

RESEARCH ARTICLE

Are the eigenvalues of the B-spline isogeometric analysis approximation of $-\Delta u = \lambda u$ known in almost closed form?

Sven-Erik Ekström¹ | Isabella Furci² | Carlo Garoni^{2,3} | Carla Manni⁴ | Stefano Serra-Capizzano^{1,2} | Hendrik Speleers⁴

¹Department of Information Technology, Division of Scientific Computing, Uppsala University, Information Technology Center, Lägerhyddsvägen 2, Hus 2, SE-751 05 Uppsala, Sweden

²Department of Science and High Technology, University of Insubria, Via Valleggio 11, 22100 Como, Italy

³Institute of Computational Science, University of Italian Switzerland (USI), Via Giuseppe Buffi 13, 6900 Lugano, Switzerland

⁴Department of Mathematics, University of Rome "Tor Vergata", Via della Ricerca Scientifica 1, 00133 Rome, Italy

Correspondence

Carlo Garoni, Department of Science and High Technology, University of Insubria, Via Valleggio 11, 22100 Como, Italy; or Institute of Computational Science, University of Italian Switzerland (USI), Via Giuseppe Buffi 13, 6900 Lugano, Switzerland.
Email: carlo.garoni@usi.ch

Funding information

Graduate School in Mathematics and Computing (FMB); Uppsala University; INdAM (Istituto Nazionale di Alta Matematica), Grant/Award Number: PCOFUND-GA-2012-600198; "Mission Sustainability" programme of the University of Rome "Tor Vergata"; INdAM GNCS (Gruppo Nazionale per il Calcolo Scientifico)

MSC Classification: 65N25, 65N30, 65D07, 65F15, 65D05, 65B05, 15B05

Summary

We consider the B-spline isogeometric analysis approximation of the Laplacian eigenvalue problem $-\Delta u = \lambda u$ over the d -dimensional hypercube $(0, 1)^d$. By using tensor-product arguments, we show that the eigenvalue–eigenvector structure of the resulting discretization matrix is completely determined by the eigenvalue–eigenvector structure of the matrix $L_n^{[p]}$ arising from the isogeometric analysis approximation based on B-splines of degree p of the unidimensional problem $-u'' = \lambda u$. Here, n is the mesh fineness parameter, and the size of $L_n^{[p]}$ is $N(n, p) = n + p - 2$. In previous works, it was established that the normalized sequence $\{n^{-2}L_n^{[p]}\}_n$ enjoys an asymptotic spectral distribution described by a function $e_p(\theta)$, the so-called spectral symbol. The contributions of this paper can be summarized as follows:

1. We prove several important analytic properties of the spectral symbol $e_p(\theta)$. In particular, we show that $e_p(\theta)$ is monotone increasing on $[0, \pi]$ for all $p \geq 1$ and that $e_p(\theta) \rightarrow \theta^2$ uniformly on $[0, \pi]$ as $p \rightarrow \infty$.
2. For $p = 1$ and $p = 2$, we show that $L_n^{[p]}$ belongs to a matrix algebra associated with a fast unitary sine transform, and we compute eigenvalues and eigenvectors of $L_n^{[p]}$. In both cases, the eigenvalues are given by $e_p(\theta_{j,n})$, $j = 1, \dots, n + p - 2$, where $\theta_{j,n} = j\pi/n$.
3. For $p \geq 3$, we provide numerical evidence of a precise asymptotic expansion for the eigenvalues of $n^{-2}L_n^{[p]}$, excluding the largest $n_p^{\text{out}} = p - 2 + \text{mod}(p, 2)$ eigenvalues (the so-called outliers). More precisely, we numerically show that for every $p \geq 3$, every integer $\alpha \geq 0$, every n , and every $j = 1, \dots, N(n, p) - n_p^{\text{out}}$,

$$\lambda_j \left(n^{-2}L_n^{[p]} \right) = e_p(\theta_{j,n}) + \sum_{k=1}^{\alpha} c_k^{[p]}(\theta_{j,n})h^k + E_{j,n,\alpha}^{[p]},$$

where

- the eigenvalues of $n^{-2}L_n^{[p]}$ are arranged in ascending order, $\lambda_1(n^{-2}L_n^{[p]}) \leq \dots \leq \lambda_{n+p-2}(n^{-2}L_n^{[p]})$;
- $\{c_k^{[p]}\}_{k=1,2,\dots}$ is a sequence of functions from $[0, \pi]$ to \mathbb{R} , which depends only on p ;
- $h = 1/n$ and $\theta_{j,n} = j\pi h$ for $j = 1, \dots, n$; and
- $E_{j,n,\alpha}^{[p]} = O(h^{\alpha+1})$ is the remainder, which satisfies $|E_{j,n,\alpha}^{[p]}| \leq C_{\alpha}^{[p]}h^{\alpha+1}$ for some constant $C_{\alpha}^{[p]}$ depending only on α and p .

We also provide a proof of this expansion for $\alpha = 0$ and $j = 1, \dots, N(n, p) - (4p - 2)$, where $4p - 2$ represents a theoretical estimate of the number of outliers n_p^{out} .

4. We show through numerical experiments that, for $p \geq 3$ and $k \geq 1$, there exists a point $\theta(p, k) \in (0, \pi)$ such that $c_k^{[p]}(\theta)$ vanishes on $[0, \theta(p, k)]$. Moreover, as it is suggested by the numerics of this paper, the infimum of $\theta(p, k)$ over all $k \geq 1$, say y_p , is strictly positive, and the equation $\lambda_j(n^{-2}L_n^{[p]}) = e_p(\theta_{j,n})$ holds numerically whenever $\theta_{j,n} < \theta(p)$, where $\theta(p)$ is a point in $(0, y_p]$ which grows with p .
5. For $p \geq 3$, based on the asymptotic expansion in the above item 3, we propose a parallel interpolation–extrapolation algorithm for computing the eigenvalues of $L_n^{[p]}$, excluding the n_p^{out} outliers. The performance of the algorithm is illustrated through numerical experiments. Note that, by the previous item 4, the algorithm is actually not necessary for computing the eigenvalues corresponding to points $\theta_{j,n} < \theta(p)$.

KEYWORDS

eigenvalues and eigenvectors, isogeometric analysis and B-splines, Laplacian eigenvalue problem, mass and stiffness matrices, polynomial interpolation and extrapolation, spectral symbol and asymptotic eigenvalue expansion

1 | INTRODUCTION

Isogeometric analysis (IgA) is a modern paradigm for analyzing problems governed by partial differential equations (PDEs); see the work of Cottrell et al.¹ Due to its capability to enhance the connection between numerical simulation and computer-aided design (CAD) systems, IgA is gaining more and more attention over time. In particular, the spectral investigation of matrices arising from the IgA discretization of PDEs has become a topic of interest in the scientific community, mainly due to the superiority of IgA over the classical finite element analysis (FEA) in approximating the spectrum of the underlying differential operator.^{2–6} It is also worth recalling that recent spectral distribution results for IgA discretization matrices^{7–13} turned out to be the keystone for designing fast IgA solvers.^{14–16}

In the present paper, motivated by the aforesaid interest, we perform a detailed spectral analysis of the matrices arising from the B-spline IgA discretization of the Laplacian eigenproblem $-\Delta u = \lambda u$. Our main results, which will be detailed in Section 1.2, complement those of other works^{7–13} and deliver a fast (parallel) interpolation–extrapolation algorithm for computing the eigenvalues of the considered IgA matrices.

1.1 | Problem setting

Consider the one-dimensional Laplacian eigenvalue problem

$$\begin{cases} -u''(x) = \lambda u(x), & x \in (0, 1), \\ u(0) = u(1) = 0. \end{cases} \quad (1)$$

The corresponding weak formulation reads as follows: Find eigenvalues $\lambda \in \mathbb{R}^+$ and eigenfunctions $u \in H_0^1(0, 1)$ such that, for all $v \in H_0^1(0, 1)$,

$$a(u, v) = \lambda(u, v),$$

where

$$a(u, v) = \int_0^1 u'(x)v'(x)dx, \quad (u, v) = \int_0^1 u(x)v(x)dx.$$

In Galerkin's method, we choose a finite-dimensional vector space $\mathcal{W} \subset H_0^1(0, 1)$, we set $N = \dim \mathcal{W}$, and we look for approximations of the exact eigenpairs

$$\lambda_j = j^2 \pi^2, \quad u_j(x) = \sin(j\pi x), \quad j \geq 1, \quad (2)$$

by solving the following Galerkin problem: Find $\lambda_{j,\mathcal{W}} \in \mathbb{R}^+$ and $u_{j,\mathcal{W}} \in \mathcal{W}$, for $j = 1, \dots, N$, such that, for all $v \in \mathcal{W}$,

$$a(u_{j,\mathcal{W}}, v) = \lambda_{j,\mathcal{W}} (u_{j,\mathcal{W}}, v). \tag{3}$$

Assuming the numerical eigenvalues $\lambda_{j,\mathcal{W}}$ are arranged in ascending order, the pair $(\lambda_{j,\mathcal{W}}, u_{j,\mathcal{W}})$ is taken as an approximation of the pair (λ_j, u_j) for all $j = 1, \dots, N$. The numbers $\lambda_{j,\mathcal{W}} / \lambda_j - 1$, $j = 1, \dots, N$, are referred to as the (relative) eigenvalue errors. If $\{\varphi_1, \dots, \varphi_N\}$ is a basis of \mathcal{W} , in view of the canonical identification between each $v \in \mathcal{W}$ and its coefficient vector with respect to $\{\varphi_1, \dots, \varphi_N\}$, solving the Galerkin problem (3) is equivalent to solving the generalized eigenvalue problem

$$K \mathbf{u}_{j,\mathcal{W}} = \lambda_{j,\mathcal{W}} M \mathbf{u}_{j,\mathcal{W}}, \tag{4}$$

where $\mathbf{u}_{j,\mathcal{W}}$ is the coefficient vector of $u_{j,\mathcal{W}}$ with respect to $\{\varphi_1, \dots, \varphi_N\}$ and

$$K = [a(\varphi_j, \varphi_i)]_{i,j=1}^N = \left[\int_0^1 \varphi_j'(x) \varphi_i'(x) dx \right]_{i,j=1}^N, \tag{5}$$

$$M = [(\varphi_j, \varphi_i)]_{i,j=1}^N = \left[\int_0^1 \varphi_j(x) \varphi_i(x) dx \right]_{i,j=1}^N. \tag{6}$$

The matrices K and M are referred to as the stiffness matrix and the mass matrix, respectively. Both K and M are always symmetric positive definite, regardless of the chosen basis functions $\varphi_1, \dots, \varphi_N$. Moreover, it is clear from (4) that the numerical eigenvalues $\lambda_{j,\mathcal{W}}$, $j = 1, \dots, N$, are just the eigenvalues of the matrix

$$L = M^{-1}K. \tag{7}$$

Now, for $p, n \geq 1$ let

$$N_{i,[p]}, \quad i = 1, \dots, n + p, \tag{8}$$

be the B-splines of degree $p \geq 1$ and smoothness $C^{p-1}(\mathbb{R})$ defined over the knot sequence

$$\underbrace{0, \dots, 0}_{p+1}, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, \underbrace{1, \dots, 1}_{p+1}.$$

The B-splines (8) form a basis for the spline space

$$\mathcal{V}_{n,[p]} = \left\{ v \in C^{p-1}[0, 1] : v|_{\left[\frac{i}{n}, \frac{i+1}{n}\right]} \in \mathbb{P}_p \text{ for } i = 0, \dots, n-1 \right\},$$

where \mathbb{P}_p is the space of polynomials of degree at most p . Except for the first and the last one, all the other B-splines vanish on the boundary of $[0, 1]$. In particular, the B-splines

$$N_{i+1,[p]}, \quad i = 1, \dots, n + p - 2, \tag{9}$$

form a basis for the space

$$\mathcal{W}_{n,[p]} = \{ v \in \mathcal{V}_{n,[p]} : v(0) = v(1) = 0 \}.$$

We refer the reader to Figure 1 for the graphs of the B-splines (8) corresponding to the degree $p = 3$. For more on B-splines, including the precise definition of the functions (8), see the works of de Boor¹⁷ and Schumaker.¹⁸

In the IgA approximation of (1) based on uniform B-splines of degree $p \geq 1$, we look for approximations of the exact eigenpairs (2) by using the Galerkin method described above, in which the basis functions $\varphi_1, \dots, \varphi_N$ are chosen as the

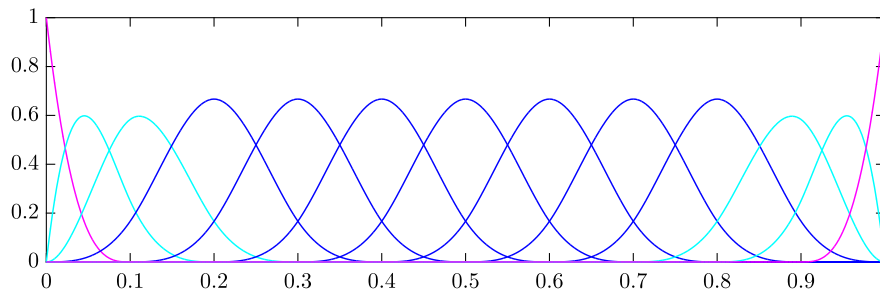


FIGURE 1 Cubic B-splines $\{N_{1,[3]}, \dots, N_{n+3,[3]}\}$ for the knot sequence $\{0, 0, 0, 0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1, 1, 1, 1\}$ ($n = 10$)

B-splines $N_{2,[p]}, \dots, N_{n+p-1,[p]}$, and consequently, the vector space \mathscr{W} is equal to $\mathscr{W}_{n,[p]}$. The resulting stiffness and mass matrices (5)–(6) are given by

$$K_n^{[p]} = \left[\int_0^1 N'_{j+1,[p]}(x) N'_{i+1,[p]}(x) dx \right]_{i,j=1}^{n+p-2}, \quad (10)$$

$$M_n^{[p]} = \left[\int_0^1 N_{j+1,[p]}(x) N_{i+1,[p]}(x) dx \right]_{i,j=1}^{n+p-2}, \quad (11)$$

and the numerical eigenvalues $\lambda_{j,n}^{[p]}, j = 1, \dots, n+p-2$, are the eigenvalues of the matrix

$$L_n^{[p]} = \left(M_n^{[p]} \right)^{-1} K_n^{[p]}. \quad (12)$$

For more details on IgA, we refer the reader to the work of Cottrell et al.¹

Let ϕ_q be the B-spline of degree $q \geq 0$ corresponding to the knot sequence $\{0, 1, \dots, q+1\}$. The function ϕ_q is usually referred to as the cardinal B-spline of degree q and is recursively defined as follows¹⁷:

$$\begin{aligned} \phi_0(t) &= \chi_{[0,1)}(t), \quad t \in \mathbb{R}, \\ \phi_q(t) &= \frac{t}{q} \phi_{q-1}(t) + \frac{q+1-t}{q} \phi_{q-1}(t-1), \quad t \in \mathbb{R}, \quad q \geq 1, \end{aligned}$$

where $\chi_{[0,1)}$ is the characteristic (indicator) function of the interval $[0, 1)$. Let

$$f_p : [0, \pi] \rightarrow \mathbb{R}, \quad f_p(\theta) = -\phi''_{2p+1}(p+1) - 2 \sum_{k=1}^p \phi''_{2p+1}(p+1-k) \cos(k\theta), \quad p \geq 1, \quad (13)$$

$$g_p : [0, \pi] \rightarrow \mathbb{R}, \quad g_p(\theta) = \phi_{2p+1}(p+1) + 2 \sum_{k=1}^p \phi_{2p+1}(p+1-k) \cos(k\theta), \quad p \geq 0, \quad (14)$$

$$e_p : [0, \pi] \rightarrow \mathbb{R}, \quad e_p(\theta) = \frac{f_p(\theta)}{g_p(\theta)}, \quad p \geq 1. \quad (15)$$

It was proved in section 3 in the work of Garoni et al.⁹ that*

$$f_p(\theta) = (2 - 2 \cos(\theta)) g_{p-1}(\theta), \quad \theta \in [0, \pi], \quad p \geq 1, \quad (16)$$

$$\left(\frac{4}{\pi^2} \right)^{p+1} \leq g_p(\theta) \leq g_p(0) = 1, \quad \theta \in [0, \pi], \quad p \geq 0, \quad (17)$$

so in particular, the function $e_p(\theta)$ is well defined. From the analysis in section 10.7 in the work of Garoni et al.¹², we know that the three sequences of matrices $\{n^{-1}K_n^{[p]}\}_n$, $\{nM_n^{[p]}\}_n$, $\{n^{-2}L_n^{[p]}\}_n$ have an asymptotic spectral distribution (in the Weyl sense) described by the functions $f_p(\theta)$, $g_p(\theta)$, $e_p(\theta)$, respectively; that is, for any sufficiently large n , up to a small number of outliers, the eigenvalues of $n^{-1}K_n^{[p]}$ (respectively, $nM_n^{[p]}$, $n^{-2}L_n^{[p]}$) are approximately given by the samples of $f_p(\theta)$ (respectively, $g_p(\theta)$, $e_p(\theta)$) over some uniform grid in $[0, \pi]$. This is illustrated in Figure 2 for the matrix $n^{-2}L_n^{[p]}$ and for $p = 1, \dots, 6$. Following the terminology in section 3.1 in the work of Garoni et al.,¹² we refer to $f_p(\theta)$, $g_p(\theta)$, $e_p(\theta)$ as the spectral symbols of $\{n^{-1}K_n^{[p]}\}_n$, $\{nM_n^{[p]}\}_n$, $\{n^{-2}L_n^{[p]}\}_n$, respectively. For more details on the spectral distribution of a sequence of matrices, see the work of Garoni et al.¹²

1.2 | Contributions of this work

The main contributions of this work can be summarized as follows. Throughout this paper, we will use the notations $n_p^{\text{out}} = p - 2 + \text{mod}(p, 2)$ and $N(n, p) = n + p - 2$.

1. We prove several important analytic properties of the spectral symbol $e_p(\theta)$. In particular, we show that $e_p(\theta)$ is monotone increasing on $[0, \pi]$ for all $p \geq 1$ and that $e_p(\theta) \rightarrow \theta^2$ uniformly on $[0, \pi]$ as $p \rightarrow \infty$. Incidentally, we also show that

*Note that in the work of Garoni et al.,⁹ the function $g_p(\theta)$ is denoted by $h_p(\theta)$.

