# Eigenvalue Isogeometric Approximations Based on B-Splines: Tools and Results 

Sven-Erik Ekström and Stefano Serra-Capizzano


#### Abstract

In this note, we focus on the spectral analysis of large matrices coming from isogeometric approximations based on B-splines of the eigenvalue problem $$
-\left(a(x) u^{\prime}(x)\right)^{\prime}=\lambda b(x) u(x), \quad x \in(0,1),
$$ where $u(0)$ and $u(1)$ are given. When considering the collocation case, global distribution results for the eigenvalues are available in the literature, despite the nonsymmetry of the related matrices. Here we complement such results by providing precise estimates for the extremal eigenvalues and hence for the spectral conditioning of the resulting matrices. In the Galerkin setting, the matrices are symmetric and positive definite and a more complete analysis has been conducted in the literature. In the latter case we furnish a further procedure that gives a highly accurate estimate of all the eigenvalues, starting from the knowledge of the spectral distribution symbol. The techniques involve dyadic decomposition arguments, tools from the theory of generalized locally Toeplitz sequences, and basic extrapolation methods.


## 1 Introduction

In this note we consider the approximation of one-dimensional elliptic eigenvalue problems by using an isogeometric either Galerkin or collocation technique with Bsplines [5]. We are interested in the eigenvalues of the large matrices stemming from the considered approximation processes. In particular, we address the problem of estimating the extremal eigenvalues and of providing efficient numerical procedures for computing a good approximation of all the eigenvalues.

[^0]In this direction, it has recently been proven that the resulting sequence of matrices, indexed with respect to the matrix size, have a canonical distribution (see [15, 17] and references therein), by using the theory of Generalized Locally Toeplitz (GLT) sequences [16].

We recall that every GLT sequence has an associated function, called the symbol, and that the uniform sampling of the symbol provides an asymptotic approximation of all the eigenvalues, if every matrix of the GLT sequence is Hermitian as it happens in our setting. However, in general the approximation is quite poor and recently some extrapolation techniques have been devised (see [9] and references therein). In the constant coefficient setting, when considering the Galerkin approach, the results presented in [10] are impressive in the sense that machine precision is obtained with very low computational cost, while for the variable coefficients further improvements are needed (see also [17]).

In the current note, when considering the Galerkin setting and variable coefficients, we propose a further numerical scheme for the computation of all the spectrum of large matrices by using the numerical computation of the eigenvalues for small matrices, underlying asymptotic expansions, and extrapolation methods as those in [9]. The numerical results are of the same quality as those produced by the best strategy in [17], especially when the problem coefficients are smooth and at least for low frequencies, which are those of highest interest for understanding the nature of the problem.

On the other hand, when dealing with the collocation approximation we obtain large nonsymmetric matrices. However, the GLT machinery can be employed and the symbol is real-valued and nonnegative [7], as in the symmetric positive definite Galerkin setting [15]. Hence we expect that most of the eigenvalues are real or with negligible imaginary part. Here we start the analysis of the collocation case, by describing a technique based on dyadic decompositions, for estimating the extreme eigenvalues and hence the asymptotic (spectral) conditioning of the involved matrix sequences. The analysis contained in Theorem 1 and Corollary 1 confirms that the conditioning grows at most as $h^{-2}, h$ being the fineness parameter, exactly as in the case of the matrices obtained with the Galerkin approximation (see e.g [14]).

Finally, it is not clear if the delicate asymptotic expansions observed in [10, 17] holds also in the collocation setting and indeed this issue will be the subject of future investigations.

## 2 Preliminaries

In the following we present the notation that we use. In particular we give the definition of eigenvalue distribution, that of rearrangement, and we briefly discuss the informal meaning behind these definitions.

A matrix-sequence is any sequence of the form $\left\{\mathbf{X}_{n}\right\}_{n}$, where $\mathbf{X}_{n}$ is a square matrix such that $\operatorname{size}\left(\mathbf{X}_{n}\right) \rightarrow \infty$ as $n \rightarrow \infty$. Let $\mu_{d}$ be the Lebesgue measure in $\mathbb{R}^{d}$ and let $C_{c}(\mathbb{C})$ be the space of continuous complex-valued functions with bounded
support defined on $\mathbb{C}$. If $\mathbf{X}$ is an $m \times m$ matrix, the eigenvalues of $\mathbf{X}$ are denoted by $\lambda_{1}(\mathbf{X}), \ldots, \lambda_{m}(\mathbf{X})$. If $\mathbf{g}: D \subset \mathbb{R}^{d} \rightarrow \mathbb{C}^{s \times s}$ is an $s \times s$ matrix-valued function, we say that $\mathbf{g}$ is measurable if its $s^{2}$ components $g_{i j}: D \rightarrow \mathbb{C}, i, j=1, \ldots, s$, are measurable.

Definition 1 Let $\left\{\mathbf{X}_{n}\right\}_{n}$ be a matrix-sequence, let $N_{n}:=\operatorname{size}\left(\mathbf{X}_{n}\right)$, and let $\mathbf{g}: D \subset$ $\mathbb{R}^{d} \rightarrow \mathbb{C}^{s \times s}$ be a measurable $s \times s$ matrix-valued function defined on a set $D$ with $0<\mu_{d}(D)<\infty$. We say that $\left\{\mathbf{X}_{n}\right\}_{n}$ has an (asymptotic) eigenvalue distribution described by $\mathbf{g}$, and we write $\left\{\mathbf{X}_{n}\right\}_{n} \sim_{\lambda} \mathbf{g}$, if

$$
\lim _{n \rightarrow \infty} \frac{1}{N_{n}} \sum_{i=1}^{N_{n}} F\left(\lambda_{i}\left(\mathbf{X}_{n}\right)\right)=\frac{1}{\mu_{d}(D)} \int_{D} \frac{\sum_{i=1}^{s} F\left(\lambda_{i}\left(\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)\right)\right)}{s} \mathrm{~d} y_{1} \ldots \mathrm{~d} y_{d}, \quad \forall F \in C_{c}(\mathbb{C}),
$$

where $\lambda_{i}\left(\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)\right), i=1, \ldots, s$, are the eigenvalues of the $s \times s$ matrix $\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)$.

