# Spectral analysis and spectral symbol of $d$-variate $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM stiffness matrices

Carlo Garoni*, Stefano Serra-Capizzano*,†, and Debora Sesana*

*University of Insubria, Department of Science and High Technology, Via Valleggio 11, 22100 Como, Italy
carlo.garoni@uninsubria.it, stefano.serrac@uninsubria.it, debora.sesana@uninsubria.it
†Department of Information Technology, Uppsala University, Box 337, SE-751 05 Uppsala, Sweden
stefano.serra@it.uu.se

**Abstract**

We study the spectral properties of the stiffness matrices coming from the $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM approximation of $d$-dimensional second order elliptic differential problems; here, $\boldsymbol{p} = (p_1, \ldots, p_d) \in \mathbb{N}^d$ and $p_j$ represents the polynomial approximation degree in the $j$-th direction. After presenting a construction of these matrices, we investigate the conditioning (behavior of the extremal eigenvalues and singular values) and the asymptotic spectral distribution in the Weyl sense, and we find out the so-called (spectral) symbol describing the asymptotic spectrum. We also study the properties of the symbol, which turns out to be a $d$-variate function taking values in the space of $D(\boldsymbol{p}) \times D(\boldsymbol{p})$ Hermitian matrices, where $D(\boldsymbol{p}) = \prod_{j=1}^{d} p_j$. Unlike the stiffness matrices coming from the $\boldsymbol{p}$-degree B-spline IgA approximation of the same differential problems, where a unique $d$-variate real-valued function describes all the spectrum, here the spectrum is described by $D(\boldsymbol{p})$ different functions, that is the $D(\boldsymbol{p})$ eigenvalues of the symbol, which are well-separated, far away, and exponentially diverging with respect to $\boldsymbol{p}$ and $d$. This very involved picture provides a clean explanation of: a) the difficulties encountered in designing robust solvers, with convergence speed independent of the matrix size, of the approximation parameters $\boldsymbol{p}$, and of the dimensionality $d$; b) the possible convergence deterioration of known iterative methods, already for moderate $\boldsymbol{p}$ and $d$.

## 1 Introduction

This paper is devoted to the analysis of the spectrum of the stiffness matrices coming from the approximation of the following second order elliptic differential problem

$$\begin{cases} -\Delta u + \boldsymbol{\beta} \cdot \nabla u + \gamma u = f & \text{in } \Omega := (0,1)^d, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $f \in L^2(\Omega)$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d)$, and $\gamma$, $\beta_j$, $j = 1, \ldots, d$, are functions in $L^\infty(\Omega)$ with $\gamma \geq 0$ over $\Omega$. We restrict our attention to the case of uniform $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian Finite Element Methods (FEM), where $\boldsymbol{p} = (p_1, \ldots, p_d) \in \mathbb{N}^d$ and $p_j$ represents the polynomial approximation degree in the $j$-th direction; see [28, 29, 10, 31, 8, 9] for a wide account on these methods and their evolution. Concerning the stiffness matrices arising from this approximation technique, we are mainly interested in studying the following items, from the perspective of (block) multilevel Toeplitz operators [6] and (block) generalized locally Toeplitz (GLT) sequences [33, 34]:

- extremal eigenvalues, conditioning and spectral localization;
- spectral clustering and spectral distribution in the Weyl sense.

The goal of this spectral analysis is the design of fast and robust iterative solvers for the linear systems coming from the $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM approximation of (1). With this in mind, we recall that the information in the first item is important from the viewpoint of the intrinsic difficulty, in terms of the inherent error, and in evaluating the convergence rate of classical stationary and non-stationary iterative solvers. On the other hand, the information contained in the second item, and especially in the spectral distribution, turned out to be the key ingredient in the design and in the convergence analysis of specialized multigrid methods (see [17, 18, 2, 1, 11, 13], [34, Subsection 3.7]) and of preconditioned Krylov solvers [3, 30] such as preconditioned conjugate gradient (PCG) and preconditioned GMRES. In fact, the knowledge of the spectral distribution is the key for explaining the superlinear convergence history of (P)CG (see [4] and references therein), thus improving the classical bounds.

Quite recently, a spectral analysis very similar to the one contained in this paper has been carried out for the stiffness/collocation matrices coming from the $\boldsymbol{p}$-degree B-spline IgA approximation of second order elliptic problems like (1); see [21, 12]. In the IgA case, the (spectral) symbol $f_{\boldsymbol{p}}$ describing the asymptotic spectral distribution is a scalar-valued $d$-variate function defined over $[-\pi,\pi]^d$, and so the eigenvalues of the IgA discretization matrices are approximated by a uniform sampling of $f_{\boldsymbol{p}}$ over $[-\pi,\pi]^d$. In this context, the surprising behavior is that, for large $\boldsymbol{p}$, $f_{\boldsymbol{p}}(\boldsymbol{\theta})$ collapses exponentially to zero when there exists a component $\theta_j$ of $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_d)$ such that $\theta_j=\pi$. In view of the interpretation based on the theory of Toeplitz matrices and matrix algebras, this phenomenon implies that the IgA matrices are ill-conditioned not only in the low frequencies (as expected) but also in the high frequencies, like in the approximation of integral operators [16]. The explicit use of this spectral information allowed us to design ad hoc iterative solvers [14, 11, 13], combining classical multigrid and preconditioned Krylov techniques, whose convergence speed is independent of the matrix size and substantially independent of $\boldsymbol{p}$ and $d$.

In the present $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM setting, we are still able to identify the spectral distribution, as for the IgA case. The associated symbol $\mathbf{f}_{\boldsymbol{p}}$ is $d$-variate and defined on $[-\pi,\pi]^d$, but the surprise is that $\mathbf{f}_{\boldsymbol{p}}$ is a $D(\boldsymbol{p})\times D(\boldsymbol{p})$ Hermitian matrix-valued function, with $D(\boldsymbol{p})=\prod_{j=1}^d p_j$ (a similar situation indeed was recently encountered in dealing with discontinuous Galerkin methods [15]). No specific pathologies are observed for large $\boldsymbol{p}$ at the points $\boldsymbol{\theta}$ such that $\theta_j=\pi$ for some $j$, implying that the source of ill-conditioning, with respect to the fineness parameters, is only in the low frequencies. However, we observe a serious problem of dimensionality, since, already for moderate $\boldsymbol{p}$ and $d$, the quantity $D(\boldsymbol{p})$ is very large. More specifically, the problem is that the spectrum of the $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM stiffness matrices is split into $D(\boldsymbol{p})$ subsets of the same cardinality and each of them is approximately a uniform sampling of the scalar-valued function $\lambda_i(\mathbf{f}_{\boldsymbol{p}})$, $i=1,\ldots,D(\boldsymbol{p})$. Furthermore, already for $d=1$, it turns out that there exist values that separate the ranges of the $D(\boldsymbol{p})$ functions so that, for large $\boldsymbol{p}$, the spectrum of the approximation matrices is described by a large number of functions whose ranges are separated and far away. For instance, as shown in Figure 2, when $d=1$ and $\boldsymbol{p}=3$, already $\lambda_1(\mathbf{f}_{\boldsymbol{p}})$ and $\lambda_3(\mathbf{f}_{\boldsymbol{p}})$ have very different behaviors, since the maximum of $\lambda_1(\mathbf{f}_{\boldsymbol{p}})$ is bounded by 3.2 and the minimum of $\lambda_3(\mathbf{f}_{\boldsymbol{p}})$ is larger than 18. This scattering of the eigenvalue functions $\lambda_i(\mathbf{f}_{\boldsymbol{p}})$ (which grows exponentially with $d$ and $\boldsymbol{p}$) provides a clean explanation of the intrinsic difficulties encountered by the solvers in the literature, already for moderate $\boldsymbol{p}$ and $d$. Indeed, it is relatively easy to design a mesh-independent solver, but the dependency on $\boldsymbol{p}$ and $d$ is generally bad. In fact, not only we observe an exponential ill-conditioning with $\boldsymbol{p}$ and $d$, as already proved in [27], but we derive another information: the subspace where this exponential ill-conditioning occurs is very large. As discussed in Subsection 6.3, the size of this subspace is approximately $D(\boldsymbol{n})/D(\boldsymbol{p})$, where $D(\boldsymbol{n})=\prod_{j=1}^d n_j$ and $n_j$ represents the fineness parameter in the $j$-th direction.

Our spectral analysis sheds new light on the numerical problem of solving linear systems associated with $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM stiffness matrices. The next step will be to use this spectral information, compactly

contained in the symbol $\mathbf{f}_{\boldsymbol{p}}$, in order to design efficient solvers, as already done in the IgA setting [14, 11, 13]: the ultimate goal is to devise numerical strategies, combining a mesh-independent convergence rate with a robustness regarding the approximation parameters $\boldsymbol{p}$ and the dimensionality $d$.

The spectral analysis contained in this paper can be extended to other settings, such as the case of FEM using either the Lagrangian basis with Gauss-Lobatto-Legendre nodes or the integrated Legendre basis [31, 9], which are widely employed in the FEM community for their better behavior in terms of the spectrum of the underlying matrices [9, 26, 27].

The paper is organized as follows. Section 2 contains preliminary results and definitions, concerning matrix analysis, asymptotic spectral analysis, tensor (Kronecker) products and multilevel block Toeplitz and circulant matrices. Section 3 is devoted to the Galerkin approach in the $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM setting, while in Section 4 we present a construction of the resulting stiffness matrices. Sections 5–6 deal with the main results: the spectral analysis of the considered matrices, the identification of the symbol, and the study of the symbol. Finally, Section 7 is devoted to conclusions and open problems: among them, we mention a conjecture that links the polynomial approximation degrees in the FEM analysis, the global regularity of the chosen approximation space, and the size of the matrix-valued symbol which is encountered in the spectral distribution.

## 2 Preliminary tools

This section is divided into five parts. First we introduce the multi-index notation, which will be systematically used in this paper, and we recall some notations and results concerning matrix analysis. Then we give the definitions of spectral distribution, symbol, clustering, and we report the statement of a useful tool for determining spectral distributions. Finally, we provide some properties of tensor products and direct sums, and we consider multilevel block Toeplitz and circulant matrices with some key properties.

### 2.1 Multi-index notation

A multi-index $\boldsymbol{i}$ is simply a vector in $\mathbb{Z}^d$; its components are denoted by $i_1, \ldots, i_d$.

- $\mathbf{0}$ and $\mathbf{1}$ are the vectors of all zeros and all ones respectively (their size will be clear from the context).
- Given multi-indices $\boldsymbol{i}, \boldsymbol{j} \in \mathbb{Z}^d$, $\boldsymbol{i} \le \boldsymbol{j}$ means that $i_\ell \le j_\ell$ for all $\ell = 1, \ldots, d$.
- Given multi-indices $\boldsymbol{h}, \boldsymbol{k} \in \mathbb{Z}^d$ with $\boldsymbol{h} \le \boldsymbol{k}$, the multi-index range $\boldsymbol{h}, \ldots, \boldsymbol{k}$ is the set $\{\boldsymbol{j} \in \mathbb{Z}^d : \boldsymbol{h} \le \boldsymbol{j} \le \boldsymbol{k}\}$. We assume for the multi-index range $\boldsymbol{h}, \ldots, \boldsymbol{k}$ the standard lexicographic ordering:

$$\left[ \ldots \left[ \left[ (j_1, \ldots, j_d) \right]_{j_d = h_d, \ldots, k_d} \right]_{j_{d-1} = h_{d-1}, \ldots, k_{d-1}} \ldots \right]_{j_1 = h_1, \ldots, k_1}. \tag{2}$$

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

  $(h_1, h_2)$, $(h_1, h_2+1)$, $\ldots$, $(h_1, k_2)$, $(h_1+1, h_2)$, $(h_1+1, h_2+1)$, $\ldots$, $(h_1+1, k_2)$, $\ldots\ldots$, $(k_1, h_2)$, $(k_1, h_2+1)$, $\ldots$, $(k_1, k_2)$.

- When a multi-index $\boldsymbol{j}$ varies over a multi-index range $\boldsymbol{h}, \ldots, \boldsymbol{k}$ (this is sometimes written as $\boldsymbol{j} = \boldsymbol{h}, \ldots, \boldsymbol{k}$ or $(j_1, \ldots, j_d) = (h_1, \ldots, h_d), \ldots, (k_1, \ldots, k_d)$), it is always understood that $\boldsymbol{j}$ varies from $\boldsymbol{h}$ to $\boldsymbol{k}$ according to the specific ordering (2). For instance, if we write $X = [x_{\boldsymbol{ij}}]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{m}}$, this means that $X$ is a matrix in $\mathbb{C}^{(m_1 \cdots m_d) \times (m_1 \cdots m_d)}$ whose components are indexed by two multi-indices $\boldsymbol{i}, \boldsymbol{j}$, both varying over the multi-index range $\mathbf{1}, \ldots, \boldsymbol{m}$ in accordance with the ordering (2). Similarly, if $\mathbf{x} = [x_{\boldsymbol{i}}]_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{m}}$ then $\mathbf{x}$ is a vector in $\mathbb{C}^{m_1 \cdots m_d}$ whose components $x_{\boldsymbol{i}}$, $\boldsymbol{i} = \mathbf{1}, \ldots, \boldsymbol{m}$, are ordered in accordance with (2).

- Given $\boldsymbol{h}, \boldsymbol{k} \in \mathbb{Z}^d$ with $\boldsymbol{h} \le \boldsymbol{k}$, the notation $\sum_{\boldsymbol{j}=\boldsymbol{h}}^{\boldsymbol{k}}$ indicates the summation over all $\boldsymbol{j}$ in the multi-index range $\boldsymbol{h}, \dots, \boldsymbol{k}$.
- For a multi-index $\boldsymbol{m} \in \mathbb{N}^d$, $D(\boldsymbol{m}) := \prod_{j=1}^d m_j$ and $\boldsymbol{m} \to \infty$ means that $\min(m_1, \dots, m_d) \to \infty$.
- Operations involving multi-indices that do not have a meaning when considering multi-indices as normal vectors must be always understood in the componentwise sense. For instance, $\boldsymbol{i}^2 = (i_1^2, \dots, i_d^2)$, $\alpha \boldsymbol{i}/\boldsymbol{j} = (\alpha i_1/j_1, \dots, \alpha i_d/j_d)$ for all $\alpha \in \mathbb{C}$ (of course, the division is defined when $j_1, \dots, j_d \neq 0$), $\boldsymbol{i} \bmod \boldsymbol{m} = (i_1 \bmod m_1, \dots, i_d \bmod m_d)$, and so on.

### 2.2 Preliminaries on matrix analysis

In this subsection we introduce some notations and recall some basic results on matrix analysis. For all $X \in \mathbb{C}^{m\times m}$ the singular values of $X$, denoted by $s_j(X)$, $j = 1, \dots, m$, are always labeled in decreasing order: $s_{\max}(X) = s_1(X) \ge \dots \ge s_m(X) = s_{\min}(X)$. If $X \in \mathbb{C}^{m\times m}$ is Hermitian, the eigenvalues of $X$, denoted by $\lambda_j(X)$, $j = 1, \dots, m$, are always labeled in decreasing order: $\lambda_{\max}(X) = \lambda_1(X) \ge \dots \ge \lambda_m(X) = \lambda_{\min}(X)$; in addition, we set $\lambda_j(X) = +\infty$ if $j < 1$ and $\lambda_j(X) = -\infty$ if $j > m$ (this convention simplify some statements). If $X, Y$ are matrices, the notation $X \ge Y$ (resp. $X > Y$) mean that $X, Y$ are Hermitian of the same dimension and $X - Y$ is non-negative definite (resp. positive definite). The identity matrix of order $m$ is denoted by $I_m$ and the conjugate transpose of a matrix $X$ is denoted by $X^*$. The $\infty$-norm and the 2-norm (spectral norm) of both vectors and matrices are denoted by $\|\cdot\|_\infty$ and $\|\cdot\|$, respectively. We recall that, for all $X \in \mathbb{C}^{m\times m}$,

$$\|X\| \le \sqrt{\|X\|_\infty \|X^T\|_\infty}. \tag{3}$$

For all matrices $X$, $\|X\| = s_{\max}(X)$. If $X$ is normal, i.e. $X^*X = XX^*$, then $\|X\| = \rho(X)$. Note that, if $X$ is Hermitian ($X = X^*$) or skew-Hermitian ($X = -X^*$), then $X$ is normal. For $X \in \mathbb{C}^{m\times m}$, $\|X\|_1$ will denote the trace norm of $X$, i.e. the sum of all the singular values of $X$. The trace norm, also called Schatten 1-norm, as well as the other Schatten $p$-norms, $1 \le p \le \infty$, are studied in [5]. Since $\operatorname{rank}(X)$ is the number of nonzero singular values of $X$ and $\|X\| = s_{\max}(X)$, we have

$$\|X\|_1 \le \operatorname{rank}(X)\|X\| \le m\|X\|, \quad \forall X \in \mathbb{C}^{m\times m}. \tag{4}$$

Both $\|\cdot\|$ and $\|\cdot\|_1$ are unitarily invariant norms, i.e. $\|PXQ\| = \|X\|$ and $\|PXQ\|_1 = \|X\|_1$ for all $P, X, Q \in \mathbb{C}^{m\times m}$ with $P, Q$ unitary.

For any square matrix $X$, denoting by $\Re(X) := \frac{X+X^*}{2}$ and $\Im(X) := \frac{X-X^*}{2\mathrm{i}}$ the real and imaginary parts of $X$ (i is the imaginary unit), and denoting by $\sigma(X)$ the spectrum of $X$, a consequence of the minimax principle [5] is that

$$\sigma(X) \subseteq [\lambda_{\min}(\Re(X)), \lambda_{\max}(\Re(X))] \times [\lambda_{\min}(\Im(X)), \lambda_{\max}(\Im(X))] \subset \mathbb{C}, \quad \forall X \in \mathbb{C}^{m\times m}. \tag{5}$$

Other consequences of the minimax principle are the following:

$$\lambda_{\min}(X + Y) \ge \lambda_{\min}(X) + \lambda_{\min}(Y), \quad \text{for all Hermitian matrices } X, Y \in \mathbb{C}^{m\times m}, \tag{6}$$

$$\lambda_{\max}(X + Y) \le \lambda_{\max}(X) + \lambda_{\max}(Y), \quad \text{for all Hermitian matrices } X, Y \in \mathbb{C}^{m\times m}, \tag{7}$$

$$\lambda_j(X) \ge \lambda_j(Y), \quad \forall j = 1, \dots, m, \quad \text{for all Hermitian matrices } X, Y \in \mathbb{C}^{m\times m} \text{ such that } X \ge Y. \tag{8}$$

An important result providing a relation between the singular values of $X$ and the eigenvalues of $\Re(X)$ is the Fan-Hoffman theorem [5, Proposition III.5.1].

**Theorem 1** (**Fan-Hoffman**). *Let $X \in \mathbb{C}^{m\times m}$, then $s_j(X) \ge \lambda_j(\Re(X))$ for all $j = 1, \dots, m$.*

The Fan-Hoffman theorem proved to be useful for estimating the spectral condition number $\kappa(X) := \|X\|\,\|X^{-1}\| = \frac{s_{\max}(X)}{s_{\min}(X)}$ of a non-singular matrix $X$ coming from the approximation of a differential problem, see [21, Theorem 11]. Even in this paper, the Fan-Hoffman theorem turns out to be effective, see Theorem 13.

We now provide the statement of two interlacing theorems. The first one is a version of the Cauchy interlacing theorem, see [5, Corollary III.1.5], while for the second one we refer to [5, p. 63].

**Theorem 2.** *Let $X \in \mathbb{C}^{m\times m}$ be Hermitian and let $Y$ be a principal submatrix of $X$ of order $\ell$. Then*

$$\lambda_j(X) \geq \lambda_j(Y) \geq \lambda_{j+m-\ell}(X), \qquad \forall j = 1,\dots,\ell.$$

**Theorem 3.** *Let $Y = X + E$, where $X, E \in \mathbb{C}^{m\times m}$ are Hermitian. Let $k^+, k^- \geq 0$ be respectively the number of positive and the number of negative eigenvalues of $E$, i.e.*

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

*Then*

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

### 2.3 Spectral distribution, symbol, clustering

We say that a matrix-valued function $\mathbf{f} : D \to \mathbb{C}^{s\times s}$, defined on a measurable set $D \subseteq \mathbb{R}^d$, is measurable (resp. continuous, in $L^p(D)$) if its components $f_{ij} : D \to \mathbb{C},\ i,j = 1,\dots,s$, are measurable (resp. continuous, in $L^p(D)$). Let $m_d$ be the Lebesgue measure on $\mathbb{R}^d$ and let $C_c(\mathbb{C})$ be the set of continuous functions with bounded support defined over $\mathbb{C}$. For $F \in C_c(\mathbb{C})$ and $X \in \mathbb{C}^{m\times m}$, we set $\Sigma_\lambda(F,X) := \frac{1}{m}\sum_{j=1}^m F(\lambda_j(X))$.

**Definition 1.** Let $\{X_n\}$ be a sequence of matrices, with $X_n$ of size $d_n$ tending to infinity, and let $\mathbf{f} : D \to \mathbb{C}^{s\times s}$ be a measurable matrix-valued function defined on the measurable set $D \subset \mathbb{R}^d$, with $0 < m_d(D) < \infty$. We say that $\{X_n\}$ is distributed like $\mathbf{f}$ in the sense of the eigenvalues, in symbols $\{X_n\} \sim_\lambda \mathbf{f}$, if

$$\lim_{n\to\infty} \Sigma_\lambda(F,X_n) = \frac{1}{m_d(D)}\int_D \frac{\sum_{i=1}^s F\left(\lambda_i(\mathbf{f}(\boldsymbol{\theta}))\right)}{s} d\boldsymbol{\theta}, \qquad \forall F \in C_c(\mathbb{C}),$$

where $\lambda_i(\mathbf{f}(\boldsymbol{\theta}))$, $i = 1,\dots,s$, are the eigenvalues of $\mathbf{f}(\boldsymbol{\theta})$. In this case, $\mathbf{f}$ is referred to as the symbol (or spectral symbol) of the sequence of matrices $\{X_n\}$.

**Remark 1.** The informal meaning behind the above definition is the following: if $\mathbf{f}$ is continuous, then a suitable ordering of the eigenvalues $\{\lambda_j(X_n)\}_{j=1,\dots,d_n}$, in correspondence of an equispaced grid on $D$, reconstructs approximately the $s$ surfaces $\boldsymbol{\theta} \to \lambda_i(\mathbf{f}(\boldsymbol{\theta}))$, $i = 1,\dots,s$. For instance, if $\mathbf{f}$ is continuous, $d = 1$, $d_n = ns$, and $D = [a,b]$, then the eigenvalues of $X_n$ are approximately equal to $\lambda_i(\mathbf{f}(a + j(b-a)/n))$, $j = 1,\dots,n,\ i = 1,\dots,s$. Analogously, if $\mathbf{f}$ is continuous, $d = 2$, $d_n = n^2 s$, and $D = [a_1,b_1]\times[a_2,b_2]$, then the eigenvalues of $X_n$ are approximately equal to $\lambda_i(\mathbf{f}(a_1 + j_1(b_1-a_1)/n, a_2 + j_2(b_2-a_2)/n))$, $j_1,j_2 = 1,\dots,n,\ i = 1,\dots,s$ (and so on in a $d$-dimensional setting).

A useful tool for proving spectral distribution results is the following theorem [22, Theorem 3.3].

**Theorem 4.** *Let $\{X_n\}$, $\{Y_n\}$ be sequences of matrices with $X_n, Y_n \in \mathbb{C}^{d_n\times d_n}$ and $d_n$ tending to infinity, and assume the following.*

1. *$\|X_n\|, \|Y_n\| \leq C$ for all $n$, with $C$ a constant independent of $n$.*
2. *Every $X_n$ is Hermitian and $\{X_n\} \sim_\lambda \mathbf{f}$, where $\mathbf{f} : D \to \mathbb{C}^{s\times s}$ is some measurable function defined on some measurable set $D \subset \mathbb{R}^d$, with $0 < m_d(D) < \infty$.*

*3.* $\|Y_n\|_1 = o(d_n)$ *as* $n \to \infty$.

*Then* $\{Z_n\} \sim_\lambda \mathbf{f}$, *where* $Z_n := X_n + Y_n$.

Now we turn to the definition of clustering. For $z \in \mathbb{C}$ and $\epsilon > 0$, we denote by $B(z, \epsilon)$ the disk with center $z$ and radius $\epsilon$, i.e. $B(z, \epsilon) := \{w \in \mathbb{C} : |w - z| < \epsilon\}$. For $S \subseteq \mathbb{C}$ and $\epsilon > 0$, we denote by $B(S, \epsilon)$ the $\epsilon$-expansion of $S$, defined as $B(S, \epsilon) := \bigcup_{z \in S} B(z, \epsilon)$.

**Definition 2.** Let $\{X_n\}$ be a sequence of matrices, with $X_n$ of size $d_n$ tending to infinity, and let $S \subseteq \mathbb{C}$ be a nonempty closed subset of $\mathbb{C}$. We say that $\{X_n\}$ is *strongly clustered* at $S$ in the sense of the eigenvalues if, for every $\epsilon > 0$, the number of eigenvalues of $X_n$ outside $B(S, \epsilon)$ is bounded by a constant $q_\epsilon$ independent of $n$. In other words

$$q_\epsilon(n, S) := \#\{j \in \{1, \dots, d_n\} : \lambda_j(X_n) \notin B(S, \epsilon)\} = O(1), \quad \text{as } n \to \infty.$$

We say that $\{X_n\}$ is *weakly clustered* at $S$ if, for every $\epsilon > 0$,

$$q_\epsilon(n, S) = o(d_n), \quad \text{as } n \to \infty.$$

If $\{X_n\}$ is strongly or weakly clustered at $S$ and $S$ is not connected, then the connected components of $S$ are called sub-clusters.