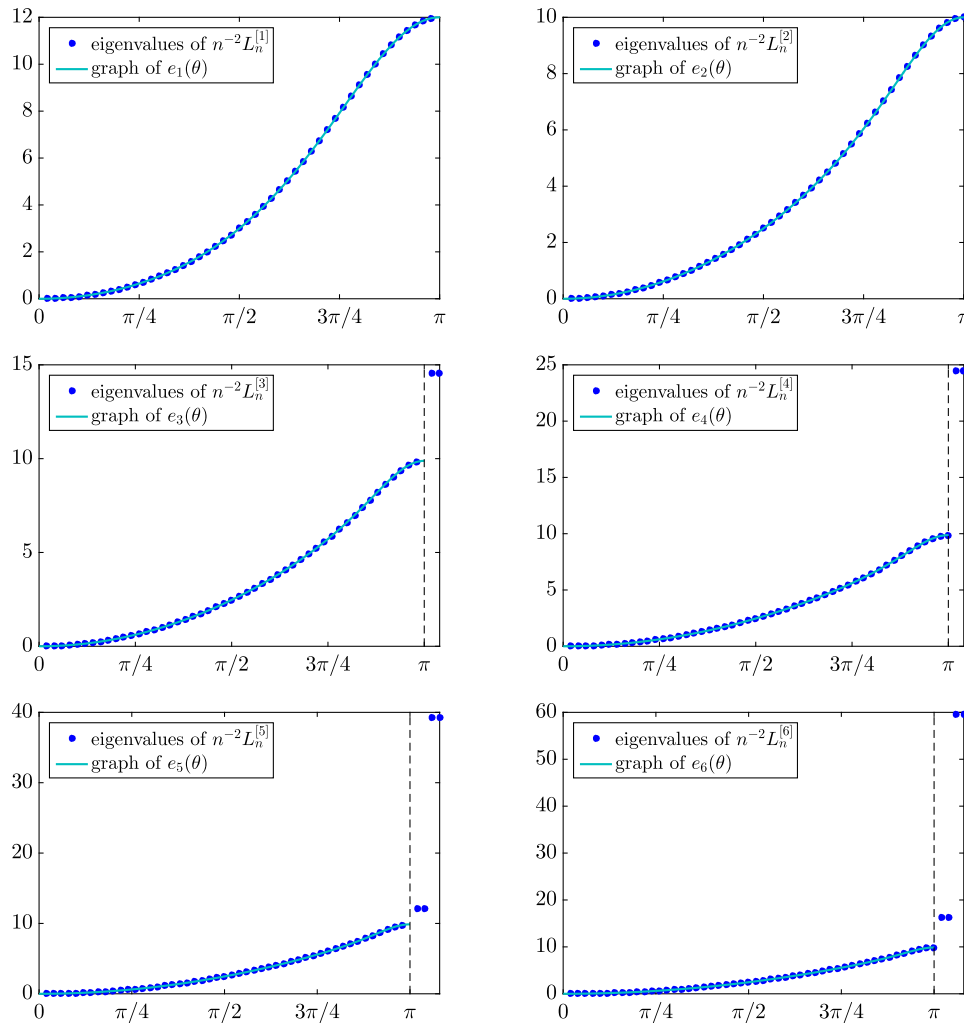


FIGURE 2 Comparison between the eigenvalues of $n^{-2}L_n^{[p]}$ and the graph of $e_p(\theta)$ for $n = 50$ and $p = 1, \dots, 6$. The eigenvalues of $n^{-2}L_n^{[p]}$ are sorted in ascending order and are represented by the thick dots placed at the points $(\theta_{j,n}, \lambda_j(n^{-2}L_n^{[p]}))$, $j = 1, \dots, n - \text{mod}(p, 2)$, where $\theta_{j,n} = j\pi/n$. The eigenvalues $\lambda_j(n^{-2}L_n^{[p]})$ for $j > n - \text{mod}(p, 2)$ are the so-called outliers and are positioned outside the domain $[0, \pi]$

the ratio $w_p(\theta) = g_p(\theta)/g_{p-1}(\theta)$ satisfies $1/3 \leq w_p(\theta) \leq 1$ for all $p \geq 1$ and $\theta \in [0, \pi]$. The latter result was already conjectured in the works of Donatelli et al.^{14,16} on the basis of numerical experiments, and it was therein exploited to design/analyze fast solvers for IgA discretization matrices.

2. For $p = 1$ and $p = 2$, we compute eigenvalues and eigenvectors of $L_n^{[p]}$. In both cases, the eigenvalues are given by $e_p(\theta_{j,n})$ for $j = 1, \dots, n + p - 2$, where $\theta_{j,n} = j\pi/n$. The exact computation of eigenvalues and eigenvectors is made possible by the fact that the matrices $K_n^{[p]}, M_n^{[p]}, L_n^{[p]}$ belong to the same matrix algebra, which is the tau algebra $\tau_{n-1}(0, 0)$ for $p = 1$ and the algebra $\tau_n(-1, -1)$ for $p = 2$ (we are using the notations from the work of Bozzo et al.¹⁹). It is worth noting that both these algebras are related to fast unitary sine transforms,¹⁹ which implies that many numerical linear algebra computations involving the matrices $K_n^{[p]}, M_n^{[p]}, L_n^{[p]}$ (matrix-vector products, solutions of linear systems, inversions, etc.) are stable and fast.
3. For $p \geq 3$, we provide numerical evidence of a precise asymptotic expansion for the eigenvalues of $n^{-2}L_n^{[p]}$. Such an expansion, which obviously begins with the spectral symbol $e_p(\theta)$, is in force for the whole of the spectrum except for the largest n_p^{out} eigenvalues (the so-called outliers; see Figure 2). To be more precise, we show through numerical experiments that for every $p \geq 3$, every integer $\alpha \geq 0$, every n , and every $j = 1, \dots, N(n, p) - n_p^{\text{out}} = n - \text{mod}(p, 2)$, we have

$$\lambda_j \left(n^{-2}L_n^{[p]} \right) = e_p(\theta_{j,n}) + \sum_{k=1}^{\alpha} c_k^{[p]}(\theta_{j,n})h^k + E_{j,n,\alpha}^{[p]}, \tag{18}$$

where

- the eigenvalues of $n^{-2}L_n^{[p]}$ are arranged in ascending order, $\lambda_1(n^{-2}L_n^{[p]}) \leq \dots \leq \lambda_{n+p-2}(n^{-2}L_n^{[p]})$;
- $\{c_k^{[p]}\}_{k=1,2,\dots}$ is a sequence of functions from $[0, \pi]$ to \mathbb{R} , which depends only on p ;
- $h = \frac{1}{n}$ and $\theta_{j,n} = \frac{j\pi}{n} = j\pi h$ for $j = 1, \dots, n$; and
- $E_{j,n,\alpha}^{[p]} = O(h^{\alpha+1})$ is the remainder (the error), which satisfies the inequality $|E_{j,n,\alpha}^{[p]}| \leq C_\alpha^{[p]} h^{\alpha+1}$ for some constant $C_\alpha^{[p]}$ depending only on α and p .

We refer the reader to Appendix B for a proof of the expansion (18) for $\alpha = 0$ and $j = 1, \dots, N(n, p) - (4p - 2)$, where $4p - 2$ represents an estimate, solely based on interlacing/rank-correction arguments, of the actual number of outliers n_p^{out} . We note that (18) is formally the same as the expansions for the eigenvalues of Toeplitz and preconditioned Toeplitz matrices, which have been conjectured and validated through numerical experiments in other works.^{20,21} In the case of Toeplitz matrices, the eigenvalue expansion has also been proved by Bogoya et al. in a sequence of recent papers.^{22–24} Furthermore, basic eigenvalue expansions (and related extrapolation techniques) have been used in other works^{25,26} in the context of finite element approximations of differential problems. In the light of these considerations, the expansion (18) is not completely unexpected, because $n^{-2}L_n^{[p]}$ is “almost” a preconditioned Toeplitz matrix as $n^{-2}L_n^{[p]} = (nM_n^{[p]})^{-1}(n^{-1}K_n^{[p]})$, and $nM_n^{[p]}$, $n^{-1}K_n^{[p]}$ are Toeplitz matrices, up to low rank corrections. To be precise, let $T_m(a)$ be the Toeplitz matrix of size m generated by the function $a \in L^1(-\pi, \pi)$, that is,

$$T_m(a) = [a_{i-j}]_{i,j=1}^m = \begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & \cdots & a_{-(m-1)} \\ a_1 & \ddots & \ddots & \ddots & & \vdots \\ a_2 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & a_{-2} \\ \vdots & & & \ddots & \ddots & a_{-1} \\ a_{m-1} & \cdots & \cdots & a_2 & a_1 & a_0 \end{bmatrix},$$

where the numbers $a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} a(\theta) e^{-ik\theta} d\theta$, $k \in \mathbb{Z}$, are the Fourier coefficients of a . Then,

$$n^{-1}K_n^{[p]} = T_{n+p-2}(f_p) + R_n^{[p]}, \quad (19)$$

$$nM_n^{[p]} = T_{n+p-2}(g_p) + S_n^{[p]}, \quad (20)$$

where f_p, g_p are defined in (13)–(14) and

$$\left(R_n^{[p]} \right)_{ij} = 0, \quad 2p \leq i \leq n - p - 1 \quad \implies \quad \text{rank} \left(R_n^{[p]} \right) \leq 4p - 2, \quad (21)$$

$$\left(S_n^{[p]} \right)_{ij} = 0, \quad 2p \leq i \leq n - p - 1 \quad \implies \quad \text{rank} \left(S_n^{[p]} \right) \leq 4p - 2; \quad (22)$$

see section 4.1 in the work of Garoni et al.⁹

4. We show through numerical experiments that, for $p \geq 3$ and $k \geq 1$, there exists a point $\theta(p, k) \in (0, \pi)$ such that $c_k^{[p]}(\theta)$ vanishes over $[0, \theta(p, k)]$. Moreover, as it is suggested by the numerics of this paper, it is very likely that $y_p = \inf_{k \geq 1} \theta(p, k) > 0$ for all $p \geq 3$. This is consistent with another crucial numerical observation, namely the fact that, for all $p \geq 3$, the equation $\lambda_j(n^{-2}L_n^{[p]}) = e_p(\theta_{j,n})$ holds numerically whenever $\theta_{j,n} < \theta(p)$, with $\theta(p)$ being a point in $(0, y_p]$. In addition, $\theta(p)$ apparently grows with p , that is, the portion of the spectrum of $\lambda_j(n^{-2}L_n^{[p]})$ which is exactly described by $e_p(\theta)$, at least from a numerical viewpoint, increases with p .
5. For $p \geq 3$, based on the expansion (18) and drawing inspiration from the work of Ekström et al.,²⁷ we propose a parallel interpolation–extrapolation algorithm for computing the eigenvalues of $L_n^{[p]}$, excluding the n_p^{out} outliers. The performance of the algorithm is illustrated through numerical experiments. Note that we actually need to compute only the eigenvalues of $L_n^{[p]}$ corresponding in the expansion (18) to points $\theta_{j,n} \geq \theta(p)$, because whenever $\theta_{j,n} < \theta(p)$, we numerically have $\lambda_j(L_n^{[p]}) = n^2 e_p(\theta_{j,n})$ by the previous item 4.
6. We present a detailed extension of the whole analysis to the general d -dimensional setting, in which problem (1) is replaced by (32). By using tensor-product arguments, we show that the eigenvalue–eigenvector structure of the matrix arising from the IgA approximation of (32) is completely determined by the eigenvalue–eigenvector structure of the matrix $L_n^{[p]}$. In short, the analysis of $L_n^{[p]}$ is enough to cover also the multidimensional case.

1.3 | Organization of the paper

The paper is organized as follows. In Section 2, we report the properties of $e_p(\theta)$ (and $w_p(\theta)$) that we shall prove in this paper; for ease of reading, the corresponding technical proofs are deferred to Appendix A. In Section 3, we compute eigenvalues and eigenvectors of the matrix $L_n^{[p]}$ for $p = 1$ and $p = 2$. In Section 4, assuming the asymptotic eigenvalue expansion (18), we present our parallel interpolation–extrapolation algorithm for computing the eigenvalues of $L_n^{[p]}$ for $p \geq 3$, excluding the n_p^{out} outliers. In Section 5, we provide numerical experiments in support of both the asymptotic eigenvalue expansion (18) and the properties described in item 4 of Section 1.2. Moreover, we numerically illustrate the performance of the algorithm presented in Section 4. In Section 6, we extend the whole analysis carried out in Sections 3–5 to the multidimensional setting by showing appropriate tensor-product arguments that the multidimensional case reduces to the unidimensional case. Finally, in Section 7, we draw conclusions and outline future lines of research.

2 | PROPERTIES OF THE SPECTRAL SYMBOL $e_p(\theta)$

The spectral symbol $e_p(\theta)$ enjoys the properties reported in Theorems 1 and 2, whose proofs are collected in Appendix A. We note that the convergence expressed in Theorem 1 was numerically observed by Ekström et al.³ and represents a starting point for the research program outlined in remark 15 in the work of Garoni et al.²⁸

Theorem 1. *The function $e_p(\theta)$ converges uniformly to θ^2 on $[0, \pi]$ as $p \rightarrow \infty$.*

Theorem 2. *The function $e_p(\theta)$ is monotone increasing on $[0, \pi]$ for all $p \geq 1$.*

As a by-product of the proofs of Theorems 1 and 2, in Appendix A, we also prove the following result for the function

$$w_p : [0, \pi] \rightarrow \mathbb{R}, \quad w_p(\theta) = \frac{g_p(\theta)}{g_{p-1}(\theta)}, \quad p \geq 1.$$

Theorem 3. *For $p \geq 1$ and $\theta \in [0, \pi]$, we have*

$$\frac{1}{3} \leq w_p(\theta) \leq 1. \quad (23)$$

Note that the bounds in (23) are sharp. Indeed, $w_p(0) = 1$ for all $p \geq 1$ and $w_1(\pi) = 1/3$. Theorem 3 provides theoretical support to the numerically observed p -robustness of the solvers devised by Donatelli et al.^{14,16} for IgA linear systems; see, in particular, section 5.5 in the work of Donatelli et al.¹⁴

3 | EIGENVALUES AND EIGENVECTORS OF $L_n^{[p]}$ FOR $p=1$ AND $p=2$

In this section, we compute the exact spectral decomposition of the matrix $L_n^{[p]}$ for $p = 1$ and $p = 2$. As a preliminary step, we recall some properties of the matrix algebras $\tau_n(\epsilon, \phi)$ introduced by Bozzo et al.¹⁹ for $\epsilon, \phi \in \{0, 1, -1\}$. It will turn out that $K_n^{[1]}, M_n^{[1]}, L_n^{[1]}$ belong to $\tau_{n-1}(0, 0)$, and $K_n^{[2]}, M_n^{[2]}, L_n^{[2]}$ belong to $\tau_n(-1, -1)$, and this will be the key for computing the eigenvalues and eigenvectors of both $L_n^{[1]}$ and $L_n^{[2]}$.

3.1 | The matrix algebras $\tau_m(\epsilon, \phi)$ for $\epsilon, \phi \in \{0, 1, -1\}$

Following the work of Bozzo et al.,¹⁹ for any $m \geq 2$ and any $\epsilon, \phi \in \{0, 1, -1\}$, we define the tridiagonal matrix

$$H_m(\epsilon, \phi) = \begin{bmatrix} \epsilon & 1 & 0 & \cdots & 0 \\ 1 & 0 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 0 & 1 \\ 0 & \cdots & 0 & 1 & \phi \end{bmatrix} = T_m(2 \cos(\theta)) + \epsilon \mathbf{e}_1 \mathbf{e}_1^T + \phi \mathbf{e}_m \mathbf{e}_m^T,$$

where \mathbf{e}_i is the i th vector of the canonical basis of \mathbb{R}^m . Because $H_m(\epsilon, \phi)$ is real and symmetric, it can be decomposed as

$$H_m(\epsilon, \phi) = Q_m(\epsilon, \phi)D_m(\epsilon, \phi)Q_m(\epsilon, \phi)^T,$$

where $Q_m(\epsilon, \phi)$ is a real unitary matrix, and $D_m(\epsilon, \phi)$ is a real diagonal matrix. The matrix algebra generated by $H_m(\epsilon, \phi)$ is denoted by $\tau_m(\epsilon, \phi)$ and is given by

$$\tau_m(\epsilon, \phi) = \{Q_m(\epsilon, \phi)D_mQ_m(\epsilon, \phi)^T : D_m \text{ is a diagonal matrix of size } m\}.$$

It turns out that the matrix $Q_m(\epsilon, \phi)$ is a fast trigonometric transform such that the matrix–vector product $Q_m(\epsilon, \phi)\mathbf{v}$ can be computed in $O(m \log m)$ operations. Moreover, the diagonal entries of the matrix $D_m(\epsilon, \phi)$ (i.e., the eigenvalues of $H_m(\epsilon, \phi)$) are equal to the samples of the function $2 \cos(\theta)$ over a uniform grid in $[0, \pi]$.

The cases of interest in this paper are $\epsilon = \phi = 0$ and $\epsilon = \phi = -1$. For $\epsilon = \phi = 0$, the matrix algebra $\tau_m(0, 0)$ is the so-called tau algebra, which was originally introduced in the work of Bini et al.²⁹ In this case, the sampling grid is

$$\frac{j\pi}{m+1}, \quad j = 1, \dots, m,$$

and we have

$$D_m(0, 0) = \text{diag}_{j=1, \dots, m} \left[2 \cos \left(\frac{j\pi}{m+1} \right) \right],$$

$$Q_m(0, 0) = \sqrt{\frac{2}{m+1}} \left[\sin \left(\frac{ij\pi}{m+1} \right) \right]_{i,j=1}^m.$$

For $\epsilon = \phi = -1$, the sampling grid is

$$\frac{j\pi}{m}, \quad j = 1, \dots, m,$$

and we have

$$D_m(-1, -1) = \text{diag}_{j=1, \dots, m} \left[2 \cos \left(\frac{j\pi}{m} \right) \right],$$

$$Q_m(-1, -1) = \sqrt{\frac{2}{m}} \left[k_j \sin \left(\frac{(2i-1)j\pi}{2m} \right) \right]_{i,j=1}^m, \quad k_j = \begin{cases} 1/\sqrt{2}, & \text{if } j = m, \\ 1, & \text{otherwise.} \end{cases}$$

For more details on the matrix algebras $\tau_m(\epsilon, \phi)$, we refer the reader to the work of Bozzo et al.¹⁹

3.2 | Eigenvalues and eigenvectors of $L_n^{[p]}$ for $p = 1$

In the case $p = 1$, the stiffness and mass matrices $K_n^{[1]}$ and $M_n^{[1]}$ have size $n - 1$, and a direct computation shows that

$$n^{-1}K_n^{[1]} = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} = T_{n-1}(f_1) = 2I_{n-1} - H_{n-1}(0, 0),$$

$$nM_n^{[1]} = \frac{1}{6} \begin{bmatrix} 4 & 1 & & & \\ 1 & 4 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 4 & 1 \\ & & & 1 & 4 \end{bmatrix} = T_{n-1}(g_1) = \frac{2}{3}I_{n-1} + \frac{1}{6}H_{n-1}(0, 0),$$

where I_m is the $m \times m$ identity matrix, and f_1, g_1 are given by (13)–(14) for $p = 1$, that is,

$$f_1(\theta) = 2 - 2 \cos(\theta),$$

$$g_1(\theta) = \frac{2}{3} + \frac{1}{3} \cos(\theta).$$

and $R_n^{[2]}, S_n^{[2]}$ are matrices of rank 4 given by

$$R_n^{[2]} = \frac{1}{6} \begin{bmatrix} 2 & 1 & & \\ 1 & & & \\ & & 1 & \\ & & & 2 \end{bmatrix},$$

$$S_n^{[2]} = \frac{1}{120} \begin{bmatrix} -26 & -1 & & \\ -1 & & & \\ & & & -1 \\ & & & -1 & -26 \end{bmatrix}.$$

We note that both $n^{-1}K_n^{[2]}$ and $nM_n^{[2]}$ are of the form

$$A_n(a, b, c) = T_n(a + 2b \cos(\theta) + 2c \cos(2\theta)) + R_n(b, c), \quad R_n(b, c) = - \begin{bmatrix} b & c & & \\ c & & & \\ & & c & \\ & & & b \end{bmatrix}. \quad (24)$$

