| \leq C \frac{n_q}{N(\mathbf{n})} \omega_w\left(\frac{1}{\min(\mathbf{n})}\right),$$

which proves (6.84). The proof of (6.85) is completely analogous (simply repeat the above steps using  $w(\mathbf{x}_j)$  in place of  $w(\mathbf{x}_i)$  and  $\text{supp}(B_{j+1,[p,k]})$  in place of  $\text{supp}(B_{i+1,[p,k]})$ ).  $\square$

**GLT analysis of the higher-order FE discretization matrices.** In what follows, we assume that  $\mathbf{n} = \gamma n$ , where  $\gamma \in \mathbb{Q}^d$  is a fixed vector with positive components and  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} = \gamma n \in \mathbb{N}^d$ . This assumption essentially says that each stepsize  $h_i = \frac{1}{n_i}$  tends to 0 with the same asymptotic speed as the others. The main result of this section is Theorem 6.16, which gives the spectral distribution of the normalized sequence  $\{B_{\mathbf{n},[p,k]}(A, \mathbf{b}, \mathbf{c}, \rho)\}_n$ , where

$$B_{\mathbf{n},[p,k]}(A, \mathbf{b}, \mathbf{c}, \rho) = B_{\mathbf{n},[p,k]}^{(n, n^{d-1})}(A, \mathbf{b}, \mathbf{c}, \rho) = \begin{bmatrix} n^{d-2} A_{\mathbf{n},[p,k]}(A) & n^{d-1} H_{\mathbf{n},[p,k]}(\mathbf{b}) \\ n^{d-1} H_{\mathbf{n},[p,k]}(\mathbf{c}) & n^d M_{\mathbf{n},[p,k]}(\rho) \end{bmatrix}$$

is the normalized version of  $S_{\mathbf{n},[p,k]}(A, \mathbf{b}, \mathbf{c}, \rho)$  defined by (6.75) for  $r = n$  and  $s = n^{d-1}$ .

**THEOREM 6.16.** *Suppose that  $a_{\ell q}, b_q, c_q, \rho \in L^1((0, 1)^d)$ , for every  $\ell, q = 1, \dots, d$ , and that the matrix  $A(\mathbf{x}) = [a_{\ell q}(\mathbf{x})]_{\ell, q=1}^d$  is symmetric for every  $\mathbf{x} \in (0, 1)^d$ . Let  $\mathbf{p} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ , let  $\gamma \in \mathbb{Q}^d$  be a vector with positive components, and assume that  $\mathbf{n} = \gamma n$  (it is understood that  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} = \gamma n \in \mathbb{N}^d$ ). Then*

$$\{B_{\mathbf{n},[p,k]}(A, \mathbf{b}, \mathbf{c}, \rho)\}_n \sim_{\sigma} \eta_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) = \begin{bmatrix} \kappa_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) & \zeta_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) \\ \zeta_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) & \omega_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) \end{bmatrix},$$

where

$$\begin{aligned} \kappa_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) &= \frac{1}{N(\gamma)} \sum_{\ell, q=1}^d \gamma_{\ell} \gamma_q a_{\ell q}(\mathbf{x}) (H_{[p,k]})_{\ell q}(\boldsymbol{\theta}), \\ \zeta_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) &= \frac{1}{N(\gamma)} \sum_{q=1}^d \gamma_q b_q(\mathbf{x}) \xi_{[p,k],q}(\boldsymbol{\theta}), \\ \zeta_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) &= \frac{1}{N(\gamma)} \sum_{q=1}^d \gamma_q c_q(\mathbf{x}) \xi_{[p,k],q}(\boldsymbol{\theta}), \\ \omega_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}) &= \frac{1}{N(\gamma)} \rho(\mathbf{x}) \mu_{[p,k]}(\boldsymbol{\theta}), \end{aligned}$$

and  $(H_{[p,k]})_{\ell q}(\boldsymbol{\theta})$ ,  $\xi_{[p,k],q}(\boldsymbol{\theta})$ ,  $\mu_{[p,k]}(\boldsymbol{\theta})$  are defined in (6.58), (6.79), (6.80), respectively. If moreover  $c_q = -b_q \in C([0, 1]^d)$ , for all  $q = 1, \dots, d$ , then we also have

$$(6.86) \quad \{B_{\mathbf{n},[p,k]}(A, \mathbf{b}, \mathbf{c}, \rho)\}_n \sim_{\lambda} \eta_{[p,k]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}).$$

*Proof.* Keeping in mind Notation 6.4, define

$$\hat{B}_{\mathbf{n},[p,k]}^0(A, \mathbf{b}, \mathbf{c}, \rho) = \begin{bmatrix} n^{d-2} \hat{A}_{\mathbf{n},[p,k]}^0(A) & n^{d-1} \hat{H}_{\mathbf{n},[p,k]}^0(\mathbf{b}) \\ n^{d-1} \hat{H}_{\mathbf{n},[p,k]}^0(\mathbf{c}) & n^d \hat{M}_{\mathbf{n},[p,k]}^0(\rho) \end{bmatrix}.$$

Since

$$\begin{aligned} A_{\mathbf{n},[p,k]}(A) &= P_{\mathbf{n},[p,k]}^* \hat{A}_{\mathbf{n},[p,k]}^0(A) P_{\mathbf{n},[p,k]}, \\ H_{\mathbf{n},[p,k]}(\mathbf{b}) &= P_{\mathbf{n},[p,k]}^* \hat{H}_{\mathbf{n},[p,k]}^0(\mathbf{b}) P_{\mathbf{n},[p,k]}, \\ H_{\mathbf{n},[p,k]}(\mathbf{c}) &= P_{\mathbf{n},[p,k]}^* \hat{H}_{\mathbf{n},[p,k]}^0(\mathbf{c}) P_{\mathbf{n},[p,k]}, \\ M_{\mathbf{n},[p,k]}(\rho) &= P_{\mathbf{n},[p,k]}^* \hat{M}_{\mathbf{n},[p,k]}^0(\rho) P_{\mathbf{n},[p,k]}, \end{aligned}$$

we have

(6.87)

$$B_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A, \mathbf{b}, \mathbf{c}, \rho) = \begin{bmatrix} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* & O \\ O & P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^* \end{bmatrix} \hat{B}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(A, \mathbf{b}, \mathbf{c}, \rho) \begin{bmatrix} P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} & O \\ O & P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} \end{bmatrix}.$$