Recall that, for a measurable function $g : D \to \mathbb{C}$, defined on a measurable set $D \subseteq \mathbb{R}^d$, the essential range of $g$ is defined as $\mathcal{ER}(g) := \{z \in \mathbb{C} : m_d(\{g \in B(z, \epsilon)\}) > 0 \text{ for all } \epsilon > 0\}$, where $\{g \in B(z, \epsilon)\} := \{x \in D : g(x) \in B(z, \epsilon)\}$. $\mathcal{ER}(g)$ is always closed; moreover, if $g$ is continuous and $D$ is sufficiently regular (say, $D$ is contained in the closure of its interior), then $\mathcal{ER}(g)$ coincides with the closure of the image of $g$.

Now, assume that $\{X_n\} \sim_\lambda \mathbf{f}$, with $\{X_n\}$, $\mathbf{f}$ as in Definition 1. Then $\{X_n\}$ is weakly clustered at the essential range of $\mathbf{f}$, defined as the union of the essential ranges of the eigenvalue functions $\lambda_i(\mathbf{f})$, $i = 1, \dots, s$, i.e. $\mathcal{ER}(\mathbf{f}) := \bigcup_{i=1}^s \mathcal{ER}(\lambda_i(\mathbf{f}))$. This result is proved in [23, Theorem 4.2].

### 2.4 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

- $X \otimes Y$ is the tensor (or Kronecker) product of $X$ and $Y$, that is the $m_1\ell_1 \times m_2\ell_2$ matrix

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

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

$$X \oplus Y := \left[\begin{array}{c|c} X & O \\ \hline O & Y \end{array}\right].$$

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

(i) Associativity: for all matrices $X, Y, Z$, $(X \otimes Y) \otimes Z = X \otimes (Y \otimes Z)$ and $(X \oplus Y) \oplus Z = X \oplus (Y \oplus Z)$. This means that we can omit parentheses in expressions like $X_1 \otimes X_2 \otimes \cdots \otimes X_d$ or $X_1 \oplus X_2 \oplus \cdots \oplus X_d$. We recall that, if $X_k \in \mathbb{C}^{m_k \times m_k}$, $k = 1, \dots, d$, then in multi-index notation we have

$$(X_1 \otimes \cdots \otimes X_d)_{\boldsymbol{ij}} = (X_1)_{i_1 j_1} \cdots (X_d)_{i_d j_d}, \qquad \forall \boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{m}. \tag{9}$$

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

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

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

From these basic properties, a lot of other interesting results follow. We recall some of them that will be used in this paper. 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 the singular values of $X \otimes Y$ (resp. $X \oplus Y$) are $\lambda_i(X)\lambda_j(Y),\ i = 1, \dots, m,\ j = 1, \dots, \ell$ and $s_i(X)s_j(Y),\ i = 1, \dots, m,\ j = 1, \dots, \ell$ (resp. $\lambda_i(X),\ \lambda_j(Y),\ i = 1, \dots, m,\ j = 1, \dots, \ell$ and $s_i(X),\ s_j(Y),\ i = 1, \dots, m,\ j = 1, \dots, \ell$). In particular, for all $X \in \mathbb{C}^{m\times m}$ and for all $Y \in \mathbb{C}^{\ell\times\ell}$,

$$\|X \otimes Y\| = \|X\|\,\|Y\|, \qquad \|X \oplus Y\| = \max(\|X\|, \|Y\|), \tag{10}$$

$$\operatorname{rank}(X \otimes Y) = \operatorname{rank}(X)\operatorname{rank}(Y),$$

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

$$\lambda_{\min}(X \otimes Y) = \lambda_{\min}(X)\lambda_{\min}(Y), \qquad \lambda_{\max}(X \otimes Y) = \lambda_{\max}(X)\lambda_{\max}(Y). \tag{11}$$

We also note that

$$X \otimes Y \geq X' \otimes Y', \quad \text{for all HPD matrices } X, Y, X', Y' \text{ such that } X \geq X' \text{ and } Y \geq Y', \tag{12}$$

because $X \otimes Y - X' \otimes Y' = (X - X') \otimes Y + X' \otimes (Y - Y')$ is a sum of two HPD matrices. We also highlight the following property: suppose we are given $2d$ 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

$$\operatorname{rank}(X_1 \otimes \cdots \otimes X_d - Y_1 \otimes \cdots \otimes Y_d) \leq \sum_{i=1}^{d} \operatorname{rank}(X_i - Y_i) m_1 \cdots m_{i-1} m_{i+1} \cdots m_d. \tag{13}$$

This is true because

$$\begin{aligned}
&\operatorname{rank}(X_1 \otimes \cdots \otimes X_d - Y_1 \otimes \cdots \otimes Y_d) = \operatorname{rank}\left(\sum_{i=1}^{d} Y_1 \otimes \cdots \otimes Y_{i-1} \otimes (X_i - Y_i) \otimes X_{i+1} \otimes \cdots \otimes X_d\right) \\
&\leq \sum_{i=1}^{d} \operatorname{rank}\left(Y_1 \otimes \cdots \otimes Y_{i-1} \otimes (X_i - Y_i) \otimes X_{i+1} \otimes \cdots \otimes X_d\right) \\
&= \sum_{i=1}^{d} \operatorname{rank}(Y_1 \otimes \cdots \otimes Y_{i-1})\operatorname{rank}(X_i - Y_i)\operatorname{rank}(X_{i+1} \otimes \cdots \otimes X_d) \leq \sum_{i=1}^{d} m_1 \cdots m_{i-1}\operatorname{rank}(X_i - Y_i) m_{i+1} \cdots m_d.
\end{aligned}$$

A property of tensor products, which can be deduced from the definition of tensor products but is not as popular as the previous ones, is given in Lemma 1; see also [24].

**Lemma 1.** *For all $\boldsymbol{m} \in \mathbb{N}^2$ there exists a permutation matrix $\Pi_{\boldsymbol{m}} \in \mathbb{C}^{(m_1m_2)\times(m_1m_2)}$ such that*

$$X_2 \otimes X_1 = \Pi_{\boldsymbol{m}}(X_1 \otimes X_2)\Pi_{\boldsymbol{m}}^T, \quad \forall X_1 \in \mathbb{C}^{m_1\times m_1},\ \forall X_2 \in \mathbb{C}^{m_2\times m_2}. \tag{14}$$

*Proof.* Let $\Pi_{\boldsymbol{m}}$ be the permutation matrix associated with the permutation $\sigma$ of $\{1,\dots,m_1m_2\}$ given by

$$\sigma := [1, m_2+1, 2m_2+1, \dots, (m_1-1)m_2+1, 2, m_2+2, 2m_2+2, \dots, (m_1-1)m_2+2, \dots\dots, m_2, 2m_2, 3m_2 \dots, m_1m_2],$$

or, equivalently, by

$$\sigma(i) := ((i-1) \bmod m_1)m_2 + \left\lfloor \frac{i-1}{m_1} \right\rfloor + 1, \qquad i = 1, \dots, m_1m_2.$$

In other words, $\Pi_{\boldsymbol{m}}$ is the matrix whose rows are, in the order, $\mathbf{e}_{\sigma(i)},\ i = 1,\dots,m_1m_2$, where $\mathbf{e}_i,\ i = 1,\dots,m_1m_2$, are the vectors of the canonical basis of $\mathbb{C}^{m_1m_2}$. It can be verified that $\Pi_{\boldsymbol{m}}$ defined in this way satisfies (14) for all $X_1 \in \mathbb{C}^{m_1\times m_1}$ and $X_2 \in \mathbb{C}^{m_2\times m_2}$. We omit the details of this verification (which is quite involved). □

Lemma 1 says that the tensor product of two matrices is 'almost' commutative. It is important to notice that the permutation matrix $\Pi_{\boldsymbol{m}}$ depend only on $\boldsymbol{m}$ and not on the specific matrices $X_1, X_2$. By induction, we now extend the result of Lemma 1 to the case of tensor products with more than two factors.

**Lemma 2.** *For all $\boldsymbol{m} \in \mathbb{N}^d$ and all permutations $\sigma$ of the set $\{1,\dots,d\}$, there exists a permutation matrix $\Pi_{\boldsymbol{m};\sigma} \in \mathbb{C}^{(m_1\cdots m_d)\times(m_1\cdots m_d)}$ such that*

$$X_{\sigma(1)} \otimes \cdots \otimes X_{\sigma(d)} = \Pi_{\boldsymbol{m};\sigma}(X_1 \otimes \cdots \otimes X_d)\Pi_{\boldsymbol{m};\sigma}^T, \quad \forall X_1 \in \mathbb{C}^{m_1\times m_1}, \dots, \forall X_d \in \mathbb{C}^{m_d\times m_d}.$$

*Proof.* The case $d = 1$ is trivial. For $d = 2$, the result is clear when $\sigma$ is the identity, and it has been proved in Lemma 1 when $\sigma = [2, 1]$. Now we fix $d \geq 3$, we assume the result is true for $d-1$, and we prove that it is true also for $d$. Let $\boldsymbol{m} \in \mathbb{N}^d$ and let $\sigma$ be a permutation of $\{1,\dots,d\}$. Denote by $i$ the index such that $\sigma(i) = d$, and let $\tau$ be the permutation of $\{1,\dots,d-1\}$ defined as $\tau(j) := \sigma(j)$ for $j = 1,\dots,i-1$ and $\tau(j) := \sigma(j+1)$ for $j = i,\dots,d-1$. Then, keeping in mind the properties of tensor products, for all $X_1,\dots,X_d$ with $X_j \in \mathbb{C}^{m_j\times m_j},\ j = 1,\dots,d$, we have

$$\begin{aligned}
&X_{\sigma(1)} \otimes \cdots \otimes X_{\sigma(d)} = X_{\sigma(1)} \otimes \cdots \otimes X_{\sigma(i-1)} \otimes X_d \otimes X_{\sigma(i+1)} \otimes \cdots \otimes X_{\sigma(d)} \\
&= X_{\sigma(1)} \otimes \cdots \otimes X_{\sigma(i-1)} \otimes \left[\Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}(X_{\sigma(i+1)} \otimes \cdots \otimes X_{\sigma(d)} \otimes X_d)\Pi^T_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right] \quad \text{(Lemma 1)} \\
&= \left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right)(X_{\sigma(1)} \otimes \cdots \otimes X_{\sigma(i-1)} \otimes X_{\sigma(i+1)} \otimes \cdots \otimes X_{\sigma(d)} \otimes X_d)\cdot \\
&\quad \cdot \left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi^T_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right) \\
&= \left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right)(X_{\tau(1)} \otimes \cdots \otimes X_{\tau(d-1)} \otimes X_d)\left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right)^T \\
&= \left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right)\left\{\left[\Pi_{(m_1,\dots,m_{d-1});\tau}(X_1 \otimes \cdots \otimes X_{d-1})\Pi^T_{(m_1,\dots,m_{d-1});\tau}\right] \otimes X_d\right\}\cdot \\
&\quad \cdot \left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right)^T \quad \text{(induction hypothesis)} \\
&= \left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right)\left(\Pi_{(m_1,\dots,m_{d-1});\tau} \otimes I_{m_d}\right)(X_1 \otimes \cdots \otimes X_{d-1} \otimes X_d)\cdot \\
&\quad \cdot \left(\Pi_{(m_1,\dots,m_{d-1});\tau} \otimes I_{m_d}\right)^T \left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right)^T = \Pi_{\boldsymbol{m};\sigma}(X_1 \otimes \cdots \otimes X_d)\Pi^T_{\boldsymbol{m};\sigma},
\end{aligned}$$

where $\Pi_{\boldsymbol{m};\sigma} := \left(I_{m_{\sigma(1)}\cdots m_{\sigma(i-1)}} \otimes \Pi_{(m_{\sigma(i+1)}\cdots m_{\sigma(d)},m_d)}\right)\left(\Pi_{(m_1,\dots,m_{d-1});\tau} \otimes I_{m_d}\right)$ is a permutation matrix, being a product of two permutation matrices. □

Now we turn to the 'distributive properties' of tensor products with respect to direct sums. Again, it turns out that these properties hold modulo permutation transformations which depend only on the dimensions of the involved matrices.

**Remark 2.** From the definition of tensor products and direct sums, for all matrices $X_1, \dots, X_d, Y$ we have

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

**Lemma 3.** *For all $\ell \in \mathbb{N},\ \boldsymbol{m} \in \mathbb{N}^2$ there exists a permutation matrix $Q_{\ell,\boldsymbol{m}} \in \mathbb{C}^{\ell(m_1+m_2)\times\ell(m_1+m_2)}$ such that*

$$X \otimes (Y_1 \oplus Y_2) = Q_{\ell,\boldsymbol{m}} \left[(X \otimes Y_1) \oplus (X \otimes Y_2)\right] Q_{\ell,\boldsymbol{m}}^T, \quad \forall X \in \mathbb{C}^{\ell\times\ell},\ \forall Y_1 \in \mathbb{C}^{m_1\times m_1},\ \forall Y_2 \in \mathbb{C}^{m_2\times m_2}.$$

*Proof.* Let $X \in \mathbb{C}^{\ell\times\ell},\ Y_1 \in \mathbb{C}^{m_1\times m_1},\ Y_2 \in \mathbb{C}^{m_2\times m_2}$. Then, keeping in mind the properties of tensor products and direct sums,

$$\begin{aligned}
X \otimes (Y_1 \oplus Y_2) &= \Pi_{(m_1+m_2,\ell)} \left[(Y_1 \oplus Y_2) \otimes X\right] \Pi_{(m_1+m_2,\ell)}^T \quad \text{(Lemma 1)} \\
&= \Pi_{(m_1+m_2,\ell)} \left[(Y_1 \otimes X) \oplus (Y_2 \otimes X)\right] \Pi_{(m_1+m_2,\ell)}^T \quad \text{(Remark 2)} \\
&= \Pi_{(m_1+m_2,\ell)} \left\{ \left[\Pi_{(\ell,m_1)}(X \otimes Y_1)\Pi_{(\ell,m_1)}^T\right] \oplus \left[\Pi_{(\ell,m_2)}(X \otimes Y_2)\Pi_{(\ell,m_2)}^T\right] \right\} \Pi_{(m_1+m_2,\ell)}^T \quad \text{(Lemma 1)} \\
&= \Pi_{(m_1+m_2,\ell)} \left\{ \left(\Pi_{(\ell,m_1)} \oplus \Pi_{(\ell,m_2)}\right) \left[(X \otimes Y_1) \oplus (X \otimes Y_2)\right] \left(\Pi_{(\ell,m_1)} \oplus \Pi_{(\ell,m_2)}\right)^T \right\} \Pi_{(m_1+m_2,\ell)}^T \\
&= Q_{\ell,\boldsymbol{m}} \left[(X \otimes Y_1) \oplus (X \otimes Y_2)\right] Q_{\ell,\boldsymbol{m}}^T,
\end{aligned}$$

where $Q_{\ell,\boldsymbol{m}} := \Pi_{(m_1+m_2,\ell)}(\Pi_{(\ell,m_1)} \oplus \Pi_{(\ell,m_2)})$ is a permutation matrix, being a product of two permutation matrices. □

**Lemma 4.** *For all $\ell \in \mathbb{N},\ \boldsymbol{m} \in \mathbb{N}^d$ there exists a permutation matrix $Q_{\ell,\boldsymbol{m}} \in \mathbb{C}^{\ell(m_1+\ldots+m_d)\times\ell(m_1+\ldots+m_d)}$ such that*

$$X \otimes (Y_1 \oplus \cdots \oplus Y_d) = Q_{\ell,\boldsymbol{m}}[(X \otimes Y_1) \oplus \cdots \oplus (X \otimes Y_d)]Q_{\ell,\boldsymbol{m}}^T, \quad \forall X \in \mathbb{C}^{\ell\times\ell},\ \forall Y_1 \in \mathbb{C}^{m_1\times m_1}, \dots,\ \forall Y_d \in \mathbb{C}^{m_d\times m_d}.$$

*Proof.* The case $d = 1$ is trivial. For $d = 2$ the result has been proved in Lemma 3. Now we fix $d \geq 3$, we assume the result is true for $d-1$, and we prove that it is true also for $d$. Let $\ell \in \mathbb{N},\ \boldsymbol{m} \in \mathbb{N}^d$. Then, for all $X \in \mathbb{C}^{\ell\times\ell}$ and all $Y_1, \dots, Y_d$ with $Y_j \in \mathbb{C}^{m_j\times m_j},\ j = 1, \dots, d$, we have

$$\begin{aligned}
&X \otimes (Y_1 \oplus \cdots \oplus Y_d) \\
&= Q_{\ell,(m_1,m_2+\ldots+m_d)} \left\{(X \otimes Y_1) \oplus [X \otimes (Y_2 \oplus \cdots \oplus Y_d)]\right\} Q_{\ell,(m_1,m_2+\ldots+m_d)}^T \quad \text{(Lemma 3)} \\
&= Q_{\ell,(m_1,m_2+\ldots+m_d)} \left\{(X \otimes Y_1) \oplus [Q_{\ell,(m_2,\ldots,m_d)}((X \otimes Y_2) \oplus \cdots \oplus (X \otimes Y_d))Q_{\ell,(m_2,\ldots,m_d)}^T]\right\} \cdot \\
&\quad \cdot Q_{\ell,(m_1,m_2+\ldots+m_d)}^T \quad \text{(induction hypothesis)} \\
&= Q_{\ell,(m_1,m_2+\ldots+m_d)} \left\{(I_{\ell m_1} \oplus Q_{\ell,(m_2,\ldots,m_d)})[(X \otimes Y_1) \oplus (X \otimes Y_2) \oplus \cdots \oplus (X \otimes Y_d)](I_{\ell m_1} \oplus Q_{\ell,(m_2,\ldots,m_d)})^T\right\} \cdot \\
&\quad \cdot Q_{\ell,(m_1,m_2+\ldots+m_d)}^T \\
&= Q_{\ell,\boldsymbol{m}}[(X \otimes Y_1) \oplus \cdots \oplus (X \otimes Y_d)]Q_{\ell,\boldsymbol{m}}^T,
\end{aligned}$$

where $Q_{\ell,\boldsymbol{m}} := Q_{\ell,(m_1,m_2+\ldots+m_d)}(I_{\ell m_1} \oplus Q_{\ell,(m_2,\ldots,m_d)})$. □

**Lemma 5.** *For all $n_1^{(k)}, n_2^{(k)} \in \mathbb{N},\ k = 1, \dots, d$, there exists a permutation matrix $P_{n_1^{(1)},n_2^{(1)},n_1^{(2)},n_2^{(2)},\ldots,n_1^{(d)},n_2^{(d)}}$ of dimension $\prod_{k=1}^d (n_1^{(k)} + n_2^{(k)})$ such that*

$$\bigotimes_{k=1}^d (X_1^{(k)} \oplus X_2^{(k)}) = P_{n_1^{(1)},n_2^{(1)},n_1^{(2)},n_2^{(2)},\ldots,n_1^{(d)},n_2^{(d)}} \left[\bigoplus_{i_1=1}^2 \cdots \bigoplus_{i_d=1}^2 (X_{i_1}^{(1)} \otimes \cdots \otimes X_{i_d}^{(d)})\right] P_{n_1^{(1)},n_2^{(1)},n_1^{(2)},n_2^{(2)},\ldots,n_1^{(d)},n_2^{(d)}}^T,$$

*for all matrices $X_1^{(k)}, X_2^{(k)},\ k = 1, \dots, d$, with $X_1^{(k)} \in \mathbb{C}^{n_1^{(k)}\times n_1^{(k)}}$ and $X_2^{(k)} \in \mathbb{C}^{n_2^{(k)}\times n_2^{(k)}}$.*

*Proof.* For $d = 1$ the result is clear. Fix $d \geq 2$, assume the result holds for $d-1$, and let us prove it for $d$. We have

$$\bigotimes_{k=1}^{d}(X_1^{(k)} \oplus X_2^{(k)}) = (X_1^{(1)} \oplus X_2^{(1)}) \otimes \left[\bigotimes_{k=2}^{d}(X_1^{(k)} \oplus X_2^{(k)})\right]$$

$$= (X_1^{(1)} \oplus X_2^{(1)}) \otimes \left\{ P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}} \left[\bigoplus_{i_2=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_{i_2}^{(2)} \otimes \cdots \otimes X_{i_d}^{(d)})\right] P^T_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}} \right\} \quad \text{(induct. hypoth.)}$$

$$= \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right) \left\{ (X_1^{(1)} \oplus X_2^{(1)}) \otimes \left[\bigoplus_{i_2=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_{i_2}^{(2)} \otimes \cdots \otimes X_{i_d}^{(d)})\right] \right\} \cdot$$

$$\cdot \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right)^T$$

$$= \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right) \left\{ \left( X_1^{(1)} \otimes \left[\bigoplus_{i_2=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_{i_2}^{(2)} \otimes \cdots \otimes X_{i_d}^{(d)})\right]\right) \right.$$

$$\left. \oplus \left( X_2^{(1)} \otimes \left[\bigoplus_{i_2=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_{i_2}^{(2)} \otimes \cdots \otimes X_{i_d}^{(d)})\right]\right) \right\} \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right)^T \quad \text{(Remark 2)}$$

$$= \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right) \left\{ \left( Q_{n_1^{(1)},\boldsymbol{\eta}} \left[\bigoplus_{i_2=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_1^{(1)} \otimes X_{i_2}^{(2)} \otimes \cdots \otimes X_{i_d}^{(d)})\right] Q^T_{n_1^{(1)},\boldsymbol{\eta}}\right) \right.$$

$$\left. \oplus \left( Q_{n_2^{(1)},\boldsymbol{\eta}} \left[\bigoplus_{i_2=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_2^{(1)} \otimes X_{i_2}^{(2)} \otimes \cdots \otimes X_{i_d}^{(d)})\right] Q^T_{n_1^{(1)},\boldsymbol{\eta}}\right) \right\} \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right)^T$$

(we used Lemma 4; $\boldsymbol{\eta} := (n_{i_2}^{(2)} \cdots n_{i_d}^{(d)})_{(i_2,\dots,i_d)=(1,\dots,1),\dots,(2,\dots,2)}$ is a multi-index, recall the multi-index notation)

$$= \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right) \left(Q_{n_1^{(1)},\boldsymbol{\eta}} \oplus Q_{n_2^{(1)},\boldsymbol{\eta}}\right) \left\{ \left[\bigoplus_{i_2=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_1^{(1)} \otimes X_{i_2}^{(2)} \otimes \cdots \otimes X_{i_d}^{(d)})\right] \right.$$

$$\left. \oplus \left[\bigoplus_{i_2=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_2^{(1)} \otimes X_{i_2}^{(2)} \otimes \cdots \otimes X_{i_d}^{(d)})\right] \right\} \left(Q_{n_1^{(1)},\boldsymbol{\eta}} \oplus Q_{n_2^{(1)},\boldsymbol{\eta}}\right)^T \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right)^T$$

$$= P_{n_1^{(1)},n_2^{(1)},n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}} \left[\bigoplus_{i_1=1}^{2}\cdots\bigoplus_{i_d=1}^{2}(X_{i_1}^{(1)} \otimes \cdots \otimes X_{i_d}^{(d)})\right] P^T_{n_1^{(1)},n_2^{(1)},n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}},$$

where $P_{n_1^{(1)},n_2^{(1)},n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}} := \left(I_{n_1^{(1)}+n_2^{(1)}} \otimes P_{n_1^{(2)},n_2^{(2)},\dots,n_1^{(d)},n_2^{(d)}}\right) \left(Q_{n_1^{(1)},\boldsymbol{\eta}} \oplus Q_{n_2^{(1)},\boldsymbol{\eta}}\right)$. □

Before concluding this subsection, we stress that a lot of other properties involving tensor products and direct sums can be proved by using techniques similar to those illustrated above. Here we have supplied only the results needed in this paper.

### 2.5 Multilevel block Toeplitz and circulant matrices

In this subsection, we provide the definition and some properties of multilevel block Toeplitz and circulant matrices. In particular, in Lemma 6 we show that a tensor product of unilevel block Toeplitz matrices associated with (matrix-valued) trigonometric polynomials coincides (modulo permutation transformations) with the multilevel block Toeplitz matrix associated with the tensor product of the trigonometric polynomials.

**Definition 3.** A matrix of the form

$$[A_{\boldsymbol{i}-\boldsymbol{j}}]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}} \in \mathbb{C}^{(m_1\cdots m_d s)\times(m_1\cdots m_d s)},$$

with blocks $A_{\boldsymbol{k}} \in \mathbb{C}^{s\times s}$, $\boldsymbol{k} = -(\boldsymbol{m}-\boldsymbol{1}), \ldots, \boldsymbol{m}-\boldsymbol{1}$, is called a multilevel block Toeplitz matrix. A matrix of the form

$$\big[A_{(\boldsymbol{i}-\boldsymbol{j}) \bmod \boldsymbol{m}}\big]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}} \in \mathbb{C}^{(m_1\cdots m_d s)\times(m_1\cdots m_d s)},$$

with blocks $A_{\boldsymbol{k}} \in \mathbb{C}^{s\times s}$, $\boldsymbol{k} = \boldsymbol{0}, \ldots, \boldsymbol{m}-\boldsymbol{1}$, is called a multilevel block circulant matrix.

Given a function $\mathbf{f} : [-\pi,\pi]^d \to \mathbb{C}^{s\times s}$ belonging to $L^1([-\pi,\pi]^d)$, we denote its Fourier coefficients by

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

where the integrals are done componentwise and $\boldsymbol{k}\cdot\boldsymbol{\theta} = k_1\theta_1 + \ldots + k_d\theta_d$, and we associate to $\mathbf{f}$ the family of multilevel block Toeplitz matrices

$$T_{\boldsymbol{m}}(\mathbf{f}) := [\mathbf{f}_{\boldsymbol{i}-\boldsymbol{j}}]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}, \qquad \boldsymbol{m}\in\mathbb{N}^d.$$

We call $\{T_{\boldsymbol{m}}(\mathbf{f})\}_{\boldsymbol{m}}$ the family of multilevel block Toeplitz matrices associated with the function $\mathbf{f}$, which is called the generating function of $\{T_{\boldsymbol{m}}(\mathbf{f})\}_{\boldsymbol{m}}$. If $\mathbf{f}$ is Hermitian matrix-valued, i.e. $\mathbf{f}(\boldsymbol{\theta})$ is Hermitian for every $\boldsymbol{\theta}$, then it can be shown that all the matrices $T_{\boldsymbol{m}}(\mathbf{f})$ are Hermitian. Theorem 5 is a fundamental result concerning multilevel block Toeplitz matrices generated by Hermitian matrix-valued functions. In particular, item 3 in Theorem 5 is the Tilli theorem [35]. Item 4 is actually a consequence of item 3, while items 1,2 can be proved by using the minimax principle.