Indeed,

$$n^{-1}K_n^{[2]} = A_n\left(1, -\frac{1}{3}, -\frac{1}{6}\right),$$

$$nM_n^{[2]} = A_n\left(\frac{11}{20}, \frac{13}{60}, \frac{1}{120}\right).$$

Now, any matrix of the form (24) is a polynomial in $H_n(-1, -1)$, and precisely,

$$A_n(a, b, c) = (a - 2c)I_n + bH_n(-1, -1) + cH_n(-1, -1)^2.$$

It follows that $A_n(a, b, c)$ belongs to the matrix algebra $\tau_n(-1, -1)$. Moreover, based on the results of Section 3.1, we have

$$A_n(a, b, c) = Q_n(-1, -1) \left(\text{diag}_{j=1, \dots, n} \left[a + 2b \cos\left(\frac{j\pi}{n}\right) + 2c \cos\left(\frac{2j\pi}{n}\right) \right] \right) Q_n(-1, -1)^T.$$

In particular, $K_n^{[2]}$ and $M_n^{[2]}$ belong to $\tau_n(-1, -1)$ and

$$n^{-1}K_n^{[2]} = Q_n(-1, -1) \left(\text{diag}_{j=1, \dots, n} \left[f_2\left(\frac{j\pi}{n}\right) \right] \right) Q_n(-1, -1)^T,$$

$$nM_n^{[2]} = Q_n(-1, -1) \left(\text{diag}_{j=1, \dots, n} \left[g_2\left(\frac{j\pi}{n}\right) \right] \right) Q_n(-1, -1)^T.$$

Given the algebra structure of $\tau_n(-1, -1)$, we obtain

$$n^{-2}L_n^{[2]} = (nM_n^{[2]})^{-1} (n^{-1}K_n^{[2]}) = Q_n(-1, -1) \left(\text{diag}_{j=1, \dots, n} \left[e_2\left(\frac{j\pi}{n}\right) \right] \right) Q_n(-1, -1)^T,$$

where

$$e_2(\theta) = \frac{f_2(\theta)}{g_2(\theta)} = \frac{20(3 - 2\cos(\theta)) - \cos(2\theta)}{33 + 26\cos(\theta) + \cos(2\theta)},$$

as defined by (15) for $p = 2$. In particular, $L_n^{[2]}$ belongs to the algebra $\tau_n(-1, -1)$ just like $K_n^{[2]}$ and $M_n^{[2]}$, and the eigenvalues and eigenvectors of $L_n^{[2]}$ are given by

$$n^2 e_2\left(\frac{j\pi}{n}\right), \quad j = 1, \dots, n,$$

$$\sqrt{\frac{2}{n}} \left[k_j \sin\left(\frac{(2i-1)j\pi}{2n}\right) \right]_{i=1}^n, \quad k_j = \begin{cases} 1/\sqrt{2}, & \text{if } j = n, \\ 1, & \text{otherwise,} \end{cases} \quad j = 1, \dots, n.$$

Remark 1. According to a private communication received by the fifth author, Tani proposed a preconditioner based on the fast sine transform $Q_n(-1, -1)$ for solving linear systems arising from the IgA discretization of unidimensional

differential problems. For the case $p = 2$, the performance of the preconditioner was extremely good: just one Krylov iteration. The theoretical explanation of such an excellent behavior lies precisely in the exact spectral decompositions obtained in this subsection, where it is shown that $Q_n(-1, -1)$ diagonalizes simultaneously the three matrices $K_n^{[2]}, M_n^{[2]}, L_n^{[2]}$. Note that decompositions of this kind can also be used for accelerating the convergence of the recently proposed iterative solvers for IgA linear systems, such as multigrid-based and preconditioned Krylov-based methods; see the works^{14,16,30} and the references therein.

Remark 2. The results of Sections 3.2 and 3.3 show that $K_n^{[p]}, M_n^{[p]}, L_n^{[p]}$ belong to the same matrix algebra for $p = 1, 2$. Does this property remains true for $p \geq 3$? The answer is “no”. Indeed, if $K_n^{[p]}, M_n^{[p]}, L_n^{[p]}$ belong to the same matrix algebra, then $K_n^{[p]}$ and $M_n^{[p]}$ commute. We numerically verified that $K_n^{[p]}$ and $M_n^{[p]}$ do not commute for $p \geq 3$.

4 | ALGORITHM FOR COMPUTING THE EIGENVALUES OF $L_n^{[p]}$ FOR $p \geq 3$

Assuming the expansion (18) and drawing inspiration from the work of Ekström et al.,²⁷ in this section, we propose a parallel interpolation–extrapolation algorithm for computing the eigenvalues of $L_n^{[p]}$, excluding the n_p^{out} outliers. In what follows, for each positive integer $n \in \mathbb{N} = \{1, 2, 3, \dots\}$ and each $p \geq 3$, we define $n^{[p]} = n - \text{mod}(p, 2)$. Moreover, with each positive integer n , we associate the step size $h = \frac{1}{n}$ and the grid points $\theta_{j,n} = j\pi h, j = 1, \dots, n$. For notational convenience, unless otherwise stated, we will always denote a positive integer and the associated step size in the same way. For example, if the positive integer is n , the associated step size is h ; if the positive integer is n_1 , the associated step size is h_1 ; if the positive integer is \bar{n} , the associated step size is \bar{h} ; etc. Throughout this section, we make the following assumptions:

- $p \geq 3$ and $n, n_1, \alpha \in \mathbb{N}$ are fixed parameters.
- $n_k = 2^{k-1}n_1$ for $k = 1, \dots, \alpha$.
- $j_k = 2^{k-1}j_1$ for $j_1 = 1, \dots, n_1$ and $k = 1, \dots, \alpha$; j_k is the index in $\{1, \dots, n_k\}$ such that $\theta_{j_k, n_k} = \theta_{j_1, n_1}$.

A graphical representation of the grids $\{\theta_{1, n_k}, \dots, \theta_{n_k, n_k}\}, k = 1, \dots, \alpha$, is reported in Figure 3 for $n_1 = 5$ and $\alpha = 4$.

For each fixed $j_1 = 1, \dots, n_1^{[p]}$, we apply α times the expansion (18) with $n = n_1, n_2, \dots, n_\alpha$ and $j = j_1, j_2, \dots, j_\alpha$. Because $\theta_{j_1, n_1} = \theta_{j_2, n_2} = \dots = \theta_{j_\alpha, n_\alpha}$ (by definition of j_2, \dots, j_α), we obtain

$$\begin{cases} E_{j_1, n_1, 0}^{[p]} = c_1^{[p]}(\theta_{j_1, n_1})h_1 + c_2^{[p]}(\theta_{j_1, n_1})h_1^2 + \dots + c_\alpha^{[p]}(\theta_{j_1, n_1})h_1^\alpha + E_{j_1, n_1, \alpha}^{[p]} \\ E_{j_2, n_2, 0}^{[p]} = c_1^{[p]}(\theta_{j_1, n_1})h_2 + c_2^{[p]}(\theta_{j_1, n_1})h_2^2 + \dots + c_\alpha^{[p]}(\theta_{j_1, n_1})h_2^\alpha + E_{j_2, n_2, \alpha}^{[p]} \\ \vdots \\ E_{j_\alpha, n_\alpha, 0}^{[p]} = c_1^{[p]}(\theta_{j_1, n_1})h_\alpha + c_2^{[p]}(\theta_{j_1, n_1})h_\alpha^2 + \dots + c_\alpha^{[p]}(\theta_{j_1, n_1})h_\alpha^\alpha + E_{j_\alpha, n_\alpha, \alpha}^{[p]} \end{cases}, \quad (25)$$

where

$$E_{j_k, n_k, 0}^{[p]} = \lambda_{j_k} \left(n_k^{-2} L_{n_k}^{[p]} \right) - e_p(\theta_{j_1, n_1}), \quad k = 1, \dots, \alpha,$$

and

$$\left| E_{j_k, n_k, \alpha}^{[p]} \right| \leq C_\alpha^{[p]} h_k^{\alpha+1}, \quad k = 1, \dots, \alpha. \quad (26)$$

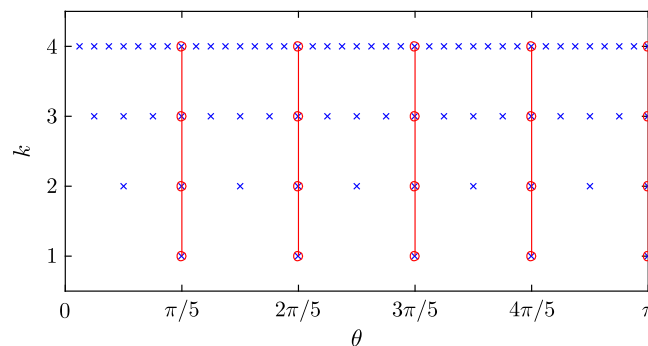


FIGURE 3 Representation of the grids $\{\theta_{1, n_k}, \dots, \theta_{n_k, n_k}\}, k = 1, \dots, \alpha$, for $n_1 = 5$ and $\alpha = 4$

Let $\tilde{c}_1^{[p]}(\theta_{j_1, n_1}), \dots, \tilde{c}_\alpha^{[p]}(\theta_{j_1, n_1})$ be the approximations of $c_1^{[p]}(\theta_{j_1, n_1}), \dots, c_\alpha^{[p]}(\theta_{j_1, n_1})$ obtained by removing all the errors $E_{j_1, n_1, \alpha}^{[p]}, \dots, E_{j_\alpha, n_\alpha, \alpha}^{[p]}$ in (25) and by solving the resulting linear system as follows:

$$\begin{cases} E_{j_1, n_1, 0}^{[p]} = \tilde{c}_1^{[p]}(\theta_{j_1, n_1})h_1 + \tilde{c}_2^{[p]}(\theta_{j_1, n_1})h_1^2 + \dots + \tilde{c}_\alpha^{[p]}(\theta_{j_1, n_1})h_1^\alpha \\ E_{j_2, n_2, 0}^{[p]} = \tilde{c}_1^{[p]}(\theta_{j_1, n_1})h_2 + \tilde{c}_2^{[p]}(\theta_{j_1, n_1})h_2^2 + \dots + \tilde{c}_\alpha^{[p]}(\theta_{j_1, n_1})h_2^\alpha \\ \vdots \\ E_{j_\alpha, n_\alpha, 0}^{[p]} = \tilde{c}_1^{[p]}(\theta_{j_1, n_1})h_\alpha + \tilde{c}_2^{[p]}(\theta_{j_1, n_1})h_\alpha^2 + \dots + \tilde{c}_\alpha^{[p]}(\theta_{j_1, n_1})h_\alpha^\alpha \end{cases} \quad (27)$$

Note that this way of computing approximations for $c_1^{[p]}(\theta_{j_1, n_1}), \dots, c_\alpha^{[p]}(\theta_{j_1, n_1})$ is completely analogous to the Richardson extrapolation procedure that is employed in the context of Romberg integration to accelerate the convergence of the trapezoidal rule (see section 3.4 in the work of Stoer et al.³¹). In this regard, the asymptotic expansion (18) plays here the same role as the Euler–Maclaurin summation formula (see section 3.3 in the work of Stoer et al.³¹). For more advanced studies on extrapolation methods, we refer the reader to Brezinski et al.³² The next theorem shows that the approximation error $|c_k^{[p]}(\theta_{j_1, n_1}) - \tilde{c}_k^{[p]}(\theta_{j_1, n_1})|$ is $O(h_1^{\alpha-k+1})$.

Theorem 4. *There exists a constant $A_\alpha^{[p]}$ depending only on α and p such that, for $j_1 = 1, \dots, n_1^{[p]}$ and $k = 1, \dots, \alpha$,*

$$\left| c_k^{[p]}(\theta_{j_1, n_1}) - \tilde{c}_k^{[p]}(\theta_{j_1, n_1}) \right| \leq A_\alpha^{[p]} h_1^{\alpha-k+1}. \quad (28)$$

Proof. It is a straightforward adaptation of the proof of theorem 1 in the work of Ekström et al.²⁷ □

Now, fix an index $j \in \{1, \dots, n^{[p]}\}$. To compute an approximation of $\lambda_j(n^{-2}L_n^{[p]})$ through the expansion (18), we would need the value $c_k^{[p]}(\theta_{j,n})$ for each $k = 1, \dots, \alpha$. Of course, $c_k^{[p]}(\theta_{j,n})$ is not available in practice, but we can approximate it by interpolating in some way the values $\tilde{c}_k^{[p]}(\theta_{j_1, n_1}), j_1 = 1, \dots, n_1^{[p]}$. For example, we may define $\tilde{c}_k^{[p]}(\theta)$ as the interpolation polynomial of the data $(\theta_{j_1, n_1}, \tilde{c}_k^{[p]}(\theta_{j_1, n_1})), j_1 = 1, \dots, n_1^{[p]}$,—so that $\tilde{c}_k^{[p]}(\theta)$ is expected to be an approximation of $c_k^{[p]}(\theta)$ over the whole interval $[0, \pi]$ —and take $\tilde{c}_k^{[p]}(\theta_{j,n})$ as an approximation to $c_k^{[p]}(\theta_{j,n})$. It is known, however, that interpolation over a large number of uniform nodes is not advisable as it may give rise to spurious oscillations (Runge's phenomenon). It is therefore better to adopt another kind of approximation. An alternative could be the following: We approximate $c_k^{[p]}(\theta)$ by the spline function $\tilde{c}_k^{[p]}(\theta)$ which is linear on each interval $[\theta_{j_1, n_1}, \theta_{j_1+1, n_1}]$ and takes the value $\tilde{c}_k^{[p]}(\theta_{j_1, n_1})$ at θ_{j_1, n_1} for all $j_1 = 1, \dots, n_1^{[p]}$. This strategy usually removes any spurious oscillation, but it is not accurate. In particular, it does not preserve the accuracy of approximation at the nodes θ_{j_1, n_1} established in Theorem 4, that is, there is no guarantee that $|c_k^{[p]}(\theta) - \tilde{c}_k^{[p]}(\theta)| \leq B_\alpha^{[p]} h_1^{\alpha-k+1}$ for $\theta \in [0, \pi]$ or $|c_k^{[p]}(\theta_{j,n}) - \tilde{c}_k^{[p]}(\theta_{j,n})| \leq B_\alpha^{[p]} h_1^{\alpha-k+1}$ for $j = 1, \dots, n^{[p]}$, with $B_\alpha^{[p]}$ being a constant depending only on α and p . As proved in Theorem 5, a local approximation strategy that preserves the accuracy (28), at least if $c_k^{[p]}(\theta)$ is sufficiently smooth, is the following: Let $\theta^{(1)}, \dots, \theta^{(\alpha-k+1)}$ be $\alpha - k + 1$ points of the grid $\{\theta_{1, n_1}, \dots, \theta_{n_1^{[p]}, n_1}\}$ which are closest to the point $\theta_{j,n}^\dagger$, and let $\tilde{c}_{k,j}^{[p]}(\theta)$ be the interpolation polynomial of the data $(\theta^{(1)}, \tilde{c}_k^{[p]}(\theta^{(1)})), \dots, (\theta^{(\alpha-k+1)}, \tilde{c}_k^{[p]}(\theta^{(\alpha-k+1)}))$; then, we approximate $c_k^{[p]}(\theta_{j,n})$ by $\tilde{c}_{k,j}^{[p]}(\theta_{j,n})$. Note that, by selecting $\alpha - k + 1$ points from $\{\theta_{1, n_1}, \dots, \theta_{n_1^{[p]}, n_1}\}$, we are implicitly assuming that $n_1^{[p]} \geq \alpha - k + 1$.

Theorem 5. *Let $p \geq 3$ and $1 \leq k \leq \alpha$, and suppose $n_1^{[p]} \geq \alpha - k + 1$ and $c_k^{[p]} \in C^{\alpha-k+1}[0, \pi]$. For $j = 1, \dots, n^{[p]}$, if $\theta^{(1)}, \dots, \theta^{(\alpha-k+1)}$ are $\alpha - k + 1$ points of $\{\theta_{1, n_1}, \dots, \theta_{n_1^{[p]}, n_1}\}$ which are closest to $\theta_{j,n}$, and if $\tilde{c}_{k,j}^{[p]}(\theta)$ is the interpolation polynomial of the data $(\theta^{(1)}, \tilde{c}_k^{[p]}(\theta^{(1)})), \dots, (\theta^{(\alpha-k+1)}, \tilde{c}_k^{[p]}(\theta^{(\alpha-k+1)}))$, then*

$$\left| c_k^{[p]}(\theta_{j,n}) - \tilde{c}_{k,j}^{[p]}(\theta_{j,n}) \right| \leq B_\alpha^{[p]} h_1^{\alpha-k+1} \quad (29)$$

for some constant $B_\alpha^{[p]}$ depending only on α and p .

Proof. It is a straightforward adaptation of the proof of theorem 2 in the work of Ekström et al.²⁷ □

[†]These $\alpha - k + 1$ points are uniquely determined by $\theta_{j,n}$ except in the following two cases: (a) $\theta_{j,n}$ coincides with a grid point θ_{j_1, n_1} , and $\alpha - k + 1$ is even; (b) $\theta_{j,n}$ coincides with the midpoint between two consecutive grid points $\theta_{j_1, n_1}, \theta_{j_1+1, n_1}$, and $\alpha - k + 1$ is odd.

We are now ready to formulate our algorithm for computing the eigenvalues of $L_n^{[p]}$, excluding the outliers. Note that this algorithm is suited for parallel implementation for the same reason why algorithm 1 in the work of Ekström et al.²⁷ is; see remark 4 in the work of Ekström et al.²⁷

Algorithm 1. Given $p \geq 3$ and $n, n_1, \alpha \in \mathbb{N}$ with $n_1^{[p]} \geq \alpha$, we compute the approximations of the eigenvalues $\lambda_j(L_n^{[p]})$ for $j = 1, \dots, n^{[p]}$ as follows.

1. For $j_1 = 1, \dots, n_1^{[p]}$, compute $\tilde{c}_1^{[p]}(\theta_{j_1, n_1}), \dots, \tilde{c}_\alpha^{[p]}(\theta_{j_1, n_1})$ by solving (27).
2. For $j = 1, \dots, n^{[p]}$
 - for $k = 1, \dots, \alpha$,
 - determine $\alpha - k + 1$ points $\theta^{(1)}, \dots, \theta^{(\alpha-k+1)} \in \{\theta_{1, n_1}, \dots, \theta_{n_1^{[p]}, n_1}\}$ which are closest to $\theta_{j, n}$;
 - compute $\tilde{c}_{k, j}^{[p]}(\theta_{j, n})$, where $\tilde{c}_{k, j}^{[p]}(\theta)$ is the interpolation polynomial of the data

$$\left(\theta^{(1)}, \tilde{c}_k^{[p]}(\theta^{(1)})\right), \dots, \left(\theta^{(\alpha-k+1)}, \tilde{c}_k^{[p]}(\theta^{(\alpha-k+1)})\right);$$
 - compute $\tilde{\lambda}_j(n^{-2}L_n^{[p]}) = e_p(\theta_{j, n}) + \sum_{k=1}^\alpha \tilde{c}_{k, j}^{[p]}(\theta_{j, n})h^k$ and $\tilde{\lambda}_j(L_n^{[p]}) = n^2 \tilde{\lambda}_j(n^{-2}L_n^{[p]})$.
3. Return $(\tilde{\lambda}_1(L_n^{[p]}), \dots, \tilde{\lambda}_{n^{[p]}}(L_n^{[p]}))$ as an approximation to $(\lambda_1(L_n^{[p]}), \dots, \lambda_{n^{[p]}}(L_n^{[p]}))$.