By Lemma 6.14 and GLT 6,

$$(6.88) \quad \begin{aligned} & \{\Pi_n \hat{B}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(A, \mathbf{b}, \mathbf{c}, \rho) \Pi_n^T\}_n \sim_{\text{GLT}} \eta_{[\mathbf{p},\mathbf{k}]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta}), \\ \Pi_n &= \Pi_{\mathbf{n}+\mathbf{k}, 2, N(\mathbf{p}-\mathbf{k})} \begin{bmatrix} \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T & O \\ O & \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \end{bmatrix}. \end{aligned}$$

Hence we obtain  $\{\hat{B}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(A, \mathbf{b}, \mathbf{c}, \rho)\}_n \sim_{\sigma} \eta_{[\mathbf{p},\mathbf{k}]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta})$  by GLT 1, and thus it follows that  $\{B_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A, \mathbf{b}, \mathbf{c}, \rho)\}_n \sim_{\sigma} \eta_{[\mathbf{p},\mathbf{k}]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta})$  by S 5. It remains to prove that  $\{B_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A, \mathbf{b}, \mathbf{c}, \rho)\}_n \sim_{\lambda} \eta_{[\mathbf{p},\mathbf{k}]}^{(\gamma)}(\mathbf{x}, \boldsymbol{\theta})$  under the assumption  $c_q = -b_q \in C([0, 1]^d)$  for all  $q = 1, \dots, d$ . This assumption ensures that  $\Pi_n \hat{B}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(A, \mathbf{b}, \mathbf{c}, \rho) \Pi_n^T$  is “almost” Hermitian. More precisely, if  $\mathbf{x}_i$  is any point in the support of the B-spline  $B_{i+1, [\mathbf{p},\mathbf{k}]}$ , by Lemma 6.15 we have

$$(6.89) \quad \Pi_n \hat{B}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(A, \mathbf{b}, \mathbf{c}, \rho) \Pi_n^T = \Pi_n C_n \Pi_n^T + \Pi_n Z_n \Pi_n^T,$$

where

$$C_n = \begin{bmatrix} n^{d-2} \hat{A}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(A, \mathbf{b}, \mathbf{c}, \rho) & n^{d-1} \sum_{q=1}^d \hat{\Delta}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(b_q) \hat{H}_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}^0(1) \\ n^{d-1} \sum_{q=1}^d \hat{H}_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}^0(1) \hat{\Delta}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(c_q) & n^d \hat{M}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(\rho) \end{bmatrix}$$

is symmetric (thanks to the relation  $\hat{H}_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}^0(1) = P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} H_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}(1) P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^*$ , the skew-symmetry of  $H_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}(1)$ , and the hypothesis  $c_q = -b_q$  for  $q = 1, \dots, d$ ) and  $Z_n$  is defined by

$$Z_n = \begin{bmatrix} O & Y_n \\ W_n & O \end{bmatrix},$$

with

$$\begin{aligned} Y_n &= n^{d-1} \sum_{q=1}^d \left( \hat{H}_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}^0(b_q) - \hat{\Delta}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(b_q) \hat{H}_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}^0(1) \right) \\ &= P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} \left( n^{d-1} \sum_{q=1}^d \left( H_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}(b_q) - \Delta_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(b_q) H_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}(1) \right) \right) P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^*, \\ W_n &= n^{d-1} \sum_{q=1}^d \left( \hat{H}_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}^0(c_q) - \hat{H}_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}^0(1) \hat{\Delta}_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^0(c_q) \right) \\ &= P_{\mathbf{n},[\mathbf{p},\mathbf{k}]} \left( n^{d-1} \sum_{q=1}^d \left( H_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}(c_q) - H_{\mathbf{n},[\mathbf{p},\mathbf{k}],q}(1) \Delta_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(c_q) \right) \right) P_{\mathbf{n},[\mathbf{p},\mathbf{k}]}^*. \end{aligned}$$

By Lemma 6.15, N 5, and the continuity of  $c_q = -b_q$ ,

$$\|Z_n\| = \max(\|Y_n\|, \|W_n\|) \leq C \sum_{q=1}^d \omega_{b_q}(n^{-1}) \rightarrow 0.$$

The thesis (6.86) now follows from GLT 2, taking into account (6.87)–(6.89).  $\square$



**6.5. Higher-order isogeometric Galerkin discretization of eigenvalue problems.** Let  $\mathbb{R}^+$  be the set of positive real numbers. Consider the following eigenvalue problem: find eigenvalues  $\lambda_j \in \mathbb{R}^+$  and eigenfunctions  $u_j$ , for  $j = 1, 2, \dots, \infty$ , such that

$$(6.90) \quad \begin{cases} -\nabla \cdot A \nabla u_j = \lambda_j b u_j, & \text{in } \Omega, \\ u_j = 0, & \text{on } \partial\Omega, \end{cases}$$

where  $A = [a_{\ell q}]_{\ell, q=1}^d$  and  $\Omega$  is a bounded open domain in  $\mathbb{R}^d$  with Lipschitz boundary. We assume that  $a_{\ell q} \in L^1(\Omega)$  for all  $\ell, q = 1, \dots, d$ , that  $A(\mathbf{x}) = [a_{\ell q}(\mathbf{x})]_{\ell, q=1}^d$  is symmetric positive definite (SPD) for almost every  $\mathbf{x} \in \Omega$ , and that  $b \in L^1(\Omega)$  with  $b > 0$  a.e. in  $\Omega$ . It can be shown that the eigenvalues  $\lambda_j$  must necessarily be real and positive. This can be formally seen by multiplying (6.90) by  $u_j$  and integrating over  $\Omega$ :

$$\lambda_j = \frac{-\int_{\Omega} (\nabla \cdot A \nabla u_j) u_j}{\int_{\Omega} b u_j^2} = \frac{\int_{\Omega} (\nabla u_j)^T A \nabla u_j}{\int_{\Omega} b u_j^2} > 0.$$

**Isogeometric Galerkin discretization.** The weak form of (6.90) reads as follows: find eigenvalues  $\lambda_j \in \mathbb{R}^+$  and eigenfunctions  $u_j \in H_0^1(\Omega)$ , for  $j = 1, 2, \dots, \infty$ , such that

$$a(u_j, w) = \lambda_j (b u_j, w), \quad \forall w \in H_0^1(\Omega),$$

where

$$a(u_j, w) = \int_{\Omega} (\nabla w)^T A \nabla u_j, \quad (b u_j, w) = \int_{\Omega} b u_j w.$$