The informal meaning behind the eigenvalue distribution $\left\{\mathbf{X}_{n}\right\}_{n} \sim_{\lambda} \mathbf{g}$ is the following: for all sufficiently large $n$, the eigenvalues of $\mathbf{X}_{n}$ can be subdivided into $s$ different subsets of approximately the same cardinality; and the eigenvalues belonging to the $i$ th subset (except possibly for $o\left(N_{n}\right)$ outliers) are approximated by the samples of the $i$ th eigenvalue function $\lambda_{i}\left(\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)\right)$ over a uniform grid in $D$ (the domain of $\mathbf{g}$ ). For example, if $d=1, N_{n}=n s$ and $D=[a, b]$, then the eigenvalues of $\mathbf{X}_{n}$ are approximately equal to

$$
\lambda_{i}\left(\mathbf{g}\left(a+j \frac{b-a}{n}\right)\right), \quad j=1, \ldots, n, \quad i=1, \ldots, s,
$$

for $n$ large enough. Likewise, if $d=2, N_{n}=n^{2} s$ and $D=\left[a_{1}, b_{1}\right] \times\left[a_{2}, b_{2}\right]$, then the eigenvalues of $\mathbf{X}_{n}$ are approximately equal to

$$
\lambda_{i}\left(\mathbf{g}\left(a_{1}+j_{1} \frac{b_{1}-a_{1}}{n}, a_{2}+j_{2} \frac{b_{2}-a_{2}}{n}\right)\right), \quad j_{1}, j_{2}=1, \ldots, n, \quad i=1, \ldots, s
$$

for $n$ large enough.
Remark 1 (rearrangement) Let $\mathbf{g}: D \subseteq \mathbb{R}^{d} \rightarrow \mathbb{C}^{s \times s}$ and suppose that $D$ is a rectangle in $\mathbb{R}^{d}$, say $D:=\left[a_{1}, b_{1}\right] \times \cdots \times\left[a_{d}, b_{d}\right]$. We also assume that the eigenvalues $\lambda_{1}\left(\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)\right), \ldots, \lambda_{s}\left(\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)\right)$ are real for all $\left(y_{1}, \ldots, y_{d}\right) \in D$. For each positive integer $r$, let $\mathcal{G}_{r}$ be the uniform grid in $D$ given by

$$
\mathcal{G}_{r}:=\left\{\left(a_{1}+\frac{i_{1}}{r}\left(b_{1}-a_{1}\right), \ldots, a_{d}+\frac{i_{d}}{r}\left(b_{d}-a_{d}\right)\right): i_{1}, \ldots, i_{d}=1, \ldots, r\right\} .
$$

Compute the samples of the eigenvalue functions

$$
\lambda_{1}\left(\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)\right), \ldots, \lambda_{s}\left(\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)\right)
$$

at the points $\left(y_{1}, \ldots, y_{d}\right) \in \mathcal{G}_{r}$, sort them in increasing order and put them in a vector $\left(q_{1}, q_{2}, \ldots, q_{s r^{d}}\right)$. Let $\eta_{r}:[0,1] \rightarrow \mathbb{R}$ be the piecewise linear non-decreasing function that interpolates the samples ( $q_{0}:=q_{1}, q_{1}, q_{2}, \ldots, q_{s r^{d}}$ ) over the nodes $\left(0, \frac{1}{s r^{d}}, \frac{2}{s r^{d}}, \ldots, 1\right)$, i.e.,

$$
\eta_{r}\left(\frac{\ell}{s r^{d}}\right):=q_{\ell}, \quad \ell=0, \ldots, s r^{d}
$$

Under certain (normally satisfied) conditions on $\mathbf{g}$, the function $\eta_{r}$ converges (a.e.) as $r \rightarrow \infty$ to a non-decreasing function $\eta:[0,1] \rightarrow \mathbb{R}$, which is referred to as the rearranged version of $\mathbf{g}$. What is important about $\eta$ is that

$$
\int_{D} \frac{\sum_{i=1}^{s} F\left(\lambda_{i}\left(\mathbf{g}\left(y_{1}, \ldots, y_{d}\right)\right)\right)}{s} \mathrm{~d} y_{1} \ldots \mathrm{~d} y_{d}=\int_{0}^{1} F(\eta(t)) \mathrm{d} t, \quad \forall F \in C_{c}(\mathbb{C}) .
$$

Therefore, if we have $\left\{\mathbf{X}_{n}\right\}_{n} \sim_{\lambda} \mathbf{g}$, then we also have $\left\{\mathbf{X}_{n}\right\}_{n} \sim_{\lambda} \eta$.

## 3 Isogeometric Galerkin Discretization of Variable-Coefficient Eigenvalue Problems

Consider the following one-dimensional variable-coefficient eigenvalue problem:

$$
\begin{cases}-\left(a(x) u_{j}^{\prime}(x)\right)^{\prime}=\lambda_{j} b(x) u_{j}(x), & x \in \Omega  \tag{1}\\ u_{j}(x)=0, & x \in \partial \Omega\end{cases}
$$

where $\Omega$ is an open interval in $\mathbb{R}$ and $a, b \in L^{1}(\Omega)$ are such that $a, b>0$ a.e. on $\Omega$. The corresponding weak formulation reads as follows: find eigenvalues $\lambda_{j} \in \mathbb{R}^{+}$ and eigenfunctions $u_{j} \in H_{0}^{1}(\Omega)$, for $j=1,2, \ldots, \infty$, such that, for all $v \in H_{0}^{1}(\Omega)$,

$$
\mathrm{a}\left(u_{j}, v\right)=\lambda_{j}\left(b u_{j}, v\right),
$$

where

$$
\mathrm{a}\left(u_{j}, v\right):=\int_{\Omega} a(x) u_{j}^{\prime}(x) v^{\prime}(x) \mathrm{d} x, \quad\left(b u_{j}, v\right):=\int_{\Omega} b(x) u_{j}(x) v(x) \mathrm{d} x .
$$

In the isogeometric Galerkin method, we assume that the physical domain $\Omega$ is described by a global geometry map $G:[0,1] \rightarrow \bar{\Omega}$, which is invertible and satisfies $G(\partial([0,1]))=\partial \bar{\Omega}$. We fix a set of basis functions $\left\{\varphi_{1}, \ldots, \varphi_{N_{n}}\right\}$ defined
on the reference (parametric) domain $[0,1]$ and vanishing on the boundary $\partial([0,1])$. We consider the basis functions

$$
\psi_{i}(x):=\varphi_{i}\left(G^{-1}(x)\right)=\varphi_{i}(t), \quad x=G(t), \quad i=1, \ldots, N_{n},
$$

which are defined on the physical domain $\Omega$, and we define the approximation space $\mathscr{W}_{n}:=\operatorname{span}\left(\psi_{1}, \ldots, \psi_{N_{n}}\right) \subset H_{0}^{1}(\Omega)$. Finally, we find approximations to the exact eigenpairs $\left(\lambda_{j}, u_{j}\right), j=1,2, \ldots, \infty$, by solving the following Galerkin problem: find $\lambda_{j, n} \in \mathbb{R}^{+}$and $u_{j, n} \in \mathscr{W}_{n}$, for $j=1, \ldots, N_{n}$, such that, for all $v_{n} \in \mathscr{W}_{n}$,

$$
\begin{equation*}
\mathrm{a}\left(u_{j, n}, v_{n}\right)=\lambda_{j, n}\left(b u_{j, n}, v_{n}\right) . \tag{2}
\end{equation*}
$$

Assuming that both the exact and numerical eigenvalues are arranged in nondecreasing order, the pair $\left(\lambda_{j, n}, u_{j, n}\right)$ is taken as an approximation of the pair $\left(\lambda_{j}, u_{j}\right)$ for all $j=1, \ldots, N_{n}$. The numbers $\lambda_{j, n} / \lambda_{j}-1, j=1, \ldots, N_{n}$, are referred to as the (relative) eigenvalue errors.