Remark 3. Algorithm 1 is specifically designed for computing the eigenvalues of $L_n^{[p]}$ in the case where n is quite large. When applying this algorithm, it is implicitly assumed that n_1 and α are small (much smaller than n), so that each $n_k = 2^{k-1}n_1$ is small as well and that the computation of the eigenvalues of $L_{n_k}^{[p]}$ —which is required in the first step—can be efficiently performed by any standard eigensolver (e.g., the MATLAB `eig` function).

The last theorem of this section provides an estimate for the approximation error made by Algorithm 1.

Theorem 6. Let $p \geq 3$, $n^{[p]} \geq n_1^{[p]} \geq \alpha$ and $c_k^{[p]} \in C^{\alpha-k+1}[0, \pi]$ for $k = 1, \dots, \alpha$. Let $(\tilde{\lambda}_1(L_n^{[p]}), \dots, \tilde{\lambda}_{n^{[p]}}(L_n^{[p]}))$ be the approximation of $(\lambda_1(L_n^{[p]}), \dots, \lambda_{n^{[p]}}(L_n^{[p]}))$ computed by Algorithm 1. Then, there exists a constant $D_\alpha^{[p]}$ depending only on α and p such that, for $j = 1, \dots, n^{[p]}$,

$$\left| \lambda_j \left(L_n^{[p]} \right) - \tilde{\lambda}_j \left(L_n^{[p]} \right) \right| \leq D_\alpha^{[p]} n h^\alpha. \tag{30}$$

Proof. By (18) and Theorem 5,

$$\begin{aligned} \left| \lambda_j \left(n^{-2} L_n^{[p]} \right) - \tilde{\lambda}_j \left(n^{-2} L_n^{[p]} \right) \right| &= \left| e_p(\theta_{j, n}) + \sum_{k=1}^\alpha c_k^{[p]}(\theta_{j, n})h^k + E_{j, n, \alpha}^{[p]} - e_p(\theta_{j, n}) - \sum_{k=1}^\alpha \tilde{c}_{k, j}^{[p]}(\theta_{j, n})h^k \right| \\ &\leq \sum_{k=1}^\alpha \left| c_k^{[p]}(\theta_{j, n}) - \tilde{c}_{k, j}^{[p]}(\theta_{j, n}) \right| h^k + \left| E_{j, n, \alpha}^{[p]} \right| \\ &\leq B_\alpha^{[p]} \sum_{k=1}^\alpha h_1^{\alpha-k+1} h^k + C_\alpha^{[p]} h^{\alpha+1} \leq D_\alpha^{[p]} h_1^\alpha h, \end{aligned}$$

where $D_\alpha^{[p]} = (\alpha + 1) \max(B_\alpha^{[p]}, C_\alpha^{[p]})$. Multiplying both sides by n^2 , we get the thesis. □

Note that the error estimate provided by Theorem 6 seems disappointing, due to the presence of the large factor n on the right-hand side of (30). However, one should take into account that (30) is an absolute error estimate that, moreover, is uniform in j . Considering that the largest nonoutlier eigenvalue of $L_n^{[p]}$, namely $\lambda_{n^{[p]}}(L_n^{[p]})$, diverges to ∞ with the same asymptotic speed as n^2 , from (30), we obtain the approximate inequality

$$\frac{\left| \lambda_{n^{[p]}} \left(L_n^{[p]} \right) - \tilde{\lambda}_{n^{[p]}} \left(L_n^{[p]} \right) \right|}{\left| \lambda_{n^{[p]}} \left(L_n^{[p]} \right) \right|} \leq D_\alpha^{[p]} h_1^\alpha h,$$

which is a good relative error estimate. We refer the reader to Section 5.2 for several numerical illustrations of the actual performance of Algorithm 1.

5 | NUMERICAL EXPERIMENTS

In Section 5.1, we implement the program described in items 3 and 4 of Section 1.2. In other words, we validate through numerical experiments the expansion (18) for $p \geq 3$; we numerically show, for $p \geq 3$ and $k \geq 1$, the existence of a point $\theta(p, k) \in (0, \pi)$ such that $c_k^{[p]}(\theta)$ vanishes over $[0, \theta(p, k)]$; and we provide numerical evidence of the fact that the infimum $y_p = \inf_{k \geq 1} \theta(p, k)$ is strictly positive and that the equation $\lambda_j(n^{-2}L_n^{[p]}) = e_p(\theta_{j,n})$ holds numerically whenever $\theta_{j,n} < \theta(p)$, with $\theta(p)$ being a point in $(0, y_p]$. In Section 5.2, we illustrate the numerical performance of Algorithm 1.

5.1 | Numerical experiments in support of the eigenvalue expansion

Fix $p \geq 3$ and $\alpha \in \mathbb{N}$. As in Section 4, for every $n_1 \in \mathbb{N}$, we set

$$\begin{aligned} n_k &= 2^{k-1} n_1, & k &= 1, \dots, \alpha, \\ j_k &= 2^{k-1} j_1, & k &= 1, \dots, \alpha, & j_1 &= 1, \dots, n_1. \end{aligned}$$

In the hypothesis that the expansion (18) holds, we can follow the derivation of Section 4 until Theorem 4 and we conclude that, for each $k = 1, \dots, \alpha$ and $j_1 = 1, \dots, n_1^{[p]}$, the value $\tilde{c}_k^{[p]}(\theta_{j_1, n_1})$ computed by solving the linear system (27) converges to the value $c_k^{[p]}(\theta_{j_1, n_1})$ as $n_1 \rightarrow \infty$ with the same asymptotic speed as $h_1^{\alpha-k+1}$. In other words, in the hypothesis that the expansion (18) holds, if we plot the values $\tilde{c}_k^{[p]}(\theta_{j_1, n_1})$ versus the points θ_{j_1, n_1} for $j_1 = 1, \dots, n_1^{[p]}$, the resulting picture should converge as $n_1 \rightarrow \infty$ to the graph of a function from $[0, \pi]$ to \mathbb{R} , which is, by definition, $c_k^{[p]}(\theta)$. The next examples show that this is in fact the case, thus providing a validation of the expansion (18). The examples also support the following conjectures:

- The limit function $c_k^{[p]}(\theta)$ vanishes over an interval $[0, \theta(p, k)]$ with $\theta(p, k) \in (0, \pi)$;
- $y_p = \inf_{k \geq 1} \theta(p, k) > 0$;
- $\lambda_j(n^{-2}L_n^{[p]}) = e_p(\theta_{j,n})$ numerically whenever $\theta_{j,n} < \theta(p)$, where $\theta(p)$ is a point in $(0, y_p]$, which grows with p .

Example 1. Fix $p = 3$ and let $\alpha = 3$. In Figure 4, we plot the pairs

$$(\theta_{j_1, n_1}, \tilde{c}_k^{[3]}(\theta_{j_1, n_1})), \quad j_1 = 1, \dots, n_1^{[3]} = n_1 - 1, \quad (31)$$

for $n_1 = 200, 300, 400$ and $k = 1, 2, 3$. We note that, for each fixed k , the graph of the pairs (31) is essentially the same for all the considered values of n_1 . In other words, this graph converges to the graph of a function $c_k^{[3]}(\theta)$ as $n_1 \rightarrow \infty$, and the convergence is essentially reached already for $n_1 = 200$, at least from the point of view of graphical visualization. Moreover, the limit function $c_k^{[3]}(\theta)$ is apparently zero over an interval $[0, \theta(3, k)]$, where $\theta(3, k) \in (0, \pi)$. An ε -approximation of $\theta(3, k)$ is obtained as the limit of $\theta_{n_1}^{(\varepsilon)}(3, k)$ for $n_1 \rightarrow \infty$, where

$$\theta_{n_1}^{(\varepsilon)}(3, k) = \max \left\{ \theta_{j_1, n_1} : 1 \leq j_1 \leq n_1 - 1, \quad \left| \tilde{c}_k^{[3]}(\theta_{i_1, n_1}) \right| \leq \varepsilon \text{ for all } i_1 < j_1 \right\},$$

and ε is a fixed threshold. Table 1 shows the values $\theta_{n_1}^{(\varepsilon)}(3, k)$ computed for $k = 1, 2, 3$ and $n_1 = 200, 300, 400, 500, 600$ with the fixed threshold $\varepsilon = 0.0005$. Both Figure 4 and Table 1 suggest that $\theta(3, k)$ grows with k . In particular, we may expect that

$$y_3 = \inf_{k \geq 1} \theta(3, k) = \theta(3, 1) > 0.$$

In Figure 5, we plot the errors $|E_{j,n,0}^{[3]}| = |\lambda_j(n^{-2}L_n^{[3]}) - e_3(\theta_{j,n})|$ versus the points $\theta_{j,n}$ for $j = 1, \dots, n^{[3]} = n - 1$ and $n = 750, 1000, 1250, 1500$. For the same values of n , in Table 2, we record the first index j such that $|E_{j,n,0}^{[3]}| > 10^{-14}$ and the corresponding grid point $\theta_{j,n}$. From Figure 5 and Table 2, we immediately see that a nontrivial portion of the spectrum of $n^{-2}L_n^{[3]}$ is exactly approximated, at least from a numerical viewpoint, by the spectral symbol $e_3(\theta)$. Moreover, the points $\theta_{j,n}$ shown in Table 2 apparently form a monotone increasing sequence; the limit of this sequence as $n \rightarrow \infty$, say $\theta(3) \approx 0.2576$, is a point such that the equation $\lambda_i(n^{-2}L_n^{[3]}) = e_3(\theta_{i,n})$ holds numerically whenever $\theta_{i,n} < \theta(3)$. In other words, the ratio $\theta(3)/\pi \approx 0.082$ represents the portion of the spectrum of $n^{-2}L_n^{[3]}$ which is exactly described by $e_3(\theta)$, at least numerically.

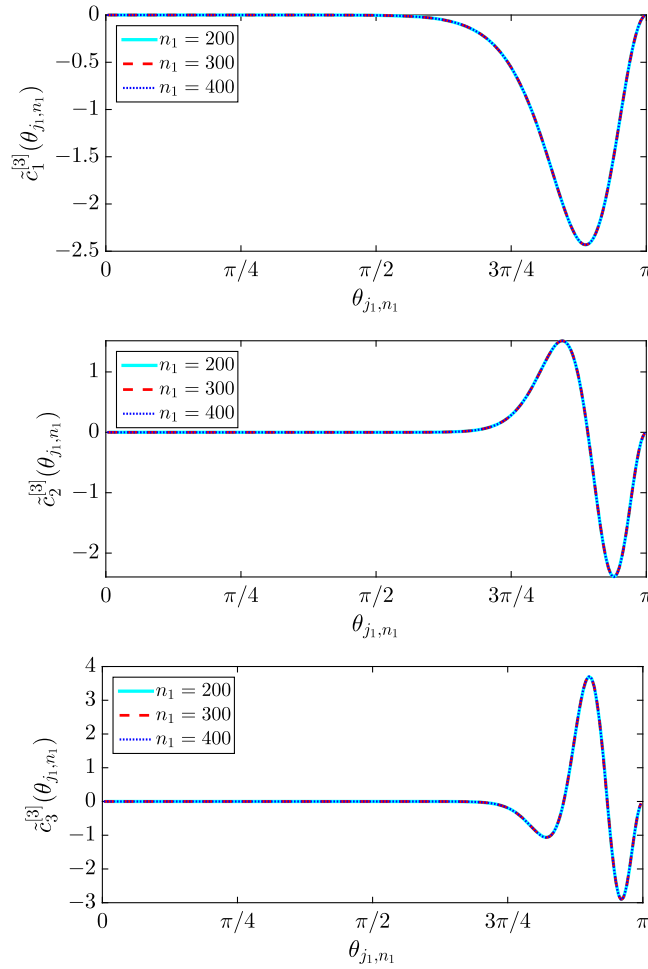


FIGURE 4 Example 1, $p = 3$: graph of the pairs $(\theta_{j_1, n_1}, \tilde{c}_k^{[3]}(\theta_{j_1, n_1}))$, $j_1 = 1, \dots, n_1 - 1$, for $n_1 = 200, 300, 400$ and $k = 1, 2, 3$

TABLE 1 Example 1, $p = 3$: values $\theta_{n_1}^{(\epsilon)}(3, k)$ for $k = 1, 2, 3$ and $n_1 = 200, 300, 400, 500, 600$, computed with the threshold $\epsilon = 0.0005$

n_1	200	300	400	500	600
$\theta_{n_1}^{(\epsilon)}(3, 1)$	$\frac{86\pi}{200} \approx 1.3509$	$\frac{129\pi}{300} \approx 1.3509$	$\frac{172\pi}{400} \approx 1.3509$	$\frac{214\pi}{500} \approx 1.3446$	$\frac{257\pi}{600} \approx 1.3456$
$\theta_{n_1}^{(\epsilon)}(3, 2)$	$\frac{115\pi}{200} \approx 1.8064$	$\frac{172\pi}{300} \approx 1.8012$	$\frac{229\pi}{400} \approx 1.7986$	$\frac{286\pi}{500} \approx 1.7970$	$\frac{343\pi}{600} \approx 1.7959$
$\theta_{n_1}^{(\epsilon)}(3, 3)$	$\frac{126\pi}{200} \approx 1.9792$	$\frac{188\pi}{300} \approx 1.9687$	$\frac{251\pi}{400} \approx 1.9713$	$\frac{313\pi}{500} \approx 1.9666$	$\frac{377\pi}{600} \approx 1.9740$

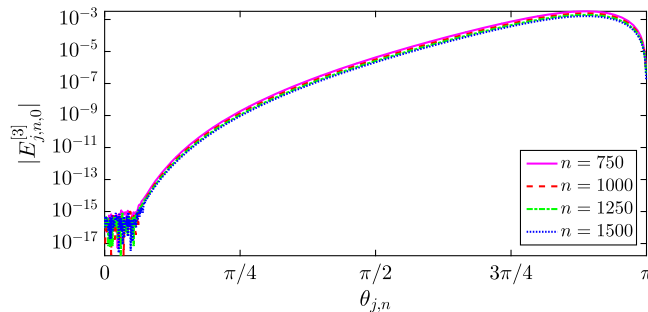


FIGURE 5 Example 1, $p = 3$: errors $|E_{j,n}^{[3]}|$ versus $\theta_{j,n}$ for $j = 1, \dots, n - 1$ and $n = 750, 1000, 1250, 1500$

TABLE 2 Example 1, $p = 3$: first index j such that $|E_{j,n,0}^{[3]}| > 10^{-14}$ and corresponding grid point $\theta_{j,n}$, for $n = 750, 1000, 1250, 1500$

n	750	1000	1250	1500
j	58	80	101	123
$\theta_{j,n}$	0.2429	0.2513	0.2538	0.2576

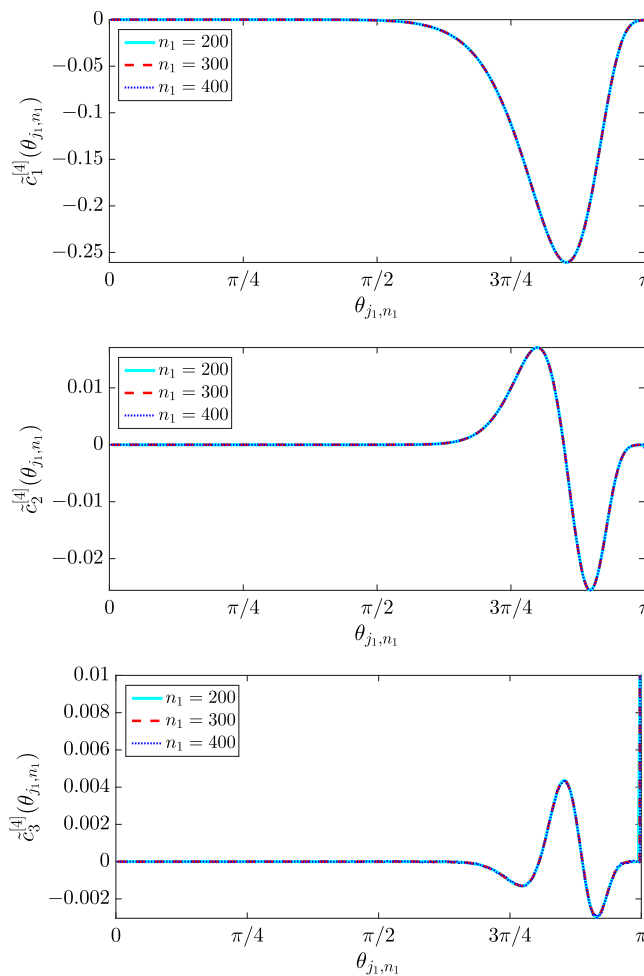


FIGURE 6 Example 2, $p = 4$: graph of the pairs $(\theta_{j_1, n_1}, \tilde{c}_k^{[4]}(\theta_{j_1, n_1}))$, $j_1 = 1, \dots, n_1$, for $n_1 = 200, 300, 400$ and $k = 1, 2, 3$

TABLE 3 Example 2, $p = 4$: values $\theta_{n_1}^{(\epsilon)}(4, k)$ for $k = 1, 2, 3$ and $n_1 = 200, 300, 400, 500, 600$, computed with the threshold $\epsilon = 0.0005$

n_1	200	300	400	500	600
$\theta_{n_1}^{(\epsilon)}(4, 1)$	$\frac{97\pi}{200} \approx 1.5237$	$\frac{146\pi}{300} \approx 1.5289$	$\frac{194\pi}{400} \approx 1.5237$	$\frac{242\pi}{500} \approx 1.5205$	$\frac{291\pi}{600} \approx 1.5237$
$\theta_{n_1}^{(\epsilon)}(4, 2)$	$\frac{129\pi}{200} \approx 2.0263$	$\frac{194\pi}{300} \approx 2.0316$	$\frac{258\pi}{400} \approx 2.0263$	$\frac{322\pi}{500} \approx 2.0232$	$\frac{387\pi}{600} \approx 2.0263$
$\theta_{n_1}^{(\epsilon)}(4, 3)$	$\frac{145\pi}{200} \approx 2.2777$	$\frac{217\pi}{300} \approx 2.2724$	$\frac{289\pi}{400} \approx 2.2698$	$\frac{362\pi}{500} \approx 2.2745$	$\frac{434\pi}{600} \approx 2.2724$

Example 2. In this example we verbatim repeat for the case $p = 4$ what we have done in Example 1 for $p = 3$. For the sake of brevity, we do not include here any comment and we limit to report the exact analogs of Figure 4, Table 1, Figure 5, and Table 2 in Figure 6, Table 3, Figure 7, and Table 4, respectively.