In the standard Galerkin method, we fix a set of basis functions  $\{\varphi_1, \dots, \varphi_N\} \subset H_0^1(\Omega)$ , we define the so-called approximation space  $\mathcal{W} = \text{span}(\varphi_1, \dots, \varphi_N)$ , and we find approximations of the exact eigenpairs  $(\lambda_j, u_j)$ ,  $j = 1, 2, \dots, \infty$ , by solving the following (Galerkin) problem: find  $\lambda_{j, \mathcal{W}} \in \mathbb{R}^+$  and  $u_{j, \mathcal{W}} \in \mathcal{W}$ , for  $j = 1, \dots, N$ , such that

$$(6.91) \quad a(u_{j, \mathcal{W}}, w) = \lambda_{j, \mathcal{W}} (b u_{j, \mathcal{W}}, w), \quad \forall w \in \mathcal{W}.$$

Assuming that the exact and numerical eigenvalues are arranged in non-decreasing order, then the pair  $(\lambda_{j, \mathcal{W}}, u_{j, \mathcal{W}})$  is taken as an approximation of the pair  $(\lambda_j, u_j)$  for all  $j = 1, \dots, N$ , as prescribed in [65, Chapter 6], where one can find an error analysis for such a choice. The numbers  $\lambda_{j, \mathcal{W}} / \lambda_j - 1$ ,  $j = 1, \dots, N$ , are referred to as the (relative) eigenvalue errors. In view of the canonical identification of each function  $w \in \mathcal{W}$  with its coefficient vector with respect to the basis  $\{\varphi_1, \dots, \varphi_N\}$ , solving the Galerkin problem (6.91) is equivalent to solving the generalized eigenvalue problem

$$(6.92) \quad A^{\text{stiff}} \mathbf{u}_{j, \mathcal{W}} = \lambda_{j, \mathcal{W}} M^{\text{mass}} \mathbf{u}_{j, \mathcal{W}},$$

where  $\mathbf{u}_{j, \mathcal{W}}$  is the coefficient vector of  $u_{j, \mathcal{W}}$  with respect to  $\{\varphi_1, \dots, \varphi_N\}$  and

$$(6.93) \quad A^{\text{stiff}} = [a(\varphi_j, \varphi_i)]_{i, j=1}^N = \left[ \int_{\Omega} (\nabla \varphi_i)^T A \nabla \varphi_j \right]_{i, j=1}^N,$$

$$(6.94) \quad M^{\text{mass}} = [(b \varphi_j, \varphi_i)]_{i, j=1}^N = \left[ \int_{\Omega} b \varphi_j \varphi_i \right]_{i, j=1}^N.$$

The matrices  $A^{\text{stiff}}$  and  $M^{\text{mass}}$  are referred to as the stiffness and mass matrices, respectively. Due to our assumption that  $A$  is SPD a.e. and  $b > 0$  a.e. in  $\Omega$ , both  $A^{\text{stiff}}$  and  $M^{\text{mass}}$  are SPD regardless of the chosen basis functions  $\varphi_1, \dots, \varphi_N$ . Moreover, it is clear from (6.92) that the numerical eigenvalues  $\lambda_j, \mathcal{W}, j = 1, \dots, N$ , are just the eigenvalues of the matrix

$$L = (M^{\text{mass}})^{-1} A^{\text{stiff}}.$$

In the isogeometric Galerkin method [23, 51], we assume that the physical domain  $\Omega$  is described by a global geometry function  $\mathbf{G} : [0, 1]^d \rightarrow \bar{\Omega}$ , which is invertible and satisfies  $\mathbf{G}(\partial([0, 1]^d)) = \partial\bar{\Omega}$ . We fix a set of basis functions  $\{\hat{\varphi}_1, \dots, \hat{\varphi}_N\}$  defined on the reference (parametric) domain  $[0, 1]^d$  and vanishing on the boundary  $\partial([0, 1]^d)$ , and we find approximations to the exact eigenpairs  $(\lambda_j, u_j), j = 1, 2, \dots, \infty$ , by using the standard Galerkin method described above, in which the basis functions  $\varphi_1, \dots, \varphi_N$  are chosen as

$$(6.95) \quad \varphi_i(\mathbf{x}) = \hat{\varphi}_i(\mathbf{G}^{-1}(\mathbf{x})) = \hat{\varphi}_i(\hat{\mathbf{x}}), \quad \mathbf{x} = \mathbf{G}(\hat{\mathbf{x}}), \quad i = 1, \dots, N.$$

The resulting stiffness and mass matrices  $A^{\text{stiff}}$  and  $M^{\text{mass}}$  are given by (6.93) and (6.94) with the basis functions  $\varphi_i$  defined as in (6.95). If we assume that  $\mathbf{G}$  and  $\hat{\varphi}_i, i = 1, \dots, N$ , are sufficiently regular, we can apply standard differential calculus to obtain for  $A^{\text{stiff}}$  and  $M^{\text{mass}}$  the following expressions:

$$(6.96) \quad A^{\text{stiff}} = \left[ \int_{(0,1)^d} |\det(J_{\mathbf{G}})| (\nabla \hat{\varphi}_i)^T ((J_{\mathbf{G}})^{-1} A(\mathbf{G}) (J_{\mathbf{G}})^{-T}) \nabla \hat{\varphi}_j \right]_{i,j=1}^N,$$

$$(6.97) \quad M^{\text{mass}} = \left[ \int_{(0,1)^d} b(\mathbf{G}) |\det(J_{\mathbf{G}})| \hat{\varphi}_j \hat{\varphi}_i \right]_{i,j=1}^N,$$

where  $J_{\mathbf{G}}$  is the Jacobian matrix of  $\mathbf{G}$ ,

$$J_{\mathbf{G}} = \left[ \frac{\partial G_i}{\partial \hat{x}_j} \right]_{i,j=1}^d = \left[ \frac{\partial x_i}{\partial \hat{x}_j} \right]_{i,j=1}^d.$$

**GLT analysis of the higher-order isogeometric Galerkin matrices.** In the higher-order isogeometric Galerkin approach considered herein, we choose the basis functions  $\hat{\varphi}_1, \dots, \hat{\varphi}_N$  as the tensor-product  $p$ -degree  $C^k$  B-splines

$$(6.98) \quad B_{2, [p,k]}, \dots, B_{n(p-k)+k, [p,k]}$$

introduced in Section 6.2. If, for any  $W = [w_{\ell q}]_{\ell, q=1}^d$  with  $w_{\ell q} \in L^1((0, 1)^d)$  and any  $w \in L^1((0, 1)^d)$ , we define

$$A_{n, [p,k]}(W) = \left[ \int_{(0,1)^d} (\nabla B_{i+1, [p,k]})^T W \nabla B_{j+1, [p,k]} \right]_{i,j=1}^{n(p-k)+k-1},$$

$$M_{n, [p,k]}(w) = \left[ \int_{(0,1)^d} w B_{j+1, [p,k]} B_{i+1, [p,k]} \right]_{i,j=1}^{n(p-k)+k-1},$$

then the stiffness and mass matrices (6.96) and (6.97) resulting from the choice of the basis functions as in (6.98) are nothing else than  $A_{n, [p,k]}(A_{\mathbf{G}})$  and  $M_{n, [p,k]}(b_{\mathbf{G}})$ , where

$$(6.99) \quad A_{\mathbf{G}} = |\det(J_{\mathbf{G}})| (J_{\mathbf{G}})^{-1} A(\mathbf{G}) (J_{\mathbf{G}})^{-T}, \quad b_{\mathbf{G}} = b(\mathbf{G}) |\det(J_{\mathbf{G}})|.$$