In view of the canonical identification of each function $v_{n} \in \mathscr{W}_{n}$ with its coefficient vector with respect to the basis $\left\{\varphi_{1}, \ldots, \varphi_{N_{n}}\right\}$, solving the Galerkin problem (2) is equivalent to solving the generalized eigenvalue problem

$$
\begin{equation*}
\mathbf{K}_{n}(a, G) \mathbf{u}_{j, n}=\lambda_{j, n} \mathbf{M}_{n}(b, G) \mathbf{u}_{j, n}, \tag{3}
\end{equation*}
$$

where $\mathbf{u}_{j, n}$ is the coefficient vector of $u_{j, n}$ with respect to $\left\{\varphi_{1}, \ldots, \varphi_{N_{n}}\right\}$ and

$$
\begin{align*}
\mathbf{K}_{n}(a, G):=\left[\mathrm{a}\left(\psi_{j}, \psi_{i}\right)\right]_{i, j=1}^{N_{n}} & =\left[\int_{\Omega} a(x) \psi_{j}^{\prime}(x) \psi_{i}^{\prime}(x) \mathrm{d} x\right]_{i, j=1}^{N_{n}} \\
& =\left[\int_{0}^{1} \frac{a(G(t))}{\left|G^{\prime}(t)\right|} \varphi_{j}^{\prime}(t) \varphi_{i}^{\prime}(t) \mathrm{d} t\right]_{i, j=1}^{N_{n}},  \tag{4}\\
\mathbf{M}_{n}(b, G):=\left[\left(b \psi_{j}, \psi_{i}\right)\right]_{i, j=1}^{N_{n}} & =\left[\int_{\Omega} b(x) \psi_{j}(x) \psi_{i}(x) \mathrm{d} x\right]_{i, j=1}^{N_{n}} \\
& =\left[\int_{0}^{1} b(G(t))\left|G^{\prime}(t)\right| \varphi_{j}(t) \varphi_{i}(t) \mathrm{d} t\right]_{i, j=1}^{N_{n}} . \tag{5}
\end{align*}
$$

The matrices $\mathbf{K}_{n}(a, G)$ and $\mathbf{M}_{n}(b, G)$ are, respectively, the stiffness and mass matrices. Due to our assumption that $a, b>0$ a.e. on $\Omega$, both $\mathbf{K}_{n}(a, G)$ and $\mathbf{M}_{n}(b, G)$ are always symmetric positive definite, regardless of the chosen basis functions $\varphi_{1}, \ldots, \varphi_{N_{n}}$ and the map $G$. Moreover, it is clear from (3) that the numerical eigenvalues $\lambda_{j, n}, j=1, \ldots, N_{n}$, are just the eigenvalues of the matrix

$$
\begin{equation*}
\mathbf{L}_{n}(a, b, G):=\left(\mathbf{M}_{n}(b, G)\right)^{-1} \mathbf{K}_{n}(a, G) . \tag{6}
\end{equation*}
$$

Therefore, if $a, b \in L^{1}(\Omega)$ and the basis functions $\varphi_{1}, \ldots, \varphi_{N_{n}}$ are chosen as the Bsplines of degree $p$ and global smoothness $C^{k}$ defined on the interval [0, 1] partitioned into $n$ equal subintervals (see [17, Sect.2.1]) and denoted by

$$
\begin{equation*}
B_{2,[p, k]}, \ldots, B_{n(p-k)+k,[p, k]}, \tag{7}
\end{equation*}
$$

then

$$
\begin{align*}
& \mathbf{K}_{n}(a, G)=\left[\int_{0}^{1} \frac{a(G(t))}{\left|G^{\prime}(t)\right|} B_{j+1,[p, k]}^{\prime}(t) B_{i+1,[p, k]}^{\prime}(t) \mathrm{d} t\right]_{i, j=1}^{n(p-k)+k-1},  \tag{8}\\
& \mathbf{M}_{n}(b, G)=\left[\int_{0}^{1} b(G(t))\left|G^{\prime}(t)\right| B_{j+1,[p, k]}(t) B_{i+1,[p, k]}(t) \mathrm{d} t\right]_{i, j=1}^{n(p-k)+k-1} . \tag{9}
\end{align*}
$$

From the $*$-algebra structure of GLT sequences, as proven in $[16,17]$, we know that

$$
\begin{array}{r}
\left\{\frac{1}{n} \mathbf{K}_{n}(a, G)\right\}_{n} \sim_{\lambda} \kappa_{[p, k]}, \\
\left\{n \mathbf{M}_{n}(b, G)\right\}_{n} \sim_{\lambda} \xi_{[p, k]}, \\
\left\{\frac{1}{n^{2}} \mathbf{L}_{n}(a, b, G)\right\}_{n} \sim_{\lambda} \zeta_{[p, k]}, \tag{12}
\end{array}
$$

where, according to Remark $1, \kappa_{[p, k]}, \xi_{[p, k]}, \zeta_{[p, k]}$ are the rearranged versions of

$$
\frac{a(G(t))}{\left|G^{\prime}(t)\right|} \mathbf{f}_{[p, k]}(\theta), \quad b(G(t))\left|G^{\prime}(t)\right| \mathbf{h}_{[p, k]}(\theta), \quad \frac{a(G(t))}{b(G(t))\left|G^{\prime}(t)\right|^{2}} \mathbf{e}_{[p, k]}(\theta),
$$

respectively, with

$$
\begin{align*}
\mathbf{f}_{[p, k]}(\theta) & :=\mathbf{K}_{[p, k]}^{[0]}+\sum_{\ell=1}^{\eta-1}\left(\mathbf{K}_{[p, k]}^{[\ell]} \mathrm{e}^{\mathbf{i} \ell \theta}+\left(\mathbf{K}_{[p, k]}^{[\ell]}\right)^{\mathrm{T}} \mathrm{e}^{-i \ell \theta}\right),  \tag{13}\\
\mathbf{h}_{[p, k]}(\theta) & :=\mathbf{M}_{[p, k]}^{[0]}+\sum_{\ell=1}^{\eta-1}\left(\mathbf{M}_{[p, k]}^{[\ell]} \mathrm{e}^{\mathbf{i} \ell \theta}+\left(\mathbf{M}_{[p, k]}^{[\ell]}\right)^{\mathrm{T}} \mathrm{e}^{-\mathbf{i} \ell \theta}\right),  \tag{14}\\
\mathbf{e}_{[p, k]}(\theta) & :=\left(\mathbf{h}_{[p, k]}(\theta)\right)^{-1} \mathbf{f}_{[p, k]}(\theta), \tag{15}
\end{align*}
$$

where the blocks $\mathbf{K}_{[p, k]}^{[\ell]}$ and $\mathbf{M}_{[p, k]}^{[\ell]}$, of size $p-k$, are defined in [17]. The analytical predictions of the eigenvalue errors are obtained through the following sampling procedure:

$$
\begin{equation*}
\frac{\lambda_{j, n}}{\lambda_{j}}-1 \approx \frac{n^{2} \zeta_{[p, k]}\left(\frac{j}{n(p-k)}\right)}{\lambda_{j}}-1, \quad j=1, \ldots, m \tag{16}
\end{equation*}
$$

where $m:=\min (n(p-k), n(p-k)+k-1)$.
However, such a procedure is not completely satisfactory especially in low frequencies (see [17]), when variable coefficients are considered. More precisely, we recall that the smallest eigenvalues are those related to slowly oscillating modes, that is the related eigenvectors belong to the subspace generated by low frequencies. Since the eigenvalues of the continuous operator appear in the denominator in formula (16), the relative error could be higher even if the absolute approximation is of high precision.