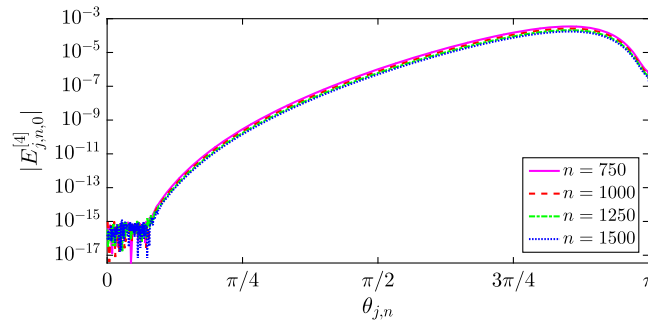


FIGURE 7 Example 2, $p = 4$: errors $|E_{j,n,0}^{[4]}|$ versus $\theta_{j,n}$ for $j = 1, \dots, n$ and $n = 750, 1000, 1250, 1500$

TABLE 4 Example 2, $p = 4$: first index j such that $|E_{j,n,0}^{[4]}| > 10^{-14}$ and corresponding grid point $\theta_{j,n}$, for $n = 750, 1000, 1250, 1500$

n	750	1000	1250	1500
j	71	97	123	152
$\theta_{j,n}$	0.2974	0.3047	0.3091	0.3183

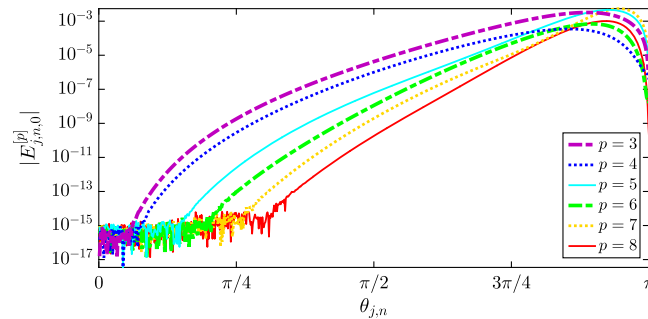


FIGURE 8 Example 3: errors $|E_{j,n,0}^{[p]}|$ versus $\theta_{j,n}$ for $j = 1, \dots, n - \text{mod}(p, 2)$ and $p = 3, \dots, 8$, with $n = 750$

Example 3. A comparison between Table 2 and Table 4 shows that the portion of the spectrum of $n^{-2}L_n^{[p]}$ which is exactly described by $e_p(\theta)$, at least from a numerical viewpoint, grows from $\theta(3)/\pi \approx 0.082$ for $p = 3$ to $\theta(4)/\pi \approx 0.101$ for $p = 4$. Actually, this spectrum portion increases more and more with p , that is, $\theta(p)$ grows with p ; see Figure 8.

5.2 | Numerical experiments illustrating the performance of Algorithm 1

Example 4. Let $p = 3$. Suppose we want to approximate the eigenvalues of $L_n^{[3]}$ (excluding the $n_3^{\text{out}} = 2$ outliers) for $n = 5000$. Let $\tilde{\lambda}_j^{(m)}(L_n^{[3]})$ be the approximation of $\lambda_j(L_n^{[3]})$ obtained by applying Algorithm 1 with $n_1 = 25 \cdot 2^{m-1}$ and $\alpha = 4$. In Figure 9, we plot the relative errors

$$\epsilon_{j,n}^{[3],m} = \frac{|\lambda_j(L_n^{[3]}) - \tilde{\lambda}_j^{(m)}(L_n^{[3]})|}{|\lambda_j(L_n^{[3]})|}$$

versus $\theta_{j,n}$ for $j = 1, \dots, n^{[3]} = n - 1$ and $m = 1, \dots, 4$. We see from the figure that the errors decrease rather quickly as m increases. A careful consideration of Figure 9 also reveals that, aside from the exceptional minima attained in a neighborhood of $\theta = 0^\ddagger$, the local minima of $\epsilon_{j,n}^{[3],m}$ are attained when $\theta_{j,n}$ is approximately equal to some of the

[‡]These minima, as well as the highly oscillatory behavior of the error around $\theta = 0$, are probably due to the fact that $e_3(\theta)$ provides a numerically exact description of the spectrum of $n^{-2}L_n^{[3]}$ around $\theta = 0$; see also Example 1.

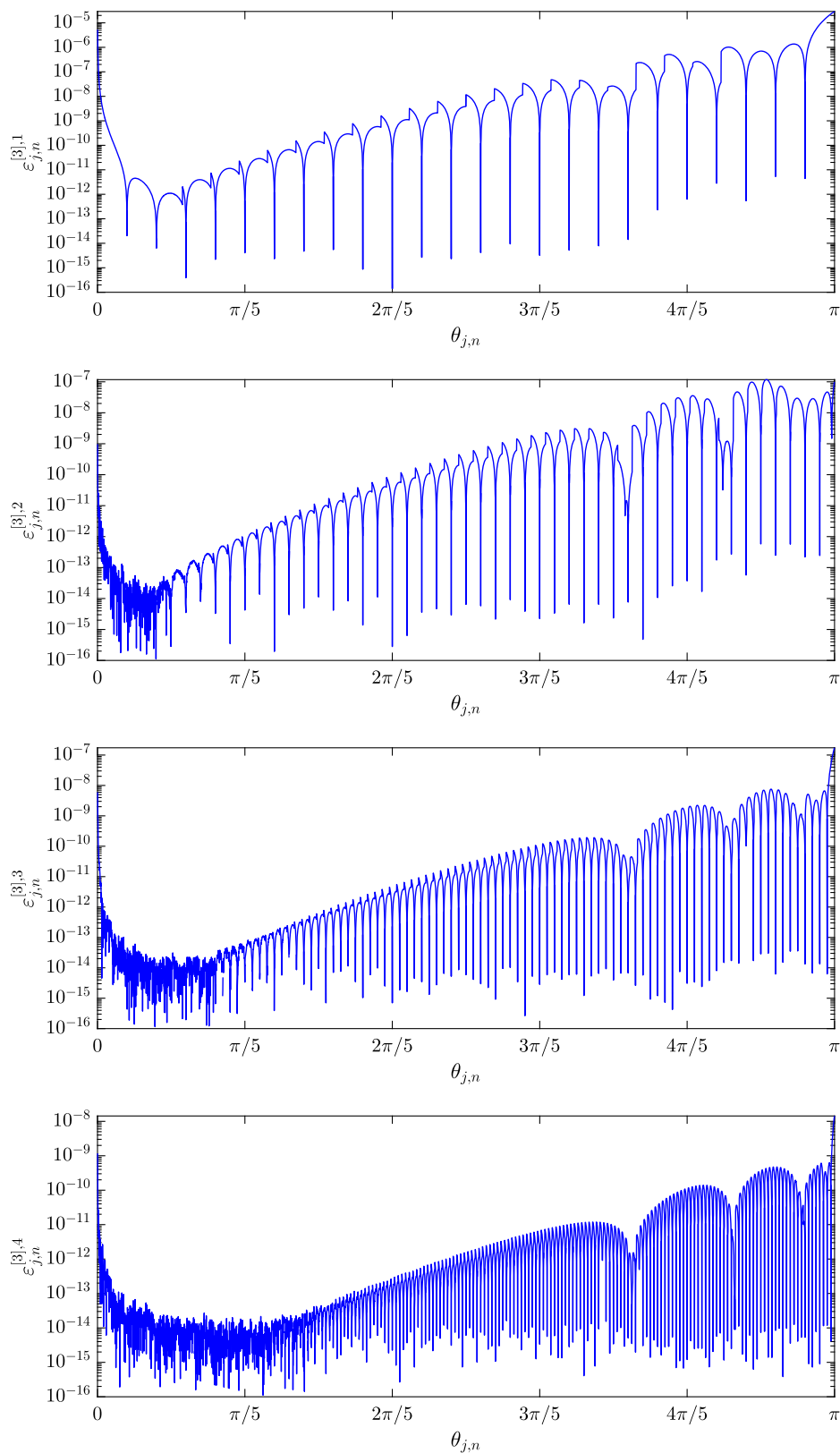


FIGURE 9 Example 4, $p = 3$: errors $\varepsilon_{j,n}^{[3],m}$ versus $\theta_{j,n}$ for $j = 1, \dots, n - 1$, in the case where $n = 5000$, $n_1 = 25 \cdot 2^{m-1}$, and $\alpha = 4$

coarse grid points $\theta_{j_1, n_1}, j_1 = 1, \dots, n_1$. This is no surprise because for $\theta_{j,n} = \theta_{j_1, n_1}$ we have $\tilde{c}_{k,j}^{[3]}(\theta_{j,n}) = \tilde{c}_k^{[3]}(\theta_{j_1, n_1})$ and $c_k^{[3]}(\theta_{j,n}) = c_k^{[3]}(\theta_{j_1, n_1})$, which means that the error of the approximation $\tilde{c}_{k,j}^{[3]}(\theta_{j,n}) \approx c_k^{[3]}(\theta_{j,n})$ reduces to the error of the approximation $\tilde{c}_k^{[3]}(\theta_{j_1, n_1}) \approx c_k^{[3]}(\theta_{j_1, n_1})$; that is, we are not introducing further error due to the interpolation process.

Example 5. Let $p = 4$. Suppose we want to approximate the eigenvalues of $L_n^{[4]}$ (excluding the $n_4^{\text{out}} = 2$ outliers) for $n = 5000$. Let $\tilde{\lambda}_j^{(m)}(L_n^{[4]})$ be the approximation of $\lambda_j(L_n^{[4]})$ obtained by applying Algorithm 1 with $n_1 = 10 \cdot 2^{m-1}$ and $\alpha = 5$. In Figure 10, we plot the relative errors

$$\varepsilon_{j,n}^{[4],m} = \frac{|\lambda_j(L_n^{[4]}) - \tilde{\lambda}_j^{(m)}(L_n^{[4]})|}{|\lambda_j(L_n^{[4]})|}$$

versus $\theta_{j,n}$ for $j = 1, \dots, n^{[4]} = n$ and $m = 1, \dots, 4$. Considerations analogous to those of Example 4 apply also in this case.

6 | EXTENSION TO THE MULTIDIMENSIONAL SETTING

We present in this section the extension to the multidimensional setting of the analysis carried out in the previous sections. In what follows, we will systematically use the multi-index notation and the properties of tensor products as described in sections 2.1.1 and 2.6.1 in the work of Garoni et al.¹³ If $w_i : D_i \rightarrow \mathbb{C}, i = 1, \dots, d$, are arbitrary functions, we will denote by $w_1 \otimes \dots \otimes w_d : D_1 \times \dots \times D_d \rightarrow \mathbb{C}$ the tensor-product function

$$(w_1 \otimes \dots \otimes w_d)(\xi_1, \dots, \xi_d) = \prod_{i=1}^d w_i(\xi_i), \quad (\xi_1, \dots, \xi_d) \in D_1 \times \dots \times D_d.$$

6.1 | Problem setting

Consider the d -dimensional Laplacian eigenvalue problem

$$\begin{cases} -\Delta u(\mathbf{x}) = \lambda u(\mathbf{x}), & \mathbf{x} \in (0, 1)^d, \\ u(\mathbf{x}) = 0, & \mathbf{x} \in \partial((0, 1)^d). \end{cases} \quad (32)$$

The corresponding weak formulation reads as follows: find eigenvalues $\lambda \in \mathbb{R}^+$ and eigenfunctions $u \in H_0^1((0, 1)^d)$ such that, for all $v \in H_0^1((0, 1)^d)$,

$$a(u, v) = \lambda(u, v),$$

where

$$a(u, v) = \int_{(0,1)^d} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) \, d\mathbf{x}, \quad (u, v) = \int_{(0,1)^d} u(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x}.$$

In the “tensor-product version” of Galerkin’s method, we choose d finite-dimensional vector spaces $\mathcal{W}_1, \dots, \mathcal{W}_d \subset H_0^1(0, 1)$ and we set

$$\mathcal{W} = \mathcal{W}_1 \otimes \dots \otimes \mathcal{W}_d = \text{span}(w_1 \otimes \dots \otimes w_d : w_1 \in \mathcal{W}_1, \dots, w_d \in \mathcal{W}_d) \subset H_0^1((0, 1)^d).$$

Then, we define $N_s = \dim \mathcal{W}_s$ for $s = 1, \dots, d$ and $\mathbf{N} = (N_1, \dots, N_d)$, and we look for approximations of the exact eigenpairs

$$\lambda_{\mathbf{j}} = \sum_{i=1}^d j_i^2 \pi^2, \quad u_{\mathbf{j}}(\mathbf{x}) = \prod_{i=1}^d \sin(j_i \pi x_i), \quad \mathbf{j} = (j_1, \dots, j_d) \in \mathbb{N}^d, \quad (33)$$

by solving the following Galerkin problem: Find $\lambda_{\mathbf{j}, \mathcal{W}} \in \mathbb{R}^+$ and $u_{\mathbf{j}, \mathcal{W}} \in \mathcal{W}$, for $\mathbf{j} = \mathbf{1}, \dots, \mathbf{N}$, such that, for all $v \in \mathcal{W}$,

$$a(u_{\mathbf{j}, \mathcal{W}}, v) = \lambda_{\mathbf{j}, \mathcal{W}} (u_{\mathbf{j}, \mathcal{W}}, v). \quad (34)$$

If $\{\varphi_{1,[s]}, \dots, \varphi_{N_s,[s]}\}$ is a basis of \mathcal{W}_s for $s = 1, \dots, d$, then

$$\varphi_{\mathbf{i}} = \varphi_{i_1,[1]} \otimes \dots \otimes \varphi_{i_d,[d]}, \quad \mathbf{i} = \mathbf{1}, \dots, \mathbf{N},$$

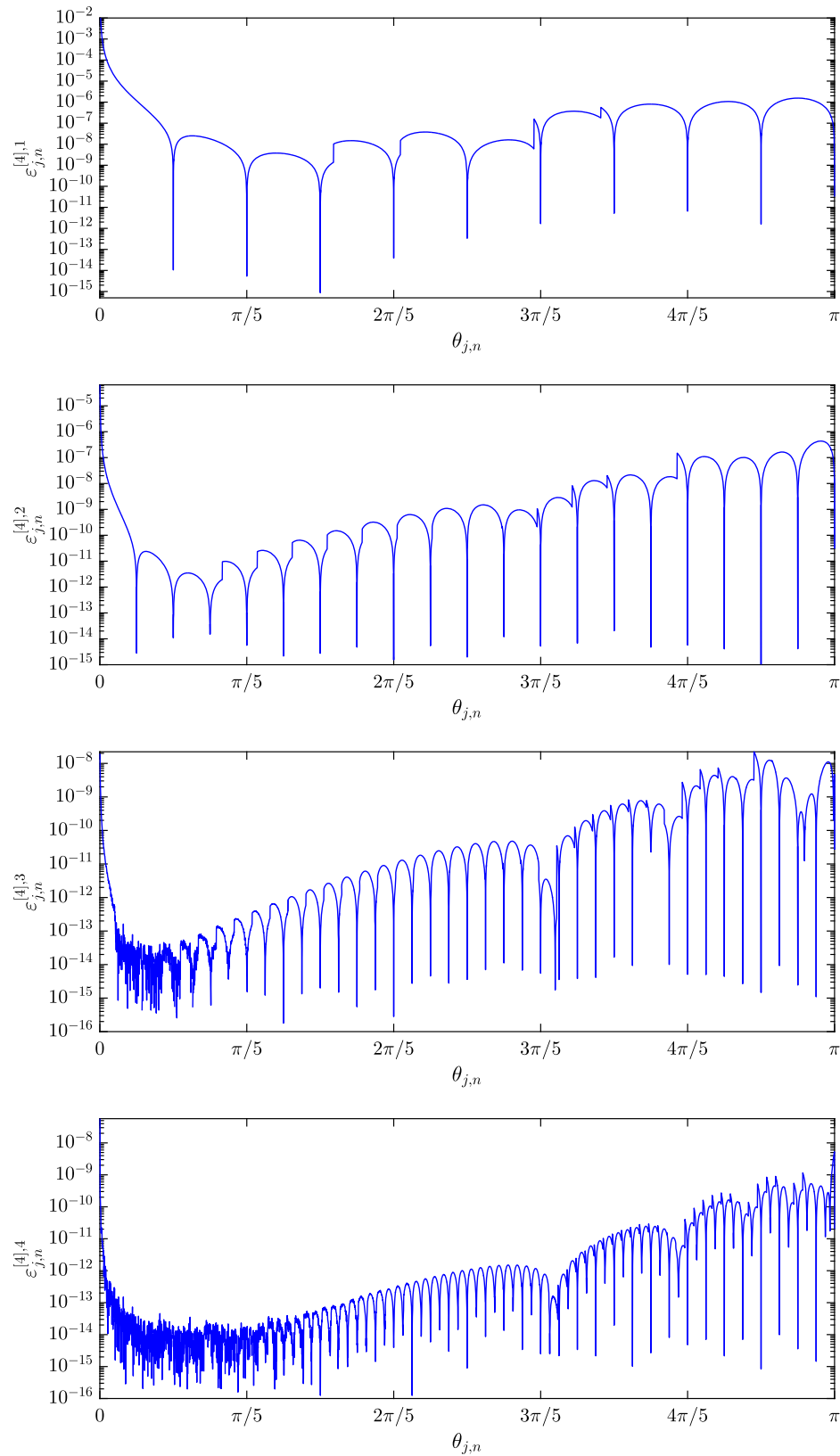


FIGURE 10 Example 5, $p = 4$: errors $\epsilon_{j,n}^{[4],m}$ versus $\theta_{j,n}$ for $j = 1, \dots, n$, in the case where $n = 5000$, $n_1 = 10 \cdot 2^{m-1}$, and $\alpha = 5$

is a basis of \mathcal{W} , and in view of the canonical identification between each $v \in \mathcal{W}$ and its coefficient vector with respect to $\{\varphi_1, \dots, \varphi_N\}$, solving the Galerkin problem (34) is equivalent to solving the generalized eigenvalue problem

$$K \mathbf{u}_{j,\mathcal{W}} = \lambda_{j,\mathcal{W}} M \mathbf{u}_{j,\mathcal{W}}, \quad (35)$$

where $\mathbf{u}_{j,\mathcal{W}}$ is the coefficient vector of $u_{j,\mathcal{W}}$ with respect to $\{\varphi_1, \dots, \varphi_N\}$,

$$K = [a(\varphi_j, \varphi_i)]_{i,j=1}^N = \left[\int_{(0,1)^d} \nabla \varphi_j(\mathbf{x}) \cdot \nabla \varphi_i(\mathbf{x}) d\mathbf{x} \right]_{i,j=1}^N = \sum_{r=1}^d \left(\bigotimes_{s=1}^{r-1} M^{(s)} \right) \otimes K^{(r)} \otimes \left(\bigotimes_{s=r+1}^d M^{(s)} \right), \quad (36)$$

$$M = [(\varphi_j, \varphi_i)]_{i,j=1}^N = \left[\int_{(0,1)^d} \varphi_j(\mathbf{x}) \varphi_i(\mathbf{x}) d\mathbf{x} \right]_{i,j=1}^N = \bigotimes_{s=1}^d M^{(s)}, \quad (37)$$

and

$$K^{(s)} = \left[\int_0^1 \varphi'_{j,[s]}(x) \varphi'_{i,[s]}(x) dx \right]_{i,j=1}^{N_s}, \quad s = 1, \dots, d,$$

$$M^{(s)} = \left[\int_0^1 \varphi_{j,[s]}(x) \varphi_{i,[s]}(x) dx \right]_{i,j=1}^{N_s}, \quad s = 1, \dots, d.$$

The matrices K and M are, respectively, the stiffness matrix and the mass matrix. Both K and M are always symmetric positive definite, regardless of the basis functions $\varphi_1, \dots, \varphi_N$. Moreover, it is clear from (35) that the numerical eigenvalues $\lambda_{j,\mathcal{W}}, \mathbf{j} = \mathbf{1}, \dots, \mathbf{N}$, are just the eigenvalues of the matrix

$$L = M^{-1}K = \sum_{r=1}^d \left(\bigotimes_{s=1}^{r-1} I_{N_s} \right) \otimes (M^{(r)})^{-1} K^{(r)} \otimes \left(\bigotimes_{s=r+1}^d I_{N_s} \right). \quad (38)$$

In the IgA approximation of (32) based on uniform tensor-product B-splines of degree $\mathbf{p} = (p_1, \dots, p_d)$, we look for approximations of the exact eigenpairs (33) by using the tensor-product version of the Galerkin method described above, in which the basis functions $\varphi_{1,[s]}, \dots, \varphi_{N_s,[s]}$ are chosen as the B-splines $N_{2,[p_s]}, \dots, N_{n_s+p_s-1,[p_s]}$ for $s = 1, \dots, d$, where the functions $N_{i_s+1,[p_s]}, i_s = 1, \dots, n_s + p_s - 2$, are defined in (8) for $n = n_s$ and $p = p_s$. Setting $\mathbf{n} = (n_1, \dots, n_d)$, the resulting stiffness and mass matrices (36)–(37) are given by

$$K_{\mathbf{n}}^{[\mathbf{p}]} = \sum_{r=1}^d \left(\bigotimes_{s=1}^{r-1} M_{n_s}^{[p_s]} \right) \otimes K_{n_r}^{[p_r]} \otimes \left(\bigotimes_{s=r+1}^d M_{n_s}^{[p_s]} \right), \quad (39)$$

$$M_{\mathbf{n}}^{[\mathbf{p}]} = \bigotimes_{s=1}^d M_{n_s}^{[p_s]}, \quad (40)$$

and the numerical eigenvalues $\lambda_{j,\mathbf{n}}^{[\mathbf{p}]}, \mathbf{j} = \mathbf{1}, \dots, \mathbf{n} + \mathbf{p} - \mathbf{2}$, are the eigenvalues of the matrix

$$L_{\mathbf{n}}^{[\mathbf{p}]} = \left(M_{\mathbf{n}}^{[\mathbf{p}]} \right)^{-1} K_{\mathbf{n}}^{[\mathbf{p}]} = \sum_{r=1}^d \left(\bigotimes_{s=1}^{r-1} I_{n_s+p_s-2} \right) \otimes L_{n_r}^{[p_r]} \otimes \left(\bigotimes_{s=r+1}^d I_{n_s+p_s-2} \right), \quad (41)$$

where the matrices $K_{\mathbf{n}}^{[\mathbf{p}]}, M_{\mathbf{n}}^{[\mathbf{p}]}, L_{\mathbf{n}}^{[\mathbf{p}]}$ are defined in (10)–(12) for all $p, n \geq 1$.