The main result of this section is Theorem 6.19. It gives the spectral distributions of the normalized sequences

$$\{n^{d-2}A_{n,[p,k]}(A_G)\}_n, \quad \{n^d M_{n,[p,k]}(b_G)\}_n, \quad \{n^{-2}L_{n,[p,k]}(A_G, b_G)\}_n,$$

where  $\mathbf{n} = \gamma n$  for some  $\gamma \in \mathbb{Q}^d$  and

$$L_{n,[p,k]}(A_G, b_G) = (M_{n,[p,k]}(b_G))^{-1} A_{n,[p,k]}(A_G)$$

is the matrix whose eigenvalues are just the numerical eigenvalues produced by the considered higher-order isogeometric Galerkin method. To prove Theorem 6.19, some preliminary work is necessary. In what follows, we systematically use Notation 6.4.

LEMMA 6.17. *Let  $\mathbf{p} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ . Suppose that  $\mathbf{n} = \gamma n$ , where  $\gamma \in \mathbb{Q}^d$  is a fixed vector with positive components and  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} = \gamma n \in \mathbb{N}^d$ . Then, for any  $W = [w_{\ell q}]_{\ell, q=1}^d$ ,  $w$  with  $w_{\ell q}, w \in L^1((0, 1)^d)$ ,*

(6.100)

$$\{n^{d-2}\Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(W) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{1}{N(\gamma)} \sum_{\ell, q=1}^d \gamma_\ell \gamma_q w_{\ell q}(\mathbf{x}) (H_{[\mathbf{p}, \mathbf{k}]})_{\ell q}(\boldsymbol{\theta}),$$

(6.101)

$$\{n^d \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(w) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \frac{1}{N(\gamma)} w(\mathbf{x}) \mu_{[\mathbf{p}, \mathbf{k}]}(\boldsymbol{\theta}),$$

where  $H_{[\mathbf{p}, \mathbf{k}]}$  and  $\mu_{[\mathbf{p}, \mathbf{k}]}$  are defined in (6.58) and (6.80), respectively.

*Proof.* By Lemma 6.14, the relations (6.100) and (6.101) hold with  $\hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(W)$  and  $\hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(w)$  in place of  $\hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(W)$  and  $\hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(w)$ . Since, by definition (see Notation 6.4),

$$\begin{aligned} \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(W) &= \hat{A}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(W) + R_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0, \\ \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(w) &= \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0(w) + R_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0, \end{aligned}$$

with  $\text{rank}(R_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^0) = o(n^d)$ , the relations (6.100) and (6.101) follow from **Z1**, **GLT 3**, and **GLT 4**.  $\square$

LEMMA 6.18. *Let  $\mathbf{p} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ . Then,  $\mu_{[\mathbf{p}, \mathbf{k}]}(\boldsymbol{\theta})$  is Hermitian positive definite for all  $\boldsymbol{\theta} \in [-\pi, \pi]^d$ .*

*Proof.* By Lemma 6.17, if we take  $\mathbf{n} = (n, \dots, n)$ , that is,  $\gamma = \mathbf{1}$ , then

$$\{n^d \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}^T \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(1) \Gamma_{\mathbf{n}+\mathbf{k}, \mathbf{p}-\mathbf{k}}\}_n \sim_{\text{GLT}} \mu_{[\mathbf{p}, \mathbf{k}]}(\boldsymbol{\theta}).$$

Since  $n^d \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(1)$  is symmetric, we infer from **GLT 1** that

$$(6.102) \quad \{n^d \hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(1)\}_n \sim_\lambda \mu_{[\mathbf{p}, \mathbf{k}]}(\boldsymbol{\theta}).$$

Now, by the discussion in Notation 6.4, the eigenvalues of  $\hat{M}_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}^1(1)$  coincide with the eigenvalues of  $M_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(1)$  plus further  $N((\mathbf{n}+\mathbf{k})(\mathbf{p}-\mathbf{k})) - N(\mathbf{n}(\mathbf{p}-\mathbf{k})+\mathbf{k}-\mathbf{1})$  eigenvalues that are equal to 1. By (6.78), **P7**, and the positive definiteness of the matrix  $M_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}$  for every  $p, n \geq 1$  and  $0 \leq k \leq p-1$ , we have

$$(6.103) \quad \lambda_{\min}(M_{\mathbf{n}, [\mathbf{p}, \mathbf{k}]}(1)) = \lambda_{\min}\left(\bigotimes_{r=1}^d M_{n, [p_r, k_r]}\right) = \prod_{r=1}^d \lambda_{\min}(M_{n, [p_r, k_r]}).$$

By (6.28), for every  $p, n \geq 1$ , every  $0 \leq k \leq p-1$ , and every  $\mathbf{y} \in \mathbb{R}^{n(p-k)+k-1}$ , we have

$$\begin{aligned} \mathbf{y}^T (nM_{n,[p,k]}) \mathbf{y} &= n \int_0^1 \left( \sum_{i=1}^{n(p-k)+k-1} y_i B_{i+1,[p,k]}(\hat{x}) \right)^2 d\hat{x} \\ &= n \left\| \sum_{i=1}^{n(p-k)+k-1} y_i B_{i+1,[p,k]} \right\|_{L^2([0,1])}^2 \geq c_p \|\mathbf{y}\|^2. \end{aligned}$$

