Eigenvalue isogeometric approximations based on B-splines: tools and results

Sven-Erik Ekström and Stefano Serra-Capizzano

Abstract In such a short note we consider the spectral analysis of large matrices coming from the numerical approximation of the eigenvalue problem

$$-(a(x)u'(x))' = \lambda b(x)u(x), \qquad x \in (0,1),$$

where $u(0)$ and $u(1)$ are given, by using isogeometric methods based on B-splines. We give precise estimates for the extremal eigenvalues and global distributional results. The techniques involve dyadic decomposition arguments, the GLT analysis, and basic extrapolation methods.

1 Introduction and Notations

In this note we consider the approximation of one-dimensional elliptic eigenvalue problems by using a isogeometric either Galerkin or collocation technique with B-splines [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.

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

S. Serra-Capizzano
Department of Science and High Technology, University of Insubria, INDAM Unit, Via Valleggio 11, 22100 Como, Italy
e-mail: stefano.serrac@uninsubria.it

S.-E. Ekström, S. Serra-Capizzano
Division of Scientific Computing, Department of Information Technology, Uppsala University, Box 337, SE-751 05 Uppsala, Sweden
e-mail: sven-erik.ekstrom@it.uu.se
e-mail: stefano.serra@it.uu.se

[16, 18] and references therein), by using the theory of the Generalized Locally Toeplitz (GLT) sequences [17].

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 [10] and references therein). In the constant coefficient setting, when considering the Galerkin approach, the results presented in [11] 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 [18]).

Here, we propose a further numerical scheme for the case of variable coefficients again based on extrapolation methods and again in the Galerkin setting. The numerical results are quite good, especially when the problem coefficients are smooth.

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 [8], as in the symmetric positive definite Galerkin setting [16]. Hence we expect that most of the eigenvalues are real or with negligible imaginary part. On the other, it is not clear if the delicate asymptotic expansions observed in [11, 18] holds also in the collocation setting and indeed this issue will be the subject of future investigations. 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 confirms that the conditioning grows as h^{-2}, h being the fineness parameter, exactly as in the case of the matrices obtained with the Galerkin approximation (see e.g [15]).

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 $\{\mathbf{X}_n\}_n$, where $\mathbf{X}_n$ is a square matrix such that $\text{size}(\mathbf{X}_n) \to \infty$ as $n \to \infty$. Let μ_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 $\mathbf{g} : D \subset \mathbb{R}^d \to \mathbb{C}^{s\times s}$ is an $s \times s$ matrix-valued function, we say that $\mathbf{g}$ is measurable if its s^2 components $g_{ij} : D \to \mathbb{C}$, $i, j = 1, \ldots, s$, are measurable.

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

$$\lim_{n\to\infty} \frac{1}{N_n} \sum_{i=1}^{N_n} F(\lambda_i(\mathbf{X}_n)) = \frac{1}{\mu_d(D)} \int_D \frac{\sum_{i=1}^{s} F(\lambda_i(\mathbf{g}(y_1, \ldots, y_d)))}{s} \mathrm{d}y_1 \ldots \mathrm{d}y_d, \ \forall F \in C_c(\mathbb{C}),$$

where $\lambda_i(\mathbf{g}(y_1, \ldots, y_d))$, $i = 1, \ldots, s$, are the eigenvalues of the $s \times s$ matrix $\mathbf{g}(y_1, \ldots, y_d)$.

The informal meaning behind the eigenvalue distribution $\{\mathbf{X}_n\}_n \sim_\lambda \mathbf{g}$ is the following: for all sufficiently large n, the eigenvaues of $\mathbf{X}_n$ can be subdivided into s different subsets of approximately the same cardinality; and the eigenvalues belonging to the ith subset (except possibly for $o(N_n)$ outliers) are approximated by the samples of the ith eigenvalue function $\lambda_i(\mathbf{g}(y_1,\ldots,y_d))$ over a uniform grid in D (the domain of $\mathbf{g}$). For example, if $d=1$, $N_n = ns$ and $D=[a,b]$, then the eigenvalues of $\mathbf{X}_n$ are approximately equal to

$$\lambda_i\Big(\mathbf{g}\Big(a+j\frac{b-a}{n}\Big)\Big), \qquad j=1,\ldots,n, \qquad i=1,\ldots,s,$$

for n large enough. Likewise, if $d=2$, $N_n = n^2 s$ and $D=[a_1,b_1]\times[a_2,b_2]$, then the eigenvalues of $\mathbf{X}_n$ are approximately equal to

$$\lambda_i\Big(\mathbf{g}\Big(a_1+j_1\frac{b_1-a_1}{n},\, a_2+j_2\frac{b_2-a_2}{n}\Big)\Big), \qquad j_1,j_2=1,\ldots,n, \qquad i=1,\ldots,s,$$

for n large enough. The informal meaning behind the singular value distribution $\{\mathbf{X}_n\}_n \sim_\sigma \mathbf{g}$ is completely analogous: for all sufficiently large n, the singular values of $\mathbf{X}_n$ can be subdivided into s different subsets of approximately the same cardinality; and the singular values belonging to the ith subset (except possibly for $o(N_n)$ outliers) are approximated by the samples of the ith singular value function $\sigma_i(\mathbf{g}(y_1,\ldots,y_d))$ over a uniform grid in the domain D.