6.2 | Eigenvalue–eigenvector structure of $L_{\mathbf{n}}^{[\mathbf{p}]}$

We now show that the eigenvalue–eigenvector structure of $L_{\mathbf{n}}^{[\mathbf{p}]}$ is determined by the eigenvalue–eigenvector structure of the matrices $L_n^{[p]}$ for $p \in \{p_1, \dots, p_d\}$. It will immediately follow that the eigenvalues and eigenvectors of $L_{\mathbf{n}}^{[\mathbf{p}]}$ are explicitly known for $\mathbf{1} \leq \mathbf{p} \leq \mathbf{2}$, due to the results of Section 3. Moreover, the parallel interpolation–extrapolation algorithm devised in Section 4 for computing the eigenvalues of $L_n^{[p]}$ also allows the computation of the eigenvalues of $L_{\mathbf{n}}^{[\mathbf{p}]}$.

For $p, n \geq 1$, let

$$L_n^{[p]} = V_n^{[p]} D_n^{[p]} \left(V_n^{[p]} \right)^{-1}, \quad D_n^{[p]} = \text{diag}_{j=1, \dots, n+p-2} \lambda_j \left(L_n^{[p]} \right), \quad (42)$$

be a spectral decomposition of $L_n^{[p]}$. Note that such a decomposition exists because $L_n^{[p]}$ is diagonalizable, due to the similarity equation

$$L_n^{[p]} = \left(M_n^{[p]}\right)^{-1} K_n^{[p]} = \left(M_n^{[p]}\right)^{-1/2} \left[\left(M_n^{[p]}\right)^{-1/2} K_n^{[p]} \left(M_n^{[p]}\right)^{-1/2}\right] \left(M_n^{[p]}\right)^{1/2}.$$

It follows from (42) and the properties of tensor products that

$$\begin{aligned} L_n^{[p]} &= \sum_{r=1}^d \left(\bigotimes_{s=1}^{r-1} I_{n_s+p_s-2} \right) \otimes L_{n_r}^{[p_r]} \otimes \left(\bigotimes_{s=r+1}^d I_{n_s+p_s-2} \right), \\ &= \left(\bigotimes_{s=1}^d V_{n_s}^{[p_s]} \right) \left[\sum_{r=1}^d \left(\bigotimes_{s=1}^{r-1} I_{n_s+p_s-2} \right) \otimes D_{n_r}^{[p_r]} \otimes \left(\bigotimes_{s=r+1}^d I_{n_s+p_s-2} \right) \right] \left(\bigotimes_{s=1}^d V_{n_s}^{[p_s]} \right)^{-1}, \end{aligned} \quad (43)$$

which is a spectral decomposition of $L_n^{[p]}$. More explicitly, let $\mathbf{v}_{1,n}^{[p]}, \dots, \mathbf{v}_{n+p-2,n}^{[p]}$ be the columns of $V_n^{[p]}$, that is, the eigenvectors of $L_n^{[p]}$,

$$L_n^{[p]} \mathbf{v}_{j,n}^{[p]} = \lambda_j \left(L_n^{[p]} \right) \mathbf{v}_{j,n}^{[p]}, \quad j = 1, \dots, n+p-2,$$

and let

$$\mathbf{v}_{j,n}^{[p]} = \bigotimes_{s=1}^d \mathbf{v}_{j_s, n_s}^{[p_s]}, \quad \mathbf{j} = \mathbf{1}, \dots, \mathbf{n+p-2}. \quad (44)$$

Then, we can rewrite (43) as

$$L_n^{[p]} \mathbf{v}_{j,n}^{[p]} = \lambda_j \left(L_n^{[p]} \right) \mathbf{v}_{j,n}^{[p]}, \quad \mathbf{j} = \mathbf{1}, \dots, \mathbf{n+p-2},$$

where

$$\lambda_j \left(L_n^{[p]} \right) = \sum_{r=1}^d \lambda_{j_r} \left(L_{n_r}^{[p_r]} \right), \quad \mathbf{j} = \mathbf{1}, \dots, \mathbf{n+p-2}. \quad (45)$$

In other words, the eigenvalue–eigenvector pairs of $L_n^{[p]}$ are

$$\left(\lambda_j \left(L_n^{[p]} \right), \mathbf{v}_{j,n}^{[p]} \right), \quad \mathbf{j} = \mathbf{1}, \dots, \mathbf{n+p-2},$$

with $\mathbf{v}_{j,n}^{[p]}$ and $\lambda_j(L_n^{[p]})$ defined as in (44) and (45), respectively.

7 | CONCLUSIONS AND PERSPECTIVES

We have considered the B-spline IgA approximation of the d -dimensional Laplacian eigenvalue problem (32). Through tensor-product arguments, we have shown that the eigenvalue–eigenvector structure of the resulting discretization matrix $L_n^{[p]}$ is completely determined by the eigenvalue–eigenvector structure of the matrix $L_n^{[p]}$ arising from the B-spline IgA approximation of the unidimensional eigenproblem (1). As for the matrix $L_n^{[p]}$, we implemented the program detailed in items 1 to 5 of Section 1.2. We conclude this work by suggesting a few possible future lines of research:

- Provide a formal proof of the asymptotic eigenvalue expansion (18). Considering that the eigenvalue expansion (18) is strongly connected with the eigenvalue expansion for preconditioned Toeplitz matrices,²⁰ a proof of the former may suggest the way to prove the latter and vice versa. Insights on how to perform these proofs might be gained from the works of Bogoya et al.^{22–24} where a completely analogous eigenvalue expansion was proved for Toeplitz matrices.
- By the results of Bogoya et al.,^{20–24} Toeplitz and preconditioned Toeplitz matrices possess asymptotic eigenvalue expansions completely analogous to (18). The matrices arising from the discretization of a linear PDE by a linear numerical method—hereinafter referred to as PDE discretization matrices—usually have a Toeplitz or Toeplitz-related structure. For example, in the case of a constant-coefficient PDE, the matrix structure is often a small perturbation of a pure Toeplitz structure, whereas in the case of a variable-coefficient PDE, the matrix structure is often a so-called (generalized) locally Toeplitz structure^{12,13,33,34}; see in particular section 7.1 in the work of Garoni et al.¹² A natural question then follows: Do we have asymptotic expansions also for the eigenvalues of PDE discretization matrices? This paper has provided a positive answer in the case where the PDE is the Laplacian eigenproblem (32) and the numerical method is the B-spline IgA. It is clear, however, that the previous question opens the doors to a series of possible future

research works, whose purpose is not only to ascertain the existence of an asymptotic eigenvalue expansion for PDE discretization matrices but also to exploit this expansion (if any) for computing the eigenvalues themselves through fast interpolation–extrapolation procedures (such as Algorithm 1).

ACKNOWLEDGEMENTS

S.-E. Ekström is supported by the Graduate School in Mathematics and Computing (FMB) and Uppsala University. C. Garoni is a Marie Curie fellow of INdAM (Istituto Nazionale di Alta Matematica) under Grant PCOFUND-GA-2012-600198. C. Manni and H. Speleers are supported by the “Mission Sustainability” programme of the University of Rome “Tor Vergata” through the project IDEAS. All the authors (except S.-E. Ekström) are supported by the INdAM GNCS (Gruppo Nazionale per il Calcolo Scientifico).

ORCID

Sven-Erik Ekström  <http://orcid.org/0000-0002-7875-7543>

Isabella Furci  <http://orcid.org/0000-0001-8864-6786>

Carlo Garoni  <http://orcid.org/0000-0001-9720-092X>

Carla Manni  <http://orcid.org/0000-0002-1519-4106>

Stefano Serra-Capizzano  <http://orcid.org/0000-0001-9477-109X>

Hendrik Speleers  <http://orcid.org/0000-0003-4110-3308>

REFERENCES

1. Cottrell JA, Hughes TJR, Bazilevs Y. *Isogeometric analysis: Toward integration of CAD and FEA*. Chichester, UK: John Wiley & Sons; 2009.
2. Cottrell JA, Reali A, Bazilevs Y, Hughes TJR. Isogeometric analysis of structural vibrations. *Comput Methods Appl Mech Eng*. 2006;195(41–43):5257–5296.
3. Ekström S-E, Garoni C, Hughes TJR, Reali A, Serra-Capizzano S, Speleers H. Symbol-based analysis of finite element and isogeometric B-spline discretizations of eigenvalue problems: exposition and review. In preparation.
4. Hughes TJR, Evans JA, Reali A. Finite element and NURBS approximations of eigenvalue, boundary-value, and initial-value problems. *Comput Methods Appl Mech Eng*. 2014;272:290–320.
5. Hughes TJR, Reali A, Sangalli G. Duality and unified analysis of discrete approximations in structural dynamics and wave propagation: comparison of p -method finite elements with k -method NURBS. *Comput Methods Appl Mech Eng*. 2008;197(49–50):4104–4124.
6. Reali A. An isogeometric analysis approach for the study of structural vibrations. *J Earthq Eng*. 2006;10:1–30.
7. Donatelli M, Garoni C, Manni C, Serra-Capizzano S, Speleers H. Spectral analysis and spectral symbol of matrices in isogeometric collocation methods. *Math Comput*. 2016;85(300):1639–1680.
8. Garoni C. Spectral distribution of PDE discretization matrices from isogeometric analysis: the case of L^1 coefficients and non-regular geometry. *J Spectr Theory*. 2018;8(1):297–313.
9. Garoni C, Manni C, Pelosi F, Serra-Capizzano S, Speleers H. On the spectrum of stiffness matrices arising from isogeometric analysis. *Numer Math*. 2014;127(4):751–799.
10. Garoni C, Manni C, Serra-Capizzano S, Sesana D, Speleers H. Spectral analysis and spectral symbol of matrices in isogeometric Galerkin methods. *Math Comput*. 2017;86(305):1343–1373.
11. Garoni C, Manni C, Serra-Capizzano S, Sesana D, Speleers H. Lusin theorem, GLT Sequences and matrix computations: an application to the spectral analysis of PDE discretization matrices. *J Math Anal Appl*. 2017;446(1):365–382.
12. Garoni C, Serra-Capizzano S. *Generalized locally Toeplitz sequences: theory and applications (Volume I)*. Cham, Switzerland: Springer; 2017.
13. Garoni C, Serra-Capizzano S. *Generalized locally Toeplitz sequences: theory and applications (Volume II)*. Cham, Switzerland: Springer; To be published.
14. Donatelli M, Garoni C, Manni C, Serra-Capizzano S, Speleers H. Robust and optimal multi-iterative techniques for IgA Galerkin linear systems. *Comput Methods Appl Mech Eng*. 2015;284:230–264.
15. Donatelli M, Garoni C, Manni C, Serra-Capizzano S, Speleers H. Robust and optimal multi-iterative techniques for IgA collocation linear systems. *Comput Methods Appl Mech Eng*. 2015;284:1120–1146.
16. Donatelli M, Garoni C, Manni C, Serra-Capizzano S, Speleers H. Symbol-based multigrid methods for Galerkin B-spline isogeometric analysis. *SIAM J Numer Anal*. 2017;55(1):31–62.
17. de Boor C. *A practical guide to splines*. New York, NY: Springer; 2001.
18. Schumaker LL. *Spline functions: basic theory*. 3rd ed. Cambridge, UK: Cambridge University Press; 2007.

19. Bozzo E, Di Fiore C. On the use of certain matrix algebras associated with discrete trigonometric transforms in matrix displacement decomposition. *SIAM J Matrix Anal Appl.* 1995;16(1):312–326.
20. Ahmad F, Al-Aidarous ES, Alrehaili DA, Ekström S-E, Furci I, Serra-Capizzano S. Are the eigenvalues of preconditioned banded symmetric Toeplitz matrices known in almost closed form? *Numer Algorithm.* 2018;78(3):867–893.
21. Ekström S-E, Garoni C, Serra-Capizzano S. Are the eigenvalues of banded symmetric Toeplitz matrices known in almost closed form? *Exper Math.* 2017. <https://doi.org/10.1080/10586458.2017.1320241>
22. Bogoya JM, Böttcher A, Grudsky SM, Maximenko EA. Eigenvalues of Hermitian Toeplitz matrices with smooth simple-loop symbols. *J Math Anal Appl.* 2015;422(2):1308–1334.
23. Bogoya JM, Grudsky SM, Maximenko EA. Eigenvalues of Hermitian Toeplitz matrices generated by simple-loop symbols with relaxed smoothness. In: *Large truncated Toeplitz matrices, Toeplitz operators, and related topics. Operator theory: Advances and applications.* Cham, Switzerland: Birkhäuser, 2017; p. 179–212.
24. Böttcher A, Grudsky SM, Maximenko EA. Inside the eigenvalues of certain Hermitian Toeplitz band matrices. *J Comput Appl Math.* 2010;233(9):2245–2264.
25. Chen H, Jia S, Xie H. Postprocessing and higher order convergence for the mixed finite element approximations of the eigenvalue problem. *Appl Numer Math.* 2011;61(4):615–629.
26. Yin X, Xie H, Jia S, Gao S. Asymptotic expansions and extrapolations of eigenvalues for the Stokes problem by mixed finite element methods. *J Comput Appl Math.* 2008;215(1):127–141.
27. Ekström S-E, Garoni C. A matrix-less and parallel interpolation–extrapolation algorithm for computing the eigenvalues of preconditioned banded symmetric Toeplitz matrices. *Numer Algorithm.* 2018. <https://doi.org/10.1007/s11075-018-0508-0>
28. Garoni C, Serra-Capizzano S. Generalized locally Toeplitz sequences: a spectral analysis tool for discretized differential equations. To appear.
29. Bini D, Capovani M. Spectral and computational properties of band symmetric Toeplitz matrices. *Linear Algebra Appl.* 1983;52–53:99–126.
30. Sangalli G, Tani M. Isogeometric preconditioners based on fast solvers for the Sylvester equation. *SIAM J Sci Comput.* 2016;38(6):A3644–A3671.
31. Stoer J, Bulirsch R. *Introduction to numerical analysis.* 3rd ed. New York, NY: Springer; 2002.
32. Brezinski C, Redivo Zaglia M. *Extrapolation methods: Theory and practice.* North Holland, The Netherlands: Elsevier; 1991.
33. Serra-Capizzano S. Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations. *Linear Algebra Appl.* 2003;366:371–402.
34. Serra-Capizzano S. The GLT class as a generalized Fourier analysis and applications. *Linear Algebra Appl.* 2006;419(1):180–233.
35. Chui CK. *An introduction to wavelets.* London, UK: Academic Press; 1992.
36. Serra-Capizzano S. On the extreme spectral properties of Toeplitz matrices generated by L^1 functions with several minima/maxima. *BIT Numer Math.* 1996;36(1):135–142.
37. Bhatia R. *Matrix analysis.* New York, NY: Springer; 1997.

How to cite this article: Ekström S-E, Furci I, Garoni C, Manni C, Serra-Capizzano S, Speleers H. Are the eigenvalues of the B-spline isogeometric analysis approximation of $-\Delta u = \lambda u$ known in almost closed form? *Numer Linear Algebra Appl.* 2018;25:e2198. <https://doi.org/10.1002/nla.2198>

APPENDIX A

Proofs of the theorems stated in Section 2

We first recall from section 3 in the work of Garoni et al.⁹ that, for every $p \geq 0$ and $\theta \in [0, \pi]$,

$$g_p(\theta) = \sum_{k \in \mathbb{Z}} \left| \widehat{\phi}_p(\theta + 2k\pi) \right|^2, \quad (\text{A1})$$

where $\widehat{\phi}_p$ is the Fourier transform of the cardinal B-spline ϕ_p , whose modulus is given by

$$\left| \widehat{\phi}_p(\theta) \right|^2 = \left(\frac{2 - 2 \cos(\theta)}{\theta^2} \right)^{p+1}; \quad (\text{A2})$$

see the work of Chui.³⁵ The next lemma is fundamental to our purposes.