**Theorem 5.** *Let $\mathbf{f} : [-\pi,\pi]^d \to \mathbb{C}^{s\times s}$ be a Hermitian matrix-valued function in $L^1([-\pi,\pi]^d)$. Put*

$$m_{\mathbf{f}} = \operatorname*{ess\,inf}_{\boldsymbol{\theta}\in[-\pi,\pi]^d} \lambda_{\min}(\mathbf{f}(\boldsymbol{\theta})), \qquad M_{\mathbf{f}} = \operatorname*{ess\,sup}_{\boldsymbol{\theta}\in[-\pi,\pi]^d} \lambda_{\max}(\mathbf{f}(\boldsymbol{\theta})).$$

*Then the following properties hold.*

1. *$\sigma(T_{\boldsymbol{m}}(\mathbf{f})) \subseteq [m_{\mathbf{f}}, M_{\mathbf{f}}]$ for every $\boldsymbol{m}\in\mathbb{N}^d$.*
2. *If $\lambda_{\min}(\mathbf{f}(\boldsymbol{\theta}))$ is not a.e. constant then $\sigma(T_{\boldsymbol{m}}(\mathbf{f})) \subset (m_{\mathbf{f}}, M_{\mathbf{f}}]$. If $\lambda_{\max}(\mathbf{f}(\boldsymbol{\theta}))$ is not a.e. constant then $\sigma(T_{\boldsymbol{m}}(\mathbf{f})) \subset [m_{\mathbf{f}}, M_{\mathbf{f}})$.*
3. *The family of matrices $\{T_{\boldsymbol{m}}(\mathbf{f})\}_{\boldsymbol{m}}$ is distributed like $\mathbf{f}$ in the sense of the eigenvalues, i.e.*

$$\lim_{\boldsymbol{m}\to\infty} \frac{1}{m_1\cdots m_d s}\sum_{k=1}^{m_1\cdots m_d s} F[\lambda_k(T_{\boldsymbol{m}}(\mathbf{f}))] = \frac{1}{(2\pi)^d}\int_{[-\pi,\pi]^d} \frac{\sum_{j=1}^s F[\lambda_j(\mathbf{f}(\boldsymbol{\theta}))]}{s}d\boldsymbol{\theta} \qquad \forall F\in C_c(\mathbb{C}). \quad (15)$$

   *Due to (15), $\mathbf{f}$ is called the symbol of the Toeplitz family $\{T_{\boldsymbol{m}}(\mathbf{f})\}_{\boldsymbol{m}}$. Note that (15) says that $\mathbf{f}$ is indeed the symbol, in the sense of Definition 1, of any sequence of matrices of the form $\{T_{\boldsymbol{m}(n)}(\mathbf{f})\}_n$, with $\boldsymbol{m}(n)\to\infty$ as $n\to\infty$ (recall from Subsection 2.1 that $\boldsymbol{m}(n)\to\infty$ means that $\min_{j=1,\ldots,d} m_j(n)\to\infty$).*
4. *For each fixed $j\geq 1$, $\lambda_j(T_{\boldsymbol{m}}(\mathbf{f}))\to M_{\mathbf{f}}$ and $\lambda_{m_1\cdots m_d s-j+1}(T_{\boldsymbol{m}}(\mathbf{f}))\to m_{\mathbf{f}}$ when $\boldsymbol{m}\to\infty$.*

*In particular, if $\mathbf{f}\geq O$ a.e. and $\mathrm{m}_d\{\boldsymbol{\theta}\in[-\pi,\pi]^d : \mathbf{f}(\boldsymbol{\theta}) > O\} > 0$, then $T_{\boldsymbol{m}}(\mathbf{f}) > O$ for all $\boldsymbol{m}\in\mathbb{N}^d$.*

The following result relates tensor products and multilevel block Toeplitz matrices. A matrix-valued function of the form $\mathbf{p}(\theta) = \sum_{k=-N}^{N} A_k \mathrm{e}^{\mathrm{i}k\theta}$, with $A_k \in \mathbb{C}^{s\times s}$, $k = -N, \ldots, N$, is called trigonometric polynomial.

**Lemma 6.** *For every $\boldsymbol{m}, \boldsymbol{s} \in \mathbb{N}^d$ there exists a permutation matrix $\Gamma_{\boldsymbol{m},\boldsymbol{s}}$ of dimension $\prod_{j=1}^{d}(m_j s_j)$ such that*

$$T_{m_1}(\mathbf{p}_1) \otimes \cdots \otimes T_{m_d}(\mathbf{p}_d) = \Gamma_{\boldsymbol{m},\boldsymbol{s}} \left[T_{\boldsymbol{m}}(\mathbf{p}_1(\theta_1) \otimes \cdots \otimes \mathbf{p}_d(\theta_d))\right] \Gamma_{\boldsymbol{m},\boldsymbol{s}}^T,$$

*for any choice of the trigonometric polynomials $\mathbf{p}_j : [-\pi, \pi] \to \mathbb{C}^{s_j \times s_j}$, $j = 1, \ldots, d$.*

*Proof.* For $\boldsymbol{k} \in \mathbb{Z}^d$ and $A \in \mathbb{C}^{s\times s}$, it can be shown by direct computation that

$$T_{\boldsymbol{m}}(A\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}}) = T_{m_1}(\mathrm{e}^{\mathrm{i}k_1\theta_1}) \otimes \cdots \otimes T_{m_d}(\mathrm{e}^{\mathrm{i}k_d\theta_d}) \otimes A.$$

Therefore, for any choice of the trigonometric polynomials

$$\mathbf{p}_j(\theta) := \sum_{k=-N_j}^{N_j} A_k^{(j)} \mathrm{e}^{\mathrm{i}k\theta}, \quad j = 1, \ldots, d \qquad (A_k^{(j)} \in \mathbb{C}^{s_j \times s_j}, \quad j = 1, \ldots, d, \ k = -N_j, \ldots, N_j),$$

we have

$$\begin{aligned} T_{\boldsymbol{m}}(\mathbf{p}_1(\theta_1) \otimes \cdots \otimes \mathbf{p}_d(\theta_d)) &= T_{\boldsymbol{m}}\left(\sum_{k_1=-N_1}^{N_1} \cdots \sum_{k_d=-N_d}^{N_d} A_{k_1}^{(1)} \otimes \cdots \otimes A_{k_d}^{(d)} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}}\right) \\ &= \sum_{k_1=-N_1}^{N_1} \cdots \sum_{k_d=-N_d}^{N_d} T_{\boldsymbol{m}}(A_{k_1}^{(1)} \otimes \cdots \otimes A_{k_d}^{(d)} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}}) \\ &= \sum_{\boldsymbol{k}=-\boldsymbol{N}}^{\boldsymbol{N}} T_{m_1}(\mathrm{e}^{\mathrm{i}k_1\theta_1}) \otimes \cdots \otimes T_{m_d}(\mathrm{e}^{\mathrm{i}k_d\theta_d}) \otimes A_{k_1}^{(1)} \otimes \cdots \otimes A_{k_d}^{(d)}. \end{aligned}$$

On the other hand,

$$\begin{aligned} T_{m_1}(\mathbf{p}_1(\theta_1)) \otimes \cdots \otimes T_{m_d}(\mathbf{p}_d(\theta_d)) &= T_{m_1}\left(\sum_{k_1=-N_1}^{N_1} A_{k_1}^{(1)} \mathrm{e}^{\mathrm{i}k_1\theta_1}\right) \otimes \cdots \otimes T_{m_d}\left(\sum_{k_d=-N_d}^{N_d} A_{k_d}^{(d)} \mathrm{e}^{\mathrm{i}k_d\theta_d}\right) \\ &= \left(\sum_{k_1=-N_1}^{N_1} T_{m_1}(\mathrm{e}^{\mathrm{i}k_1\theta_1}) \otimes A_{k_1}^{(1)}\right) \otimes \cdots \otimes \left(\sum_{k_d=-N_d}^{N_d} T_{m_d}(\mathrm{e}^{\mathrm{i}k_d\theta_d}) \otimes A_{k_d}^{(d)}\right) \\ &= \sum_{\boldsymbol{k}=-\boldsymbol{N}}^{\boldsymbol{N}} T_{m_1}(\mathrm{e}^{\mathrm{i}k_1\theta_1}) \otimes A_{k_1}^{(1)} \otimes \cdots \otimes T_{m_d}(\mathrm{e}^{\mathrm{i}k_d\theta_d}) \otimes A_{k_d}^{(d)}. \end{aligned}$$

By Lemma 2, there exists the permutation matrix $\Gamma_{\boldsymbol{m},\boldsymbol{s}} := \Pi_{(\boldsymbol{m},\boldsymbol{s});\sigma}$, where $\sigma := [1, d+1, 2, d+2, \ldots, d, 2d]$, which depends only on $\boldsymbol{m}, \boldsymbol{s}$ and satisfies

$$T_{m_1}(\mathrm{e}^{\mathrm{i}k_1\theta_1}) \otimes A_{k_1}^{(1)} \otimes \cdots \otimes T_{m_d}(\mathrm{e}^{\mathrm{i}k_d\theta_d}) \otimes A_{k_d}^{(d)} = \Gamma_{\boldsymbol{m},\boldsymbol{s}} \left[T_{m_1}(\mathrm{e}^{\mathrm{i}k_1\theta_1}) \otimes \cdots \otimes T_{m_d}(\mathrm{e}^{\mathrm{i}k_d\theta_d}) \otimes A_{k_1}^{(1)} \otimes \cdots \otimes A_{k_d}^{(d)}\right] \Gamma_{\boldsymbol{m},\boldsymbol{s}}^T.$$

Hence,

$$T_{m_1}(\mathbf{p}_1(\theta_1)) \otimes \cdots \otimes T_{m_d}(\mathbf{p}_d(\theta_d)) = \sum_{\boldsymbol{k}=-\boldsymbol{N}}^{\boldsymbol{N}} \Gamma_{\boldsymbol{m},\boldsymbol{s}} \left[T_{m_1}(\mathrm{e}^{\mathrm{i}k_1\theta_1}) \otimes \cdots \otimes T_{m_d}(\mathrm{e}^{\mathrm{i}k_d\theta_d}) \otimes A_{k_1}^{(1)} \otimes \cdots \otimes A_{k_d}^{(d)}\right] \Gamma_{\boldsymbol{m},\boldsymbol{s}}^T$$

$$= \Gamma_{\boldsymbol{m},\boldsymbol{s}} \left\{ \sum_{\boldsymbol{k}=-\boldsymbol{N}}^{\boldsymbol{N}} \left[ T_{m_1}(\mathrm{e}^{\mathrm{i}k_1\theta_1}) \otimes \cdots \otimes T_{m_d}(\mathrm{e}^{\mathrm{i}k_d\theta_d}) \otimes A^{(1)}_{k_1} \otimes \cdots \otimes A^{(d)}_{k_d} \right] \right\} \Gamma^T_{\boldsymbol{m},\boldsymbol{s}}$$
$$= \Gamma_{\boldsymbol{m},\boldsymbol{s}} T_{\boldsymbol{m}}(\mathbf{p}_1(\theta_1) \otimes \cdots \otimes \mathbf{p}_d(\theta_d)) \Gamma^T_{\boldsymbol{m},\boldsymbol{s}}.$$

□

The next results are known in the literature. For $\boldsymbol{m} \in \mathbb{N}^d$ we denote by $F_{\boldsymbol{m}}$ the unitary $d$-level Fourier transform, i.e. $F_{\boldsymbol{m}} := F_{m_1} \otimes \cdots \otimes F_{m_d}$, where $F_m := \frac{1}{\sqrt{m}} \left(\mathrm{e}^{-2\pi \mathrm{i} jk/m}\right)_{j,k=0}^{m-1}$ is the standard unitary Fourier transform of order $m$.

**Proposition 1.** [32] *Let $\mathbf{f}, \mathbf{g} : [-\pi,\pi]^d \to \mathbb{C}^{s\times s}$ be Hermitian matrix-valued functions in $L^1([-\pi,\pi]^d)$ with $\mathbf{f}(\boldsymbol{\theta}) \geq \mathbf{g}(\boldsymbol{\theta})$ a.e. Then $T_{\boldsymbol{m}}(\mathbf{f}) \geq T_{\boldsymbol{m}}(\mathbf{g})$ for all $\boldsymbol{m} \in \mathbb{N}^d$.*

**Theorem 6.** [25] *Let $C = \left[A_{(\boldsymbol{i}-\boldsymbol{j}) \bmod \boldsymbol{m}}\right]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{m}}$ be a multilevel block circulant matrix with blocks $A_{\boldsymbol{k}} \in \mathbb{C}^{s\times s}$, $\boldsymbol{k} = \boldsymbol{0}, \ldots, \boldsymbol{m}-\boldsymbol{1}$. Then $C$ has the following block spectral decomposition:*

$$C = (F_{\boldsymbol{m}} \otimes I_s) \underset{\boldsymbol{j}=\boldsymbol{0},\ldots,\boldsymbol{m}-\boldsymbol{1}}{\operatorname{diag}} \left[ \mathbf{g}\left(\frac{2\pi \boldsymbol{j}}{\boldsymbol{m}}\right) \right] (F_{\boldsymbol{m}} \otimes I_s)^*, \tag{16}$$

*where $\mathbf{g}(\boldsymbol{\theta}) = \sum_{\boldsymbol{k}=\boldsymbol{0}}^{\boldsymbol{m}-\boldsymbol{1}} A_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}}$. In particular, the spectrum of $C$ is given by the union of the spectra of the diagonal blocks $\mathbf{g}(2\pi \boldsymbol{j}/\boldsymbol{m})$, $\boldsymbol{j} = \boldsymbol{0}, \ldots, \boldsymbol{m}-\boldsymbol{1}$.*

## 3 Problem setting and $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM

Let $H_0^1(\Omega)$ be the Sobolev space of functions defined on $\Omega$, vanishing on the boundary $\partial\Omega$, and possessing weak (or Sobolev) partial derivatives of first order [7]. We consider the elliptic differential equation (1), whose weak form can be stated in the following way: find $u \in H_0^1(\Omega)$ such that

$$a(u,v) = \langle f, v\rangle, \quad \forall v \in H_0^1(\Omega), \tag{17}$$

where $a(u,v) := \int_\Omega (\nabla u \cdot \nabla v + \boldsymbol{\beta} \cdot \nabla u\, v + \gamma uv)$ and $\langle f, v\rangle := \int_\Omega fv$.

In the standard Galerkin approach, we find an approximation of $u$ by choosing a finite dimensional subspace $W \subset H_0^1(\Omega)$, called the approximation space, and by solving the following (Galerkin) problem: find $u_W \in W$ such that

$$a(u_W, v) = \langle f, v\rangle, \quad \forall v \in W. \tag{18}$$

If $\dim W = N$ and we fix a basis $\{\varphi_1, \ldots, \varphi_N\}$ for $W$, then the computation of $u_W = \sum_{j=1}^N u_j \varphi_j$ of (18) is reduced to solving the linear system

$$A\mathbf{u} = \mathbf{f}, \tag{19}$$

where $A = [a(\varphi_j, \varphi_i)]_{i,j=1}^N$ is the stiffness matrix and $\mathbf{f} = [\langle f, \varphi_i\rangle]_{i=1}^N$. Once we find $\mathbf{u}$, we know $u_W = \sum_{j=1}^N u_j \varphi_j$.

In the context of $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM, $W$ is chosen as a space of continuous piecewise polynomial functions vanishing on the boundary of $\Omega$. More precisely, let $\mathbb{P}_p$ be the space of polynomials of degree less than or equal to $p$, and define for $p, n \geq 1$ the spaces

$$V_n^{(p)} := \left\{ s \in C([0,1]) : s_{|\left[\frac{i}{n}, \frac{i+1}{n}\right)} \in \mathbb{P}_p \ \ \forall i = 0, \ldots, n-1 \right\},$$
$$W_n^{(p)} := \{ s \in V_n^{(p)} : s(0) = s(1) = 0 \} \subset H_0^1(0,1).$$

It is known that $\dim V_n^{(p)} = np + 1$ and $\dim W_n^{(p)} = np - 1$. Consider for $V_n^{(p)}$ the Lagrangian basis $\{\ell_{0,(p)}, \dots, \ell_{np,(p)}\}$ on the uniform knot sequence $\xi_i = \frac{i}{np},\ i = 0, \dots, np$. This means that $\ell_{j,(p)}$ is the unique function in $V_n^{(p)}$ taking the value 1 at $\xi_j$ and 0 at $\xi_i$ for $i \neq j$:

$$\ell_{j,(p)}(\xi_i) = \delta_{ij}, \quad \forall i, j = 0, \dots, np.$$

Since $\ell_{1,(p)}, \dots, \ell_{np-1,(p)}$ vanish at the boundary of $[0, 1]$, we infer that $\{\ell_{1,(p)}, \dots, \ell_{np-1,(p)}\}$ is a basis for $W_n^{(p)}$ (the Lagrangian basis of $W_n^{(p)}$). For later purposes, we report the explicit expressions of the basis functions $\ell_{1,(p)}, \dots, \ell_{np-1,(p)}$ and of their (Sobolev) derivatives in terms of the Lagrange polynomials $L_0, \dots, L_p$ associated with the knots $t_k = \frac{k}{p},\ k = 0, \dots, p$, which are given by

$$L_h(t) = \prod_{\substack{k=0 \\ k \neq h}}^{p} \frac{t - t_k}{t_h - t_k} = \prod_{\substack{k=0 \\ k \neq h}}^{p} \frac{pt - k}{h - k}, \quad \forall h = 0, \dots, p, \qquad L_h(t_k) = \delta_{hk}, \quad \forall h, k = 0, \dots, p. \tag{20}$$

If $j$ is a multiple of $p$, then the support of $\ell_{j,(p)}$ is $\mathrm{supp}(\ell_{j,(p)}) = [\xi_{j-p}, \xi_{j+p}]$,

$$\ell_{j,(p)}(x) = \begin{cases} L_p\left(\dfrac{x - \xi_{j-p}}{\xi_j - \xi_{j-p}}\right) & \xi_{j-p} \leq x \leq \xi_j, \\ L_0\left(\dfrac{x - \xi_j}{\xi_{j+p} - \xi_j}\right) & \xi_j \leq x \leq \xi_{j+p}, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} L_p(nx - n\xi_{j-p}) & \xi_{j-p} \leq x \leq \xi_j, \\ L_0(nx - n\xi_j) & \xi_j \leq x \leq \xi_{j+p}, \\ 0 & \text{otherwise,} \end{cases} \tag{21}$$

and the derivative of $\ell_{j,(p)}$ is

$$\ell'_{j,(p)}(x) = \begin{cases} nL'_p(nx - n\xi_{j-p}) & \xi_{j-p} < x < \xi_j, \\ nL'_0(nx - n\xi_j) & \xi_j < x < \xi_{j+p}, \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

If $j$ is not a multiple of $p$, let $j_p = j \bmod p \in \{1, \dots, p-1\}$. Then $\mathrm{supp}(\ell_{j,(p)}) = [\xi_{j-j_p}, \xi_{j-j_p+p}]$,

$$\ell_{j,(p)}(x) = \begin{cases} L_{j_p}\left(\dfrac{x - \xi_{j-j_p}}{\xi_{j-j_p+p} - \xi_{j-j_p}}\right) & \xi_{j-j_p} \leq x \leq \xi_{j-j_p+p}, \\ 0 & \text{otherwise,} \end{cases} = \begin{cases} L_{j_p}(nx - n\xi_{j-j_p}) & \xi_{j-j_p} \leq x \leq \xi_{j-j_p+p}, \\ 0 & \text{otherwise,} \end{cases} \tag{23}$$

and the derivative of $\ell_{j,(p)}$ is

$$\ell'_{j,(p)}(x) = \begin{cases} nL'_{j_p}(nx - n\xi_{j-j_p}) & \xi_{j-j_p} < x < \xi_{j-j_p+p}, \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Figure 1 reports the graph of $\ell_{1,(p)}, \dots, \ell_{np-1,(p)}$ in the case $p = 2,\ n = 3$, together with the graph of the Lagrange polynomials $L_0, L_1, L_2$ (20) for $p = 2$. Now, for any pair of multi-indices $\boldsymbol{p}, \boldsymbol{n} \in \mathbb{N}^d$, we define

$$W_{\boldsymbol{n}}^{(\boldsymbol{p})} := W_{n_1}^{(p_1)} \otimes \cdots \otimes W_{n_d}^{(p_d)} := \mathrm{span}\left(\ell_{\boldsymbol{j},(\boldsymbol{p})} : \boldsymbol{j} = \boldsymbol{1}, \dots, \boldsymbol{np - 1}\right), \tag{25}$$

where $\ell_{\boldsymbol{j},(\boldsymbol{p})} := \ell_{j_1,(p_1)} \otimes \cdots \otimes \ell_{j_d,(p_d)}$, and the tensor product function $w_1 \otimes \cdots \otimes w_d : [0, 1]^d \to \mathbb{R}$ is defined in terms of the 'components' $w_i : [0, 1] \to \mathbb{R},\ i = 1, \dots, d$, by $(w_1 \otimes \cdots \otimes w_d)(x_1, \dots, x_d) := w_1(x_1) \cdots w_d(x_d)$.

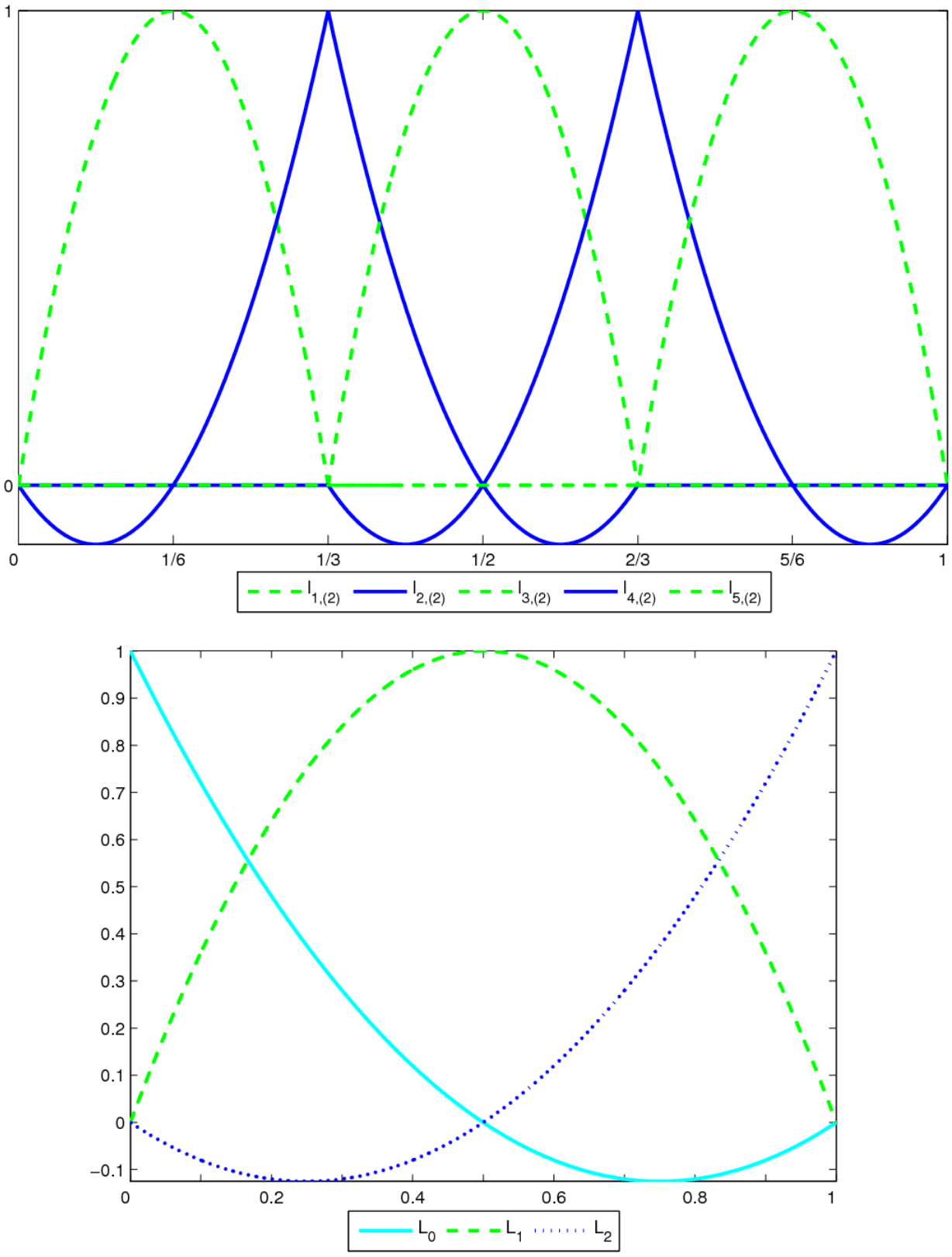


**Figure 1:** graph of the Lagrangian basis functions $\ell_{1,(p)}, \ldots, \ell_{np-1,(p)}$ in the case $p = 2,\ n = 3$, and of the Lagrange polynomials $L_0, \ldots, L_p$ in (20) for $p = 2$.