Hence, by the minimax principle for eigenvalues [14, Corollary III.1.2],

$$\lambda_{\min}(nM_{n,[p,k]}) = \min_{\mathbf{y} \neq \mathbf{0}} \frac{\mathbf{y}^T (nM_{n,[p,k]}) \mathbf{y}}{\|\mathbf{y}\|^2} \geq c_p$$

for all  $n$ . In view of (6.103), this implies that  
(6.104)

$$\lambda_{\min}(n^d M_{n,[p,k]}(1)) \geq \prod_{r=1}^d c_{p_r} \implies \lambda_{\min}(n^d \hat{M}_{n,[p,k]}^1(1)) \geq \min\left(1, \prod_{r=1}^d c_{p_r}\right)$$

for all  $n$ . Taking into account that  $\lambda_{\min}(\mu_{[p,k]}(\boldsymbol{\theta}))$  is a continuous function of  $\boldsymbol{\theta}$  just as  $\mu_{[p,k]}(\boldsymbol{\theta})$ , by (6.102), (6.104), and S 3 we have

$$\lambda_{\min}(\mu_{[p,k]}(\boldsymbol{\theta})) \geq \min\left(1, \prod_{r=1}^d c_{p_r}\right)$$

for almost every  $\boldsymbol{\theta} \in [-\pi, \pi]^d$ , that is, for all  $\boldsymbol{\theta} \in [-\pi, \pi]^d$ , thanks to the continuity of  $\lambda_{\min}(\mu_{[p,k]}(\boldsymbol{\theta}))$ . We then conclude that  $\mu_{[p,k]}(\boldsymbol{\theta})$  is Hermitian positive definite for all  $\boldsymbol{\theta} \in [-\pi, \pi]^d$ .  $\square$

**THEOREM 6.19.** *Let  $\Omega$  be a bounded open domain in  $\mathbb{R}^d$  with Lipschitz boundary and suppose that the following conditions on the PDE coefficients and the geometry map are satisfied:*

- $a_{\ell q} \in L^1(\Omega)$  for all  $\ell, q = 1, \dots, d$ ;
- $b \in L^1(\Omega)$  and  $b > 0$  a.e. in  $\Omega$ ;
- $A(\mathbf{x}) = [a_{\ell q}(\mathbf{x})]_{\ell, q=1}^d$  is SPD for a.e.  $\mathbf{x} \in \Omega$ ;
- $|\det(J_{\mathbf{G}})| > 0$  a.e. in  $[0, 1]^d$  and  $(A_{\mathbf{G}})_{\ell q} \in L^1((0, 1)^d)$  for all  $\ell, q = 1, \dots, d$ , where  $A_{\mathbf{G}}$  is defined in (6.99).

Let  $\mathbf{p} \geq \mathbf{1}$  and  $\mathbf{0} \leq \mathbf{k} \leq \mathbf{p} - \mathbf{1}$ , let  $\boldsymbol{\gamma} \in \mathbb{Q}^d$  be a vector with positive components, and assume that  $\mathbf{n} = \boldsymbol{\gamma}n$  (it is understood that  $n$  varies in the infinite subset of  $\mathbb{N}$  such that  $\mathbf{n} = \boldsymbol{\gamma}n \in \mathbb{N}^d$ ). Then,

(6.105)

$$\{n^{d-2} A_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A_{\mathbf{G}})\}_n \sim_{\sigma,\lambda} \kappa_{\mathbf{G},[\mathbf{p},\mathbf{k}]}^{(\boldsymbol{\gamma})}(\hat{\mathbf{x}}, \boldsymbol{\theta}) = \frac{1}{N(\boldsymbol{\gamma})} \sum_{\ell, q=1}^d \gamma_{\ell} \gamma_q (A_{\mathbf{G}})_{\ell q}(\hat{\mathbf{x}}) (H_{[\mathbf{p},\mathbf{k}]} \mu_{[\mathbf{p},\mathbf{k}]}(\boldsymbol{\theta})),$$

(6.106)

$$\{n^d M_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(b_{\mathbf{G}})\}_n \sim_{\sigma,\lambda} \mu_{\mathbf{G},[\mathbf{p},\mathbf{k}]}^{(\boldsymbol{\gamma})}(\hat{\mathbf{x}}, \boldsymbol{\theta}) = \frac{1}{N(\boldsymbol{\gamma})} b_{\mathbf{G}}(\hat{\mathbf{x}}) \mu_{[\mathbf{p},\mathbf{k}]}(\boldsymbol{\theta}),$$

(6.107)

$$\{n^{-2} L_{\mathbf{n},[\mathbf{p},\mathbf{k}]}(A_{\mathbf{G}}, b_{\mathbf{G}})\}_n \sim_{\sigma,\lambda} (\mu_{\mathbf{G},[\mathbf{p},\mathbf{k}]}^{(\boldsymbol{\gamma})}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1} \kappa_{\mathbf{G},[\mathbf{p},\mathbf{k}]}^{(\boldsymbol{\gamma})}(\hat{\mathbf{x}}, \boldsymbol{\theta}),$$

where  $b_{\mathbf{G}}$  is defined in (6.99), while  $H_{[p,k]}$  and  $\mu_{[p,k]}$  are defined in (6.58) and (6.80), respectively.

*Proof.* We first note that it is enough to prove (6.105)–(6.107) with  $A_{n,[p,k]}$ ,  $M_{n,[p,k]}$ ,  $L_{n,[p,k]}$  replaced by, respectively,  $\hat{A}_{n,[p,k]}^1$ ,  $\hat{M}_{n,[p,k]}^1$ ,  $\hat{L}_{n,[p,k]}^1$ , that is,

$$(6.108) \quad \{n^{d-2} \hat{A}_{n,[p,k]}^1(A_{\mathbf{G}})\}_n \sim_{\sigma,\lambda} \kappa_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}),$$

$$(6.109) \quad \{n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}})\}_n \sim_{\sigma,\lambda} \mu_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}),$$

$$(6.110) \quad \{n^{-2} \hat{L}_{n,[p,k]}^1(A_{\mathbf{G}}, b_{\mathbf{G}})\}_n \sim_{\sigma,\lambda} (\mu_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1} \kappa_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}).$$

Moreover, (6.108) and (6.109) follow immediately from Lemma 6.17 taking into account the symmetry of the matrices  $\hat{A}_{n,[p,k]}^1(A_{\mathbf{G}})$  and  $\hat{M}_{n,[p,k]}^1(b_{\mathbf{G}})$ . It only remains to prove (6.110). Keeping in mind the discussion in Notation 6.4, we note that

$$\begin{aligned} & (\hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1} (\hat{A}_{n,[p,k]}^1(A_{\mathbf{G}})) = (\hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1} \hat{A}_{n,[p,k]}^1(A_{\mathbf{G}}) \\ &= (P_{n,[p,k]} M_{n,[p,k]}(b_{\mathbf{G}}) P_{n,[p,k]}^* + R_{n,[p,k]}^0)^{-1} (P_{n,[p,k]} A_{n,[p,k]}(A_{\mathbf{G}}) P_{n,[p,k]}^* + R_{n,[p,k]}^0) \\ &= (P_{n,[p,k]} (M_{n,[p,k]}(b_{\mathbf{G}}))^{-1} P_{n,[p,k]}^* + R_{n,[p,k]}^0) (P_{n,[p,k]} A_{n,[p,k]}(A_{\mathbf{G}}) P_{n,[p,k]}^* + R_{n,[p,k]}^0) \\ &= (P_{n,[p,k]} (M_{n,[p,k]}(b_{\mathbf{G}}))^{-1} A_{n,[p,k]}(A_{\mathbf{G}}) P_{n,[p,k]}^* + R_{n,[p,k]}^0) \\ &= \hat{L}_{n,[p,k]}^1(A_{\mathbf{G}}, b_{\mathbf{G}}). \end{aligned}$$