Lemma 1. For $p \geq 1$ and $\theta \in [0, \pi]$, we have

$$\frac{9}{5}\pi(\pi - \theta)\left(\frac{\theta}{2\pi - \theta}\right)^{2p+2} \leq e_p(\theta) - \theta^2 \leq 4\pi(\pi - \theta)\left(\frac{\theta}{2\pi - \theta}\right)^{2p+2} + 5\theta^2\left(\frac{\theta}{2\pi + \theta}\right)^{2p}. \tag{A3}$$

Proof. From (16) and (A1)–(A2), we obtain

$$f_p(\theta) = (2 - 2 \cos(\theta))^{p+1} \sum_{k \in \mathbb{Z}} \frac{1}{(\theta + 2k\pi)^{2p}} = (2 - 2 \cos(\theta))^{p+1} \left[\frac{1}{\theta^{2p}} + \sum_{k \neq 0} \frac{1}{(\theta + 2k\pi)^{2p}} \right],$$

$$g_p(\theta) = (2 - 2 \cos(\theta))^{p+1} \sum_{k \in \mathbb{Z}} \frac{1}{(\theta + 2k\pi)^{2p+2}} = (2 - 2 \cos(\theta))^{p+1} \left[\frac{1}{\theta^{2p+2}} + \sum_{k \neq 0} \frac{1}{(\theta + 2k\pi)^{2p+2}} \right].$$

By setting

$$r_p(\theta) = \theta^{2p} \sum_{k \neq 0} \frac{1}{(\theta + 2k\pi)^{2p}} \geq 0,$$

we see that

$$e_p(\theta) - \theta^2 = \frac{f_p(\theta)}{g_p(\theta)} - \theta^2 = \theta^2 \frac{1 + r_p(\theta)}{1 + r_{p+1}(\theta)} - \theta^2 = \theta^2 \frac{r_p(\theta) - r_{p+1}(\theta)}{1 + r_{p+1}(\theta)}. \tag{A4}$$

Furthermore,

$$r_p(\theta) - r_{p+1}(\theta) = \theta^{2p} (A_{p,+}(\theta) + A_{p,-}(\theta)), \tag{A5}$$

where

$$A_{p,+}(\theta) = \sum_{k \geq 1} \frac{1}{(2k\pi + \theta)^{2p}} \left(1 - \frac{\theta^2}{(2k\pi + \theta)^2} \right), \tag{A6}$$

$$A_{p,-}(\theta) = \sum_{k \geq 1} \frac{1}{(2k\pi - \theta)^{2p}} \left(1 - \frac{\theta^2}{(2k\pi - \theta)^2} \right). \tag{A7}$$

Assume $\theta \in [0, \pi]$. We observe that

$$0 \leq 1 - \frac{\theta^2}{(2k\pi + \theta)^2} \leq 1, \quad k \geq 1,$$

which implies that

$$A_{p,+}(\theta) \leq \frac{1}{(2\pi + \theta)^{2p}} + \sum_{k \geq 2} \frac{1}{(2k\pi + \theta)^{2p}} \leq \frac{1}{(2\pi + \theta)^{2p}} + \int_1^{+\infty} \frac{d\kappa}{(2\pi\kappa + \theta)^{2p}}$$

$$= \frac{1}{(2\pi + \theta)^{2p}} + \frac{1}{2\pi(2p-1)(2\pi + \theta)^{2p-1}} \leq \frac{5}{2} \frac{1}{(2\pi + \theta)^{2p}}.$$

Similarly,

$$A_{p,-}(\theta) \leq \frac{4\pi(\pi - \theta)}{(2\pi - \theta)^{2p+2}} + \frac{8\pi(2\pi - \theta)}{(4\pi - \theta)^{2p+2}} + \sum_{k \geq 3} \frac{1}{(2k\pi - \theta)^{2p}}$$

$$\leq \frac{4\pi(\pi - \theta)}{(2\pi - \theta)^{2p+2}} + \frac{8\pi(2\pi - \theta)}{(4\pi - \theta)^{2p+2}} + \int_2^{+\infty} \frac{d\kappa}{(2\pi\kappa - \theta)^{2p}}$$

$$= \frac{4\pi(\pi - \theta)}{(2\pi - \theta)^{2p+2}} + \frac{8\pi(2\pi - \theta)}{(4\pi - \theta)^{2p+2}} + \frac{1}{2\pi(2p-1)(4\pi - \theta)^{2p-1}}$$

$$\leq \frac{4\pi(\pi - \theta)}{(2\pi - \theta)^{2p+2}} + \frac{5}{2} \frac{1}{(2\pi + \theta)^{2p}},$$

where we have exploited the fact that

$$4\pi - \theta \geq 2\pi + \theta, \quad \frac{8\pi(2\pi - \theta)}{(4\pi - \theta)^2} \leq 1.$$

By combining (A4) and (A5) with the obtained upper bounds for $A_{p,+}$ and $A_{p,-}$, we get the upper bound in (A3).

To prove the lower bound in (A3), we use the inequality

$$r_{p+1}(\theta) \leq \theta^{2p+2} \left(\frac{1}{(2\pi + \theta)^{2p+2}} + \frac{1}{(2\pi - \theta)^{2p+2}} + \int_1^{+\infty} \left[\frac{1}{(2\pi\kappa + \theta)^{2p+2}} + \frac{1}{(2\pi\kappa - \theta)^{2p+2}} \right] d\kappa \right)$$

$$= \theta^{2p+2} \left(\frac{1}{(2\pi + \theta)^{2p+2}} + \frac{1}{(2\pi - \theta)^{2p+2}} + \frac{1}{2\pi(2p+1)} \left[\frac{1}{(2\pi + \theta)^{2p+1}} + \frac{1}{(2\pi - \theta)^{2p+1}} \right] \right).$$

Note that

$$\frac{1}{(2\pi + \theta)^q} + \frac{1}{(2\pi - \theta)^q} \leq \frac{1}{(3\pi)^q} + \frac{1}{\pi^q}, \quad q \geq 1,$$

because the function on the left-hand side is monotone increasing for $\theta \in [0, \pi]$. Therefore, for $p \geq 1$,

$$r_{p+1}(\theta) \leq \left(\frac{\theta}{\pi}\right)^{2p+2} \left(\frac{1}{3^{2p+2}} + 1 + \frac{1}{2(2p+1)} \left[\frac{1}{3^{2p+1}} + 1\right]\right) \leq \frac{1}{81} + 1 + \frac{1}{6} \left[\frac{1}{27} + 1\right] = \frac{32}{27}.$$

Moreover, from (A6) and (A7), we deduce that

$$A_{p,+}(\theta) + A_{p,-}(\theta) = \sum_{k \geq 1} \frac{4k\pi(k\pi + \theta)}{(2k\pi + \theta)^{2p+2}} + \frac{4k\pi(k\pi - \theta)}{(2k\pi - \theta)^{2p+2}} \geq \frac{4\pi(\pi - \theta)}{(2\pi - \theta)^{2p+2}}.$$

Taking into account (A5), we arrive at

$$\frac{r_p(\theta) - r_{p+1}(\theta)}{1 + r_{p+1}(\theta)} \geq \frac{4\pi(\pi - \theta)}{(2\pi - \theta)^2} \left(\frac{\theta}{2\pi - \theta}\right)^{2p} \frac{27}{59}.$$

In view of (A4), this immediately gives the lower bound in (A3). □

We are now ready to prove Theorems 1 and 3.

Proof of Theorem 1. From the upper bound in (A3), we have

$$\max_{\theta \in [0, \pi]} |e_p(\theta) - \theta^2| \leq \max_{\theta \in [0, \pi]} \left[4\pi(\pi - \theta) \left(\frac{\theta}{2\pi - \theta}\right)^{2p+2} + 5\theta^2 \left(\frac{\theta}{2\pi + \theta}\right)^{2p} \right].$$

By setting $z = \frac{\theta}{\pi} \in [0, 1]$, we obtain

$$\begin{aligned} \max_{\theta \in [0, \pi]} |e_p(\theta) - \theta^2| &\leq \max_{z \in [0, 1]} \left[4\pi^2(1 - z) \left(\frac{z}{2 - z}\right)^{2p+2} + 5\pi^2 z^2 \left(\frac{z}{2 + z}\right)^{2p} \right] \\ &\leq \max_{z \in [0, 1]} 5\pi^2 \left[\left(\frac{z}{2 - z}\right)^{2p+2} \left(1 - \frac{z}{2 - z}\right) + \frac{1}{3^{2p}} \right]. \end{aligned}$$

Finally, by setting $y = \frac{z}{2 - z} \in [0, 1]$ and observing that

$$\max_{y \in [0, 1]} y^{2p+2}(1 - y) = \left(1 - \frac{1}{2p + 3}\right)^{2p+2} \frac{1}{2p + 3} \leq \frac{1}{2p + 3},$$

we get

$$\max_{\theta \in [0, \pi]} |e_p(\theta) - \theta^2| \leq 5\pi^2 \left(\frac{1}{2p + 3} + \frac{1}{3^{2p}}\right).$$

This concludes the proof. □

Proof of Theorem 3. For $p = 1$, the bounds $1/3 \leq w_p(\theta) \leq 1$ stated in the theorem hold because we know from (14) that

$$g_0(\theta) = 1, \quad g_1(\theta) = \frac{2}{3} + \frac{1}{3} \cos(\theta).$$

In the following, we focus on the case $p \geq 2$. From (17), it is clear that the bounds hold for $\theta = 0$. From (16) and (A3), we deduce that, for $\theta \in (0, \pi]$,

$$\begin{aligned} 1 \leq \frac{1}{\theta^2} \frac{f_p(\theta)}{g_p(\theta)} &= \frac{2 - 2\cos(\theta)}{\theta^2} \frac{g_{p-1}(\theta)}{g_p(\theta)} \leq 1 + \frac{4\pi(\pi - \theta)}{(2\pi - \theta)^2} \left(\frac{\theta}{2\pi - \theta}\right)^{2p} + 5 \left(\frac{\theta}{2\pi + \theta}\right)^{2p} \\ &\leq 1 + \frac{4\pi(\pi - \theta)}{(2\pi - \theta)^2} \left(\frac{\theta}{2\pi - \theta}\right)^4 + 5 \left(\frac{1}{3}\right)^4 \leq 1 + \frac{3}{20} + \frac{5}{81} < \frac{12}{\pi^2}. \end{aligned}$$

Because

$$1 \leq \frac{\theta^2}{2 - 2\cos(\theta)} \leq \frac{\pi^2}{4}, \quad \theta \in (0, \pi],$$

we obtain

$$1 \leq \frac{g_{p-1}(\theta)}{g_p(\theta)} < 3,$$

which is equivalent to $1/3 < w_p(\theta) \leq 1$. □

In order to prove Theorem 2, further work is needed. In particular, we shall need to analyze the auxiliary functions

$$R_{k,p}(\omega) = \left(\frac{\omega}{k\pi + \omega}\right)^{2p+1} - \left(\frac{\omega}{k\pi - \omega}\right)^{2p+1}, \quad k, p \geq 1, \quad \omega \in \left[0, \frac{\pi}{2}\right]. \tag{A8}$$

The next three technical lemmas are devoted to this purpose.

Lemma 2. For $p \geq 1$ and $k \geq 2$, the function

$$R_{k,p+1}(\omega) - R_{k,p}(\omega) \tag{A9}$$

is nonnegative, monotone increasing, and convex for $\omega \in [0, \frac{\pi}{2}]$.

Proof. Assume $\omega \in [0, \frac{\pi}{2}]$. We have

$$R_{k,p+1}(\omega) - R_{k,p}(\omega) = z_k^{2p+3} - z_k^{2p+1} + y_k^{2p+1} - y_k^{2p+3}, \quad k \geq 1,$$

where

$$y_k = \frac{\omega}{k\pi - \omega}, \quad z_k = \frac{\omega}{k\pi + \omega}. \tag{A10}$$

It is easy to check that y_k is a monotone increasing and convex function of ω . Similarly, z_k is a monotone increasing and concave function of ω . Moreover,

$$\frac{z_k}{y_k} = \frac{1}{1 + 2y_k}, \quad \frac{z'_k}{y'_k} = \left(\frac{z_k}{y_k}\right)^2, \quad \frac{z''_k}{y''_k} = -\left(\frac{z_k}{y_k}\right)^3, \quad k \geq 1, \tag{A11}$$

and

$$0 \leq z_k \leq y_k \leq \frac{1}{2k-1} \leq \frac{1}{3}, \quad k \geq 2. \tag{A12}$$

Proving the nonnegativity of the function in (A9) is equivalent to showing that

$$y_k^{2p+1} (1 - y_k^2) \geq z_k^{2p+1} (1 - z_k^2).$$

In view of (A11), this is equivalent to

$$\frac{1 - y_k^2}{1 - z_k^2} \geq \frac{1}{(1 + 2y_k)^{2p+1}}.$$

Because

$$\frac{1 - y_k^2}{1 - z_k^2} \geq 1 - y_k^2,$$

it suffices to prove that

$$1 - y_k^2 \geq \frac{1}{(1 + 2y_k)^{2p+1}}.$$

A direct computation shows that the above inequality holds for $y_k \in [0, 1/3]$ (it is enough to verify it for $p = 1$, as the right-hand side decreases with p). Taking into account (A12), this proves the nonnegativity of (A9) for $k \geq 2$.

We now show that the function (A9) is convex. With some elementary manipulations, we obtain

$$R''_{k,p+1}(\omega) - R''_{k,p}(\omega) = A_k + B_k - C_k - D_k, \tag{A13}$$

where

$$A_k = 2y_k^{2p-1} (y'_k)^2 [p(2p+1) - (p+1)(2p+3)y_k^2], \quad B_k = y_k^{2p} y''_k [2p+1 - (2p+3)y_k^2],$$

$$C_k = 2z_k^{2p-1} (z'_k)^2 [p(2p+1) - (p+1)(2p+3)z_k^2], \quad D_k = z_k^{2p} z''_k [2p+1 - (2p+3)z_k^2].$$

From (A12), it follows that, for $p \geq 1$ and $k \geq 2$,

$$p(2p+1) - (p+1)(2p+3)x_k^2 > 0, \quad 2p+1 - (2p+3)x_k^2 > 0, \quad x_k = y_k, z_k.$$

As a consequence, we have $B_k \geq 0$ and $D_k \leq 0$ because $y''_k \geq 0$ and $z''_k \leq 0$. In the following, we show that $A_k \geq C_k$. Taking into account (A11), this is equivalent to proving that

$$\frac{p(2p+1) - (p+1)(2p+3)y_k^2}{p(2p+1) - (p+1)(2p+3)z_k^2} \geq \left(\frac{z_k}{y_k}\right)^{2p-1} \left(\frac{z'_k}{y'_k}\right)^2 = \frac{1}{(1 + 2y_k)^{2p+3}}.$$

Because

$$\frac{p(2p+1) - (p+1)(2p+3)y_k^2}{p(2p+1) - (p+1)(2p+3)z_k^2} \geq \frac{p(2p+1) - (p+1)(2p+3)y_k^2}{p(2p+1)},$$

it suffices to prove that

$$1 - \frac{(p+1)(2p+3)}{p(2p+1)}y_k^2 \geq \frac{1}{(1+2y_k)^{2p+3}}.$$

The above inequality holds for $p \geq 1$ and $y_k \in [0, 1/3]$ (it is enough to verify it for $p = 1$). Recalling (A12), this shows the convexity of (A9).

Finally, the monotonicity of the function (A9) follows from convexity by observing that the first derivative vanishes at $\omega = 0$. \square

Lemma 3. For $p \geq 1$, the function

$$R_{1,p+1}(\omega) - R_{1,p}(\omega) \tag{A14}$$

is nonnegative for $\omega \in [0, \omega_p^*]$ and concave for $\omega \in [\omega_p^*, \frac{\pi}{2}]$, where

$$\omega_p^* = \frac{\pi}{2} \left(1 - \frac{1}{48p-1} \right). \tag{A15}$$

Proof. Along the proof, we use the same notation as in the proof of Lemma 2. We first address the nonnegativity. With the same line of arguments as in the proof of Lemma 2, we deduce that the function in (A14) is nonnegative if

$$1 - y_1^2 \geq \frac{1}{(1+2y_1)^{2p+1}}.$$

The above inequality holds for $p \geq 1$ whenever

$$0 \leq y_1 \leq 1 - \frac{1}{24p} = y_{1,p}^*. \tag{A16}$$

In view of (A10) and (A15), this is equivalent to $0 \leq \omega \leq \omega_p^*$.

We now prove concavity. Similar to (A13), we have

$$R''_{1,p+1}(\omega) - R''_{1,p}(\omega) = A_1 + B_1 - C_1 - D_1,$$

where

$$\begin{aligned} A_1 &= 2y_1^{2p-1}(y_1')^2 [p(2p+1) - (p+1)(2p+3)y_1^2], & B_1 &= y_1^{2p}y_1'' [2p+1 - (2p+3)y_1^2], \\ C_1 &= 2z_1^{2p-1}(z_1')^2 [p(2p+1) - (p+1)(2p+3)z_1^2], & D_1 &= z_1^{2p}z_1'' [2p+1 - (2p+3)z_1^2]. \end{aligned}$$

Because $0 \leq z_1 \leq \frac{1}{3}$, we have

$$p(2p+1) - (p+1)(2p+3)z_1^2 > 0, \quad 2p+1 - (2p+3)z_1^2 > 0.$$

Moreover, for $y_1 \geq y_{1,p}^*$,

$$p(2p+1) - (p+1)(2p+3)y_1^2 < 0, \quad 2p+1 - (2p+3)y_1^2 < 0.$$

Hence, $A_1 < 0$ and $C_1 > 0$ for $\omega \in [\omega_p^*, \frac{\pi}{2}]$. In the following, for $\omega \in [\omega_p^*, \frac{\pi}{2}]$, we prove that $B_1 \leq D_1$ or, equivalently,

$$\frac{2p+1 - (2p+3)y_1^2}{2p+1 - (2p+3)z_1^2} \leq \left(\frac{z_1}{y_1} \right)^{2p} \frac{z_1''}{y_1''}.$$

By (A11), this is equivalent to

$$\frac{(2p+3)y_1^2 - (2p+1)}{2p+1 - (2p+3)z_1^2} \geq \frac{1}{(1+2y_1)^{2p+3}}.$$

Because

$$\frac{(2p+3)y_1^2 - (2p+1)}{2p+1 - (2p+3)z_1^2} \geq \frac{(2p+3)y_1^2 - (2p+1)}{2p+1 - (2p+1)z_1^2} = \left(\frac{2p+3}{2p+1} y_1^2 - 1 \right) \frac{1}{1-z_1^2} \geq \frac{2p+3}{2p+1} y_1^2 - 1,$$

it suffices to prove that, for $\omega \in [\omega_p^*, \frac{\pi}{2}]$,

$$\frac{2p+3}{2p+1} y_1^2 - 1 \geq \frac{1}{(1+2y_1)^{2p+3}}. \quad (\text{A17})$$

Note that the left-hand side in (A17) is monotone increasing, whereas the right-hand side is monotone decreasing. Thus, the observation that the inequality (A17) holds for $y_1 = y_{1,p}^*$ and $p \geq 1$ concludes the proof. \square

Lemma 4. For $p \geq 1$ and $\omega \in [0, \frac{\pi}{2}]$, we have

$$1 + (p+1) \sum_{k \geq 1} R_{k,p+1}(\omega) - p \sum_{k \geq 1} R_{k,p}(\omega) \geq 0. \quad (\text{A18})$$

Proof. Assume $\omega \in [0, \frac{\pi}{2}]$. When taking the derivative of $R_{k,p}$,

$$R'_{k,p}(\omega) = (2p+1) \left[\left(\frac{\omega}{k\pi + \omega} \right)^{2p} \frac{k\pi}{(k\pi + \omega)^2} - \left(\frac{\omega}{k\pi - \omega} \right)^{2p} \frac{k\pi}{(k\pi - \omega)^2} \right] \leq 0, \quad (\text{A19})$$

we see that $R_{k,p}(\omega)$ is a monotone decreasing function with $R_{k,p}(0) = 0$. In addition,

$$\sum_{k \geq 1} R_{k,p} \left(\frac{\pi}{2} \right) = \sum_{k \geq 1} \left[\frac{1}{(2k+1)^{2p+1}} - \frac{1}{(2k-1)^{2p+1}} \right] = -1, \quad (\text{A20})$$

so

$$1 + \sum_{k \geq 1} R_{k,p+1}(\omega) \geq 1 + \sum_{k \geq 1} R_{k,p} \left(\frac{\pi}{2} \right) = 0. \quad (\text{A21})$$

In the following, we prove that the sum of the remaining terms in (A18) is nonnegative as well, that is,

$$p \sum_{k \geq 1} [R_{k,p+1}(\omega) - R_{k,p}(\omega)] \geq 0.$$

From Lemmas 2 and 3, it follows that this is true for $\omega \in [0, \omega_p^*]$. Therefore, it remains to show that

$$S_p(\omega) = \sum_{k \geq 2} [R_{k,p+1}(\omega) - R_{k,p}(\omega)] \geq R_{1,p}(\omega) - R_{1,p+1}(\omega), \quad \omega \in \left[\omega_p^*, \frac{\pi}{2} \right]. \quad (\text{A22})$$