In the framework of $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM, the subspace $W$ in the Galerkin problem (18) is chosen as $W_{\boldsymbol{n}}^{(\boldsymbol{p})}$ for some $\boldsymbol{p}, \boldsymbol{n} \in \mathbb{N}^d$ (usually $\boldsymbol{p} = (p, \dots, p)$ for some $p \geq 1$), and for $W_{\boldsymbol{n}}^{(\boldsymbol{p})}$ we choose the tensor Lagrangian basis in (25), ordered according to the standard lexicographic ordering (2) for the multi-index range $\mathbf{1}, \dots, \boldsymbol{np} - \mathbf{1}$. With these choices, we obtain in (19) a stiffness matrix $A$, which henceforth will be denoted by $A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ in order to emphasize its dependence on $\boldsymbol{p}$ and $\boldsymbol{n}$:

$$A_{\boldsymbol{n}}^{(\boldsymbol{p})} := \left[a(\ell_{\boldsymbol{j},(\boldsymbol{p})}, \ell_{\boldsymbol{i},(\boldsymbol{p})})\right]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{np}-\mathbf{1}}. \tag{26}$$

Let us consider the following split of the matrix, according to the diffusion, advection and reaction terms, respectively:

$$A_{\boldsymbol{n}}^{(\boldsymbol{p})} = \left[\int_\Omega \nabla \ell_{\boldsymbol{j},(\boldsymbol{p})} \cdot \nabla \ell_{\boldsymbol{i},(\boldsymbol{p})}\right]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{np}-\mathbf{1}} + \left[\int_\Omega \boldsymbol{\beta} \cdot \nabla \ell_{\boldsymbol{j},(\boldsymbol{p})}\, \ell_{\boldsymbol{i},(\boldsymbol{p})}\right]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{np}-\mathbf{1}} + \left[\int_\Omega \gamma\, \ell_{\boldsymbol{j},(\boldsymbol{p})} \ell_{\boldsymbol{i},(\boldsymbol{p})}\right]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{np}-\mathbf{1}}. \tag{27}$$

For obvious reasons, the first matrix in the righthand side of (27) is called diffusion matrix, the second advection matrix, and the third reaction matrix. With expressive notation, we denote these three matrices by, respectively, $A_{\boldsymbol{n},D}^{(\boldsymbol{p})}, A_{\boldsymbol{n},A}^{(\boldsymbol{p})}, A_{\boldsymbol{n},R}^{(\boldsymbol{p})}$, i.e.:

$$A_{\boldsymbol{n},D}^{(\boldsymbol{p})} := \left[\int_\Omega \nabla \ell_{\boldsymbol{j},(\boldsymbol{p})} \cdot \nabla \ell_{\boldsymbol{i},(\boldsymbol{p})}\right]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{np}-\mathbf{1}}, \quad A_{\boldsymbol{n},A}^{(\boldsymbol{p})} := \left[\int_\Omega \boldsymbol{\beta} \cdot \nabla \ell_{\boldsymbol{j},(\boldsymbol{p})}\, \ell_{\boldsymbol{i},(\boldsymbol{p})}\right]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{np}-\mathbf{1}}, \quad A_{\boldsymbol{n},R}^{(\boldsymbol{p})} := \left[\int_\Omega \gamma\, \ell_{\boldsymbol{j},(\boldsymbol{p})} \ell_{\boldsymbol{i},(\boldsymbol{p})}\right]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{np}-\mathbf{1}}. \tag{28}$$

The diffusion matrix is symmetric positive definite (SPD), the reaction matrix is symmetric positive semidefinite (SPSD) and it is SPD if $\gamma \neq 0$ a.e., while the advection matrix is not symmetric and is responsible for the non-symmetry of $A_{\boldsymbol{n}}^{(\boldsymbol{p})}$. The following lemma provides an upper bound for the spectral norm $\|A_{\boldsymbol{n},A}^{(\boldsymbol{p})}\|$. In the following, the symbol $\gamma_*$ will denote a nonnegative constant such that $\gamma \geq \gamma_*$ a.e. on $\Omega$. Moreover, $\|\boldsymbol{\beta}\|_{L^\infty(\Omega)} := \max_{j=1,\dots,d} \|\beta_j\|_{L^\infty(\Omega)}$.

**Lemma 7.** *Let $\boldsymbol{p} \in \mathbb{N}^d$, then there is a constant $B_{\boldsymbol{p}}$, depending only on $\boldsymbol{p}$, such that*

$$\|A_{\boldsymbol{n},A}^{(\boldsymbol{p})}\| \leq B_{\boldsymbol{p}} \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \frac{\sum_{k=1}^d n_k}{n_1 \cdots n_d}, \quad \forall \boldsymbol{n} \in \mathbb{N}^d. \tag{29}$$

*Proof.* By (3) we have $\|A_{\boldsymbol{n},A}^{(\boldsymbol{p})}\| \leq \sqrt{\|A_{\boldsymbol{n},A}^{(\boldsymbol{p})}\|_\infty \|(A_{\boldsymbol{n},A}^{(\boldsymbol{p})})^T\|_\infty}$. Recalling that the $\infty$-norm of a matrix is the maximum 1-norm of its row vectors, if we show that

(a) each entry of $A_{\boldsymbol{n},A}^{(\boldsymbol{p})}$ is bounded from above by $\tilde{B}_{\boldsymbol{p}} \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \frac{\sum_{k=1}^d n_k}{n_1 \cdots n_d}$ for some constant $\tilde{B}_{\boldsymbol{p}}$ depending only on $\boldsymbol{p}$,

(b) each row and column of $A_{\boldsymbol{n},A}^{(\boldsymbol{p})}$ contains a number of nonzero entries bounded from above by some constant $\hat{B}_{\boldsymbol{p}}$ depending only on $\boldsymbol{p}$,

then the thesis follows with $B_{\boldsymbol{p}} = \hat{B}_{\boldsymbol{p}} \tilde{B}_{\boldsymbol{p}}$.

For all $p \geq 1$, set $U_p := \max\{\|L_j\|_{L^\infty(0,1)}, \|L_j'\|_{L^\infty(0,1)} : j = 0, \dots, p\}$, where $L_0, \dots, L_p$ are the Lagrange polynomials (20). From the expressions of $\ell_{1,(p)}, \dots, \ell_{np-1,(p)}$ and of their derivatives given in (21)–(24), and taking into account the supports of $\ell_{1,(p)}, \dots, \ell_{np-1,(p)}$, for all $p, n \geq 1$ and for all $i, j = 1, \dots, np-1$ we have

$$\int_{(0,1)} |\ell_{j,(p)}||\ell_{i,(p)}| \leq \begin{cases} 2U_p^2/n & \text{if } |i-j| \leq p, \\ 0 & \text{otherwise,} \end{cases} \qquad \int_{(0,1)} |\ell_{j,(p)}'||\ell_{i,(p)}| \leq \begin{cases} 2U_p^2 & \text{if } |i-j| \leq p, \\ 0 & \text{otherwise.} \end{cases}$$

Now, for $\boldsymbol{p}, \boldsymbol{n} \in \mathbb{N}^d$, for $\boldsymbol{i}, \boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{np} - \mathbf{1}$ and for $k = 1, \dots, d$, since $\ell_{\boldsymbol{j},(\boldsymbol{p})} = \ell_{j_1,(p_1)} \otimes \cdots \otimes \ell_{j_d,(p_d)}$ and $\Omega = (0,1)^d$ is rectangular, we have

$$\begin{aligned}
\frac{\partial \ell_{\boldsymbol{j},(\boldsymbol{p})}}{\partial x_k} &= \ell_{j_1,(p_1)} \otimes \cdots \otimes \ell_{j_{k-1},(p_{k-1})} \otimes \ell'_{j_k,(p_k)} \otimes \ell_{j_{k+1},(p_{k+1})} \otimes \cdots \otimes \ell_{j_d,(p_d)}, \\
\int_\Omega \left| \frac{\partial \ell_{\boldsymbol{j},(\boldsymbol{p})}}{\partial x_k} \right| |\ell_{\boldsymbol{i},(\boldsymbol{p})}| &= \int_{(0,1)} |\ell_{j_1,(p_1)}||\ell_{i_1,(p_1)}| \cdots \int_{(0,1)} |\ell_{j_{k-1},(p_{k-1})}||\ell_{i_{k-1},(p_{k-1})}| \int_{(0,1)} |\ell'_{j_k,(p_k)}||\ell_{i_k,(p_k)}| \\
&\quad \cdot \int_{(0,1)} |\ell_{j_{k+1},(p_{k+1})}||\ell_{i_{k+1},(p_{k+1})}| \cdots \int_{(0,1)} |\ell_{j_d,(p_d)}||\ell_{i_d,(p_d)}| \\
&\le \begin{cases} \frac{2U_{p_1}^2}{n_1} \cdots \frac{2U_{p_{k-1}}^2}{n_{k-1}} \cdot 2U_{p_k}^2 \cdot \frac{2U_{p_{k+1}}^2}{n_{k+1}} \cdots \frac{2U_{p_d}^2}{n_d} & \text{if } |i_1 - j_1| \le p_1, \dots, |i_d - j_d| \le p_d \\ 0 & \text{otherwise} \end{cases} \\
&\le \begin{cases} U_{\boldsymbol{p}} \frac{n_k}{n_1 \cdots n_d} & \text{if } \|\boldsymbol{i} - \boldsymbol{j}\|_\infty \le \|\boldsymbol{p}\|_\infty \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

where $U_{\boldsymbol{p}} := 2^d \prod_{i=1}^d U_{p_i}^2$. Hence,

$$\begin{aligned}
\left| [A^{(\boldsymbol{p})}_{\boldsymbol{n},A}]_{\boldsymbol{i},\boldsymbol{j}} \right| &= \left| \int_\Omega \boldsymbol{\beta} \cdot \nabla \ell_{\boldsymbol{j},(\boldsymbol{p})}\, \ell_{\boldsymbol{i},(\boldsymbol{p})} \right| \le \int_\Omega \left| \boldsymbol{\beta} \cdot \nabla \ell_{\boldsymbol{j},(\boldsymbol{p})}\, \ell_{\boldsymbol{i},(\boldsymbol{p})} \right| \le \int_\Omega \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \sum_{k=1}^d \left| \frac{\partial \ell_{\boldsymbol{j},(\boldsymbol{p})}}{\partial x_k} \right| |\ell_{\boldsymbol{i},(\boldsymbol{p})}| \\
&= \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \sum_{k=1}^d \int_\Omega \left| \frac{\partial \ell_{\boldsymbol{j},(\boldsymbol{p})}}{\partial x_k} \right| |\ell_{\boldsymbol{i},(\boldsymbol{p})}| \le \begin{cases} \tilde{B}_{\boldsymbol{p}} \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \frac{\sum_{k=1}^d n_k}{n_1 \cdots n_d} & \text{if } \|\boldsymbol{i} - \boldsymbol{j}\|_\infty \le \|\boldsymbol{p}\|_\infty \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

where $\tilde{B}_{\boldsymbol{p}} := U_{\boldsymbol{p}}$. This implies that, for a fixed $\boldsymbol{i} \in \{\mathbf{1}, \dots, \boldsymbol{np} - \mathbf{1}\}$, the $\boldsymbol{i}$-th row of the matrix $A^{(\boldsymbol{p})}_{\boldsymbol{n},A}$ contains at most $\hat{B}_{\boldsymbol{p}} := \prod_{i=1}^d (2p_i + 1)$ nonzero entries (those corresponding to the column multi-indices $\boldsymbol{j}$ such that $\|\boldsymbol{i} - \boldsymbol{j}\|_\infty \le \|\boldsymbol{p}\|_\infty$), and every nonzero entry is bounded from above by $\tilde{B}_{\boldsymbol{p}} \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \frac{\sum_{k=1}^d n_k}{n_1 \cdots n_d}$. Similarly, for a fixed $\boldsymbol{j} \in \{\mathbf{1}, \dots, \boldsymbol{np} - \mathbf{1}\}$, the $\boldsymbol{j}$-th column of $A^{(\boldsymbol{p})}_{\boldsymbol{n},A}$ contains at most $\hat{B}_{\boldsymbol{p}}$ nonzero entries (those corresponding to the row multi-indices $\boldsymbol{i}$ such that $\|\boldsymbol{i} - \boldsymbol{j}\|_\infty \le \|\boldsymbol{p}\|_\infty$), and every nonzero entry is bounded from above by $\tilde{B}_{\boldsymbol{p}} \|\boldsymbol{\beta}\|_{L^\infty(\Omega)} \frac{\sum_{k=1}^d n_k}{n_1 \cdots n_d}$. The conditions (a) and (b) are then satisfied and the thesis follows. □

Now we introduce the mass matrix

$$N^{(\boldsymbol{p})}_{\boldsymbol{n}} := \left[ \int_\Omega \ell_{\boldsymbol{j},(\boldsymbol{p})} \ell_{\boldsymbol{i},(\boldsymbol{p})} \right]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{np}-\mathbf{1}}.$$

This matrix is of interest because

$$\gamma_* N^{(\boldsymbol{p})}_{\boldsymbol{n}} \le A^{(\boldsymbol{p})}_{\boldsymbol{n},R} \le \|\gamma\|_{L^\infty(\Omega)} N^{(\boldsymbol{p})}_{\boldsymbol{n}}. \tag{30}$$

Since all the matrices in (30) are SPSD, their spectral norm equals their maximal eigenvalue. Therefore

$$\gamma_* \|N^{(\boldsymbol{p})}_{\boldsymbol{n}}\| \le \|A^{(\boldsymbol{p})}_{\boldsymbol{n},R}\| \le \|\gamma\|_{L^\infty(\Omega)} \|N^{(\boldsymbol{p})}_{\boldsymbol{n}}\|. \tag{31}$$

Taking into account the tensor structure of the $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian basis $\{\ell_{\boldsymbol{j},(\boldsymbol{p})} : \boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{np} - \mathbf{1}\}$ and the rectangularity of the domain $\Omega$, it can be proved that

$$A^{(\boldsymbol{p})}_{\boldsymbol{n},D} = \sum_{k=1}^d \frac{1}{n_1} M^{(p_1)}_{n_1} \otimes \cdots \otimes \frac{1}{n_{k-1}} M^{(p_{k-1})}_{n_{k-1}} \otimes n_k K^{(p_k)}_{n_k} \otimes \frac{1}{n_{k+1}} M^{(p_{k+1})}_{n_{k+1}} \otimes \cdots \otimes \frac{1}{n_d} M^{(p_d)}_{n_d}, \tag{32}$$

$$N_{\boldsymbol{n}}^{(\boldsymbol{p})} = \frac{1}{n_1} M_{n_1}^{(p_1)} \otimes \cdots \otimes \frac{1}{n_d} M_{n_d}^{(p_d)}, \tag{33}$$

where, for $p, n \geq 1$, $K_n^{(p)}$ and $M_n^{(p)}$ are the SPD matrices given by

$$nK_n^{(p)} := \left[ \int_{(0,1)} \ell'_{j,(p)} \ell'_{i,(p)} \right]_{i,j=1}^{np-1}, \qquad \frac{1}{n} M_n^{(p)} := \left[ \int_{(0,1)} \ell_{j,(p)} \ell_{i,(p)} \right]_{i,j=1}^{np-1};$$

see [19, Theorem 3.1] for details.

## 4 Construction of $K_n^{(p)}, M_n^{(p)}$

This section is devoted to the proof of the following theorem. From now on, the symbol $\langle \, , \rangle$ will be used to denote the scalar product in $L^2([0,1])$, i.e. $\langle \varphi, \psi \rangle := \int_{(0,1)} \varphi\psi$ for all $\varphi, \psi \in L^2([0,1])$.

**Theorem 7.** *Let $p, n \geq 1$. Then*

$$K_n^{(p)} = \begin{bmatrix} K_0 & K_1^T & & \\ K_1 & \ddots & \ddots & \\ & \ddots & \ddots & K_1^T \\ & & K_1 & K_0 \end{bmatrix}_-, \qquad M_n^{(p)} = \begin{bmatrix} M_0 & M_1^T & & \\ M_1 & \ddots & \ddots & \\ & \ddots & \ddots & M_1^T \\ & & M_1 & M_0 \end{bmatrix}_- \tag{34}$$

*where the subscripts '$-$' mean that the last row and column of the matrices in square brackets are deleted, while $K_0, K_1, M_0, M_1$ are $p \times p$ blocks given by*

$$K_0 = \left[\begin{array}{ccc|c} \langle L'_1, L'_1 \rangle & \cdots & \langle L'_{p-1}, L'_1 \rangle & \langle L'_p, L'_1 \rangle \\ \vdots & & \vdots & \vdots \\ \langle L'_1, L'_{p-1} \rangle & \cdots & \langle L'_{p-1}, L'_{p-1} \rangle & \langle L'_p, L'_{p-1} \rangle \\ \hline \langle L'_1, L'_p \rangle & \cdots & \langle L'_{p-1}, L'_p \rangle & \langle L'_p, L'_p \rangle + \langle L'_0, L'_0 \rangle \end{array}\right] \qquad K_1 = \left[\begin{array}{ccc|c} 0 & \cdots & 0 & \langle L'_0, L'_1 \rangle \\ 0 & \cdots & 0 & \langle L'_0, L'_2 \rangle \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & \langle L'_0, L'_p \rangle \end{array}\right] \tag{35}$$

$$M_0 = \left[\begin{array}{ccc|c} \langle L_1, L_1 \rangle & \cdots & \langle L_{p-1}, L_1 \rangle & \langle L_p, L_1 \rangle \\ \vdots & & \vdots & \vdots \\ \langle L_1, L_{p-1} \rangle & \cdots & \langle L_{p-1}, L_{p-1} \rangle & \langle L_p, L_{p-1} \rangle \\ \hline \langle L_1, L_p \rangle & \cdots & \langle L_{p-1}, L_p \rangle & \langle L_p, L_p \rangle + \langle L_0, L_0 \rangle \end{array}\right] \qquad M_1 = \left[\begin{array}{ccc|c} 0 & \cdots & 0 & \langle L_0, L_1 \rangle \\ 0 & \cdots & 0 & \langle L_0, L_2 \rangle \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & \langle L_0, L_p \rangle \end{array}\right] \tag{36}$$

*where $L_0, \ldots, L_p$ are the Lagrange polynomials* (20)*. In particular, $K_n^{(p)}, M_n^{(p)}$ are the leading principal submatrices of order $np-1$ of the block Toeplitz matrices $T_n(\mathbf{f}_p), T_n(\mathbf{h}_p)$, respectively, where $\mathbf{f}_p, \mathbf{h}_p : [-\pi, \pi] \to \mathbb{C}^{p \times p}$ are Hermitian matrix-valued functions given by*

$$\begin{aligned} \mathbf{f}_p(\theta) &:= K_0 + K_1 \mathrm{e}^{\mathrm{i}\theta} + K_1^T \mathrm{e}^{-\mathrm{i}\theta} \\ &= \left[\begin{array}{ccc|c} \langle L'_1, L'_1 \rangle & \cdots & \langle L'_{p-1}, L'_1 \rangle & \langle L'_p, L'_1 \rangle + \langle L'_0, L'_1 \rangle \mathrm{e}^{\mathrm{i}\theta} \\ \vdots & & \vdots & \vdots \\ \langle L'_1, L'_{p-1} \rangle & \cdots & \langle L'_{p-1}, L'_{p-1} \rangle & \langle L'_p, L'_{p-1} \rangle + \langle L'_0, L'_{p-1} \rangle \mathrm{e}^{\mathrm{i}\theta} \\ \hline \langle L'_1, L'_p \rangle + \langle L'_0, L'_1 \rangle \mathrm{e}^{-\mathrm{i}\theta} & \cdots & \langle L'_{p-1}, L'_p \rangle + \langle L'_0, L'_{p-1} \rangle \mathrm{e}^{-\mathrm{i}\theta} & \langle L'_p, L'_p \rangle + \langle L'_0, L'_0 \rangle + 2\langle L'_0, L'_p \rangle \cos\theta \end{array}\right] \\ &= \left[\begin{array}{c|c} [\langle L'_j, L'_i \rangle]_{i,j=1}^{p-1} & \big[\langle L'_p, L'_i \rangle + \langle L'_0, L'_i \rangle \mathrm{e}^{\mathrm{i}\theta}\big]_{i=1}^{p-1} \\ \hline \big[\langle L'_p, L'_i \rangle + \langle L'_0, L'_i \rangle \mathrm{e}^{-\mathrm{i}\theta}\big]_{i=1}^{p-1} & \langle L'_p, L'_p \rangle + \langle L'_0, L'_0 \rangle + 2\langle L'_0, L'_p \rangle \cos\theta \end{array}\right], \end{aligned} \tag{37}$$

$$\mathbf{h}_p(\theta) := M_0 + M_1 \mathrm{e}^{\mathrm{i}\theta} + M_1^T \mathrm{e}^{-\mathrm{i}\theta}$$

$$= \left[\begin{array}{ccc|c} \langle L_1, L_1\rangle & \cdots & \langle L_{p-1}, L_1\rangle & \langle L_p, L_1\rangle + \langle L_0, L_1\rangle \mathrm{e}^{\mathrm{i}\theta} \\ \vdots & & \vdots & \vdots \\ \langle L_1, L_{p-1}\rangle & \cdots & \langle L_{p-1}, L_{p-1}\rangle & \langle L_p, L_{p-1}\rangle + \langle L_0, L_{p-1}\rangle \mathrm{e}^{\mathrm{i}\theta} \\ \hline \langle L_1, L_p\rangle + \langle L_0, L_1\rangle \mathrm{e}^{-\mathrm{i}\theta} & \cdots & \langle L_{p-1}, L_p\rangle + \langle L_0, L_{p-1}\rangle \mathrm{e}^{-\mathrm{i}\theta} & \langle L_p, L_p\rangle + \langle L_0, L_0\rangle + 2\langle L_0, L_p\rangle \cos\theta \end{array}\right]$$

$$= \left[\begin{array}{c|c} [\langle L_j, L_i\rangle]_{i,j=1}^{p-1} & \left[\langle L_p, L_i\rangle + \langle L_0, L_i\rangle \mathrm{e}^{\mathrm{i}\theta}\right]_{i=1}^{p-1} \\ \hline \left[\langle L_p, L_i\rangle + \langle L_0, L_i\rangle \mathrm{e}^{-\mathrm{i}\theta}\right]_{i=1}^{p-1} & \langle L_p, L_p\rangle + \langle L_0, L_0\rangle + 2\langle L_0, L_p\rangle \cos\theta \end{array}\right]. \tag{38}$$

*Proof.* We only give the construction of $K_n^{(p)}$, since the construction of $M_n^{(p)}$ is similar. For convenience, denote by $K$ the matrix in the righthand side of the first equality in (34): we have to show that $K_n^{(p)} = K$. Since both $K_n^{(p)}$ and $K$ are symmetric, it suffices to show that

$$(K_n^{(p)})_{ij} = K_{ij} \qquad \forall i, j = 1, \ldots, np-1 \ \text{ with } \ i \geq j. \tag{39}$$

As in Section 3, set $\xi_i = \frac{i}{np}$ for $i = 0, \ldots, np$ and let $\{\ell_{1,(p)}, \ldots, \ell_{np-1,(p)}\}$ be the Lagrangian basis for $W_n^{(p)}$. For all integers $j$, let $j_p = j \bmod p \in \{0, \ldots, p-1\}$. To prove (39), we first notice that, for all $i, j = 1, \ldots, np-1$ with $i \geq j$,

$$K_{ij} = \begin{cases} \langle L_p', L_p'\rangle + \langle L_0', L_0'\rangle & \text{if } j \text{ is a multiple of } p \text{ and } i = j \\ \langle L_0', L_{i_p}'\rangle & \text{if } j \text{ is a multiple of } p \text{ and } j < i < j + p \\ \langle L_0', L_p'\rangle & \text{if } j \text{ is a multiple of } p \text{ and } i = j + p \\ 0 & \text{if } j \text{ is a multiple of } p \text{ and } i > j + p \\ \langle L_{j_p}', L_{i_p}'\rangle & \text{if } j \text{ is not a multiple of } p \text{ and } j \leq i < j - j_p + p \\ \langle L_{j_p}', L_p'\rangle & \text{if } j \text{ is not a multiple of } p \text{ and } i = j - j_p + p \\ 0 & \text{if } j \text{ is not a multiple of } p \text{ and } i > j - j_p + p \end{cases} \tag{40}$$

We verify that (39) holds by considering the seven cases in (40). The verification is plain: it suffices to use the expressions of $\ell_{1,(p)}, \ldots, \ell_{np-1,(p)}$ and of their derivatives given in (21)–(24). For completeness, we include the verification in the paper.