We address this issue in the next section.
For notational completeness, we stress that in the following the geometrical map $G$ is always the identity and hence the matrices we consider are those denoted as $\mathbf{L}_{n}(a, b):=\mathbf{L}_{n}(a, b, I)$ with $\mathbf{L}_{n}(a, b, G)$ as in (6); furthermore, the regularity will be maximal that is $k=p-1$ and hence $e_{p}(\theta), f_{p}(\theta), h_{p}(\theta)$ will denote $\mathbf{e}_{[p, p-1]}(\theta), \mathbf{f}_{[p, p-1]}(\theta), \mathbf{h}_{[p, p-1]}(\theta)$, respectively.

## 4 Global Distribution Results and Extrapolation

As discussed in [17], the mismatch between the analytical predictions and the eigenvalue errors essentially occurs only for small eigenvalues and a way to significantly reduce it has already been illustrated in that paper. Drawing inspiration from [1, 9, $10,12,13]$, we here describe-in the case of an isogeometric Galerkin discretization based on B-splines of the eigenvalue problem (1)-an alternative interpolationextrapolation procedure to considerably improve the analytical predictions for small eigenvalues. We also illustrate the performance of this procedure by two examples.

1. We assume that there exists a function $c_{p}:[0, \pi] \rightarrow \mathbb{R}$, depending only on $p$ and the coefficients $a, b$ of the considered eigenvalue problem (1), such that, when using an isogeometric $p$-degree $C^{p-1} \mathrm{~B}$-spline discretization, the following property holds: independently of the parameter $n$, all the eigenvalues of the matrix $\frac{1}{n^{2}} \mathbf{L}_{n}(a, b)$, sorted in increasing order, satisfy

$$
\begin{equation*}
\frac{\lambda_{j, n}}{n^{2}} \approx c_{p}\left(\theta_{j, n}\right) e_{p}\left(\theta_{j, n}\right), \quad j=1, \ldots, \min (n+p-2, n) \tag{17}
\end{equation*}
$$

where

$$
\begin{aligned}
e_{p}=\frac{f_{p}(\theta)}{h_{p}(\theta)} & =\frac{(2-2 \cos \theta) h_{p-1}(\theta)}{h_{p}(\theta)} \\
& =\frac{(2-2 \cos \theta)\left(\phi_{[2 p-1]}(p)+2 \sum_{\ell=1}^{p-1} \phi_{[2 p-1]}(p-\ell) \cos (\ell \theta)\right)}{\phi_{[2 p+1]}(p+1)+2 \sum_{\ell=1}^{p} \phi_{[2 p+1]}(p+1-\ell) \cos (\ell \theta)}
\end{aligned}
$$

is defined in [17], and

$$
\theta_{j, n}=\frac{j \pi}{n}, \quad j=1, \ldots, n
$$

2. Compute the eigenvalues of the matrix $\frac{1}{n_{1}^{2}} \mathbf{L}_{n_{1}}(a, b)$ corresponding to a small value of $n_{1}$, i.e.,

$$
\frac{\lambda_{j_{1}, n_{1}}}{n_{1}^{2}}, \quad j_{1}=1, \ldots, n_{1}+p-2
$$

Since $n_{1}$ is small, this computation can be efficiently performed by any standard eigensolver (e.g., the MATLAB eig function).
3. Under the assumption in item 1 ,

$$
c_{p}\left(\theta_{j_{1}, n_{1}}\right) \approx \frac{\lambda_{j_{1}, n_{1}}}{n_{1}^{2} e_{p}\left(\theta_{j_{1}, n_{1}}\right)}, \quad j_{1}=1, \ldots, \min \left(n_{1}+p-2, n_{1}\right)
$$

This means that we have an approximation of the unknown function $c_{p}$ over the coarse uniform grid consisting of the points $\theta_{j_{1}, n_{1}}$.
4. Interpolate the data

$$
\left(\theta_{j_{1}, n_{1}}, \frac{\lambda_{j_{1}, n_{1}}}{n_{1}^{2} e_{p}\left(\theta_{j_{1}, n_{1}}\right)}\right), \quad j_{1}=1, \ldots, \min \left(n_{1}+p-2, n_{1}\right)
$$

by using, e.g., the MATLAB interp1 function with the 'spline' option, so as to obtain an approximation of $c_{p}$ over the whole interval $[0, \pi]$. We call this approximation $\tilde{c}_{p}$. Note that, by construction,

$$
\tilde{c}_{p}\left(\theta_{j_{1}, n_{1}}\right)=\frac{\lambda_{j_{1}, n_{1}}}{n_{1}^{2} e_{p}\left(\theta_{j_{1}, n_{1}}\right)}, \quad j_{1}=1, \ldots, \min \left(n_{1}+p-2, n_{1}\right)
$$

5. Given a large $n$, compute approximations of the numerical eigenvalues $\lambda_{j, n}$ by replacing $c_{p}$ with $\tilde{c}_{p}$ in (17):

$$
\lambda_{j, n} \approx n^{2} \tilde{c}_{p}\left(\theta_{j, n}\right) e_{p}\left(\theta_{j, n}\right), \quad j=1, \ldots, \min (n+p-2, n)
$$

6. Compute analytical predictions for the eigenvalue errors as follows:

$$
\frac{\lambda_{j, n}}{\lambda_{j}}-1 \approx \frac{n^{2} \tilde{c}_{p}\left(\theta_{j, n}\right) e_{p}\left(\theta_{j, n}\right)}{\lambda_{j}}-1, \quad j=1, \ldots, \min (n+p-2, n)
$$

The next two examples show that the interpolation-extrapolation algorithm defined by items 1-6 may lead to analytical predictions of the eigenvalue errors which perform better (for small eigenvalues) than the analytical predictions proposed in (16).

Table 1 Example 1 [Linear $C^{0}$ B-spline discretization with $a(x)=2+0.5 x$ and $b(x)=1$ ]: The five smallest eigenvalues and eigenvalue approximations, scaled by $10^{2}$, for $n=200, n^{\prime}=1500$, $p^{\prime}=5$, and $r=10000$

| $p$ | $\lambda_{j, n^{\prime}}$ | $\lambda_{j, n}$ | $n^{2} \zeta_{r}\left(\frac{j}{n}\right)$ | $n^{2} \tilde{c}_{1}\left(\theta_{j, n}\right) e_{1}\left(\theta_{j, n}\right)$ |
| :--- | :--- | :--- | :--- | :--- |
| 1 | 0.221308917999822 | 0.221313437064828 | 0.225759339802366 | 0.221202202766985 |
|  | 0.885445782386009 | 0.885517680038136 | 0.894411370606914 | 0.884899608283264 |
|  | 1.992343949797817 | 1.992707562710680 | 2.006135775082463 | 1.991307942612523 |
| 3.54200822229894 | 3.543157060383101 | 3.560895188355945 | 3.540749097918976 |  |
| 5.534445409962895 | 5.537249919330117 | 5.559588944750918 | 5.533651310530473 |  |