Thus,

$$(6.111) \quad \begin{aligned} n^{-2} \hat{L}_{n,[p,k]}^1(A_{\mathbf{G}}, b_{\mathbf{G}}) &= (n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1} (n^{d-2} \hat{A}_{n,[p,k]}^1(A_{\mathbf{G}})) \\ &\sim (n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2} (n^{d-2} \hat{A}_{n,[p,k]}^1(A_{\mathbf{G}})) (n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2}, \end{aligned}$$

where  $X \sim Y$  means that the matrix  $X$  is similar to  $Y$ ; note that  $\hat{M}_{n,[p,k]}^1(b_{\mathbf{G}})$  is positive definite because  $b_{\mathbf{G}} > 0$  a.e. in  $[0, 1]^d$  by the assumptions on  $b$  and  $\mathbf{G}$ , hence  $(\hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2}$  is well-defined. By combining the equality in (6.111) with Lemmas 6.17, 6.18, and GLT 4, we immediately obtain

$$(6.112) \quad \{n^{-2} \Gamma_{n+k,p-k}^T \hat{L}_{n,[p,k]}^1(A_{\mathbf{G}}, b_{\mathbf{G}}) \Gamma_{n+k,p-k}\}_n \sim_{\text{GLT}} (\mu_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1} \kappa_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}).$$

The singular value distribution in (6.110) follows from (6.112) and GLT 1. Moreover, by Lemmas 6.17, 6.18, GLT 4, and GLT 5 (applied with  $f(z) = |z|^{1/2}$ ), we have<sup>3</sup>

$$\begin{aligned} & \{\Gamma_{n+k,p-k}^T (n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2} (n^{d-2} \hat{A}_{n,[p,k]}^1(A_{\mathbf{G}})) (n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2} \Gamma_{n+k,p-k}\}_n \\ & \sim_{\text{GLT}} (\mu_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1/2} (\kappa_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta})) (\mu_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1/2}. \end{aligned}$$

Considering that  $(n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2} (n^{d-2} \hat{A}_{n,[p,k]}^1(A_{\mathbf{G}})) (n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2}$  is symmetric, from GLT 1 we get

$$\begin{aligned} & \{(n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2} (n^{d-2} \hat{A}_{n,[p,k]}^1(A_{\mathbf{G}})) (n^d \hat{M}_{n,[p,k]}^1(b_{\mathbf{G}}))^{-1/2}\}_n \\ & \sim_{\lambda} (\mu_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1/2} (\kappa_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta})) (\mu_{\mathbf{G},[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1/2}, \end{aligned}$$

<sup>3</sup>Recall that  $Yg(X)Y^{-1} = g(YXY^{-1})$ ; see [50, Theorem 1.13].

which is equivalent to

$$\begin{aligned} & \{(n^d \hat{M}_{n,[p,k]}^1(b_G))^{-1/2} (n^{d-2} \hat{A}_{n,[p,k]}^1(A_G)) (n^d \hat{M}_{n,[p,k]}^1(b_G))^{-1/2}\}_n \\ & \sim_\lambda (\mu_{G,[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1} (\kappa_{G,[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta})) \end{aligned}$$

by definition of the spectral distribution since

$$\begin{aligned} & (\mu_{G,[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1} (\kappa_{G,[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta})) \\ & \sim (\mu_{G,[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1/2} (\kappa_{G,[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta})) (\mu_{G,[p,k]}^{(\gamma)}(\hat{\mathbf{x}}, \boldsymbol{\theta}))^{-1/2} \end{aligned}$$

for all  $(\hat{\mathbf{x}}, \boldsymbol{\theta}) \in [0, 1]^d \times [-\pi, \pi]^d$ . In view of the similarity in (6.111), we conclude that the eigenvalue distribution in (6.110) is satisfied.  $\square$