(i) If $j$ is a multiple of $p$ and $i = j$, then

$$(K_n^{(p)})_{ij} = \frac{1}{n}\int_0^1 \ell_{j,(p)}'(x)^2 dx = n\int_{\xi_{j-p}}^{\xi_j} L_p'(nx - n\xi_{j-p})^2 dx + n\int_{\xi_j}^{\xi_{j+p}} L_0'(nx - n\xi_j)^2 dx \quad \text{(by (22))}$$
$$= \int_0^1 L_p'(t)^2 dt + \int_0^1 L_0'(t)^2 dt = \langle L_p', L_p'\rangle + \langle L_0', L_0'\rangle = K_{ij}.$$

(ii) If $j$ is a multiple of $p$ and $j < i < j + p$, then $i$ is not a multiple of $p$, $i - i_p = j$, $\mathrm{supp}(\ell_{i,(p)}) = [\xi_j, \xi_{j+p}]$ and

$$(K_n^{(p)})_{ij} = \frac{1}{n}\int_0^1 \ell_{j,(p)}'(x)\ell_{i,(p)}'(x) dx = n\int_{\xi_j}^{\xi_{j+p}} L_0'(nx - n\xi_j) L_{i_p}'(nx - n\xi_j) dx \quad \text{(by (22) and (24))}$$
$$= \int_0^1 L_0'(t) L_{i_p}'(t) dt = \langle L_0', L_{i_p}'\rangle = K_{ij}.$$

(iii) If $j$ is a multiple of $p$ and $i = j + p$, then $i$ is a multiple of $p$, $\mathrm{supp}(\ell_{i,(p)}) \cap \mathrm{supp}(\ell_{j,(p)}) = [\xi_{i-p}, \xi_{i+p}] \cap [\xi_{j-p}, \xi_{j+p}] = [\xi_j, \xi_{j+2p}] \cap [\xi_{j-p}, \xi_{j+p}] = [\xi_j, \xi_{j+p}]$ and

$$(K_n^{(p)})_{ij} = \frac{1}{n}\int_0^1 \ell'_{j,(p)}(x)\ell'_{i,(p)}(x)dx = n\int_{\xi_j}^{\xi_{j+p}} L'_0(nx - n\xi_j)L'_p(nx - n\xi_j)dx \quad \text{(by (22))}$$
$$= \int_0^1 L'_0(t)L'_p(t)dt = \langle L'_0, L'_p\rangle = K_{ij}.$$

(iv) If $j$ is a multiple of $p$ and $i > j + p$, then $\mathrm{supp}(\ell_{i,(p)}) \subseteq [\xi_{j+p}, 1]$ and $\mathrm{supp}(\ell_{j,(p)}) = [\xi_{j-p}, \xi_{j+p}]$, and so

$$(K_n^{(p)})_{ij} = \frac{1}{n}\int_0^1 \ell'_{j,(p)}(x)\ell'_{i,(p)}(x)dx = 0 = K_{ij}.$$

(v) If $j$ is not a multiple of $p$ and $j \le i < j - j_p + p$, then $i$ is not a multiple of $p$, $i - i_p = j - j_p$, $\mathrm{supp}(\ell_{i,(p)}) = [\xi_{i-i_p}, \xi_{i-i_p+p}] = [\xi_{j-j_p}, \xi_{j-j_p+p}] = \mathrm{supp}(\ell_{j,(p)})$ and

$$(K_n^{(p)})_{ij} = \frac{1}{n}\int_0^1 \ell'_{j,(p)}(x)\ell'_{i,(p)}(x)dx = n\int_{\xi_{j-j_p}}^{\xi_{j-j_p+p}} L'_{j_p}(nx - n\xi_{j-j_p})L'_{i_p}(nx - n\xi_{j-j_p})dx \quad \text{(by (24))}$$
$$= \int_0^1 L'_{j_p}(t)L'_{i_p}(t)dt = \langle L'_{j_p}, L'_{i_p}\rangle = K_{ij}.$$

(vi) If $j$ is not a multiple of $p$ and $i = j - j_p + p$, then $i$ is a multiple of $p$, $i - p = j - j_p$, $\mathrm{supp}(\ell_{i,(p)}) \cap \mathrm{supp}(\ell_{j,(p)}) = [\xi_{i-p}, \xi_{i+p}] \cap [\xi_{j-j_p}, \xi_{j-j_p+p}] = [\xi_{j-j_p}, \xi_{j-j_p+2p}] \cap [\xi_{j-j_p}, \xi_{j-j_p+p}] = [\xi_{j-j_p}, \xi_{j-j_p+p}]$ and

$$(K_n^{(p)})_{ij} = \frac{1}{n}\int_0^1 \ell'_{j,(p)}(x)\ell'_{i,(p)}(x)dx = n\int_{\xi_{j-j_p}}^{\xi_{j-j_p+p}} L'_{j_p}(nx - n\xi_{j-j_p})L'_p(nx - n\xi_{j-j_p})dx \quad \text{(by (24))}$$
$$= \int_0^1 L'_{j_p}(t)L'_p(t)dt = \langle L'_{j_p}, L'_p\rangle = K_{ij}.$$

(vii) If $j$ is not a multiple of $p$ and $i > j - j_p + p$, then $\mathrm{supp}(\ell_{i,(p)}) \subseteq [\xi_{j-j_p+p}, 1]$ and $\mathrm{supp}(\ell_{j,(p)}) = [\xi_{j-j_p}, \xi_{j-j_p+p}]$, and so

$$(K_n^{(p)})_{ij} = \frac{1}{n}\int_0^1 \ell'_{j,(p)}(x)\ell'_{i,(p)}(x)dx = 0 = K_{ij}.$$

□

**Remark 3.** Let $p \ge 1$ and let $L_0, \dots, L_p$ be the Lagrange polynomials (20). Then, for every $h = 0, \dots, p$ and every $t \in \mathbb{R}$, a direct verification shows that $L_h(1 - t) = L_{p-h}(t)$. As a consequence, the equalities $\langle L_i, L_j\rangle = \langle L_{p-i}, L_{p-j}\rangle$ and $\langle L'_i, L'_j\rangle = \langle L'_{p-i}, L'_{p-j}\rangle$ hold for all $i, j = 0, \dots, p$. These relations may be used to give alternative expressions for the entries of the blocks $K_0, K_1, M_0, M_1$ in (35)–(36).

## 5 Properties of $\mathbf{f}_p(\theta)$ and $\mathbf{h}_p(\theta)$

In this section we derive some properties of the Hermitian matrix-valued functions $\mathbf{f}_p(\theta)$, $\mathbf{h}_p(\theta)$ defined in (37)–(38). We need some results concerning the Lagrange polynomials.

**Lemma 8.** *Let $p \geq 1$ and let $L_0, \ldots, L_p$ be the Lagrange polynomials* (20)*. Then*

$$\sum_{j=1}^{p} jL'_j = p \quad \textit{identically,} \tag{41}$$

$$\sum_{j=0}^{p} L'_j = 0 \quad \textit{identically,} \tag{42}$$

*while every proper subset of $\{L'_0, \ldots, L'_p\}$ is linearly independent.*

*Proof.* (41) holds because $\sum_{j=1}^{p} jL_j = \sum_{j=0}^{p} jL_j$ is the interpolating polynomial which takes the value $j$ over the knot $t_j = \frac{j}{p}$, for $j = 0, \ldots, p$, and hence $\sum_{j=1}^{p} jL_j(t) = pt$ identically. (42) holds because $\sum_{j=0}^{p} L_j$ is the interpolating polynomial which takes the value 1 over the uniform knots $t_k = \frac{k}{p},\ k = 0, \ldots, p$, and hence $\sum_{j=0}^{p} L_j = 1$ identically.

We prove that every proper subset of $\{L'_0, \ldots, L'_p\}$ is linearly independent. To this end, it suffices to prove that every proper subset of $\{L'_0, \ldots, L'_p\}$ with cardinality $p$ is linearly independent. Actually, we will only prove that $\{L'_1, \ldots, L'_p\}$ is linearly independent, since the proof for the other subsets is similar. Let $\alpha_1, \ldots, \alpha_p$ be numbers such that $\sum_{i=1}^{p} \alpha_i L'_i = (\sum_{i=1}^{p} \alpha_i L_i)' = 0$ identically. Then there exists a constant $C$ such that

$$\sum_{i=1}^{p} \alpha_i L_i = C \quad \text{identically.} \tag{43}$$

By evaluating (43) in $t_k = \frac{k}{p},\ k = 0, \ldots, p$, and by remembering (20), we find that $C = 0$ and $\alpha_1 = \ldots = \alpha_p = C$, which yields $\alpha_1 = \ldots = \alpha_p = 0$. Thus $L'_1, \ldots, L'_p$ are linearly independent. □

**Lemma 9.** *Let $p \geq 1$ and set $d_p := \det([\langle L'_j, L'_i \rangle]_{i,j=1}^{p})$, where $L_0, \ldots, L_p$ are the Lagrange polynomials* (20)*. Then $d_p > 0$ and $d_p = \det([\langle L'_j, L'_i \rangle]_{i,j=1}^{p-1})$.* [1]

*Proof.* The lemma is true if $p = 1$, because $L_1(t) = t$, $L'_1(t) = 1$, and $d_1 = \langle L'_1, L'_1 \rangle = 1$. In the following we assume $p \geq 2$. We have $d_p > 0$ because the matrix $[\langle L'_j, L'_i \rangle]_{i,j=1}^{p}$ is SPD, due to the fact that $L'_1, \ldots, L'_p$ are linearly independent (Lemma 8).

We want to show that $d_p = \det(\mathcal{L})$, where $\mathcal{L} := [\langle L'_j, L'_i \rangle]_{i,j=1}^{p-1}$. To this end, we perform the block Gauss transformation that creates zeros in the first $p-1$ components of the last row and column of $[\langle L'_j, L'_i \rangle]_{i,j=1}^{p}$. Setting

$$G = \left[\begin{array}{c|c} I_{p-1} & \mathbf{0} \\ \hline -[\langle L'_1, L'_p \rangle \ \cdots \ \langle L'_{p-1}, L'_p \rangle]\mathcal{L}^{-1} & 1 \end{array}\right],$$

we have

$$G[\langle L'_j, L'_i \rangle]_{i,j=1}^{p} G^T = G \left[\begin{array}{c|c} \mathcal{L} & \begin{matrix} \langle L'_p, L'_1 \rangle \\ \vdots \\ \langle L'_p, L'_{p-1} \rangle \end{matrix} \\ \hline \langle L'_1, L'_p \rangle \ \cdots \ \langle L'_{p-1}, L'_p \rangle & \langle L'_p, L'_p \rangle \end{array}\right] G^T = \left[\begin{array}{c|c} \mathcal{L} & \mathbf{0} \\ \hline \mathbf{0}^T & s \end{array}\right] =: Z,$$

where $s := \langle L'_p, L'_p \rangle - [\langle L'_1, L'_p \rangle \ \cdots \ \langle L'_{p-1}, L'_p \rangle]\mathcal{L}^{-1} \begin{bmatrix} \langle L'_p, L'_1 \rangle \\ \vdots \\ \langle L'_p, L'_{p-1} \rangle \end{bmatrix}$ is the Schur complement of $\mathcal{L}$. Since $\det(G) = \det(G^T) = 1$, we have $d_p = \det(Z) = \det(\mathcal{L})s$, and so the lemma is proved provided we show

[1] We use the (standard) convention that the determinant of the empty matrix is 1, so that the latter formula gives $d_1 = 1$.

that $s = 1$. To prove this, we note that $\mathcal{L}^{-1}\begin{bmatrix} \langle L'_p, L'_1\rangle \\ \vdots \\ \langle L'_p, L'_{p-1}\rangle \end{bmatrix}$ is the solution of the linear system $\mathcal{L}\mathbf{u} = \begin{bmatrix} \langle L'_p, L'_1\rangle \\ \vdots \\ \langle L'_p, L'_{p-1}\rangle \end{bmatrix}$, which is easily seen to be $\mathbf{u} := \left[-\frac{1}{p}, -\frac{2}{p}, \ldots, -\frac{p-1}{p}\right]^T$. Indeed, by Lemma 8, for all $i = 1, \ldots, p-1$ we have

$$(\mathcal{L}\mathbf{u})_i = \sum_{j=1}^{p-1}\langle L'_j, L'_i\rangle u_j = -\frac{1}{p}\sum_{j=1}^{p-1} j\langle L'_j, L'_i\rangle = -\frac{1}{p}\left\langle \sum_{j=1}^{p-1} jL'_j, L'_i\right\rangle = -\frac{1}{p}\langle p - pL'_p, L'_i\rangle = -\langle 1, L'_i\rangle + \langle L'_p, L'_i\rangle$$
$$= -\int_0^1 L'_i(t)dt + \langle L'_p, L'_i\rangle = -L_i(1) + L_i(0) + \langle L'_p, L'_i\rangle = \langle L'_p, L'_i\rangle,$$

where the last equality is due to the fact that $L_i(0) = L_i(1) = 0$ for $i = 1, \ldots, p-1$. Using again Lemma 8, we obtain

$$s = \langle L'_p, L'_p\rangle - [\langle L'_1, L'_p\rangle \ \cdots \ \langle L'_{p-1}, L'_p\rangle]\mathcal{L}^{-1}\begin{bmatrix} \langle L'_p, L'_1\rangle \\ \vdots \\ \langle L'_p, L'_{p-1}\rangle \end{bmatrix} = \langle L'_p, L'_p\rangle - [\langle L'_1, L'_p\rangle \ \cdots \ \langle L'_{p-1}, L'_p\rangle]\mathbf{u}$$
$$= \langle L'_p, L'_p\rangle - \sum_{j=1}^{p-1}\langle L'_j, L'_p\rangle u_j = \langle L'_p, L'_p\rangle + \left\langle \sum_{j=1}^{p-1}\frac{j}{p}L'_j, L'_p\right\rangle = \left\langle \sum_{j=1}^{p}\frac{j}{p}L'_j, L'_p\right\rangle = \langle 1, L'_p\rangle$$
$$= \int_0^1 L'_p(t)dt = L_p(1) - L_p(0) = 1,$$

which concludes the proof. □

**Theorem 8.** *Let $p \geq 1$, then*

$$\det(\mathbf{f}_p(\theta)) = d_p(2 - 2\cos\theta), \tag{44}$$

*where $d_p$ is defined in Lemma 9.*

*Proof.* The theorem is true if $p = 1$, because $d_1 = 1$ and $\mathbf{f}_1(\theta) = 2 - 2\cos\theta$. In the following we assume $p \geq 2$. By (37) and by the linearity of the determinant with respect to each row and column, we have

$$\det(\mathbf{f}_p(\theta)) = \left|\begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_p, L'_i\rangle + \langle L'_0, L'_i\rangle \mathrm{e}^{\mathrm{i}\theta}]_{i=1}^{p-1} \\ \hline \left[\langle L'_p, L'_i\rangle + \langle L'_0, L'_i\rangle \mathrm{e}^{-\mathrm{i}\theta}\right]_{i=1}^{p-1} & \langle L'_p, L'_p\rangle + \langle L'_0, L'_0\rangle + 2\langle L'_0, L'_p\rangle\cos\theta \end{array}\right|$$
$$= \left|\begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_p, L'_i\rangle]_{i=1}^{p-1} \\ \hline \left[\langle L'_p, L'_i\rangle + \langle L'_0, L'_i\rangle \mathrm{e}^{-\mathrm{i}\theta}\right]_{i=1}^{p-1} & \langle L'_p, L'_p\rangle + \langle L'_0, L'_p\rangle \mathrm{e}^{-\mathrm{i}\theta} \end{array}\right|$$
$$+ \left|\begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_0, L'_i\rangle \mathrm{e}^{\mathrm{i}\theta}]_{i=1}^{p-1} \\ \hline \left[\langle L'_p, L'_i\rangle + \langle L'_0, L'_i\rangle \mathrm{e}^{-\mathrm{i}\theta}\right]_{i=1}^{p-1} & \langle L'_0, L'_0\rangle + \langle L'_0, L'_p\rangle \mathrm{e}^{\mathrm{i}\theta} \end{array}\right|$$
$$= \left|\begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_p, L'_i\rangle]_{i=1}^{p-1} \\ \hline \left[\langle L'_p, L'_i\rangle\right]_{i=1}^{p-1} & \langle L'_p, L'_p\rangle \end{array}\right| + \left|\begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_p, L'_i\rangle]_{i=1}^{p-1} \\ \hline \left[\langle L'_0, L'_i\rangle \mathrm{e}^{-\mathrm{i}\theta}\right]_{i=1}^{p-1} & \langle L'_0, L'_p\rangle \mathrm{e}^{-\mathrm{i}\theta} \end{array}\right|$$
$$+ \left|\begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_0, L'_i\rangle \mathrm{e}^{\mathrm{i}\theta}]_{i=1}^{p-1} \\ \hline \left[\langle L'_p, L'_i\rangle\right]_{i=1}^{p-1} & \langle L'_0, L'_p\rangle \mathrm{e}^{\mathrm{i}\theta} \end{array}\right| + \left|\begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_0, L'_i\rangle \mathrm{e}^{\mathrm{i}\theta}]_{i=1}^{p-1} \\ \hline \left[\langle L'_0, L'_i\rangle \mathrm{e}^{-\mathrm{i}\theta}\right]_{i=1}^{p-1} & \langle L'_0, L'_0\rangle \end{array}\right|$$

$$= d_p + \mathrm{e}^{-\mathrm{i}\theta} \left| \begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_p, L'_i\rangle]_{i=1}^{p-1} \\ \hline [\langle L'_0, L'_i\rangle]_{i=1}^{p-1} & \langle L'_0, L'_p\rangle \end{array} \right| + \mathrm{e}^{\mathrm{i}\theta} \left| \begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_0, L'_i\rangle]_{i=1}^{p-1} \\ \hline \big[\langle L'_p, L'_i\rangle\big]_{i=1}^{p-1} & \langle L'_0, L'_p\rangle \end{array} \right|$$
$$+ \left| \begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_0, L'_i\rangle]_{i=1}^{p-1} \\ \hline [\langle L'_0, L'_i\rangle]_{i=1}^{p-1} & \langle L'_0, L'_0\rangle \end{array} \right| =: d_p + \mathrm{e}^{-\mathrm{i}\theta} d'_p + \mathrm{e}^{\mathrm{i}\theta} d'_p + d''_p = d_p + 2d'_p \cos\theta + d''_p \qquad (45)$$

We prove that

$$\det(\mathbf{f}_p(0)) = d_p + 2d'_p + d''_p = 0, \qquad (46)$$
$$d_p + d'_p = 0, \qquad (47)$$

after which (44) follows from (45). By (37) we have

$$\mathbf{f}_p(0) = \left[ \begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_0 + L'_p, L'_i\rangle]_{i=1}^{p-1} \\ \hline [\langle L'_0 + L'_p, L'_i\rangle]_{i=1}^{p-1} & \langle L'_0 + L'_p, L'_0 + L'_p\rangle \end{array} \right] =: [\langle N'_j, N'_i\rangle]_{i,j=1}^{p},$$

where $N_i := L_i$ for $i = 1, \ldots, p-1$ and $N_p := L_0 + L_p$. Since $\sum_{i=1}^{p} N'_i = \sum_{i=0}^{p} L'_i = 0$ identically (Lemma 8), it follows that $N'_1, \ldots, N'_p$ are linearly dependent, $\mathbf{f}_p(0)$ is singular, and (46) holds. To prove (47) we simply note that

$$d_p + d'_p = \left| \begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_p, L'_i\rangle]_{i=1}^{p-1} \\ \hline [\langle L'_p, L'_i\rangle]_{i=1}^{p-1} & \langle L'_p, L'_p\rangle \end{array} \right| + \left| \begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_0, L'_i\rangle]_{i=1}^{p-1} \\ \hline [\langle L'_p, L'_i\rangle]_{i=1}^{p-1} & \langle L'_0, L'_p\rangle \end{array} \right|$$
$$= \left| \begin{array}{c|c} [\langle L'_j, L'_i\rangle]_{i,j=1}^{p-1} & [\langle L'_p + L'_0, L'_i\rangle]_{i=1}^{p-1} \\ \hline [\langle L'_p, L'_i\rangle]_{i=1}^{p-1} & \langle L'_p + L'_0, L'_p\rangle \end{array} \right| = 0,$$

where the latter is a consequence of the fact that, by Lemma 8, $L'_p + L'_0$ is a linear combination of $L'_1, \ldots, L'_{p-1}$, which implies that the last column of $d_p + d'_p$ is a linear combination of the others. □

**Theorem 9.** *Let $p \in \{1, \ldots, 15\}$, then*

$$\det(\mathbf{h}_p(\theta)) = a_p \left(1 + \frac{(-1)^{p+1}}{p+1} \cos\theta\right), \qquad (48)$$

*where $a_p = \det(\mathbf{h}_p(\frac{\pi}{2})) > 0$.*

*Proof.* The proof is computer assisted and indeed the result has been verified by direct computation using the symbolic program MAPLE. □

Although the result of Theorem 9 has not been proved for all $p \geq 1$, we can certainly formulate the following conjecture.

**Conjecture 1.** *Theorem 9 holds for all $p \geq 1$.*

**Remark 4.** Assuming we are able to prove that formula (48) holds with some constant $a_p$, we do not need to prove that $a_p = \det(\mathbf{h}_p(\frac{\pi}{2})) > 0$. Indeed, if (48) holds with some constant $a_p$, then, by evaluating both sides at $\theta = \frac{\pi}{2}$, we get $a_p = \det(\mathbf{h}_p(\frac{\pi}{2}))$. Moreover, $a_p > 0$. Indeed, we have $\mathbf{h}_p(0) = [\langle N_j, N_i\rangle]_{i,j=1}^{p}$ with $N_i := L_i$ for $i = 1, \ldots, p-1$ and $N_p := L_0 + L_p$. Since $N_1, \ldots, N_p$ are linearly independent, due to the linear independence of $L_0, \ldots, L_p$, it follows that $\mathbf{h}_p(0) > O$. Hence, $\det(\mathbf{h}_p(0)) > 0$ and

$$a_p = \frac{\det(\mathbf{h}_p(0))}{\left(1 + \frac{(-1)^{p+1}}{p+1}\right)} > 0.$$

In this paper, we will assume that Conjecture 1 holds. The results relying on this conjecture are certainly true for $p = 1, \dots, 15$.

In the following, for $p \geq 2$ we denote by $\mu_1^{(p)} \geq \dots \geq \mu_{p-1}^{(p)} > 0$ and $\eta_1^{(p)} \geq \dots \geq \eta_{p-1}^{(p)} > 0$ the eigenvalues of the SPD matrices $[\langle L_j', L_i' \rangle]_{i,j=1}^{p-1}$ and $[\langle L_j, L_i \rangle]_{i,j=1}^{p-1}$, respectively, where $L_0, \dots, L_p$ are the Lagrange polynomials (20). Moreover, we define

$$m_{\mathbf{f}_p} := \min_{\theta \in [-\pi,\pi]} \lambda_{\min}(\mathbf{f}_p(\theta)), \qquad M_{\mathbf{f}_p} := \max_{\theta \in [-\pi,\pi]} \lambda_{\max}(\mathbf{f}_p(\theta)),$$

$$m_{\mathbf{h}_p} := \min_{\theta \in [-\pi,\pi]} \lambda_{\min}(\mathbf{h}_p(\theta)), \qquad M_{\mathbf{h}_p} := \max_{\theta \in [-\pi,\pi]} \lambda_{\max}(\mathbf{h}_p(\theta)).$$

**Corollary 1.** *The following properties hold.*

1. *Let $p \geq 2$, then $\lambda_1(\mathbf{f}_p(\theta)) \geq \mu_1^{(p)} \geq \lambda_2(\mathbf{f}_p(\theta)) \geq \mu_2^{(p)} \geq \dots \geq \lambda_{p-1}(\mathbf{f}_p(\theta)) \geq \mu_{p-1}^{(p)} \geq \lambda_p(\mathbf{f}_p(\theta))$ for all $\theta$.*
2. *Let $p \geq 1$, then there exists a constant $c_p > 0$ such that, for all $\theta$,*

$$c_p(2 - 2\cos\theta) \leq \lambda_{\min}(\mathbf{f}_p(\theta)) \leq 2 - 2\cos\theta. \tag{49}$$

*In (49) we can take $c_1 = 1$ and $c_p = \dfrac{\mu_{p-1}^{(p)}}{M_{\mathbf{f}_p}}$ for $p \geq 2$. In particular, $m_{\mathbf{f}_p} = 0$, $\mathbf{f}_p(\theta) \geq O$ for all $\theta \in [-\pi, \pi]$, and $\mathbf{f}_p(\theta) > O$ for all nonzero $\theta \in [-\pi, \pi]$.*

*Proof.* For $p = 1$ the corollary can be directly verified, because $\mathbf{f}_1(\theta) = 2 - 2\cos\theta$. Assume $p \geq 2$. Item 1 follows from the Cauchy interlacing theorem and from the fact that $[\langle L_j', L_i' \rangle]_{i,j=1}^{p-1}$ is the leading principal submatrix of $\mathbf{f}_p(\theta)$ for all $\theta$. To prove item 2, observe that, by Theorem 8,

$$\lambda_1(\mathbf{f}_p(\theta)) \cdots \lambda_p(\mathbf{f}_p(\theta)) = \det(\mathbf{f}_p(\theta)) = d_p(2 - 2\cos\theta) \;\Rightarrow\; \lambda_{\min}(\mathbf{f}_p(\theta)) = \frac{d_p}{\lambda_1(\mathbf{f}_p(\theta)) \cdots \lambda_{p-1}(\mathbf{f}_p(\theta))}(2 - 2\cos\theta).$$

Furthermore, by item 1 and Lemma 9, for all $\theta$ we have

$$\lambda_1(\mathbf{f}_p(\theta)) \cdots \lambda_{p-1}(\mathbf{f}_p(\theta)) \geq \mu_1^{(p)} \cdots \mu_{p-1}^{(p)} = \det([\langle L_j', L_i' \rangle]_{i,j=1}^{p-1}) = d_p,$$

$$\lambda_1(\mathbf{f}_p(\theta)) \cdots \lambda_{p-1}(\mathbf{f}_p(\theta)) \leq M_{\mathbf{f}_p} \mu_1^{(p)} \cdots \mu_{p-2}^{(p)} = \frac{M_{\mathbf{f}_p} \mu_1^{(p)} \cdots \mu_{p-1}^{(p)}}{\mu_{p-1}^{(p)}} = \frac{M_{\mathbf{f}_p} d_p}{\mu_{p-1}^{(p)}},$$

and item 2 follows. □

**Corollary 2.** *The following properties hold.*

1. *Let $p \geq 2$, then $\lambda_1(\mathbf{h}_p(\theta)) \geq \eta_1^{(p)} \geq \lambda_2(\mathbf{h}_p(\theta)) \geq \eta_2^{(p)} \geq \dots \geq \lambda_{p-1}(\mathbf{h}_p(\theta)) \geq \eta_{p-1}^{(p)} \geq \lambda_p(\mathbf{h}_p(\theta))$ for all $\theta$.*
2. *Let $p \geq 1$, then $m_{\mathbf{h}_p} > 0$. In particular, $\mathbf{h}_p(\theta) > O$ for all $\theta$. In addition, $m_{\mathbf{h}_1} = \frac{1}{3}$, while for $p \geq 2$ we have $m_{\mathbf{h}_p} \geq \dfrac{p a_p \eta_{p-1}^{(p)}}{(p+1)\eta_1^{(p)} \cdots \eta_{p-1}^{(p)}}$, where $a_p = \det(\mathbf{h}_p(\frac{\pi}{2})) > 0$.*