Fig. 1 Example 1 [Linear $C^{0}$ B-spline discretization with $a(x)=2+0.5 x$ and $\left.b(x)=1\right]$ : analytical predictions $n^{2} \zeta_{r}\left(\frac{j}{n}\right) / \lambda_{j, n^{\prime}}-1$ and eigenvalue errors $\lambda_{j, n} / \lambda_{j, n^{\prime}}-1$ versus $j / N_{n}, j=1, \ldots, N_{n}$ ( $N_{n}=n-1, n=200, n^{\prime}=1500, p^{\prime}=5, r=10000$ )

Example 1 Let $p=1, n=200, a(x)=2+0.5 x$ and $b(x)=1$. Let $n^{\prime}=1500 \gg$ $n, p^{\prime}=5$, and take the first $n-1$ eigenvalues of $\mathbf{L}_{n^{\prime}}(a, b)$, namely $\lambda_{1, n^{\prime}}, \ldots, \lambda_{n-1, n^{\prime}}$, as approximations of the unknown exact eigenvalues $\lambda_{1}, \ldots, \lambda_{n-1}$. In Table 1 we report the five smallest eigenvalues and approximated eigenvalues for $p=1$. As is clear the new approximation, $n^{2} \tilde{c}_{1}\left(\theta_{j, n}\right) e_{1}\left(\theta_{j, n}\right)$, performs better than $n^{2} \zeta_{r}\left(\frac{j}{n}\right)$ to approximate $\lambda_{j, n}$, where the function $\zeta_{r}$ introduced in [17, Sect.3.1] is defined in the following way. Sample $\frac{a(x)}{b(x)} e_{1}(\theta)$ at the grid points $(x, \theta) \in \mathcal{G}_{r}$, for a chosen $r$, where

$$
\mathcal{G}_{r}=\left\{\left(\frac{i}{r}, \frac{j \pi}{r}\right): \quad i, j=1, \ldots, r\right\} .
$$

The samples are ordered in increasing order in a vector $\left(z_{1}, z_{2}, \ldots, z_{r^{2}}\right)$. Let $\zeta_{r}$ : $[0,1] \rightarrow \mathbb{R}$ be a piecewise linear non-decreasing function that interpolate the samples $\left(z_{0}:=z_{1}, z_{2}, \ldots, z_{r^{2}}\right)$ over the nodes $\left(0, \frac{1}{r^{2}}, \frac{2}{r^{2}}, \ldots, 1\right)$.

In Fig. 1 we present the (approximate) analytical predictions $n^{2} \zeta_{r}\left(\frac{j}{n}\right) / \lambda_{j, n^{\prime}}-1$, with $r=10000$, together with the (approximate) eigenvalue errors $\lambda_{j, n} / \lambda_{j, n^{\prime}}-1$.


Fig. 2 Example 1 [Linear $C^{0}$ B-spline discretization with $a(x)=2+0.5 x$ and $\left.b(x)=1\right]$ : analytical predictions $n^{2} \tilde{c}_{1}\left(\theta_{j, n}\right) e_{1}\left(\theta_{j, n}\right) / \lambda_{j, n^{\prime}}-1$ and eigenvalue errors $\lambda_{j, n} / \lambda_{j, n^{\prime}}-1$ versus $j / N_{n}$, $j=1, \ldots, N_{n}\left(N_{n}=n-1, n=200, n^{\prime}=1500, p^{\prime}=5, n_{1}=10\right)$

Table 2 Example 2 [ $p$-degree $C^{p-1} \mathrm{~B}$-spline discretization for $p=2,3,4,5$ with $a(x)=2.1$. $10^{9}+1.05 \cdot 10^{9} x$ and $\left.b(x)=8000\right]$ : The five smallest eigenvalues and eigenvalue approximations, scaled by $10^{7}$, for $n=200, n^{\prime}=1500, p^{\prime}=5$, and $r=10000$

| $p$ | $\lambda_{j, n^{\prime}}$ | $\lambda_{j, n}$ | $n^{2} \zeta_{r,[p, p-1]}\left(\frac{j}{n}\right)$ | $n^{2} \tilde{c}_{p}\left(\theta_{j, n}\right) e_{p}\left(\theta_{j, n}\right)$ |
| :---: | :---: | :---: | :---: | :---: |
| 2 | 0.320242073856735 | 0.320242073906214 | 0.326994616633570 | 0.319596822473787 |
|  | 1.281965164064426 | 1.281965165832994 | 1.295209495587940 | 1.278576943997320 |
|  | 2.884839278092422 | 2.884839298406320 | 2.904497248558355 | 2.877209798346387 |
|  | 5.128863345308352 | 5.128863459872621 | 5.154973083201467 | 5.115745284850879 |
|  | 8.014037230782240 | 8.014037668616439 | 8.046649299586466 | 7.994414445665448 |
| 3 | 0.320242073856735 | 0.320242073878847 | 0.326994616613888 | 0.319613348810547 |
|  | 1.281965164064426 | 1.281965164062972 | 1.295209493290653 | 1.278636625771722 |
|  | 2.884839278092422 | 2.884839278083705 | 2.904497230696836 | 2.877330529920427 |
|  | 5.128863345308352 | 5.128863345316139 | 5.154972936807892 | 5.115937346593204 |
|  | 8.014037230782240 | 8.014037230847546 | 8.046648648420712 | 7.994681508249463 |
| 4 | 0.320242073856735 | 0.320242073879743 | 0.326994616613781 | 0.319603273140503 |
|  | 1.281965164064426 | 1.281965164063456 | 1.295209493290564 | 1.278600044346683 |
|  | 2.884839278092422 | 2.884839278081923 | 2.904497230695864 | 2.877256125504316 |
|  | 5.128863345308352 | 5.128863345304529 | 5.154972936792049 | 5.115818311034016 |
|  | 8.014037230782240 | 8.014037230780470 | 8.046648648302227 | 7.994514965707993 |
| 5 | 0.320242073856735 | 0.320242073879344 | 0.326994616614055 | 0.319620210926112 |
|  | 1.281965164064426 | 1.281965164062812 | 1.295209493290633 | 1.278661525969442 |
|  | 2.884839278092422 | 2.884839278081491 | 2.904497230695859 | 2.877381146957867 |
|  | 5.128863345308352 | 5.128863345303335 | 5.154972936792168 | 5.116018280099667 |
|  | 8.014037230782240 | 8.014037230779344 | 8.046648648302266 | 7.994794677407651 |




Fig. 3 Example 2 [ $p$-degree $C^{p-1} \mathrm{~B}$-spline discretization for $p=2,3,4,5$ with $a(x)=2.1$. $10^{9}+1.05 \cdot 10^{9} x$ and $\left.b(x)=8000\right]$ : analytical predictions $n^{2} \zeta_{r,[p, p-1]}\left(\frac{j}{n}\right) / \lambda_{j, n^{\prime}}-1$ and eigenvalue errors $\lambda_{j, n} / \lambda_{j, n^{\prime}}-1$ versus $j / N_{n}, j=1, \ldots, N_{n}\left(N_{n}=n+p-2, n=200, n^{\prime}=1500\right.$, $p^{\prime}=5, r=10000$ )