To this end, we first deduce from (A20) that

$$\sum_{k \geq 1} \left[R_{k,p+1} \left(\frac{\pi}{2} \right) - R_{k,p} \left(\frac{\pi}{2} \right) \right] = 0,$$

implying that

$$S_p \left(\frac{\pi}{2} \right) = R_{1,p} \left(\frac{\pi}{2} \right) - R_{1,p+1} \left(\frac{\pi}{2} \right) = \frac{1}{3^{2p+1}} - \frac{1}{3^{2p+3}} \geq 0.$$

Moreover, from (A19), we get

$$R'_{k,p} \left(\frac{\pi}{2} \right) = (2p+1) \frac{4k}{\pi} \left[\frac{1}{(2k+1)^{2p+2}} - \frac{1}{(2k-1)^{2p+2}} \right],$$

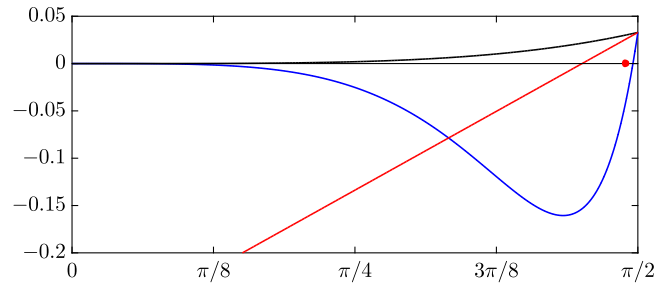


FIGURE A1 Graphs of $S_1(\omega)$ (black), $R_{1,1}(\omega) - R_{1,2}(\omega)$ (blue), and $T_1(\omega)$ (red). The value ω_1^* is marked with a red dot

which gives[§]

$$\begin{aligned}
 S'_p\left(\frac{\pi}{2}\right) &= R'_{2,p+1}\left(\frac{\pi}{2}\right) - R'_{2,p}\left(\frac{\pi}{2}\right) + \sum_{k \geq 3} \left[R'_{k,p+1}\left(\frac{\pi}{2}\right) - R'_{k,p}\left(\frac{\pi}{2}\right) \right] \\
 &\leq \frac{16(p+1)}{3^{2p+2}\pi} + \frac{4}{\pi} \sum_{k \geq 3} \left[\frac{(2p+3)k}{(2k+1)^{2p+4}} + \frac{(2p+1)k}{(2k-1)^{2p+2}} \right] \\
 &\leq \frac{16(p+1)}{3^{2p+2}\pi} + \frac{12}{5\pi} \sum_{k \geq 3} \left[\frac{2p+3}{(2k+1)^{2p+3}} + \frac{2p+1}{(2k-1)^{2p+1}} \right] \\
 &\leq \frac{16(p+1)}{3^{2p+2}\pi} + \frac{12}{5\pi} \int_2^{+\infty} \left[\frac{2p+3}{(2\kappa+1)^{2p+3}} + \frac{2p+1}{(2\kappa-1)^{2p+1}} \right] d\kappa \\
 &= \frac{16(p+1)}{3^{2p+2}\pi} + \frac{6}{5\pi} \left[\frac{2p+3}{2(p+1)5^{2p+2}} + \frac{2p+1}{2p3^{2p}} \right] \\
 &\leq \frac{p+1}{3^{2p-1}\pi} = m_p.
 \end{aligned}$$

From Lemma 2, it follows that $S_p(\omega)$ is convex on $[0, \frac{\pi}{2}]$, so

$$S_p(\omega) \geq \left(\omega - \frac{\pi}{2}\right) m_p + S_p\left(\frac{\pi}{2}\right) = T_p(\omega). \quad (\text{A23})$$

The straight line $T_p(\omega)$ vanishes at

$$\hat{\omega}_p = \frac{\pi}{2} - S_p\left(\frac{\pi}{2}\right) \frac{1}{m_p} = \frac{\pi}{2} - \left(\frac{1}{3^{2p+1}} - \frac{1}{3^{2p+3}}\right) \frac{3^{2p-1}\pi}{p+1} = \frac{\pi}{2} - \frac{8\pi}{81(p+1)},$$

and

$$\hat{\omega}_p = \frac{\pi}{2} \left(1 - \frac{16}{81(p+1)}\right) < \omega_p^*.$$

From Lemma 3, we know that $R_{1,p}(\omega) - R_{1,p+1}(\omega)$ is convex on $[\omega_p^*, \frac{\pi}{2}]$, and hence,

$$R_{1,p}(\omega) - R_{1,p+1}(\omega) \leq T_p(\omega), \quad \omega \in \left[\omega_p^*, \frac{\pi}{2}\right]. \quad (\text{A24})$$

These functions are illustrated in Figure A1. By combining (A23) and (A24), we get (A22). \square

We are now ready to prove Theorem 2.

Proof of Theorem 2. Assume $\theta \in [0, \pi]$. From lemma A.2 in the work of Donatelli et al.,¹⁶ we know that $g'_p(\theta) \leq 0$.[¶] Moreover, by (16)–(17), we have $g_p(\theta), f_p(\theta) \geq 0$ and $f_p(\theta) = (2 - 2\cos(\theta))g_{p-1}(\theta)$. Finally, from the lower bound

[§]The equality holds due to the uniform convergence of the series.

[¶]Note that in the work of Donatelli et al.,¹⁶ the function $g_p(\theta)$ is denoted by $h_p(\theta)$.

in (A3), we deduce that $f_p(\theta) \geq \theta^2 g_p(\theta)$. Therefore,

$$\begin{aligned} e'_p(\theta) &= \frac{f'_p(\theta)g_p(\theta) - f_p(\theta)g'_p(\theta)}{(g_p(\theta))^2} \geq \frac{f'_p(\theta) - \theta^2 g'_p(\theta)}{g_p(\theta)} \\ &= \frac{2 \sin(\theta)g_{p-1}(\theta) + (2 - 2 \cos(\theta))g'_{p-1}(\theta) - \theta^2 g'_p(\theta)}{g_p(\theta)}. \end{aligned}$$

This means that, in order to prove the monotonicity of e_p , it suffices to show that

$$G_p(\theta) = 2 \sin(\theta)g_{p-1}(\theta) + (2 - 2 \cos(\theta))g'_{p-1}(\theta) - \theta^2 g'_p(\theta) \geq 0, \quad \theta \in [0, \pi]. \tag{A25}$$

From (14), it follows that $g'_p(0) = g'_p(\pi) = 0$ for $p \geq 0$, so that $G_p(0) = G_p(\pi) = 0$ for $p \geq 1$. It remains to prove the inequality in (A25) for $\theta \in (0, \pi)$.

Let $\omega = \frac{\theta}{2} \in (0, \frac{\pi}{2})$. From the proof of lemma A.2 in the work of Donatelli et al.,¹⁶ we know that

$$g_p(\theta) = \sum_{k \in \mathbb{Z}} \left(\frac{\sin(\omega)}{\omega + k\pi} \right)^{2p+2},$$

and

$$g'_p(\theta) = (p + 1)(\sin(\omega))^{2p+1} \cos(\omega) \sum_{k \in \mathbb{Z}} \left[\frac{1}{(\omega + k\pi)^{2p+2}} - \frac{\tan(\omega)}{(\omega + k\pi)^{2p+3}} \right].$$

Therefore, recalling that $2 - 2 \cos(\theta) = 4(\sin(\omega))^2$ and $\sin(\theta) = 2 \sin(\omega) \cos(\omega)$, with some manipulations we obtain

$$\begin{aligned} G_p(\theta) &= 4(\sin(\omega))^{2p+1} \left(\cos(\omega)(p + 1) \sum_{k \in \mathbb{Z}} \frac{1}{(\omega + k\pi)^{2p}} \left[1 - \left(\frac{\omega}{\omega + k\pi} \right)^2 \right] \right. \\ &\quad \left. + \sin(\omega) \sum_{k \in \mathbb{Z}} \frac{1}{(\omega + k\pi)^{2p+1}} \left[(p + 1) \left(\frac{\omega}{\omega + k\pi} \right)^2 - p \right] \right). \end{aligned} \tag{A26}$$

Considering the positivity of the first sum in (A26), it suffices to show that

$$\begin{aligned} &\frac{4(\sin(\omega))^{2p+2}}{\omega^{2p+1}} \left(1 + \sum_{k \geq 1} \left(\frac{\omega}{k\pi + \omega} \right)^{2p+1} \left[(p + 1) \left(\frac{\omega}{k\pi + \omega} \right)^2 - p \right] \right. \\ &\quad \left. - \sum_{k \geq 1} \left(\frac{\omega}{k\pi - \omega} \right)^{2p+1} \left[(p + 1) \left(\frac{\omega}{k\pi - \omega} \right)^2 - p \right] \right) \geq 0. \end{aligned}$$

This inequality follows from (A18). □

APPENDIX B

Proof of the eigenvalue expansion for $\alpha = 0$

This Appendix is devoted to the proof of the following theorem, that is, the expansion (18) for $\alpha = 0$ and $j = 1, \dots, N(n, p) - (4p - 2)$.

Theorem 7. For every $p \geq 3$, every n , and every $j = 1, \dots, N(n, p) - (4p - 2) = n - 3p$, we have

$$\lambda_j \left(n^{-2} L_n^{[p]} \right) = e_p(\theta_{j,n}) + E_{j,n,0}^{[p]}, \tag{B1}$$

where

- the eigenvalues of $n^{-2} L_n^{[p]}$ are arranged in ascending order, $\lambda_1(n^{-2} L_n^{[p]}) \leq \dots \leq \lambda_{n+p-2}(n^{-2} L_n^{[p]})$;
- e_p is the function defined in (15);
- $h = \frac{1}{n}$ and $\theta_{j,n} = \frac{j\pi}{n} = j\pi h$ for $j = 1, \dots, n$; and
- $|E_{j,n,0}^{[p]}| \leq C^{[p]} h$ for some constant $C^{[p]}$ depending only on p .

Proof. Throughout this proof, we will use the simplified notations $N = N(p, n)$ and $\rho = 4p - 2$. Moreover, we will write $V \subseteq_{\text{sp.}} \mathbb{C}^N$ to indicate that V is a vector subspace of \mathbb{C}^N . If A is an $N \times N$ matrix and $V \subseteq_{\text{sp.}} \mathbb{C}^N$, the symbol $A(V)$

$\dim(U \cap F^\perp) = \dim U + \dim F^\perp - \dim(U + F^\perp) \geq u + (N - \rho) - N = u - \rho$. Thus, for $j = 1, \dots, N - \rho$, from (B5)–(B6) and (B9), we obtain

$$\begin{aligned}
\lambda_j \left(n^{-2} L_n^{[p]} \right) &\leq \max_{\substack{U \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim U = N-j+1}} \min_{\substack{y \in U \cap F^\perp \\ y \neq 0}} \frac{y^* \left(\tau_N(f_p) + \hat{R}_n^{[p]} \right) y}{y^* \left(\tau_N(g_p) + \hat{S}_n^{[p]} \right) y} \\
&= \max_{\substack{U \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim U = N-j+1}} \min_{y \neq 0} \frac{y^* \tau_N(f_p) y}{y^* \tau_N(g_p) y} \\
&\leq \max_{\substack{W \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim W \geq N-(j+\rho)+1}} \min_{y \in W} \frac{y^* \tau_N(f_p) y}{y^* \tau_N(g_p) y} \\
&= \max_{\substack{W \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim W \geq N-(j+\rho)+1}} \min_{\substack{x \in (\tau_N(g_p))^{1/2}(W) \\ x \neq 0}} \frac{x^* \left(\tau_N(g_p) \right)^{-1/2} \tau_N(f_p) \left(\tau_N(g_p) \right)^{-1/2} x}{x^* x} \\
&= \max_{\substack{V \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim V \geq N-(j+\rho)+1}} \min_{\substack{x \in V \\ x \neq 0}} \frac{x^* \tau_N(e_p) x}{x^* x} \\
&= \max_{\substack{V \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim V = N-(j+\rho)+1}} \min_{\substack{x \in V \\ x \neq 0}} \frac{x^* \tau_N(e_p) x}{x^* x} \\
&= \lambda_{j+\rho}(\tau_N(e_p)) = e_p \left(\frac{(j + \rho)\pi}{N + 1} \right), \tag{B10}
\end{aligned}$$

where the last equality is due to the monotonicity of e_p (Theorem 2). Similarly, using again the minimax principle for Hermitian matrices, for $j = \rho + 1, \dots, N$, we obtain

$$\begin{aligned}
\lambda_j \left(n^{-2} L_n^{[p]} \right) &= \lambda_j \left(n^{-2} \left(M_n^{[p]} \right)^{-1/2} K_n^{[p]} \left(M_n^{[p]} \right)^{-1/2} \right) \\
&= \min_{\substack{V \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim V = j}} \max_{\substack{x \in V \\ x \neq 0}} \frac{n^{-2} x^* \left(M_n^{[p]} \right)^{-1/2} K_n^{[p]} \left(M_n^{[p]} \right)^{-1/2} x}{x^* x} \\
&= \min_{\substack{V \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim V = j}} \max_{\substack{y \in \left(M_n^{[p]} \right)^{-1/2}(V) \\ y \neq 0}} \frac{n^{-2} y^* K_n^{[p]} y}{y^* M_n^{[p]} y} \\
&= \min_{\substack{U \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim U = j}} \max_{y \neq 0} \frac{y^* \left(n^{-1} K_n^{[p]} \right) y}{y^* \left(n M_n^{[p]} \right) y} \\
&\geq \min_{\substack{U \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim U = j}} \max_{y \neq 0} \frac{y^* \left(\tau_N(f_p) + \hat{R}_n^{[p]} \right) y}{y^* \left(\tau_N(g_p) + \hat{S}_n^{[p]} \right) y} \\
&= \min_{\substack{U \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim U = j}} \max_{y \neq 0} \frac{y^* \tau_N(f_p) y}{y^* \tau_N(g_p) y} \\
&\geq \min_{\substack{W \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim W \geq j-\rho}} \max_{y \in W} \frac{y^* \tau_N(f_p) y}{y^* \tau_N(g_p) y} \\
&= \min_{\substack{W \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim W \geq j-\rho}} \max_{\substack{x \in (\tau_N(g_p))^{1/2}(W) \\ x \neq 0}} \frac{x^* \left(\tau_N(g_p) \right)^{-1/2} \tau_N(f_p) \left(\tau_N(g_p) \right)^{-1/2} x}{x^* x} \\
&= \min_{\substack{V \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim V \geq j-\rho}} \max_{\substack{x \in V \\ x \neq 0}} \frac{x^* \tau_N(e_p) x}{x^* x} \\
&= \min_{\substack{V \subseteq_{\text{sp.}} \mathbb{C}^N \\ \dim V = j-\rho}} \max_{\substack{x \in V \\ x \neq 0}} \frac{x^* \tau_N(e_p) x}{x^* x} \\
&= \lambda_{j-\rho}(\tau_N(e_p)) = e_p \left(\frac{(j - \rho)\pi}{N + 1} \right). \tag{B11}
\end{aligned}$$

Putting together (B10) and (B11), we get

$$e_p \left(\frac{(j-\rho)\pi}{N+1} \right) \leq \lambda_j \left(n^{-2}L_n^{[p]} \right) \leq e_p \left(\frac{(j+\rho)\pi}{N+1} \right), \quad j = \rho+1, \dots, N-\rho. \quad (\text{B12})$$

From (B12), we immediately obtain

$$\begin{aligned} \left| \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right) \right| &\leq \max \left(\left| e_p \left(\frac{(j-\rho)\pi}{N+1} \right) - e_p \left(\frac{j\pi}{N+1} \right) \right|, \left| e_p \left(\frac{(j+\rho)\pi}{N+1} \right) - e_p \left(\frac{j\pi}{N+1} \right) \right| \right) \\ &\leq \|e'_p\|_\infty \frac{\rho\pi}{N+1} \leq \|e'_p\|_\infty \rho\pi h, \quad j = \rho+1, \dots, N-\rho. \end{aligned} \quad (\text{B13})$$

Moreover, since the eigenvalues of $n^{-2}L_n^{[p]}$ are positive (because of the similarity between $L_n^{[p]}$ and the symmetric positive definite matrix $(M_n^{[p]})^{-1/2}K_n^{[p]}(M_n^{[p]})^{-1/2}$) and $e_p(0) = 0 = \min_{\theta \in [0, \pi]} e_p(\theta)$ (by (16)–(17)), for $j = 1, \dots, \rho$, we have

$$\begin{aligned} \left| \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right) \right| &= \begin{cases} \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right), & \text{if } \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right) \geq 0, \\ e_p \left(\frac{j\pi}{N+1} \right) - \lambda_j \left(n^{-2}L_n^{[p]} \right), & \text{otherwise,} \end{cases} \\ &\leq \begin{cases} \lambda_{\rho+1} \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right), & \text{if } \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right) \geq 0, \\ e_p \left(\frac{j\pi}{N+1} \right), & \text{otherwise,} \end{cases} \\ &\leq \begin{cases} \left| \lambda_{\rho+1} \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{(\rho+1)\pi}{N+1} \right) \right| + e_p \left(\frac{(\rho+1)\pi}{N+1} \right) - e_p \left(\frac{j\pi}{N+1} \right), & \text{if } \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right) \geq 0, \\ e_p \left(\frac{\rho\pi}{N+1} \right) - e_p(0), & \text{otherwise,} \end{cases} \\ &\leq \begin{cases} \|e'_p\|_\infty \rho\pi h + \|e'_p\|_\infty \rho\pi h, & \text{if } \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right) \geq 0, \\ \|e'_p\|_\infty \rho\pi h, & \text{otherwise,} \end{cases} \\ &\leq 2\|e'_p\|_\infty \rho\pi h. \end{aligned} \quad (\text{B14})$$

Combining (B13) and (B14), we obtain

$$\left| \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p \left(\frac{j\pi}{N+1} \right) \right| \leq 2\|e'_p\|_\infty \rho\pi h, \quad j = 1, \dots, N-\rho. \quad (\text{B15})$$

To conclude the proof, we note that the step sizes $h = \frac{1}{n}$ and $H = \frac{1}{N+1}$ are such that

$$0 < h - H = \frac{N+1-n}{n(N+1)} = \frac{p-1}{n(n+p-1)} < \frac{p}{n^2}, \quad (\text{B16})$$

and consequently, the grid points $\theta_{j,n} = j\pi h$ and $\Theta_{j,n} = j\pi H$ satisfy

$$0 < \theta_{j,n} - \Theta_{j,n} < \frac{p\pi}{n}, \quad j = 1, \dots, n. \quad (\text{B17})$$

Thus, the inequality (B15) yields the thesis (B1) with

$$\begin{aligned} |E_{j,n,0}^{[p]}| &= \left| \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p(\theta_{j,n}) \right| \leq \left| \lambda_j \left(n^{-2}L_n^{[p]} \right) - e_p(\Theta_{j,n}) \right| + |e_p(\Theta_{j,n}) - e_p(\theta_{j,n})| \\ &\leq 2\|e'_p\|_\infty \rho\pi h + \|e'_p\|_\infty p\pi h = C^{[p]}h, \quad j = 1, \dots, N-\rho, \end{aligned}$$

where $C^{[p]} = (2\rho + p)\pi\|e'_p\|_\infty$. \square