*Proof.* For $p = 1$ the corollary can be directly verified, because $\mathbf{h}_1(\theta) = \frac{2}{3} + \frac{1}{3}\cos\theta$. Assume $p \geq 2$. Item 1 follows from the Cauchy interlacing theorem and from the fact that $[\langle L_j, L_i \rangle]_{i,j=1}^{p-1}$ is the leading principal submatrix of $\mathbf{h}_p(\theta)$ for all $\theta$. To prove item 2, we simply note that, by item 1 and Conjecture 1,

$$\lambda_1(\mathbf{h}_p(\theta)) \cdots \lambda_p(\mathbf{h}_p(\theta)) = \det(\mathbf{h}_p(\theta)) = a_p \left(1 + \frac{(-1)^{p+1}}{p+1}\cos\theta\right)$$

$$\Rightarrow\; \lambda_{\min}(\mathbf{h}_p(\theta)) = \frac{a_p}{\lambda_1(\mathbf{h}_p(\theta)) \cdots \lambda_{p-1}(\mathbf{h}_p(\theta))}\left(1 + \frac{(-1)^{p+1}}{p+1}\cos\theta\right) \geq \frac{a_p}{M_{\mathbf{h}_p}\eta_1^{(p)} \cdots \eta_{p-2}^{(p)}}\left(1 - \frac{1}{p+1}\right).$$

□

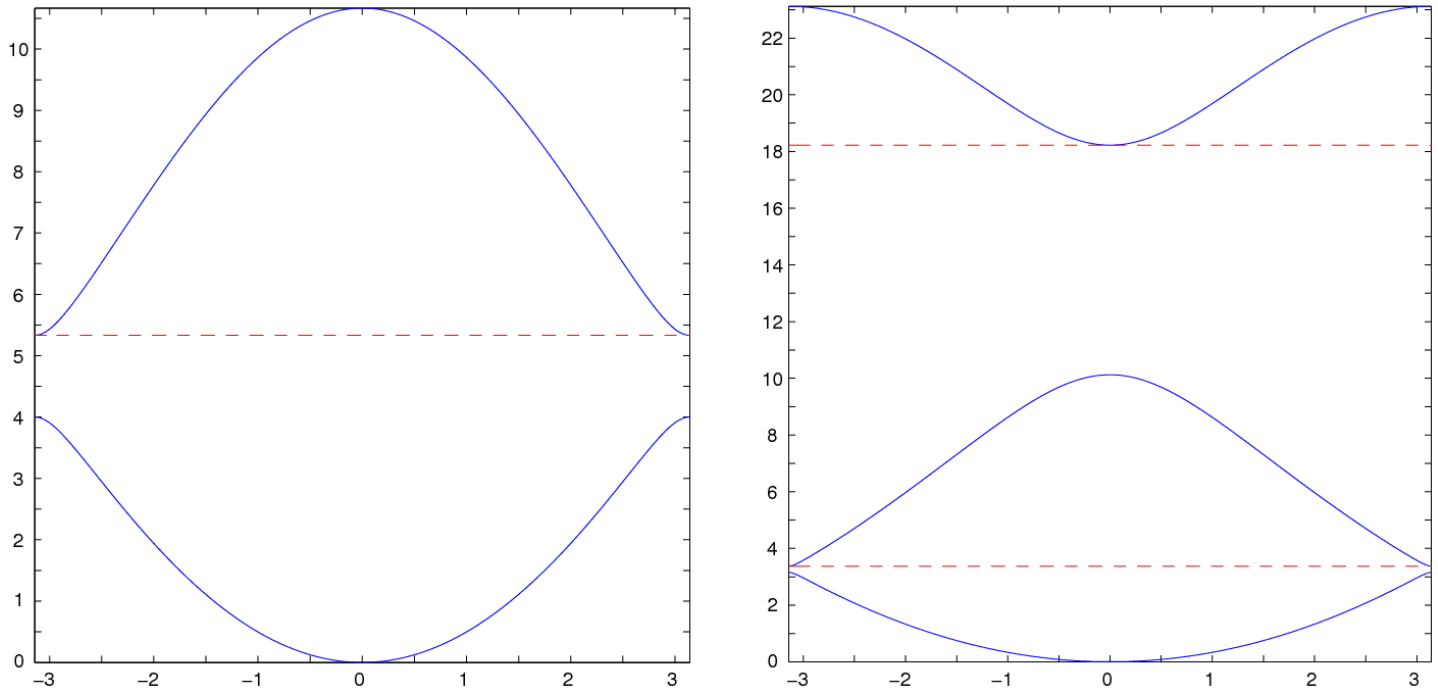


**Figure 2:** left: graph of the eigenvalue functions $\theta \mapsto \lambda_j(\mathbf{f}_2(\theta)),\ j = 1, 2$ (solid lines), and of the separating line with ordinate $\frac{16}{3}$ (dashed line); right: graph of the eigenvalue functions $\theta \mapsto \lambda_j(\mathbf{f}_3(\theta)),\ j = 1, 2, 3$ (solid lines), and of the separating lines with ordinates $\frac{729}{40}$ and $\frac{27}{8}$ (dashed lines).

Item 1 in Corollary 1 has the following geometric interpretation: the $p-1$ horizontal lines in the plane with ordinates $\mu_j^{(p)},\ j = 1, \dots, p-1$, are 'separating lines' for the eigenvalues of $\mathbf{f}_p(\theta)$. This is illustrated in Figure 2 for the cases $p = 2, 3$. Item 1 in Corollary 2 has an analogous geometric interpretation.

## 6 Spectral analysis and spectral symbol

In this section we study the spectral properties of the stiffness matrix $A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ in (26), focusing on the asymptotic behavior as the fineness parameters $\boldsymbol{n} \to \infty$. In particular, we give estimates for the eigenvalues and for the spectral condition number $\kappa(A_{\boldsymbol{n}}^{(\boldsymbol{p})})$. Moreover, assuming $\boldsymbol{n} = \boldsymbol{\nu} n = (\nu_1 n, \dots, \nu_d n) \in \mathbb{N}^d$ for a fixed $\boldsymbol{\nu} \in \mathbb{Q}^d_{>0} := \{(\nu_1, \dots, \nu_d) \in \mathbb{Q}^d : \nu_1, \dots, \nu_d > 0\}$, we prove that the sequence $\{n^{d-2} A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ has an asymptotic spectral distribution characterized by the Hermitian matrix-valued function

$$\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta}) : [-\pi, \pi]^d \to \mathbb{C}^{D(\boldsymbol{p}) \times D(\boldsymbol{p})}$$

$$\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta}) := \sum_{k=1}^{d} c_k(\boldsymbol{\nu})\, \mathbf{h}_{p_1}(\theta_1) \otimes \cdots \otimes \mathbf{h}_{p_{k-1}}(\theta_{k-1}) \otimes \mathbf{f}_{p_k}(\theta_k) \otimes \mathbf{h}_{p_{k+1}}(\theta_{k+1}) \otimes \cdots \otimes \mathbf{h}_{p_d}(\theta_d), \tag{50}$$

where $D(\boldsymbol{p}) = \prod_{j=1}^d p_j$ (see Subsection 2.1), $\mathbf{f}_p$ and $\mathbf{h}_p$ are given in (37)–(38), and

$$c_k(\boldsymbol{\nu}) := \frac{\nu_k}{\nu_1 \cdots \nu_{k-1} \nu_{k+1} \cdots \nu_d}, \quad k = 1, \dots, d. \tag{51}$$

Unfortunately, it turns out that the spectrum of $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$ presents an exponential scattering with respect to $\boldsymbol{p}$ and $d$, and this implies a substantial numerical difficulty in treating the linear systems associated with the matrix $n^{d-2} A_{\boldsymbol{n}}^{(\boldsymbol{p})}$, already for moderate $\boldsymbol{p}$ and $d$. In the last subsection, still assuming that $\boldsymbol{n} = \boldsymbol{\nu} n$ for some $\boldsymbol{\nu} \in \mathbb{Q}^d_{>0}$, we investigate the clustering properties of the sequence $\{n^{d-2} A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ and we show that $\{n^{d-2} A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ is strongly clustered at $[0, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}]$, where $M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}} := \max_{\boldsymbol{\theta} \in [-\pi, \pi]^d} \lambda_{\max}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta}))$.

### 6.1 Estimates for the eigenvalues, localization of the spectrum and conditioning of $A_n^{(p)}$

Let us provide estimates for the eigenvalues of $K_n^{(p)}, M_n^{(p)}$. This is fundamental for estimating the condition number $\kappa(A_n^{(p)})$. By Theorem 7, the matrices $K_n^{(p)}, M_n^{(p)}$ are the leading principal submatrices of order $np-1$ of the Hermitian block Toeplitz matrices $T_n(\mathbf{f}_p), T_n(\mathbf{h}_p)$, respectively. Moreover, $T_{n-1}(\mathbf{f}_p), T_{n-1}(\mathbf{h}_p)$ are the leading principal submatrices of order $np-p$ of $K_n^{(p)}, M_n^{(p)}$, respectively. Hence, by Theorem 2 we have, for all $j$,

$$\lambda_j(T_n(\mathbf{f}_p)) \geq \lambda_j(K_n^{(p)}) \geq \lambda_{j+1}(T_n(\mathbf{f}_p)), \quad \lambda_j(K_n^{(p)}) \geq \lambda_j(T_{n-1}(\mathbf{f}_p)) \geq \lambda_{j+p-1}(K_n^{(p)}), \tag{52}$$

$$\lambda_j(T_n(\mathbf{h}_p)) \geq \lambda_j(M_n^{(p)}) \geq \lambda_{j+1}(T_n(\mathbf{h}_p)), \quad \lambda_j(M_n^{(p)}) \geq \lambda_j(T_{n-1}(\mathbf{h}_p)) \geq \lambda_{j+p-1}(M_n^{(p)}). \tag{53}$$

By (52)–(53) and by Theorem 5, recalling from Corollary 1 that $m_{\mathbf{f}_p} = 0$, we have

$$\sigma(K_n^{(p)}) \subset (0, M_{\mathbf{f}_p}], \qquad \sigma(M_n^{(p)}) \subset [m_{\mathbf{h}_p}, M_{\mathbf{h}_p}]. \tag{54}$$

Note that the point 0 is excluded from $\sigma(K_n^{(p)})$ either because $K_n^{(p)}$ is positive definite (see Section 3), or because, by Corollary 1, $\lambda_{\min}(\mathbf{f}_p(\theta))$ is not constant and so Theorem 5 excludes 0 from $\sigma(T_n(\mathbf{f}_p))$. Furthermore, Theorem 5 and (52)–(53) imply that, for each fixed $j \geq 1$, when $n \to \infty$ we have

$$\begin{aligned} &\lambda_j(K_n^{(p)}) \nearrow M_{\mathbf{f}_p}, \qquad \lambda_j(M_n^{(p)}) \nearrow M_{\mathbf{h}_p}, \\ &\lambda_{np-j}(K_n^{(p)}) \searrow 0, \qquad \lambda_{np-j}(M_n^{(p)}) \searrow m_{\mathbf{h}_p}, \end{aligned} \tag{55}$$

where the convergence is monotone by Theorem 2 and by the fact that $K_n^{(p)}$ (resp. $M_n^{(p)}$) is a leading principal submatrix of $K_{n+1}^{(p)}$ (resp. $M_{n+1}^{(p)}$) for every $n$. Relation (55) says that, for fixed $p$, the matrix $K_n^{(p)}$ is ill-conditioned for large $n$, while $M_n^{(p)}$ is not (recall that $m_{\mathbf{h}_p} > 0$ by Corollary 2). Theorem 10 allows us to understand 'how much' $K_n^{(p)}$ is ill-conditioned.

**Remark 5.** Using the same argument in [21, Theorem 8], it can be shown that $K_n^{(p)} \geq \frac{\pi^2}{n^2} M_n^{(p)}$ for all $p, n \geq 1$.

**Theorem 10.** *Let $p, n \geq 1$ and let $c_p > 0$ be a constant satisfying (49). Then the following properties hold.*

1. *We have*

$$\lambda_j(K_n^{(p)}) \geq \max\left(\frac{\pi^2}{n^2}\lambda_j(M_n^{(p)}), c_p\lambda_{j+1}(T_n(2-2\cos\theta)\otimes I_p)\right) \qquad \forall j = 1, \ldots, np-1, \tag{56}$$

$$\lambda_{\min}(K_n^{(p)}) \geq \max\left(\frac{\pi^2}{n^2}m_{\mathbf{h}_p}, 4c_p\sin^2\left(\frac{\pi}{2n+2}\right)\right) \overset{n\to\infty}{\sim} \frac{\pi^2\max(m_{\mathbf{h}_p}, c_p)}{n^2}. \tag{57}$$

2. *If $n \geq 3$, we have*

$$\lambda_{j+2}(C_n^{(p)}) \leq \lambda_j(K_n^{(p)}) \leq \lambda_{j-1}(C_n^{(p)}) \qquad \forall j = 1, \ldots, np-1, \tag{58}$$

$$\lambda_{\min}(K_n^{(p)}) \leq 4\sin^2\left(\frac{\pi}{n}\right) \overset{n\to\infty}{\sim} \frac{4\pi^2}{n^2}, \tag{59}$$

*where $C_n^{(p)}$ is the Hermitian block circulant matrix of order $np$ defined in the forthcoming proof, see (60).*

*Proof.* 1. By Remark 5 and by the minimax principle, $\lambda_j(K_n^{(p)}) \geq \frac{\pi^2}{n^2}\lambda_j(M_n^{(p)})$ for all $j = 1, \ldots, np-1$. Moreover, by (49), for all $\theta$ we have

$$\mathbf{f}_p(\theta) \geq c_p(2 - 2\cos\theta)I_p.$$

By Proposition 1, this implies that

$$T_n(\mathbf{f}_p) \geq T_n(c_p(2 - 2\cos\theta)I_p) = c_p T_n(2 - 2\cos\theta) \otimes I_p,$$

where the last equality follows from the definitions of tensor product and $T_n(c_p(2 - 2\cos\theta)I_p)$. By the minimax principle, see (8), we infer

$$\lambda_j(T_n(\mathbf{f}_p)) \geq c_p\lambda_j(T_n(2 - 2\cos\theta) \otimes I_p) \qquad \forall j = 1, \ldots, np$$

and consequently, by (52),

$$\lambda_j(K_n^{(p)}) \geq c_p\lambda_{j+1}(T_n(2 - 2\cos\theta) \otimes I_p) \qquad \forall j = 1, \ldots, np-1.$$

This completes the proof of (56). Relation (57) is obtained from (56) by setting $j = np-1$. To see this, note that $\lambda_{\min}(M_n^{(p)}) \geq m_{\mathbf{h}_p}$ by (54); moreover,

$$\lambda_{\min}(T_n(2 - 2\cos\theta) \otimes I_p) = \lambda_{\min}(T_n(2 - 2\cos\theta)) = 4\sin^2\left(\frac{\pi}{2n+2}\right),$$

where the last equality holds because the eigenvalues of $T_n(2 - 2\cos\theta)$ are known and, in particular, the minimal eigenvalue equals $2 - 2\cos\frac{\pi}{n+1} = 4\sin^2\left(\frac{\pi}{2n+2}\right)$.

2. Let $n \geq 3$. With the notation of Theorem 7, we have

$$T_n(\mathbf{f}_p) = \begin{bmatrix} K_0 & K_1^T & & \\ K_1 & \ddots & \ddots & \\ & \ddots & \ddots & K_1^T \\ & & K_1 & K_0 \end{bmatrix} = \begin{bmatrix} K_0 & K_1^T & & K_1 \\ K_1 & \ddots & \ddots & \\ & \ddots & \ddots & K_1^T \\ K_1^T & & K_1 & K_0 \end{bmatrix} - \begin{bmatrix} & & & K_1 \\ & & & \\ & & & \\ K_1^T & & & \end{bmatrix} =: C_n^{(p)} - E_n^{(p)}, \qquad (60)$$

where $C_n^{(p)}$ is a block circulant matrix, while $E_n^{(p)}$ is Hermitian with $\operatorname{rank}(E_n^{(p)}) = 2$. The latter is true because $\operatorname{rank}(K_1) = 1$.[2] Therefore, $E_n^{(p)}$ has exactly two nonzero eigenvalues $\lambda, \mu$, which are one the opposite of the other because $\lambda + \mu = \operatorname{trace}(E_n^{(p)}) = 0$. Thus, we can apply Theorem 3 with $k^+ = k^- = 1$ and we obtain

$$\lambda_{j-1}(C_n^{(p)}) \geq \lambda_j(T_n(\mathbf{f}_p)) \geq \lambda_{j+1}(C_n^{(p)}) \qquad \forall j = 1, \ldots, np. \qquad (61)$$

The inequalities (58) follow from (61),(52). To obtain (59), note that the spectral decomposition of $C_n^{(p)}$ is known (Theorem 6) and, when applying Theorem 6 to $C_n^{(p)}$, the function $\mathbf{g}$ in (16) satisfies $\mathbf{g}\left(\frac{2\pi j}{n}\right) = \mathbf{f}_p\left(\frac{2\pi j}{n}\right)$ for all $j = 0, \ldots, n-1$. Moreover, by Corollary 1, $\lambda_{\min}(\mathbf{f}_p(\theta)) \leq 2 - 2\cos\theta$ for all $\theta$, and $\lambda_{\min}(\mathbf{f}_p(\theta))$ is 'well-separated' from the other eigenvalue functions $\lambda_j(\mathbf{f}_p(\theta))$, $j = 1, \ldots, p-1$, by the separating line $\mu_{p-1}^{(p)}$. Hence, for $j = np-1$, from (58) we obtain

$$\begin{aligned} \lambda_{\min}(K_n^{(p)}) &\leq \lambda_{np-2}(C_n^{(p)}) = \text{the third smallest number in the set } \left\{\lambda_{\min}(\mathbf{f}_p(\tfrac{2\pi j}{n}))\right\}_{j=0,\ldots,n-1} \\ &\leq \text{the third smallest number in the set } \left\{2 - 2\cos\tfrac{2\pi j}{n}\right\}_{j=0,\ldots,n-1} \\ &= 2 - 2\cos\tfrac{2\pi}{n} = 4\sin^2(\tfrac{\pi}{n}). \end{aligned}$$

□

[2] Note that $K_1 \neq O$, otherwise we would have $\langle L_0', L_1'\rangle = \ldots = \langle L_0', L_p'\rangle = 0$, implying $\langle L_0', L_1' + \ldots + L_p'\rangle = 0$ and, by Lemma 8, $-\langle L_0', L_0'\rangle = 0$: this is impossible, because $L_0'$ is not identically 0.

| $n$ | $2n^2\lambda_{2n-1}(K_n^{(2)})$ | $3n^2\lambda_{3n-1}(K_n^{(3)})$ | $2n^2\lambda_{2n-2}(K_n^{(2)})$ | $3n^2\lambda_{3n-2}(K_n^{(3)})$ | $2n^2\lambda_{2n-3}(K_n^{(2)})$ | $3n^2\lambda_{3n-3}(K_n^{(3)})$ |
|---|---|---|---|---|---|---|
| 20 | 9.8683332 | 9.8693541 | 39.4579402 | 39.4744220 | 88.7216045 | 88.8062922 |
| 40 | 9.8692871 | 9.8695418 | 39.4733327 | 39.4774163 | 88.8006247 | 88.8213755 |
| 80 | 9.8695251 | 9.8695887 | 39.4771485 | 39.4781671 | 88.8200104 | 88.8251718 |
| 160 | 9.8695846 | 9.8696005 | 39.4781005 | 39.4783550 | 88.8248338 | 88.8261226 |
| 320 | 9.8695994 | 9.8696034 | 39.4783383 | 39.4784019 | 88.8260383 | 88.8263603 |
| 640 | 9.8696032 | 9.8696042 | 39.4783978 | 39.4784137 | 88.8263393 | 88.8264198 |

**Table 1:** computation of $pn^2\lambda_{pn-j}(K_n^{(p)})$ for $p = 2, 3$, for $j = 1, 2, 3$ and for increasing values of $n$.

**Remark 6.** The argument used for proving (61) can be generalized to the case where $\mathbf{f}_p$ is replaced by any Hermitian matrix-valued trigonometric polynomial. To be precise, assume that $\mathbf{q}(\theta) = \sum_{k=-m}^{m} \mathbf{q}_k \mathrm{e}^{\mathrm{i}k\theta} : [-\pi, \pi] \to \mathbb{C}^{p\times p}$ is a Hermitian matrix-valued trigonometric polynomial. Then $\mathbf{q}_{-j} = \mathbf{q}_j^*$ for every $j = 0, \dots, m$ and $T_n(\mathbf{q})$ is Hermitian for all $n \geq 1$ (see Section 2). For every $n \geq 2m + 1$ we can write $T_n(\mathbf{q}) = C_n - E_n$, where $C_n := T_n(\mathbf{q}) + E_n$ is a block circulant matrix and the matrix $E_n$, given by

$$E_n := \begin{bmatrix} O & O & B \\ O & O & O \\ B^* & O & O \end{bmatrix}, \qquad B := \begin{bmatrix} \mathbf{q}_m & \cdots & \mathbf{q}_1 \\ & \ddots & \vdots \\ & & \mathbf{q}_m \end{bmatrix},$$

is Hermitian with $\mathrm{rank}(E_n) \leq 2mp$. It can be shown that the nonzero eigenvalues of $E_n$ coincide with the nonzero singular values of $B$ together with their negatives, see [5, p. 35]. Hence, $E_n$ has the same number of positive and negative eigenvalues and, by Theorem 3, we get

$$\lambda_{j-mp}(C_n) \geq \lambda_j(T_n(\mathbf{q})) \geq \lambda_{j+mp}(C_n), \qquad \forall j = 1, \dots, np.$$

Notice also that the spectral decomposition of $C_n$ for $n \geq 2m + 1$ is given by (16) with

$$\mathbf{g}(\theta) = \sum_{k=0}^{m} \mathbf{q}_k \mathrm{e}^{\mathrm{i}k\theta} + \sum_{k=n-m}^{n-1} \mathbf{q}_{k-n} \mathrm{e}^{\mathrm{i}k\theta} = \sum_{k=0}^{m} \mathbf{q}_k \mathrm{e}^{\mathrm{i}k\theta} + \sum_{\ell=-m}^{-1} \mathbf{q}_\ell \mathrm{e}^{\mathrm{i}(\ell+n)\theta} = \sum_{k=0}^{m} \mathbf{q}_k \mathrm{e}^{\mathrm{i}k\theta} + \mathrm{e}^{\mathrm{i}n\theta} \sum_{\ell=-m}^{-1} \mathbf{q}_\ell \mathrm{e}^{\mathrm{i}\ell\theta},$$

hence $\mathbf{g}\left(\frac{2\pi j}{n}\right) = \mathbf{q}\left(\frac{2\pi j}{n}\right)$ for every $j = 0, \dots, n - 1$.

Table 1 shows the results of some numerical experiments, which confirm that $\lambda_{\min}(K_n^{(p)})$ goes to 0 as $1/n^2$ when $n \to \infty$, in accordance with Theorem 10. They also show, at least for $p = 2, 3$ and $j = 1, 2, 3$, that

$$\lim_{n\to\infty} pn^2\lambda_{np-j}(K_n^{(p)}) = j^2\pi^2. \tag{62}$$

This limit relation, which we guess to hold for every $p \geq 1$ and every $j \geq 1$, is analogous to the one in [21, Eq. (55)] and has the same motivation given there.[3] In particular, the numbers $j^2\pi^2$, $j = 1, 2, \dots$, are the eigenvalues of the 1D boundary value problem

$$\begin{cases} -u''(x) = f(x) & \text{on } (0, 1), \\ u(0) = u(1) = 0, \end{cases}$$

with corresponding eigenvectors $\sin(j\pi x)$, $j = 1, 2, \dots$ Note that (62) certainly holds for $p = 1$ and $j \geq 1$, since $K_n^{(1)} = T_{n-1}(2 - 2\cos\theta)$ and it is known that $\lambda_j(K_n^{(1)}) = 2 - 2\cos\frac{(n-j)\pi}{n}$, $j = 1, \dots, n - 1$.

[3] Note that, contrary to (62), in [21, Eq. (55)] the eigenvalues are arranged in increasing order.

We now provide a localization of the spectrum of $A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ and an estimate of its condition number under the assumption that $\boldsymbol{\beta} \in \mathbb{R}^d$ is constant. In this case, the advection matrix $A_{\boldsymbol{n},A}^{(\boldsymbol{p})}$ in (28) is skew-symmetric and, consequently, the real and imaginary parts of $A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ are explicitly given by

$$\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) = A_{\boldsymbol{n},D}^{(\boldsymbol{p})} + A_{\boldsymbol{n},R}^{(\boldsymbol{p})}, \tag{63}$$

$$\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) = -\mathrm{i}A_{\boldsymbol{n},A}^{(\boldsymbol{p})}. \tag{64}$$

Note that, from (63), (30), (32)–(33), we obtain

$$\begin{aligned}\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) \geq \sum_{k=1}^{d} \frac{1}{n_1} M_{n_1}^{(p_1)} \otimes \cdots \otimes \frac{1}{n_{k-1}} M_{n_{k-1}}^{(p_{k-1})} \otimes n_k K_{n_k}^{(p_k)} \otimes \frac{1}{n_{k+1}} M_{n_{k+1}}^{(p_{k+1})} \otimes \cdots \otimes \frac{1}{n_d} M_{n_d}^{(p_d)} \\ + \gamma_* \frac{1}{n_1} M_{n_1} \otimes \cdots \otimes \frac{1}{n_d} M_{n_d},\end{aligned} \tag{65}$$

$$\begin{aligned}\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) \leq \sum_{k=1}^{d} \frac{1}{n_1} M_{n_1}^{(p_1)} \otimes \cdots \otimes \frac{1}{n_{k-1}} M_{n_{k-1}}^{(p_{k-1})} \otimes n_k K_{n_k}^{(p_k)} \otimes \frac{1}{n_{k+1}} M_{n_{k+1}}^{(p_{k+1})} \otimes \cdots \otimes \frac{1}{n_d} M_{n_d}^{(p_d)} \\ + \|\gamma\|_{L^\infty(\Omega)} \frac{1}{n_1} M_{n_1} \otimes \cdots \otimes \frac{1}{n_d} M_{n_d}.\end{aligned} \tag{66}$$

In particular, (12) combined with Remark 5 yields $\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) \geq \dfrac{\pi^2 d + \gamma_*}{n_1 \cdots n_d} M_{n_1}^{(p_1)} \otimes \cdots \otimes M_{n_d}^{(p_d)}$.