Fig. 4 Example 2 [ $p$-degree $C^{p-1} \mathrm{~B}$-spline discretization for $p=2,3,4,5$ with $a(x)=2.1$. $10^{9}+1.05 \cdot 10^{9} x$ and $\left.b(x)=8000\right]$ : analytical predictions $n^{2} \tilde{c}_{p}\left(\theta_{j, n}\right) e_{p}\left(\theta_{j, n}\right) / \lambda_{j, n^{\prime}}-1$ and eigenvalue errors $\lambda_{j, n} / \lambda_{j, n^{\prime}}-1$ versus $j / N_{n}, j=1, \ldots, N_{n}\left(N_{n}=n+p-2, n=200\right.$, $n^{\prime}=$ $1500, p^{\prime}=5, n_{1}=10$ )

We note the mismatch for the smallest eigenvalues. In Fig. 2 we plot the (approximate) analytical predictions $n^{2} \tilde{c}_{1}\left(\theta_{j, n}\right) e_{1}\left(\theta_{j, n}\right) / \lambda_{j, n^{\prime}}-1$, obtained from the above interpolation-extrapolation algorithm for $n_{1}=10$, and the (approximate) eigenvalue errors $\lambda_{j, n} / \lambda_{j, n^{\prime}}-1$ versus $j /(n-1)$, for $j=1, \ldots, n-1$. We clearly see that in Fig. 2 the slight mismatch for small eigenvalues observed Fig. 1 has been reduced.

Example 2 Let $p=2,3,4,5, n=200, a(x)=2.1 \cdot 10^{9}+1.05 \cdot 10^{9} x$ and $b(x)=$ 8000. The approximation parameters $n^{\prime}$ and $n_{1}$ have been chosen as $n^{\prime}=1500$ and $n_{1}=10$, respectively, and the eigenvalues $\lambda_{j, n^{\prime}}$ have been chosen correspondingly, in the sense that $\lambda_{j, n^{\prime}}$ is computed as follows: we compute the eigenvalues $\lambda_{j, n^{\prime}}$ for a matrix $\mathbf{L}_{n^{\prime}}(a, b)$ of order $n^{\prime}=1500$ and B -spline degree $p^{\prime}=5$. Then we assume that the first $n+p-2$ (for $n=200$ ) of these eigenvalues, that is $j=1, \ldots, n+p-2$, are accurate representations of the true eigenvalues $\lambda_{j}$; see [17] for a motivation and a discussion on this matter.

In Table 2 we present the five smallest eigenvalues and approximated eigenvalues for $p=2,3,4,5$. As is clear the new approximation, $n^{2} \tilde{c}_{p}\left(\theta_{j, n}\right) e_{p}\left(\theta_{j, n}\right)$, performs better than $n^{2} \zeta_{r,[p, p-1]}\left(\frac{j}{n}\right)$ to approximate $\lambda_{j, n}$. Figure 3 shows the comparison between the (approximate) analytical predictions

$$
\frac{n^{2} \zeta_{r,[p, p-1]}\left(\frac{j}{n}\right)}{\lambda_{j, n^{\prime}}}-1, \quad j=1, \ldots, n,
$$

and the (approximate) eigenvalue errors

$$
\frac{\lambda_{j, n}}{\lambda_{j, n^{\prime}}}-1, \quad j=1, \ldots, n+p-2,
$$

whereas Fig. 4 shows the comparison between the (approximate) analytical predictions

$$
\frac{n^{2} \tilde{c}_{p}\left(\theta_{j, n}\right) e_{p}\left(\theta_{j, n}\right)}{\lambda_{j, n^{\prime}}}-1, \quad j=1, \ldots, n,
$$

with the (approximate) eigenvalue errors. We see from Fig. 4 that the mismatch for small eigenvalues observed in Fig. 3 has been lowered.

## 5 Dyadic Decomposition Argument and Extreme Eigenvalues

While the distribution results are available both in the Galerkin [15] and in the collocation setting [7], the use of extrapolation methods, as those described in the previous section and in $[10,11]$, has been developed only in the Galerkin setting. The reason is the inherent nonsymmetry of the collocation matrices. However, this issue has to be
investigated in the future, even if few preliminary experiments seem quite promising. In this section, we start by analyzing the extreme eigenvalues of the (nonsymmetric) stiffness matrices in the collocation setting. We use a dyadic decomposition argument already employed for symmetric structures in several contexts (see [3, 20] and references therein).

Consider the one-dimensional Poisson problem

$$
\begin{cases}-u^{\prime \prime}(x)=\mathrm{f}(x), & x \in(0,1),  \tag{18}\\ u(x)=0, & x \in\{0,1\},\end{cases}
$$

where $\mathrm{f} \in C([0,1])$. Suppose we approximate (18) by using the isogeometric collocation method based on uniform B-splines (see [7, Sect.4] for the details on this method). Then, the resulting discretization matrix is

$$
\begin{equation*}
n^{2} \tilde{\mathbf{K}}_{n}^{[p]}:=\left[-N_{j,[p]}^{\prime \prime}\left(\xi_{i,[p]}\right)\right]_{i, j=2}^{n+p-1}, \tag{19}
\end{equation*}
$$

where $\xi_{i,[p]}, i=2, \ldots, n+p-1$, are the Greville abscissae defined in [7,Eq. (4.6)], while $N_{j,[p]}, j=2, \ldots, n+p-1$, are the usual B-spline basis functions of maximal regularity $k=p-1$. In other words, with reference to (7), we have

$$
N_{j,[p]}=B_{j,[p, p-1]}, \quad j=2, \ldots, n+p-1 .
$$

We denote by $\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)$ the real part of $\tilde{\mathbf{K}}_{n}^{[p]}$ (the real part of a square complex matrix $X$ [4] is by definition $\left(X+X^{\mathrm{H}}\right) / 2, X^{\mathrm{H}}$ being the transpose conjugate). In Theorem 1 we prove for $p=3$ the following result:

$$
\begin{equation*}
\exists c_{p}>0: \mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[p]}\right) \geq c_{p} \tau_{n+p-2}(2-2 \cos \theta), \quad \forall n \geq 2, \tag{20}
\end{equation*}
$$

where
and

$$
\tau_{n+1}(2-2 \cos \theta)=\left[\begin{array}{rrrrr}
2-1 & & & \\
-1 & \ddots & \ddots & & \\
\ddots & & \ddots & \\
& & \ddots & \ddots & -1 \\
& & & -1 & 2
\end{array}\right]
$$

Theorem 1 There exists a constant $c_{3}>0$ such that

$$
\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[3]}\right) \geq c_{3} \tau_{n+1}(2-2 \cos \theta), \quad \forall n \geq 2 .
$$

Proof For $n=2, \ldots, 5$, a direct verification shows that $\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[3]}\right)$ is positive definite. Therefore, the theorem is proved if there exists a constant $c>0$ such that

$$
\begin{equation*}
\Re\left(\tilde{\mathbf{K}}_{n}^{[3]}\right) \geq c \tau_{n+1}(2-2 \cos \theta), \quad \forall n \geq 6 \tag{21}
\end{equation*}
$$