#### REFERENCES

- [1] A. S. AL-FHAID, S. SERRA-CAPIZZANO, D. SESANA, AND M. Z. ULLAH, *Singular-value (and eigenvalue) distribution and Krylov preconditioning of sequences of sampling matrices approximating integral operators*, Numer. Linear Algebra Appl., 21 (2014), pp. 722–743.
- [2] F. AVRAM, *On bilinear forms in Gaussian random variables and Toeplitz matrices*, Probab. Theory Related Fields, 79 (1988), pp. 37–45.
- [3] G. BARBARINO, *Equivalence between GLT sequences and measurable functions*, Linear Algebra Appl., 529 (2017), pp. 397–412.
- [4] ———, *Spectral measures*, in Structured Matrices in Numerical Linear Algebra, D. Bini, F. Di Benedetto, E. Tyrtyshnikov, and M. Van Barel, eds., Springer INdAM Series 30, Springer, Cham, 2019, pp. 1–24.
- [5] ———, *Normal form for GLT sequences*, Preprint on arXiv, 2018.  
<https://arxiv.org/abs/1805.08708v2>
- [6] ———, *Diagonal matrix sequences and their spectral symbols*, Preprint on arXiv, 2017.  
<https://arxiv.org/abs/1710.00810>
- [7] G. BARBARINO AND C. GARONI, *From convergence in measure to convergence of matrix-sequences through concave functions and singular values*, Electron. J. Linear Algebra, 32 (2017), pp. 500–513.
- [8] G. BARBARINO, C. GARONI, AND S. SERRA-CAPIZZANO, *Block generalized locally Toeplitz sequences: theory and applications in the unidimensional case.*, Electron. Trans. Numer. Anal., 53 (2020), pp. 28–112.  
<http://etna.ricam.oeaw.ac.at/vol.53.2020/pp28-112.dir/pp28-112.pdf>
- [9] G. BARBARINO AND S. SERRA-CAPIZZANO, *Non-Hermitian perturbations of Hermitian matrix-sequences and applications to the spectral analysis of approximated PDEs*, Numer. Linear Algebra Appl., to appear.
- [10] B. BECKERMANN AND A. B. J. KUIJLAARS, *Superlinear convergence of conjugate gradients*, SIAM J. Numer. Anal., 39 (2001), pp. 300–329.
- [11] B. BECKERMANN AND S. SERRA-CAPIZZANO, *On the asymptotic spectrum of finite element matrix sequences*, SIAM J. Numer. Anal., 45 (2007), pp. 746–769.
- [12] P. BENEDUSI, C. GARONI, R. KRAUSE, X. LI, AND S. SERRA-CAPIZZANO, *Space-time FE-DG discretization of the anisotropic diffusion equation in any dimension: the spectral symbol*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 1383–1420.
- [13] D. BERTACCINI, M. DONATELLI, F. DURASTANTE, AND S. SERRA-CAPIZZANO, *Optimizing a multigrid Runge-Kutta smoother for variable-coefficient convection-diffusion equations*, Linear Algebra Appl., 533 (2017), pp. 507–535.
- [14] R. BHATIA, *Matrix Analysis*, Springer, New York, 1997.
- [15] V. I. BOGACHEV, *Measure Theory. Vol. I.*, Springer, Berlin, 2007.
- [16] A. BÖTTCHER, C. GARONI, AND S. SERRA-CAPIZZANO, *Exploration of Toeplitz-like matrices with unbounded symbols is not a purely academic journey*, Sb. Math., 208 (2017), pp. 1602–1627.
- [17] A. BÖTTCHER AND S. M. GRUDSKY, *Toeplitz Matrices, Asymptotic Linear Algebra, and Functional Analysis*, Birkhäuser, Basel, 2000.
- [18] ———, *Spectral Properties of Banded Toeplitz Matrices*, SIAM, Philadelphia, 2005.
- [19] A. BÖTTCHER AND B. SILBERMANN, *Introduction to Large Truncated Toeplitz Matrices*, Springer, New York, 1999.
- [20] ———, *Analysis of Toeplitz Operators*, 2nd ed., Springer, Berlin, 2006.
- [21] H. BREZIS, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, 2011.
- [22] O. CAMELLO, *Theories, Sites, Toposes*, Oxford University Press, Oxford, 2018.

- [23] J. A. COTTRELL, T. J. R. HUGHES, AND Y. BAZILEVS, *Isogeometric Analysis*, Wiley, Chichester, 2009.
- [24] C. DE BOOR, *A Practical Guide to Splines*, revised ed., Springer, New York, 2001.
- [25] M. DONATELLI, C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, AND H. SPELEERS, *Robust and optimal multi-iterative techniques for IgA Galerkin linear systems*, *Comput. Methods Appl. Mech. Engrg.*, 284 (2015), pp. 230–264.
- [26] ———, *Robust and optimal multi-iterative techniques for IgA collocation linear systems*, *Comput. Methods Appl. Mech. Engrg.*, 284 (2015), pp. 1120–1146.
- [27] ———, *Spectral analysis and spectral symbol of matrices in isogeometric collocation methods*, *Math. Comp.*, 85 (2016), pp. 1639–1680.
- [28] ———, *Symbol-based multigrid methods for Galerkin B-spline isogeometric analysis*, *SIAM J. Numer. Anal.*, 55 (2017), pp. 31–62.
- [29] M. DONATELLI, C. GARONI, M. MAZZA, S. SERRA-CAPIZZANO, AND D. SESANA, *Spectral behavior of preconditioned non-Hermitian multilevel block Toeplitz matrices with matrix-valued symbol*, *Appl. Math. Comput.*, 245 (2014), pp. 158–173.
- [30] ———, *Preconditioned HSS method for large multilevel block Toeplitz linear systems via the notion of matrix-valued symbol*, *Numer. Linear Algebra Appl.*, 23 (2016), pp. 83–119.
- [31] M. DONATELLI, M. MAZZA, AND S. SERRA-CAPIZZANO, *Spectral analysis and structure preserving preconditioners for fractional diffusion equations*, *J. Comput. Phys.*, 307 (2016), pp. 262–279.
- [32] M. DONATELLI, M. NEYTCHIEVA, AND S. SERRA-CAPIZZANO, *Canonical eigenvalue distribution of multilevel block Toeplitz sequences with non-Hermitian symbols*, in *Spectral Theory, Mathematical System Theory, Evolution Equations, Differential and Difference Equations*, W. Arendt, J. A. Ball, J. Behrndt, K.-H. Förster, V. Mehrmann, and C. Trunk, eds., *Oper. Theory Adv. Appl.*, 221, Birkhäuser, Basel, 2012, pp. 269–291.
- [33] A. DOROSTKAR, M. NEYTCHIEVA, AND S. SERRA-CAPIZZANO, *Spectral analysis of coupled PDEs and of their Schur complements via generalized locally Toeplitz sequences in 2D*, *Comput. Methods Appl. Mech. Engrg.*, 309 (2016), pp. 74–105.
- [34] M. DUMBSER, F. FAMBRI, I. FURCI, M. MAZZA, S. SERRA-CAPIZZANO, AND M. TAVELLI, *Staggered discontinuous Galerkin methods for the incompressible Navier-Stokes equations: spectral analysis and computational results*, *Numer. Linear Algebra Appl.*, 25 (2018), Art. e2151, 31 pages.
- [35] C. GARONI, *Spectral distribution of PDE discretization matrices from isogeometric analysis: the case of  $L^1$  coefficients and non-regular geometry*, *J. Spectr. Theory*, 8 (2018), pp. 297–313.
- [36] C. GARONI, C. MANNI, F. PELOSI, S. SERRA-CAPIZZANO, AND H. SPELEERS, *On the spectrum of stiffness matrices arising from isogeometric analysis*, *Numer. Math.*, 127 (2014), pp. 751–799.
- [37] C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, D. SESANA, AND H. SPELEERS, *Lusin theorem, GLT sequences and matrix computations: an application to the spectral analysis of PDE discretization matrices*, *J. Math. Anal. Appl.*, 446 (2017), pp. 365–382.
- [38] ———, *Spectral analysis and spectral symbol of matrices in isogeometric Galerkin methods*, *Math. Comp.*, 86 (2017), pp. 1343–1373.
- [39] C. GARONI, M. MAZZA, AND S. SERRA-CAPIZZANO, *Block generalized locally Toeplitz sequences: from the theory to the applications*, *Axioms*, 7 (2018), Art. 49, 29 pages.
- [40] C. GARONI AND S. SERRA-CAPIZZANO, *Generalized Locally Toeplitz Sequences: Theory and Applications. Vol. I.*, Springer, Cham, 2017.
- [41] ———, *Generalized Locally Toeplitz Sequences: Theory and Applications. Vol. II.*, Springer, Cham, 2018.
- [42] ———, *Block generalized locally Toeplitz sequences: the case of matrix functions and an engineering application*, *Electron. J. Linear Algebra*, 35 (2019), pp. 204–222.
- [43] C. GARONI, S. SERRA-CAPIZZANO, AND D. SESANA, *Tools for determining the asymptotic spectral distribution of non-Hermitian perturbations of Hermitian matrix-sequences and applications*, *Integral Equat. Oper. Theory*, 81 (2015), pp. 213–225.
- [44] ———, *Spectral analysis and spectral symbol of  $d$ -variate  $\mathbb{Q}_p$  Lagrangian FEM stiffness matrices*, *SIAM J. Matrix Anal. Appl.*, 36 (2015), pp. 1100–1128.
- [45] ———, *Block locally Toeplitz sequences: construction and properties*, in *Structured Matrices in Numerical Linear Algebra*, D. Bini, F. Di Benedetto, E. Tyrtyshnikov, and M. Van Barel, eds., Springer INdAM Series 30, Springer, Cham, 2019, pp. 25–58.
- [46] ———, *Block generalized locally Toeplitz sequences: topological construction, spectral distribution results, and star-algebra structure*, in *Structured Matrices in Numerical Linear Algebra*, D. Bini, F. Di Benedetto, E. Tyrtyshnikov, and M. Van Barel, eds., Springer INdAM Series 30, Springer, Cham, 2019, pp. 59–79.
- [47] C. GARONI, H. SPELEERS, S.-E. EKSTRÖM, A. REALI, S. SERRA-CAPIZZANO, AND T. J. R. HUGHES, *Symbol-based analysis of finite element and isogeometric B-spline discretizations of eigenvalue problems: exposition and review*, *Arch. Comput. Methods Engrg.*, 26 (2019), pp. 1639–1690.
- [48] L. GOLINSKII AND S. SERRA-CAPIZZANO, *The asymptotic properties of the spectrum of nonsymmetrically perturbed Jacobi matrix sequences*, *J. Approx. Theory*, 144 (2007), pp. 84–102.
- [49] U. GRENANDER AND G. SZEGŐ, *Toeplitz Forms and Their Applications*, 2nd ed., Chelsea, New York, 1984.