**Theorem 11** (**localization of the spectrum of** $\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})$)**.** *Assume that $\boldsymbol{\beta} \in \mathbb{R}^d$ is constant and, for $p, n \geq 1$, define $\zeta_{n,p} := \max\left(\pi^2, \dfrac{4c_p}{m_{\mathbf{h}_p}} n^2 \sin\left(\dfrac{\pi}{2n+2}\right)\right)$, where $c_p > 0$ is a constant satisfying* (49)*. Then, for every $\boldsymbol{n}, \boldsymbol{p} \in \mathbb{N}^d$,*

$$\lambda_{\min}\left(\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})\right) \geq \frac{\sum_{k=1}^{d} \zeta_{n_k,p_k} + \gamma_*}{n_1 \cdots n_d} G_{\boldsymbol{p}} \geq \frac{\pi^2 d + \gamma_*}{n_1 \cdots n_d} G_{\boldsymbol{p}}, \tag{67}$$

$$\lambda_{\max}(\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})) \leq \frac{\sum_{k=1}^{d} n_k^2 (M_{\mathbf{f}_{p_k}}/M_{\mathbf{h}_{p_k}}) + \|\gamma\|_{L^\infty(\Omega)}}{n_1 \cdots n_d} S_{\boldsymbol{p}}, \tag{68}$$

*where $G_{\boldsymbol{p}} := m_{\mathbf{h}_{p_1}} \cdots m_{\mathbf{h}_{p_d}}$ and $S_{\boldsymbol{p}} := M_{\mathbf{h}_{p_1}} \cdots M_{\mathbf{h}_{p_d}}$.*

*Proof.* Apply (6),(11),(54),(57) in (65) to obtain (67). Then, apply (7),(11),(54) in (66) to obtain (68). □

**Theorem 12** (**localization of the spectrum of** $A_{\boldsymbol{n}}^{(\boldsymbol{p})}$)**.** *Assume that $\boldsymbol{\beta} \in \mathbb{R}^d$ is constant and, for $p, n \geq 1$, let $\zeta_{n,p}$ be as in Theorem 11. Then, for every $\boldsymbol{n}, \boldsymbol{p} \in \mathbb{N}^d$,*

$$\begin{aligned}&\sigma(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) \subseteq [\lambda_{\min}(\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})), \lambda_{\max}(\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})}))] \times [\lambda_{\min}(\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})), \lambda_{\max}(\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})}))] \\ &\subseteq \left[\frac{\sum_{k=1}^{d} \zeta_{n_k,p_k} + \gamma_*}{n_1 \cdots n_d} G_{\boldsymbol{p}}, \frac{\sum_{k=1}^{d} n_k^2 (M_{\mathbf{f}_{p_k}}/M_{\mathbf{h}_{p_k}}) + \|\gamma\|_{L^\infty(\Omega)}}{n_1 \cdots n_d} S_{\boldsymbol{p}}\right] \times \left[-B_{\boldsymbol{p}} \|\boldsymbol{\beta}\|_\infty \frac{\sum_{k=1}^{d} n_k}{n_1 \cdots n_d}, B_{\boldsymbol{p}} \|\boldsymbol{\beta}\|_\infty \frac{\sum_{k=1}^{d} n_k}{n_1 \cdots n_d}\right],\end{aligned}$$

*where $G_{\boldsymbol{p}} := m_{\mathbf{h}_{p_1}} \cdots m_{\mathbf{h}_{p_d}}$, $S_{\boldsymbol{p}} := M_{\mathbf{h}_{p_1}} \cdots M_{\mathbf{h}_{p_d}}$, and $B_{\boldsymbol{p}}$ is a constant satisfying* (29)*.*

*Proof.* From Theorem 11 we have

$$\frac{\sum_{k=1}^{d} \zeta_{n_k,p_k} + \gamma_*}{n_1 \cdots n_d} G_{\boldsymbol{p}} \leq \lambda_{\min}(\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})) \leq \lambda_{\max}(\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})) \leq \frac{\sum_{k=1}^{d} n_k^2 (M_{\mathbf{f}_{p_k}}/M_{\mathbf{h}_{p_k}}) + \|\gamma\|_{L^\infty(\Omega)}}{n_1 \cdots n_d} S_{\boldsymbol{p}},$$

| $n$ | $\kappa(A_{\boldsymbol{n}}^{(\boldsymbol{p})})/n^2$ | $\kappa(A_{\boldsymbol{n}})/n^2$ |
|---|---|---|
| 8 | 1.2597 | 0.2278 |
| 16 | 1.2573 | 0.2173 |
| 32 | 1.2553 | 0.2101 |
| 64 | 1.2543 | 0.2065 |
| 128 | 1.2539 | 0.2049 |
| 256 | 1.2538 | 0.2041 |

**Table 2:** computation of $\kappa(A_{\boldsymbol{n}}^{(\boldsymbol{p})})/n^2$ and $\kappa(A_{\boldsymbol{n}})/n^2$ in the case $d = 2$, $\boldsymbol{\beta} = \mathbf{0}$, $\gamma = 0$, $\boldsymbol{p} = (2, 2)$, $\boldsymbol{n} = (n, \log_2 n)$, for increasing values of $n$. Note that we are in the presence of a non-uniform mesh refinement.

and from Lemma 7, combined with (64) and with the fact that $\boldsymbol{\beta}$ is constant, we have

$$-B_{\boldsymbol{p}}\|\boldsymbol{\beta}\|_\infty \frac{\sum_{k=1}^d n_k}{n_1 \cdots n_d} \leq -\|\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})\| \leq \lambda_{\min}(\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})) \leq \lambda_{\max}(\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})) \leq \|\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})\| \leq B_{\boldsymbol{p}}\|\boldsymbol{\beta}\|_\infty \frac{\sum_{k=1}^d n_k}{n_1 \cdots n_d}.$$

The thesis follows from (5). □

**Theorem 13** (**conditioning**). *Assume that $\boldsymbol{\beta}$ is constant. Then, for every $\boldsymbol{p} \in \mathbb{N}^d$ there exists a constant $\alpha_{\boldsymbol{p}}$ such that, for all $\boldsymbol{n} \in \mathbb{N}^d$,*

$$\kappa(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) \leq \alpha_{\boldsymbol{p}} \sum_{k=1}^d n_k^2. \tag{69}$$

*Proof.* From $A_{\boldsymbol{n}}^{(\boldsymbol{p})} = \Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) + \mathrm{i}\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})$ and from the fact that $\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})$, $\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})$ are Hermitian, we have

$$s_{\max}(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) = \|A_{\boldsymbol{n}}^{(\boldsymbol{p})}\| \leq \|\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})\| + \|\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})\| = \rho(\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})) + \rho(\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})).$$

Hence, by Theorem 12 we see that

$$\|A_{\boldsymbol{n}}^{(\boldsymbol{p})}\| \leq \hat{\alpha}_{\boldsymbol{p}} \frac{\sum_{k=1}^d n_k^2}{n_1 \cdots n_d},$$

for some constant $\hat{\alpha}_{\boldsymbol{p}}$ independent of $\boldsymbol{n}$. Furthermore, by Theorem 11 and by the Fan-Hoffman theorem,

$$s_{\min}\left(A_{\boldsymbol{n}}^{(\boldsymbol{p})}\right) \geq \lambda_{\min}\left(\Re(A_{\boldsymbol{n}}^{(\boldsymbol{p})})\right) \geq \frac{\tilde{\alpha}_{\boldsymbol{p}}}{n_1 \cdots n_d},$$

for some constant $\tilde{\alpha}_{\boldsymbol{p}} > 0$ independent of $\boldsymbol{n}$. Thus, $\kappa(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) = \frac{s_{\max}(A_{\boldsymbol{n}}^{(\boldsymbol{p})})}{s_{\min}(A_{\boldsymbol{n}}^{(\boldsymbol{p})})} \leq \alpha_{\boldsymbol{p}} \sum_{k=1}^d n_k^2$, with $\alpha_{\boldsymbol{p}} = \hat{\alpha}_{\boldsymbol{p}}/\tilde{\alpha}_{\boldsymbol{p}}$. □

(69) says that $\kappa(A_{\boldsymbol{n}}^{(\boldsymbol{p})})$ is bounded from above by $\max(\boldsymbol{n}^2) = \max(n_1^2, \ldots, n_d^2)$ multiplied by some constant independent of $\boldsymbol{n}$ (for instance $\alpha_{\boldsymbol{p}} d$). This upper bound is the sharpest possible, as shown by the numerical experiments in Table 2, where we fixed $d = 2$, $\boldsymbol{\beta} = \mathbf{0}$, $\gamma = 0$, $\boldsymbol{p} = (2, 2)$, and we computed $\kappa(A_{\boldsymbol{n}}^{(\boldsymbol{p})}) = \kappa(A_{\boldsymbol{n},D}^{(\boldsymbol{p})})$ (normalized by $n^2$) for $\boldsymbol{n} = (n, \log_2 n)$ and for increasing values of $n$. For a nice comparison with Finite Differences (FD), in the third column of Table 2 we reported the values of $\kappa(A_{\boldsymbol{n}})/n^2$ for $d = 2$, where

$$A_{\boldsymbol{n}} := \sum_{k=1}^d n_k^2 \, I_{n_1-1} \otimes \cdots \otimes I_{n_{k-1}-1} \otimes T_{n_k-1}(2 - 2\cos\theta) \otimes I_{n_{k+1}-1} \otimes \cdots \otimes I_{n_d-1}$$

is the (diffusion) matrix coming from the standard centered FD approximation of (1) on the mesh $\boldsymbol{j}/\boldsymbol{n}$, $\boldsymbol{j} = \mathbf{0}, \ldots, \boldsymbol{n}$, in the case $\boldsymbol{\beta} = \mathbf{0}$, $\gamma = 0$.

### 6.2 Spectral distribution and symbol of the normalized sequence $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$

In this subsection we assume that $n_j = \nu_j n$ for all $j = 1, \dots, d$, i.e. $\boldsymbol{n} = \boldsymbol{\nu} n = (\nu_1 n, \dots, \nu_d n) \in \mathbb{N}^d$, where $\boldsymbol{\nu} \in \mathbb{Q}^d_{>0}$ is fixed and $n$ varies in the set of natural numbers such that $\boldsymbol{n} \in \mathbb{N}^d$. Under this assumption, from (27), (32) we have

$$\begin{aligned} n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})} &= n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})} + n^{d-2}A_{\boldsymbol{n},A}^{(\boldsymbol{p})} + n^{d-2}A_{\boldsymbol{n},R}^{(\boldsymbol{p})} \\ &= \sum_{k=1}^{d} c_k(\boldsymbol{\nu})\, M_{n_1}^{(p_1)} \otimes \cdots \otimes M_{n_{k-1}}^{(p_{k-1})} \otimes K_{n_k}^{(p_k)} \otimes M_{n_{k+1}}^{(p_{k+1})} \otimes \cdots \otimes M_{n_d}^{(p_d)} + n^{d-2}A_{\boldsymbol{n},A}^{(\boldsymbol{p})} + n^{d-2}A_{\boldsymbol{n},R}^{(\boldsymbol{p})}, \quad (70) \end{aligned}$$

where the $c_k(\boldsymbol{\nu})$, $k = 1, \dots, d$, are given in (51). Recall from (26) that $A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ is of size $(n_1p_1 - 1)\cdots(n_dp_d - 1)$.

In Theorem 14 we prove that the sequence of matrices $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ in (70) is distributed, in the sense of the eigenvalues, like the Hermitian matrix-valued function $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$ in (50), which is therefore the symbol of the sequence $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$. Note that $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ is really a sequence of matrices, due to the assumption $\boldsymbol{n} = \boldsymbol{\nu} n$. This assumption must be kept in mind while reading this subsection.

Before stating and proving Theorem 14, let us observe that, by the properties of $\mathbf{f}_p(\theta)$ and $\mathbf{h}_p(\theta)$, see Corollaries 1–2, and by the properties of tensor products, see Subsection 2.4, $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta}) \geq O$ for all $\boldsymbol{\theta} \in [-\pi, \pi]^d$ and $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta}) > O$ for all $\boldsymbol{\theta} \in [-\pi, \pi]^d \backslash \{\mathbf{0}\}$.

**Theorem 14.** *Let $\boldsymbol{p} \in \mathbb{N}^d$, $\boldsymbol{\nu} \in \mathbb{Q}^d_{>0}$ and $\boldsymbol{n} = \boldsymbol{\nu} n$, then $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n \sim_\lambda \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$. In particular, $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ is weakly clustered at the essential range of $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$ (see the discussion at the end of Subsection 2.3).*

*Proof.* For all $p, n \geq 1$, define the following matrices, of size $np$:

$$\widetilde{K}_n^{(p)} := K_n^{(p)} \oplus [0], \qquad \widetilde{M}_n^{(p)} := M_n^{(p)} \oplus [0].$$

Let $n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})}$ be the matrix of size $n_1p_1 \cdots n_dp_d = (\nu_1p_1 \cdots \nu_dp_d)n^d$ obtained from $n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}$ by replacing $K, M$ with $\widetilde{K}, \widetilde{M}$, i.e.

$$n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})} = \sum_{k=1}^{d} c_k(\boldsymbol{\nu})\, \widetilde{M}_{n_1}^{(p_1)} \otimes \cdots \otimes \widetilde{M}_{n_{k-1}}^{(p_{k-1})} \otimes \widetilde{K}_{n_k}^{(p_k)} \otimes \widetilde{M}_{n_{k+1}}^{(p_{k+1})} \otimes \cdots \otimes \widetilde{M}_{n_d}^{(p_d)}.$$

By Lemma 5, there exists the permutation matrix $P_{\boldsymbol{n},\boldsymbol{p}} := P_{n_1p_1-1,1,n_2p_2-1,1,\dots,n_dp_d-1,1}$, depending only on $\boldsymbol{n}, \boldsymbol{p}$, such that

$$n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})} = P_{\boldsymbol{n},\boldsymbol{p}}[(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) \oplus O]P_{\boldsymbol{n},\boldsymbol{p}}^T,$$

where $O$ is the zero matrix of order $n_1p_1 \cdots n_dp_d - (n_1p_1 - 1)\cdots(n_dp_d - 1) = o(n^d)$. Hence,

$$\begin{aligned} n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})} &:= P_{\boldsymbol{n},\boldsymbol{p}}\left[(n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}) \oplus O\right] P_{\boldsymbol{n},\boldsymbol{p}}^T = P_{\boldsymbol{n},\boldsymbol{p}}[n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})} \oplus O + n^{d-2}A_{\boldsymbol{n},A}^{(\boldsymbol{p})} \oplus O + n^{d-2}A_{\boldsymbol{n},R}^{(\boldsymbol{p})} \oplus O]P_{\boldsymbol{n},\boldsymbol{p}}^T \\ &= n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})} + n^{d-2}\widetilde{A}_{\boldsymbol{n},A}^{(\boldsymbol{p})} + n^{d-2}\widetilde{A}_{\boldsymbol{n},R}^{(\boldsymbol{p})}, \end{aligned}$$

where $n^{d-2}\widetilde{A}_{\boldsymbol{n},A}^{(\boldsymbol{p})} := P_{\boldsymbol{n},\boldsymbol{p}}[(n^{d-2}A_{\boldsymbol{n},A}^{(\boldsymbol{p})}) \oplus O]P_{\boldsymbol{n},\boldsymbol{p}}^T$ and $n^{d-2}\widetilde{A}_{\boldsymbol{n},R}^{(\boldsymbol{p})} := P_{\boldsymbol{n},\boldsymbol{p}}[(n^{d-2}A_{\boldsymbol{n},R}^{(\boldsymbol{p})}) \oplus O]P_{\boldsymbol{n},\boldsymbol{p}}^T$. The eigenvalues of $n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})}$ are those of $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ with only $o(n^d)$ extra eigenvalues equal to 0. Consequently, by Definition 1, if we prove that $\{n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n \sim_\lambda \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$ then $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n \sim_\lambda \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$.

Now, let

$$T_{\boldsymbol{n}}^{(\boldsymbol{p})} := \sum_{k=1}^{d} c_k(\boldsymbol{\nu})\, T_{n_1}(\mathbf{h}_{p_1}) \otimes \cdots \otimes T_{n_{k-1}}(\mathbf{h}_{p_{k-1}}) \otimes T_{n_k}(\mathbf{f}_{p_k}) \otimes T_{n_{k+1}}(\mathbf{h}_{p_{k+1}}) \otimes \cdots \otimes T_{n_d}(\mathbf{h}_{p_d}). \quad (71)$$

To show that $\{n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n \sim_\lambda \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$, we prove that the hypotheses of Theorem 4 are satisfied with $X_n := T_{\boldsymbol{n}}^{(\boldsymbol{p})}$, $Y_n := n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})} - T_{\boldsymbol{n}}^{(\boldsymbol{p})}$ and $\mathbf{f} = \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$.

Note that each $T_{\boldsymbol{n}}^{(\boldsymbol{p})}$ is Hermitian because $\mathbf{f}_p$, $\mathbf{h}_p$ are Hermitian matrix-valued functions for all $p \geq 1$. By Lemma 6, $T_{\boldsymbol{n}}^{(\boldsymbol{p})}$ is also similar to $T_{\boldsymbol{n}}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})$, and, by Theorem 5, $\{T_{\boldsymbol{n}}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})\}_n \sim_\lambda \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$, implying $\{T_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n \sim_\lambda \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$. Now observe that, since $K_n^{(p)}$, $M_n^{(p)}$, $\widetilde{K}_n^{(p)}$, $\widetilde{M}_n^{(p)}$, $T_n(\mathbf{f}_p)$, $T_n(\mathbf{h}_p)$ are normal for all $p, n \geq 1$, we have

$$\|\widetilde{K}_n^{(p)}\| = \rho(\widetilde{K}_n^{(p)}) = \rho(K_n^{(p)}) = \|K_n^{(p)}\| \leq M_{\mathbf{f}_p}, \qquad \|T_n(\mathbf{f}_p)\| = \rho(T_n(\mathbf{f}_p)) \leq M_{\mathbf{f}_p}, \tag{72}$$

$$\|\widetilde{M}_n^{(p)}\| = \rho(\widetilde{M}_n^{(p)}) = \rho(M_n^{(p)}) = \|M_n^{(p)}\| \leq M_{\mathbf{h}_p}, \qquad \|T_n(\mathbf{h}_p)\| = \rho(T_n(\mathbf{h}_p)) \leq M_{\mathbf{h}_p}. \tag{73}$$

From (72)–(73), from the triangle inequality, and from (10), it follows that the norms $\|T_{\boldsymbol{n}}^{(\boldsymbol{p})}\|$, $\|n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})}\| = \|n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}\|$ are bounded from above by some constant independent of $n$. Moreover, from Lemma 7, (31), (33), (73), (10), we have

$$\|n^{d-2}\widetilde{A}_{\boldsymbol{n},A}^{(\boldsymbol{p})}\| = \|n^{d-2}A_{\boldsymbol{n},A}^{(\boldsymbol{p})}\| \leq \frac{n^{d-2}B_{\boldsymbol{p}}\|\boldsymbol{\beta}\|_{L^\infty(\Omega)}\sum_{k=1}^d n_k}{n_1\cdots n_d} = \frac{B_{\boldsymbol{p}}\|\boldsymbol{\beta}\|_{L^\infty(\Omega)}\sum_{k=1}^d \nu_k}{\nu_1\cdots\nu_d n}, \tag{74}$$

$$\|n^{d-2}\widetilde{A}_{\boldsymbol{n},R}^{(\boldsymbol{p})}\| = \|n^{d-2}A_{\boldsymbol{n},R}^{(\boldsymbol{p})}\| \leq \frac{n^{d-2}\|\gamma\|_{L^\infty(\Omega)}S_{\boldsymbol{p}}}{n_1\cdots n_d} = \frac{\|\gamma\|_{L^\infty(\Omega)}S_{\boldsymbol{p}}}{\nu_1\cdots\nu_d n^2}, \tag{75}$$

where $S_{\boldsymbol{p}} := M_{\mathbf{h}_{p_1}}\cdots M_{\mathbf{h}_{p_d}}$. Therefore, taking into account the triangle inequality

$$\|n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})}\| \leq \|n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})}\| + \|n^{d-2}\widetilde{A}_{\boldsymbol{n},A}^{(\boldsymbol{p})}\| + \|n^{d-2}\widetilde{A}_{\boldsymbol{n},R}^{(\boldsymbol{p})}\|,$$

we conclude that $\|n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})}\|$ is bounded from above by some constant independent of $n$. Hence,

$$\|T_{\boldsymbol{n}}^{(\boldsymbol{p})}\|,\ \|n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})}\|,\ \|n^{d-2}\widetilde{A}_{\boldsymbol{n},A}^{(\boldsymbol{p})}\|,\ \|n^{d-2}\widetilde{A}_{\boldsymbol{n},R}^{(\boldsymbol{p})}\|,\ \|n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})}\|,\ \|n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})} - T_{\boldsymbol{n}}^{(\boldsymbol{p})}\| \leq C, \tag{76}$$

for some $C$ independent of $n$. To finish the proof, we have to show that $\|n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})} - T_{\boldsymbol{n}}^{(\boldsymbol{p})}\|_1 = o(n^d)$ as $n \to \infty$. Note that, for all $p, n \geq 1$,

$$\operatorname{rank}(\widetilde{K}_n^{(p)} - T_n(\mathbf{f}_p)) \leq 2, \qquad \operatorname{rank}(\widetilde{M}_n^{(p)} - T_n(\mathbf{h}_p)) \leq 2.$$

Therefore, by (4) and by the property (13) of tensor products we infer

$$\begin{aligned}
\left\|n^{d-2}\widetilde{A}_{\boldsymbol{n}}^{(\boldsymbol{p})} - T_{\boldsymbol{n}}^{(\boldsymbol{p})}\right\|_1 &\leq \left\|n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})} - T_{\boldsymbol{n}}^{(\boldsymbol{p})}\right\|_1 + \left\|n^{d-2}\widetilde{A}_{\boldsymbol{n},A}^{(\boldsymbol{p})}\right\|_1 + \left\|n^{d-2}\widetilde{A}_{\boldsymbol{n},R}^{(\boldsymbol{p})}\right\|_1 \\
&\leq \left(d\sum_{i=1}^d 2n_1p_1\cdots n_{i-1}p_{i-1}n_{i+1}p_{i+1}\cdots n_dp_d\right)\left\|n^{d-2}\widetilde{A}_{\boldsymbol{n},D}^{(\boldsymbol{p})} - T_{\boldsymbol{n}}^{(\boldsymbol{p})}\right\| \\
&\qquad + n_1p_1\cdots n_dp_d\left\|n^{d-2}\widetilde{A}_{\boldsymbol{n},A}^{(\boldsymbol{p})}\right\| + n_1p_1\cdots n_dp_d\left\|n^{d-2}\widetilde{A}_{\boldsymbol{n},R}^{(\boldsymbol{p})}\right\|,
\end{aligned}$$

and the latter is $o(n^d)$, thanks to (74)–(76). □

### 6.3 Exponential scattering and ill-conditioning of the symbol

The discussion on the exponential ill-conditioning of the symbol contained in this subsection is based on the informal meaning behind the definition of spectral distribution. According to Remark 1, the spectral information contained in the symbol $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$ can be summarized as follows: *the eigenvalues of* $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ *are*

| $p$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| $\phi_{p,1}^{(1)}$ | 1.33 | 5.78 | $1.84 \cdot 10$ | $5.45 \cdot 10$ | $1.59 \cdot 10^2$ | $4.84 \cdot 10^2$ | $1.54 \cdot 10^3$ | $5.12 \cdot 10^3$ | $1.77 \cdot 10^4$ |
| $\left[\phi_{p,1}^{(1)}\right]^{1/p}$ | 1.15 | 1.79 | 2.07 | 2.22 | 2.33 | 2.42 | 2.50 | 2.58 | 2.66 |

**Table 3:** computation of $\phi_{\boldsymbol{p},d}^{(\boldsymbol{\nu})}$ and $\left[\phi_{\boldsymbol{p},d}^{(\boldsymbol{\nu})}\right]^{1/D(\boldsymbol{p})}$ in the case $d = 1$, $\boldsymbol{p} = p$, $\boldsymbol{\nu} = 1$, for $p = 2, \ldots, 10$. Note that in this case $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta})$ is nothing else than $\mathbf{f}_p(\theta)$.

| $p$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| $\phi_{(p,p),2}^{(1,1)}$ | 2.13 | 9.72 | $3.44 \cdot 10$ | $1.54 \cdot 10^2$ | $7.47 \cdot 10^2$ | $4.39 \cdot 10^3$ | $3.01 \cdot 10^4$ | $2.42 \cdot 10^5$ | $2.17 \cdot 10^6$ |
| $\left[\phi_{(p,p),2}^{(1,1)}\right]^{1/p^2}$ | 1.21 | 1.29 | 1.25 | 1.22 | 1.20 | 1.19 | 1.17 | 1.17 | 1.16 |

**Table 4:** computation of $\phi_{\boldsymbol{p},d}^{(\boldsymbol{\nu})}$ and $\left[\phi_{\boldsymbol{p},d}^{(\boldsymbol{\nu})}\right]^{1/D(\boldsymbol{p})}$ in the case $d = 2$, $\boldsymbol{p} = (p, p)$, $\boldsymbol{\nu} = (1, 1)$, for $p = 2, \ldots, 10$.