For $n \geq 6$ and $c>0$, the matrix $\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[3]}\right)-c \tau_{n+1}(2-2 \cos \theta)$ is explicitly given by

$$
\begin{aligned}
& {\left[\begin{array}{lll}
\frac{11}{2}-2 c & -\frac{4}{3}+c & -\frac{1}{6}
\end{array}\right.} \\
& -\frac{4}{3}+c \quad \frac{5}{2}-2 c-1+c \\
& -\frac{1}{6}-1+c \quad 2-2 c-1+c \\
& -1+c \quad 2-2 c-1+c \\
& \ddots \quad \ddots \quad \ddots \text {, } \\
& -1+c \quad 2-2 c-1+c \\
& -1+c \quad 2-2 c \quad-1+c \quad-\frac{1}{6} \\
& -1+c \quad \frac{5}{2}-2 c \quad-\frac{4}{3}+c \\
& \left.-\frac{1}{6}-\frac{4}{3}+c \quad \frac{11}{2}-2 c\right]
\end{aligned}
$$

and can be decomposed as follows:
\(\left.$$
\begin{array}{cc}r & -\frac{4}{3}+c \\
-\frac{4}{3}+c & s\end{array}
$$\right]_{2}^{1}+{\left.\begin{array}{ccc}\frac{11}{2}-2 c-r \& 0 \& -\frac{1}{6} <br>
0 \& 0 \& 0 <br>

-\frac{1}{6} \& 0 \& t\end{array}\right]_{3}^{1}+\)| $\frac{5}{2}-2 c-s$ | $-1+c$ |
| :---: | :---: |
| $-1+c$ | $1-c-t$ |$}_{3}$



In the above decomposition, $r, s$, and $t$ are arbitrary real numbers, whereas each term of the summation is a $(n+1) \times(n+1)$ matrix whose only nonzero entries are shown in the associated box and are contained in the principal submatrix corresponding to the rows from the superscript to the subscript. For instance,

If each term of the above summation is a nonnegative definite matrix, then $\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[3]}\right)-$ $c \tau_{n+1}(2-2 \cos \theta)$ is nonnegative definite as well. The following conditions ensure that each term of the summation is nonnegative definite:

$$
\left\{\begin{array}{l}
r>0, s>0, r s \geq\left(\frac{4}{3}-c\right)^{2} \\
\frac{11}{2}-2 c-r>0, t>0,\left(\frac{11}{2}-2 c-r\right) t \geq \frac{1}{36} \\
\frac{5}{2}-2 c-s>0,1-c-t>0 \\
\left(\frac{5}{2}-2 c-s\right)(1-c-t) \geq(1-c)^{2} \\
c \leq 1
\end{array}\right.
$$

These conditions are satisfied, for instance, with

$$
c=\frac{58}{59}, \quad s=\frac{1}{2}, \quad r=\frac{\left(\frac{4}{3}-c\right)^{2}}{s}, \quad t=\frac{\frac{1}{36}}{\frac{11}{2}-2 c-r} .
$$

Hence, (21) holds with $c=58 / 59$.
We verified that the dyadic decomposition argument used in the proof of Theorem 1 can also be used to prove (20) for $p=2$, 4. Although this argument becomes quite difficult to apply for $p \geq 5$, we have reason to believe that a careful application of it could prove (20) for any given $p \geq 2$.

The above theorem and the relations in (20) have interesting consequences on the conditioning measured with respect to the induced Euclidean norm as reported in Corollary 1 , whose proof requires a technical but interesting in itself lemma.

Lemma 1 Given $X$ square matrix, assuming $\mathfrak{i}(X)$ positive definite, we have

$$
\lambda_{\min }(\Re(X)) \leq \sigma_{\min }(X) .
$$

Proof As a first step define the following matrix $\hat{X}$ of order $2 n$

$$
\hat{X}=\left[\begin{array}{ll}
0 & X \\
X^{\mathrm{H}} & 0
\end{array}\right] .
$$

By taking into account the standard singular value decomposition, it is easy to see (and it is well-known) that the singular values $\sigma_{j}(X)$ of $X, j=1, \ldots, n$, coincide with the $n$ largest eigenvalues of $\hat{X}$, since the eigenvalues of the Hermitian matrix $\hat{X}$ are $\pm \sigma_{j}(X), j=1, \ldots, n$. Therefore, we infer that

$$
\sigma_{\min }(X)=\lambda_{n+1}(\hat{X}),
$$

with $\lambda_{1}(\hat{X}) \leq \lambda_{2}(\hat{X}) \leq \cdots \leq \lambda_{2 n}(\hat{X})$. From the minimax characterization (see [4]) of the $(n+1)$-th eigenvalue of $\hat{X}$, we obtain

$$
\begin{equation*}
\sigma_{\min }(X)=\max _{\operatorname{dim}(U)=n} \min _{u \in U} \frac{u^{\mathrm{H}} \hat{X} u}{u^{\mathrm{H}} u}, \tag{22}
\end{equation*}
$$

where the maximum is taken over all subspaces $U$ with codimension $n$ in $\mathbb{C}^{2 n}$ or, equivalently, over all subspaces $U$ of dimension $n$. Let $U_{*}$ be the $n$-dimensional subspace of $\mathbb{C}^{2 n}$ made of all vectors $y$ of the form

$$
y=\binom{u}{u}, \quad u \in \mathbb{C}^{n}, \quad u=\left(\begin{array}{l}
u_{1} \\
u_{2} \\
\vdots \\
u_{n} .
\end{array}\right) .
$$

If $y \in U_{*}$ is partitioned into blocks as above, it holds

$$
\begin{equation*}
\frac{y^{\mathrm{H}} \hat{X} y}{y^{\mathrm{H}} y}=\frac{u^{\mathrm{H}} \Re(X) u}{u^{\mathrm{H}} u}, \tag{23}
\end{equation*}
$$

and since $U_{*}$ is a particular subspace of dimension $n$, from (22) the thesis follows.

Corollary 1 Assume that (20) is satisfied for a given p. Then

$$
\begin{equation*}
\mu_{2}\left(\Re\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)\right) \sim n^{2}, \tag{24}
\end{equation*}
$$

and

$$
\begin{equation*}
\mu_{2}\left(\tilde{\mathbf{K}}_{n}^{[p]}\right) \leq C_{p} n^{2}, \tag{25}
\end{equation*}
$$

with $C_{p}$ a positive constant depending only on $p$ and $\mu(X)=\|X\|\left\|X^{-1}\right\|$ being the conditioning of an invertible matrix $X$ measured with respect to the induced Euclidean norm \|| \|.

Proof First of all we observe that both $\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)$ and $\tilde{\mathbf{K}}_{n}^{[p]}$ are banded matrices with coefficients not depending on the parameter $n$. Hence we have that their induced Euclidean norm is asymptotic to a a constant not depending on $n$ that is

$$
\begin{gather*}
\left\|\Re\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)\right\| \sim 1,  \tag{26}\\
\left\|\tilde{\mathbf{K}}_{n}^{[p]}\right\| \sim 1 . \tag{27}
\end{gather*}
$$