- [50] N. J. HIGHAM, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.
- [51] T. J. R. HUGHES, J. A. COTTRELL, AND Y. BAZILEVS, *Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement*, *Comput. Methods Appl. Mech. Engrg.*, 194 (2005), pp. 4135–4195.
- [52] M. MAZZA, A. RATNANI, AND S. SERRA-CAPIZZANO, *Spectral analysis and spectral symbol for the 2D curl-curl (stabilized) operator with applications to the related iterative solutions*, *Math. Comp.*, 88 (2019), pp. 1155–1188.
- [53] S. V. PARTER, *On the distribution of the singular values of Toeplitz matrices*, *Linear Algebra Appl.*, 80 (1986), pp. 115–130.
- [54] A. QUARTERONI, *Numerical Models for Differential Problems*, 3rd. ed., Springer, Cham, 2017.
- [55] R. RAHLA, S. SERRA-CAPIZZANO, C. TABLINO-POSSIO, *Spectral analysis of  $\mathbb{P}_k$  finite element matrices in the case of Friedrichs-Keller triangulations via GLT technology*, *Numer. Linear Algebra Appl.*, to appear.
- [56] F. ROMAN, C. MANNI, AND H. SPELEERS, *Spectral analysis of matrices in Galerkin methods based on generalized B-splines with high smoothness*, *Numer. Math.*, 135 (2017), pp. 169–216.
- [57] H. L. ROYDEN AND P. M. FITZPATRICK, *Real Analysis*, 4th ed., Pearson, New York, 2010.
- [58] W. RUDIN, *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill, New York, 1976.
- [59] E. SALINELLI, S. SERRA-CAPIZZANO, AND D. SESANA, *Eigenvalue-eigenvector structure of Schoenmakers-Coffey matrices via Toeplitz technology and applications*, *Linear Algebra Appl.*, 491 (2016), pp. 138–160.
- [60] L. L. SCHUMAKER, *Spline Functions: Basic Theory*, 3rd ed., Cambridge University Press, Cambridge, 2007.
- [61] S. SERRA-CAPIZZANO, *More inequalities and asymptotics for matrix valued linear positive operators: the noncommutative case*, in *Toeplitz Matrices and Singular Integral Equations*, A. Böttcher, I. Gohberg, and P. Junghanns, eds., *Oper. Theory Adv. Appl.*, 135, Birkhäuser, Basel, 2002, pp. 293–315.
- [62] ———, *Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations*, *Linear Algebra Appl.*, 366 (2003), pp. 371–402.
- [63] ———, *The GLT class as a generalized Fourier analysis and applications*, *Linear Algebra Appl.*, 419 (2006), pp. 180–233.
- [64] S. SERRA-CAPIZZANO AND P. TILLI, *On unitarily invariant norms of matrix-valued linear positive operators*, *J. Inequal. Appl.*, 7 (2002), pp. 309–330.
- [65] G. STRANG AND G. FIX, *An Analysis of the Finite Element Method*, 2nd. ed., Wellesley-Cambridge, Wellesley, 2008.
- [66] P. TILLI, *Locally Toeplitz sequences: spectral properties and applications*, *Linear Algebra Appl.*, 278 (1998), pp. 91–120.
- [67] ———, *A note on the spectral distribution of Toeplitz matrices*, *Linear and Multilinear Algebra*, 45 (1998), pp. 147–159.
- [68] ———, *Some results on complex Toeplitz eigenvalues*, *J. Math. Anal. Appl.*, 239 (1999), pp. 390–401.
- [69] E. E. TYRTYSHNIKOV, *A unifying approach to some old and new theorems on distribution and clustering*, *Linear Algebra Appl.*, 232 (1996), pp. 1–43.
- [70] E. E. TYRTYSHNIKOV AND N. L. ZAMARASHKIN, *Spectra of multilevel Toeplitz matrices: advanced theory via simple matrix relationships*, *Linear Algebra Appl.*, 270 (1998), pp. 15–27.
- [71] N. L. ZAMARASHKIN AND E. E. TYRTYSHNIKOV, *Distribution of eigenvalues and singular values of Toeplitz matrices under weakened conditions on the generating function*, *Sb. Math.*, 188 (1997), pp. 1191–1201.