*approximately given by a uniform sampling of the eigenvalue functions* $\lambda_i(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})$ *over an equispaced grid in the domain* $[-\pi, \pi]^d$. To fix the ideas, assume that the equispaced grid is

$$-\boldsymbol{\pi} + \frac{2\boldsymbol{j}\boldsymbol{\pi}}{\boldsymbol{n}} = \left(-\pi + \frac{2j_1\pi}{n_1}, \ldots, -\pi + \frac{2j_d\pi}{n_d}\right), \quad \boldsymbol{j} = \mathbf{0}, \ldots, \boldsymbol{n} - \mathbf{1},$$

where $\boldsymbol{\pi} := (\pi, \ldots, \pi)$. Then, the eigenvalues of $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ are approximately given by [4]

$$\lambda_i\left(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}\left(-\boldsymbol{\pi} + \frac{2\boldsymbol{j}\boldsymbol{\pi}}{\boldsymbol{n}}\right)\right), \quad \boldsymbol{j} = \mathbf{0}, \ldots, \boldsymbol{n} - \mathbf{1}, \quad i = 1, \ldots, D(\boldsymbol{p}). \tag{77}$$

From (77) we infer that the ratio

$$\phi_{\boldsymbol{p},d}^{(\boldsymbol{\nu})} := \frac{\min_{\boldsymbol{\theta}\in[-\pi,\pi]^d} \lambda_{\max}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta}))}{\max_{\boldsymbol{\theta}\in[-\pi,\pi]^d} \lambda_{\min}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta}))}$$

is an index of the scattering of the eigenvalues of $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$. Indeed, if $\phi_{\boldsymbol{p},d}^{(\boldsymbol{\nu})}$ is large (resp. small), then the eigenvalues of $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ obtained from (77) for $i = 1$, which correspond to the maximal eigenvalue of the symbol, are far away from (resp. very close to) the eigenvalues obtained for $i = D(\boldsymbol{p})$, which correspond to the minimal eigenvalue of the symbol. Furthermore, in the case where $\phi_{\boldsymbol{p},d}^{(\boldsymbol{\nu})}$ is large, the 'ill-conditioned subspace', that is the subspace corresponding to the largest eigenvalues of $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ obtained by setting $i = 1$ in (77), is very large: its dimension is about

$$\#\left\{\lambda_{\max}\left(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}\left(-\boldsymbol{\pi} + \frac{2\boldsymbol{j}\boldsymbol{\pi}}{\boldsymbol{n}}\right)\right) : \boldsymbol{j} = \mathbf{0}, \ldots, \boldsymbol{n} - \mathbf{1}\right\} = \frac{D(\boldsymbol{n})}{D(\boldsymbol{p})} = \frac{n_1 \cdots n_d}{p_1 \cdots p_d}. \tag{78}$$

Tables 3–4 shows, for $d = 1, 2$, the behavior of $\phi_{\boldsymbol{p},d}^{(\boldsymbol{\nu})}$ in the case $\boldsymbol{p} = (p, \ldots, p)$, $\boldsymbol{\nu} = (1, \ldots, 1)$, for different values of $p$. Not only we observe an exponential ill-conditioning with $p$ and $d$, as already proved in [27], but we can also predict, on the base of (78), that the subspace where this exponential ill-conditioning occurs is very large: for the case displayed in Tables 3–4, the size of such subspace is approximately $n^d/p^d$. This involved picture shows that the numerical solution of the linear systems associated with the matrix $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ is intractable for large $\boldsymbol{p}$ and $d$, not only because of the exponential ill-conditioning, but also for the large size of the subspace where this ill-conditioning is attained.

[4] Ignore the mismatch with the size of $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$: the reasoning that we are following in this subsection is heuristic. Think of $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ as if it were *exactly* the Toeplitz matrix $T_{\boldsymbol{n}}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})$ generated by the symbol $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$.

### 6.4 Clustering of the normalized sequence $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$

In this subsection, we still assume that $\boldsymbol{n} = \boldsymbol{\nu} n$, where $\boldsymbol{\nu} \in \mathbb{Q}_{>0}^d$ is fixed and $n$ varies in the set of natural numbers such that $\boldsymbol{n} \in \mathbb{N}^d$. In this situation, we have seen in Theorem 14 that $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n \sim_\lambda \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$ and, consequently, $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ is weakly clustered at the essential range of $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$, given by the union of the essential ranges of the eigenvalue functions $\lambda_i(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})$, $i = 1, \ldots, D(\boldsymbol{p})$, that is $\mathcal{ER}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}) = \bigcup_{i=1}^{D(\boldsymbol{p})} \mathcal{ER}(\lambda_i(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}))$. Note that $\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$ is continuous over $[-\pi, \pi]^d$, hence the eigenvalue functions are continuous over $[-\pi, \pi]^d$, which means that their essential ranges coincide exactly with their images. Being weakly clustered at $\mathcal{ER}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})$, the sequence $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ is a fortiori weakly clustered at the convex hull of $\mathcal{ER}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})$, which is given by $[0, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}]$, $M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}} := \max_{\boldsymbol{\theta} \in [-\pi,\pi]^d} \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}(\boldsymbol{\theta})$. We are going to see that actually $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ is strongly clustered at $[0, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}]$, in the case where $\boldsymbol{\beta}$ is constant.

**Theorem 15.** *We have $\sigma(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) \subset (0, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}]$, and, moreover, for each fixed $j$ and for $n \to \infty$ we have*

$$\lambda_{D(\boldsymbol{np-1})-j}(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) \to 0, \qquad \lambda_j(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) \to M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}. \tag{79}$$

*Proof.* Since $A_{\boldsymbol{n},D}^{(\boldsymbol{p})}$ is SPD, $\lambda_{\min}(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) > 0$. To prove the inclusion $\sigma(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) \subset (0, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}]$, recall that in the proof of Theorem 14 we have defined the matrix $T_{\boldsymbol{n}}^{(\boldsymbol{p})}$, see Eq. (71), and we have noticed that $T_{\boldsymbol{n}}^{(\boldsymbol{p})}$ is similar to $T_{\boldsymbol{n}}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})$. We show that for every $\mathbf{x} \in \mathbb{C}^{D(\boldsymbol{np-1})}$ there exists $\mathbf{y} \in \mathbb{C}^{D(\boldsymbol{np})}$ with $\|\mathbf{y}\| = \|\mathbf{x}\|$ such that

$$\mathbf{x}^*(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})\mathbf{x} = \mathbf{y}^* T_{\boldsymbol{n}}^{(\boldsymbol{p})} \mathbf{y}, \tag{80}$$

which implies, by the minimax principle,

$$\lambda_{\max}(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) = \max_{\|\mathbf{x}\|=1}(\mathbf{x}^*(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})\mathbf{x}) \le \max_{\|\mathbf{y}\|=1}(\mathbf{y}^* T_{\boldsymbol{n}}^{(\boldsymbol{p})}\mathbf{y}) = \lambda_{\max}(T_{\boldsymbol{n}}^{(\boldsymbol{p})}) = \lambda_{\max}(T_{\boldsymbol{n}}(\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})})) \le M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}},$$

the last inequality being justified by Theorem 5.

In order to prove (80), it is convenient to index vectors and matrices using multi-indices in $\mathbb{N}^d$, with the standard lexicographic ordering on them, see Subsection 2.1. With this convention, for every $\mathbf{x} \in \mathbb{C}^{D(\boldsymbol{np-1})}$ we have

$$\mathbf{x}^*(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})\mathbf{x} = \sum_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{np-1}} \overline{x}_{\boldsymbol{i}}(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})_{\boldsymbol{ij}} x_{\boldsymbol{j}} = \sum_{\boldsymbol{i},\boldsymbol{j} \in \{\boldsymbol{1},\ldots,\boldsymbol{np-1}\}} \overline{x}_{\boldsymbol{i}}(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})_{\boldsymbol{ij}} x_{\boldsymbol{j}}.$$

Define $\mathbf{y} \in \mathbb{C}^{D(\boldsymbol{np})}$ in the following way:

$$y_{\boldsymbol{i}} = x_{\boldsymbol{i}} \quad \text{if } \boldsymbol{i} \in \{\boldsymbol{1}, \ldots, \boldsymbol{np-1}\}, \qquad y_{\boldsymbol{i}} = 0 \quad \text{if } \boldsymbol{i} \in \{\boldsymbol{1}, \ldots, \boldsymbol{np}\} \backslash \{\boldsymbol{1}, \ldots, \boldsymbol{np-1}\}.$$

Then $\|\mathbf{y}\| = \|\mathbf{x}\|$ and, moreover,

$$\mathbf{x}^*(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})\mathbf{x} = \sum_{\boldsymbol{i},\boldsymbol{j} \in \{\boldsymbol{1},\ldots,\boldsymbol{np-1}\}} \overline{x}_{\boldsymbol{i}}(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})_{\boldsymbol{ij}} x_{\boldsymbol{j}} = \sum_{\boldsymbol{i},\boldsymbol{j} \in \{\boldsymbol{1},\ldots,\boldsymbol{np}\}} \overline{y}_{\boldsymbol{i}}(T_{\boldsymbol{n}}^{(\boldsymbol{p})})_{\boldsymbol{ij}} y_{\boldsymbol{j}} = \mathbf{y}^* T_{\boldsymbol{n}}^{(\boldsymbol{p})} \mathbf{y}. \tag{81}$$

This concludes the proof of the inclusion $\sigma(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) \subset (0, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}]$, but we wish to prove in some more detail the central equality in (81).

- If $\boldsymbol{i} \in \{\boldsymbol{1}, \ldots, \boldsymbol{np}\} \backslash \{\boldsymbol{1}, \ldots, \boldsymbol{np-1}\}$ or $\boldsymbol{j} \in \{\boldsymbol{1}, \ldots, \boldsymbol{np}\} \backslash \{\boldsymbol{1}, \ldots, \boldsymbol{np-1}\}$, the $(\boldsymbol{i}, \boldsymbol{j})$ term in the righthand side of the central equality is 0.

- If $\boldsymbol{i} \in \{\mathbf{1}, \dots, \boldsymbol{np}-\mathbf{1}\}$ and $\boldsymbol{j} \in \{\mathbf{1}, \dots, \boldsymbol{np}-\mathbf{1}\}$, the $(\boldsymbol{i},\boldsymbol{j})$ term in the righthand side of the central equality is $\overline{y}_{\boldsymbol{i}}(T_{\boldsymbol{n}}^{(\boldsymbol{p})})_{\boldsymbol{ij}}y_{\boldsymbol{j}} = \overline{x}_{\boldsymbol{i}}(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})_{\boldsymbol{ij}}x_{\boldsymbol{j}}$, because $y_{\boldsymbol{i}} = x_{\boldsymbol{i}},\ y_{\boldsymbol{j}} = x_{\boldsymbol{j}}$, and, recalling (9) and the fact that $K_n^{(p)}$ and $M_n^{(p)}$ are the leading principal submatrices of order $np-1$ of $T_n(\mathbf{f}_p)$ and $T_n(\mathbf{h}_p)$, respectively, we have

$$
\begin{aligned}
(T_{\boldsymbol{n}}^{(\boldsymbol{p})})_{\boldsymbol{ij}} &= \sum_{k=1}^{d} c_k(\boldsymbol{\nu})\,[T_{n_1}(\mathbf{h}_{p_1}) \otimes \cdots \otimes T_{n_{k-1}}(\mathbf{h}_{p_{k-1}}) \otimes T_{n_k}(\mathbf{f}_{p_k}) \otimes T_{n_{k+1}}(\mathbf{h}_{p_{k+1}}) \otimes \cdots \otimes T_{n_d}(\mathbf{h}_{p_d})]_{\boldsymbol{ij}} \\
&= \sum_{k=1}^{d} c_k(\boldsymbol{\nu})\,[T_{n_1}(\mathbf{h}_{p_1})]_{i_1j_1} \cdots [T_{n_{k-1}}(\mathbf{h}_{p_{k-1}})]_{i_{k-1}j_{k-1}}[T_{n_k}(\mathbf{f}_{p_k})]_{i_kj_k}[T_{n_{k+1}}(\mathbf{h}_{p_{k+1}})]_{i_{k+1}j_{k+1}} \cdots [T_{n_d}(\mathbf{h}_{p_d})]_{i_dj_d} \\
&= \sum_{k=1}^{d} c_k(\boldsymbol{\nu})\,[M_{n_1}^{(p_1)}]_{i_1j_1} \cdots [M_{n_{k-1}}^{(p_{k-1})}]_{i_{k-1}j_{k-1}}[K_{n_k}^{(p_k)}]_{i_kj_k}[M_{n_{k+1}}^{(p_{k+1})}]_{i_{k+1}j_{k+1}} \cdots [M_{n_d}^{(p_d)}]_{i_dj_d} \\
&= \sum_{k=1}^{d} c_k(\boldsymbol{\nu})\,[M_{n_1}^{(p_1)} \otimes \cdots \otimes M_{n_{k-1}}^{(p_{k-1})} \otimes K_{n_k}^{(p_k)} \otimes M_{n_{k+1}}^{(p_{k+1})} \otimes \cdots \otimes M_{n_d}^{(p_d)}]_{\boldsymbol{ij}} = (n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})})_{\boldsymbol{ij}}.
\end{aligned}
$$

This concludes the proof of the central equality in (81) and the proof of the inclusion $\sigma(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) \subset (0, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}]$. Relation (79) follows from this inclusion and the fact that $\{n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}\}_n \sim_\lambda \mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}$ (by Theorem 14 applied with $\boldsymbol{\beta} = \mathbf{0}$ and $\gamma = 0$). We omit the formal proof of (79), which is based on the same argument that is employed for proving that items $1, 3$ in Theorem 5 imply item 4. □

**Theorem 16.** *Assume that $\boldsymbol{\beta}$ is constant. Then*

$$
\sigma(n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}) \subset \left[\frac{\sum_{k=1}^{d}\zeta_{n_k,p_k}+\gamma_*}{n^2}G_{\boldsymbol{p}}, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}} + \frac{\|\gamma\|_{L^\infty(\Omega)}}{\nu_1\cdots\nu_d n^2}S_{\boldsymbol{p}}\right] \times \left[-\frac{B_{\boldsymbol{p}}\|\boldsymbol{\beta}\|_{L^\infty(\Omega)}\sum_{k=1}^{d}\nu_k}{\nu_1\cdots\nu_d n}, \frac{B_{\boldsymbol{p}}\|\boldsymbol{\beta}\|_{L^\infty(\Omega)}\sum_{k=1}^{d}\nu_k}{\nu_1\cdots\nu_d n}\right],
$$

*with $\zeta_{n,p}$, $G_{\boldsymbol{p}}$, $S_{\boldsymbol{p}}$, $B_{\boldsymbol{p}}$ as in Theorem 12. In particular, $\{n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}\}_n$ is strongly clustered at $[0, M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}}]$.*

*Proof.* The real and imaginary parts of $n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}$ are

$$
\Re(n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}) = n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})} + n^{d-2}A_{\boldsymbol{n},R}^{(\boldsymbol{p})}, \qquad \Im(n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})}) = -\mathrm{i}\,n^{d-2}A_{\boldsymbol{n},A}^{(\boldsymbol{p})};
$$

cf. (63)–(64). By Theorem 12, Theorem 15, (7) and (75) we have

$$
\begin{aligned}
\frac{\sum_{k=1}^{d}\zeta_{n_k,p_k}+\gamma_*}{\nu_1\cdots\nu_d n^2}G_{\boldsymbol{p}} \le \lambda_{\min}(\Re(n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})})) \le \lambda_{\max}(\Re(n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})})) &\le \lambda_{\max}(n^{d-2}A_{\boldsymbol{n},D}^{(\boldsymbol{p})}) + \lambda_{\max}(n^{d-2}A_{\boldsymbol{n},R}^{(\boldsymbol{p})}) \\
&\le M_{\mathbf{f}_{\boldsymbol{p}}^{(\boldsymbol{\nu})}} + \frac{\|\gamma\|_{L^\infty(\Omega)}S_{\boldsymbol{p}}}{\nu_1\cdots\nu_d n^2}.
\end{aligned}
$$

By (74) we have

$$
-\frac{B_{\boldsymbol{p}}\|\boldsymbol{\beta}\|_\infty\sum_{k=1}^{d}\nu_k}{\nu_1\cdots\nu_d n} \le \lambda_{\min}(\Im(n^{d-2}A_{\boldsymbol{n}}^{(\boldsymbol{p})})) \le \lambda_{\max}(\Im(A_{\boldsymbol{n}}^{(\boldsymbol{p})})) \le \frac{B_{\boldsymbol{p}}\|\boldsymbol{\beta}\|_\infty\sum_{k=1}^{d}\nu_k}{\nu_1\cdots\nu_d n}.
$$

The thesis follows from (5). □

## 7 Concluding remarks, open problems, and a conjecture

We studied the spectral properties of the stiffness matrices that arise in the context of $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM for the numerical solution of the model equation (1). In particular, we investigated the conditioning and the asymptotic spectral distribution in the Weyl sense, and we identified the symbol describing the asymptotic spectrum. In view of the applications to the fast solution of the related linear systems by means of iterative solvers like preconditioned Krylov or multigrid methods, we also studied the properties of the symbol, which turned out to be a $D(\boldsymbol{p}) \times D(\boldsymbol{p})$ Hermitian matrix-valued function in $d$ variables. Unlike the $\boldsymbol{p}$-degree B-spline IgA approach, where a unique scalar-valued $d$-variate function describes all the spectrum, here the spectrum is described by $D(\boldsymbol{p})$ different functions, that is the $D(\boldsymbol{p})$ eigenvalues of the symbol. This very involved picture provides a clean explanation of the difficulties encountered in designing optimal and robust solvers (with convergence speed independent of $\boldsymbol{n}$, $\boldsymbol{p}$, $d$), and of the convergence deterioration of classical iterative solvers, already for moderate $\boldsymbol{p}$ and $d$.

We notice that the case of $\boldsymbol{p}$-degree B-spline IgA and the case of $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM are in a sense two extremes of the same class: in the IgA setting, given the polynomial approximation degree $p_j$ in direction $x_j$, the smoothness in that direction is maximal i.e. $s_j = p_j - 1$; conversely, in the $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM setting, if $p_j$ is the polynomial approximation degree in direction $x_j$, then the smoothness is minimal, that is only continuity or, equivalently, $s_j = 0$. A general conjecture, formulated by the second author and discussed in detail in [20], can be stated as follows:

**Conjecture 2.** *When considering uniform rectangular Finite Elements approximating problem* (1)*, if* $p_j \geq 1$ *is the polynomial approximation degree in direction* $x_j$ *and if* $s_j$ *is the global smoothness in that direction,* $0 \leq s_j \leq p_j - 1$*, then the resulting sequence of (normalized) stiffness matrices is distributed in Weyl sense and the symbol is a* $D(\boldsymbol{k})$ *matrix-valued function, where* $\boldsymbol{k} = (k_1, \ldots, k_d)$*,* $D(\boldsymbol{k}) = \prod_{j=1}^{d} k_j$*, and* $k_j = p_j - s_j$*.*

As a consequence of this conjecture, the structure of the approximation space would have interesting (and unexpected) consequences on the spectral complexity of the resulting approximation matrices: the larger is the gap between the polynomial approximation degree and the imposed smoothness and the more complicate is the spectral structure. This shows that the $\mathbb{Q}_{\boldsymbol{p}}$ Lagrangian FEM approach is the worst from this numerical linear algebra viewpoint, while the IgA approach induces the simplest and most handful spectral structure; see [21].

We plan to extend the analysis in this paper to other cases, such as either FEM with Lagrangian basis accompanied by Gauss-Lobatto-Legendre nodes or FEM with integrated Legendre basis [31, 9], which are widely employed for their better behavior in terms of the spectrum of the underlying matrices [9, 26, 27].

## 8 Acknowledgments

We are deeply indebted with Annalisa Buffa and Micol Pennacchio for helpful discussions and for their careful guidance in the Finite Element world. This work was partially supported by INdAM-GNCS Gruppo Nazionale per il Calcolo Scientifico and by the Program 'Becoming the Number One – Sweden (2014)' of the Knut and Alice Wallenberg Foundation.

## References

[1] Aricò A., Donatelli M. *A V-cycle multigrid for multilevel matrix algebras: proof of optimality.* Numerische Mathematik **105** (2007) 511–547.

[2] Aricò A., Donatelli M., Serra-Capizzano S. *V-cycle optimal convergence for certain (multilevel) structured linear systems.* Siam Journal on Matrix Analysis and Applications **26** (2004) 186–214.

[3] Axelsson O., Barker V. A. *Finite Element solution of boundary value problems: theory and computation.* Siam (2001).

[4] Beckermann B., Kuijlaars A. B. J. *Superlinear convergence of Conjugate Gradients.* Siam Journal on Numerical Analysis **39** (2001) 300–329.

[5] Bhatia R. *Matrix analysis.* Springer-Verlag, New York (1997).

[6] Böttcher A., Silbermann B. *Introduction to large truncated Toeplitz matrices.* Springer-Verlag, New York (1999).

[7] Brezis H. *Functional Analysis, Sobolev spaces and partial differential equations.* Springer-Verlag, New York (2011).

[8] Brezzi F., Fortin M. *Mixed and hybrid Finite Element Methods.* Springer-Verlag, New York (1991).

[9] Canuto C., Hussaini M. Y., Quarteroni A., Zang T. A. *Spectral methods: evolution to complex geometries and applications to fluid dynamics.* Springer-Verlag, Berlin Heidelberg (2007).

[10] Ciarlet P. *The Finite Element Method for elliptic problems.* Siam (2002).

[11] Donatelli M., Garoni C., Manni C., Serra-Capizzano S., Speleers H. *Robust and optimal multi-iterative techniques for IgA Galerkin linear systems.* Computer Methods in Applied Mechanics and Engineering **284** (2015) 230–264.

[12] Donatelli M., Garoni C., Manni C., Serra-Capizzano S., Speleers H. *Spectral analysis and spectral symbol of matrices in isogeometric collocation methods.* Mathematics of Computation (2014), under revision.

[13] Donatelli M., Garoni C., Manni C., Serra-Capizzano S., Speleers H. *Robust and optimal multi-iterative techniques for IgA collocation linear systems.* Computer Methods in Applied Mechanics and Engineering (2014), under revision.

[14] Donatelli M., Garoni C., Manni C., Serra-Capizzano S., Speleers H. *Symbol-based multigrid (and multi-iterative) methods for Galerkin B-spline isogeometric analysis.* Siam Journal on Numerical Analysis (2014), submitted.

[15] Donatelli M., Molteni M., Pennati V., Serra-Capizzano S. *Multigrid methods for cubic spline solution of two points (and 2D) boundary value problems.* Applied Numerical Mathematics (2014) DOI 10.1016/j.apnum.2014.04.004.

[16] Engl H., Hanke M., Neubauer A. *Regularization of inverse problems.* Kluwer Academic Publishers, Dordrecht, The Netherlands (1996).

[17] Fiorentino G., Serra S. *Multigrid methods for Toeplitz matrices.* Calcolo **28** (1991) 283–305.

[18] Fiorentino G., Serra S. *Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions.* Siam Journal on Scientific Computing **17** (1996) 1068–1081.

[19] Garoni C. *Structured matrices coming from PDE Approximation Theory: spectral analysis, spectral symbol and design of fast iterative solvers.* Ph.D. Thesis in Mathematics of Computation, University of Insubria, Como, Italy (2014).

[20] Garoni C., Hughes T. J. R., Reali A., Serra-Capizzano S., Speleers H. *Smoothness vs polynomial degree: why IgA outperforms FEA in the spectral approximation.* In preparation.

[21] Garoni C., Manni C., Pelosi F., Serra-Capizzano S., Speleers H. *On the spectrum of stiffness matrices arising from Isogeometric Analysis.* Numerische Mathematik **127** (2014) 751–799. A wider version is available as Technical Report: Report TW632, Dept. Computer Science, K. U. Leuven (2013).

[22] Garoni C., Serra-Capizzano S., Sesana D. *Tools for determining the asymptotic spectral distribution of non-Hermitian perturbations of Hermitian matrix-sequences and applications.* Integral Equations and Operator Theory (2014) http://dx.doi.org/10.1007/s00020-014-2157-6

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

[24] Graham A. *Kronecker products and matrix calculus: with applications.* Ellis Horwood Limited, Chichester (1981).

[25] Jin X. Q. *Developments and applications of block Toeplitz iterative solvers.* Kluwer Academic Publishers, Dordrecht (2002).

[26] Melenk J. M. *On condition numbers in hp-FEM with Gauss-Lobatto based shape functions.* Journal of Computational and Applied Mathematics **139** (2002) 21–48

[27] Olsen E., Douglas J. *Bounds on spectral condition numbers of matrices arising in the p-version of the Finite Element Method.* Numerische Mathematik **69** (1995) 333–352.

[28] Quarteroni A. *Numerical models for differential problems.* Springer-Verlag Italia, Milan (2009).

[29] Quarteroni A., Valli A. *Numerical approximation of partial differential equations.* Springer-Verlag, Berlin Heidelberg (2008).

[30] Saad Y. *Iterative methods for sparse linear systems.* Siam (2003).

[31] Schwab C. *p- and hp- Finite Element Methods.* Clarendon Press, Oxford (1998).

[32] Serra S. *Asymptotic results on the spectra of block Toeplitz preconditioned matrices.* SIAM Journal on Matrix Analysis and Applications **20** (1998) 31–44.

[33] Serra-Capizzano S. *Generalized Locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations.* Linear Algebra and its Applications **366** (2003) 371–402.

[34] Serra-Capizzano S. *GLT sequences as a generalized Fourier Analysis and applications.* Linear Algebra and its Applications **419** (2006) 180–233.

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