We first prove (24). By (20) we have

$$
\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[p]}\right) \geq c_{p} \tau_{n+p-2}(2-2 \cos \theta), \quad \forall n \geq 2
$$

with $c_{p}>0$ and hence $\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)$ is positive definite and its minimal eigenvalue is bounded from below by

$$
c_{p} 4 \sin ^{2}\left(\frac{\pi}{2(n+p-1)}\right)
$$

$4 \sin ^{2}\left(\frac{\pi}{2(n+p-1)}\right)$ being the minimal eigenvalue of $\tau_{n+p-2}(2-2 \cos \theta)$ (see [19] and references therein). On the other hand, $\mathfrak{\Re ( \tilde { \mathbf { K } } _ { n } ^ { [ p ] } ) \text { contains as a principal submatrix a }}$ Toeplitz matrix of size proportional to $n$ and generated by $f_{p}$ being nonnegative and having a unique zero of order two (see [7,14]). Hence such a Toeplitz matrix has the minimal eigenvalue $l_{n}$ asymptotic to $n^{-2}$ (see again [19] and references therein) and by the Cauchy interlacing theorem the minimal eigenvalue of $\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)$ is bounded from above by $l_{n} \sim n^{-2}$. Therefore

$$
\lambda_{\min }\left(\Re\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)\right) \sim n^{-2}
$$

But $\mathfrak{R}\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)$ is positive definite and hence

$$
\left\|\left[\Re\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)\right]^{-1}\right\|=\left[\lambda_{\min }\left(\Re\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)\right)\right]^{-1} \sim n^{2}
$$

from which, using (26), we deduce

$$
\mu_{2}\left(\Re\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)\right) \sim n^{2} .
$$

We now prove (25). Given the known fact $\left\|X^{-1}\right\|=\left[\sigma_{\text {min }}(X)\right]^{-1}$, statement (25) is a consequence of (27), of (20), of Lemma 1 , and of the fact that the minimal eigenvalue of $\tau_{n+p-2}(2-2 \cos \theta)$ is

$$
\sin ^{2}\left(\frac{\pi}{2(n+p-1)}\right) \sim n^{-2}
$$

Furthermore, we have reasons to conjecture that the spectral conditioning of the collocation stiffness matrices grows asymptotically as $n^{2}$, exactly as in the Galerkin setting [14], since $\tilde{\mathbf{K}}_{n}^{[p]}$ contains as a principal submatrix a Toeplitz matrix of size proportional to $n$ and generated by $f_{p}$ having a unique zero of order two (as $\Re\left(\tilde{\mathbf{K}}_{n}^{[p]}\right)$, see the proof of Corollary 1 ).

We finally stress that (20) is also important in a multigrid context for proving the optimality of the related two-grid and multigrid techniques (see [2, 6, 8, 19]), by following the theory reported by Ruge and Stüben in [18].

## 6 Conclusions

In this note, we have considered the spectral analysis of large matrices coming from the isogeometric approximations based on B-splines of the eigenvalue problem

$$
-\left(a(x) u^{\prime}(x)\right)^{\prime}=\lambda b(x) u(x), \quad x \in(0,1),
$$

where $u(0)$ and $u(1)$ are given. In the collocation setting, we complemented global eigenvalue distribution results, available in the literature [7], with precise estimates for the extremal eigenvalues and hence for the spectral conditioning of the resulting matrices. In the Galerkin setting, we have designed an efficient matrix-less procedure (see [9]) that gives a highly accurate estimate of the all the eigenvalues, starting from the knowledge of the spectral GLT distribution symbol. Possible extensions include a more systematic treatment of the collocation case, both via the use of dyadic decomposition arguments and via the use of proper matrix-less extrapolation techniques. This last item is completely new and represents a real challenge, thus it will be the subject of a future investigation.

## References

1. Ahmad, F., Al-Aidarous, E.S., Abdullah Alrehaili, D., Ekström, S.-E., Furci, I., SerraCapizzano, S.: Are the eigenvalues of preconditioned banded symmetric Toeplitz matrices known in almost closed form? Numer. Algorithms 78, 867-893 (2018)
2. Aricò, A., Donatelli, M., Serra-Capizzano, S.: V-cycle optimal convergence for certain (multilevel) structured linear systems. SIAM J. Matrix Anal. Appl. 26, 186-214 (2004)
3. Beckermann, B., Serra-Capizzano, S.: On the asymptotic spectrum of finite element matrix sequences. SIAM J. Numer. Anal. 45, 746-769 (2007)
4. Bhatia, R.: Matrix Analysis. Springer, New York (1997)
5. Cottrell, J.A., Hughes, T.J.R., Bazilevs, Y.: Isogeometric Analysis: Toward Integration of CAD and FEA. Wiley, Chichester (2009)
6. Donatelli, M., Garoni, C., Manni, C., Serra-Capizzano, S., Speleers, H.: Two-grid optimality for Galerkin linear systems based on B-splines. Comput. Vis. Sci. 17, 119-133 (2015)
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. 85, 16391680 (2016)
8. 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. 55, 31-62 (2017)
9. Ekström, S.-E.: Matrix-less methods for computing eigenvalues of large structured matrices. Ph.D. Thesis, Uppsala University (2018)
10. 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 Algebr. Appl. 25, e2151 (2018)
11. Ekström, S.-E., Furci, I., Serra-Capizzano, S.: Exact formulae and matrix-less eigensolvers for block banded symmetric Toeplitz matrices. BIT Numer. Math. 58, 937-968 (2018)
12. Ekström, S.-E., Garoni, C.: A matrix-less and parallel interpolation-extrapolation algorithm for computing the eigenvalues of preconditioned banded symmetric Toeplitz matrices. Numer. Algorithms 80(3), 819-848 (2019)
13. Ekström, S.-E., Garoni, C., Serra-Capizzano, S.: Are the eigenvalues of banded symmetric Toeplitz matrices known in almost closed form? Exp. Math. 27, 478-487 (2018)
14. Garoni, C., Manni, C., Pelosi, F., Serra-Capizzano, S., Speleers, H.: On the spectrum of stiffness matrices arising from isogeometric analysis. Numer. Math. 127, 751-799 (2014)
15. Garoni, C., Manni, C., Serra-Capizzano, S., Sesana, D., Speleers, H.: Spectral analysis and spectral symbol of matrices in isogeometric Galerkin methods. Math. Comput. 86, 1343-1373 (2017)
16. Garoni, C., Serra-Capizzano, S.: Generalized Locally Toeplitz Sequences: Theory and Applications. Springer Monographs in Mathematics, vol. 1. Springer, Cham (2017)
17. Garoni, C., Speleers, H., Ekström, S.-E., Hughes, T.J.R., Reali, A., Serra-Capizzano, S.: Symbol-based analysis of finite element and isogeometric B-spline discretizations of eigenvalue problems: Exposition and review. Arch. Comput. Methods Eng. (In press) https://doi. org/10.1007/s11831-018-9295-y
18. Ruge, J.W., Stüben, K.: Algebraic multigrid. In: McCormick, S. (ed.) Multigrid Methods, pp. 73-130. SIAM Publications, Philadelphia (1987)
19. Serra-Capizzano, S.: Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences. Numer. Math. 92, 433-465 (2002)
20. Serra-Capizzano, S., Tablino-Possio, C.: Spectral and structural analysis of high precision finite difference matrices for elliptic operators. Linear Algebra Appl. 293, 85-131 (1999)

[^0]:    S.-E. Ekström ( $\boxtimes$ ) • S. Serra-Capizzano

    Department of Information Technology, Uppsala University, Uppsala, Sweden
    e-mail: sven-erik.ekstrom@it.uu.se
    S. Serra-Capizzano

    Department of Humanities and Innovation, University of Insubria, Varese, Italy
    e-mail: stefano.serra@it.uu.se; stefano.serrac@uninsubria.it