*Remark 1 (**rearrangement**).* Let $\mathbf{g}: D\subseteq\mathbb{R}^d \to \mathbb{C}^{s\times s}$ and suppose that D is a rectangle in $\mathbb{R}^d$, say $D:=[a_1,b_1]\times\cdots\times[a_d,b_d]$. We also assume that the eigenvalues $\lambda_1(\mathbf{g}(y_1,\ldots,y_d)),\ldots,\lambda_s(\mathbf{g}(y_1,\ldots,y_d))$ are real for all $(y_1,\ldots,y_d)\in D$. For each positive integer r, let $\mathscr{G}_r$ be the uniform grid in D given by

$$\mathscr{G}_r := \left\{ \Big(a_1+\frac{i_1}{r}(b_1-a_1),\ldots,a_d+\frac{i_d}{r}(b_d-a_d)\Big) :\ i_1,\ldots,i_d=1,\ldots,r \right\}.$$

Compute the samples of the eigenvalue functions

$$\lambda_1(\mathbf{g}(y_1,\ldots,y_d)),\ldots,\lambda_s(\mathbf{g}(y_1,\ldots,y_d))$$

at the points $(y_1,\ldots,y_d)\in\mathscr{G}_r$, sort them in increasing order and put them in a vector $(q_1,q_2,\ldots,q_{sr^d})$. Let $\eta_r:[0,1]\to\mathbb{R}$ be the piecewise linear non-decreasing function that interpolates the samples $(q_0:=q_1,q_1,q_2,\ldots,q_{sr^d})$ over the nodes $(0,\frac{1}{sr^d},\frac{2}{sr^d},\ldots,1)$, i.e.,

$$\eta_r\Big(\frac{\ell}{sr^d}\Big) := q_\ell, \qquad \ell=0,\ldots,sr^d.$$

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

$$\int_D \frac{\sum_{i=1}^{s} F(\lambda_i(\mathbf{g}(y_1,\ldots,y_d)))}{s}\mathrm{d}y_1\ldots\mathrm{d}y_d = \int_0^1 F(\eta(t))\mathrm{d}t, \qquad \forall F \in C_c(\mathbb{C}).$$

Therefore, if we have $\{\mathbf{X}_n\}_n \sim_\lambda \mathbf{g}$, then we also have $\{\mathbf{X}_n\}_n \sim_\lambda \eta$.

2 Isogeometric Galerkin discretization of variable-coefficient eigenvalue problems

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

$$\begin{cases} -(a(x)u_j'(x))' = \lambda_j b(x)u_j(x), & x \in \Omega, \\ u_j(x) = 0, & x \in \partial\Omega, \end{cases} \tag{1}$$

where Ω is an open interval in $\mathbb{R}$ and $a,b \in L^1(\Omega)$ are such that $a,b > 0$ a.e. on Ω. 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 $v \in H_0^1(\Omega)$,

$$\mathrm{a}(u_j,v) = \lambda_j(b\,u_j,v),$$

where

$$\mathrm{a}(u_j,v) := \int_\Omega a(x)u_j'(x)v'(x)\mathrm{d}x, \qquad (b\,u_j,v) := \int_\Omega b(x)u_j(x)v(x)\mathrm{d}x.$$

In the isogeometric Galerkin method, we assume that the physical domain Ω is described by a global geometry map $G : [0,1] \to \overline{\Omega}$, which is invertible and satisfies $G(\partial([0,1])) = \partial\overline{\Omega}$. We fix a set of basis functions $\{\varphi_1,\ldots,\varphi_{N_n}\}$ defined on the reference (parametric) domain $[0,1]$ and vanishing on the boundary $\partial([0,1])$. We consider the basis functions

$$\psi_i(x) := \varphi_i(G^{-1}(x)) = \varphi_i(t), \qquad x = G(t), \qquad i = 1,\ldots,N_n,$$

which are defined on the physical domain Ω, and we define the approximation space $\mathscr{W}_n := \mathrm{span}(\psi_1,\ldots,\psi_{N_n}) \subset H_0^1(\Omega)$. Finally, we find approximations to the exact eigenpairs (λ_j,u_j), $j = 1,2,\ldots,\infty$, 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 $v_n \in \mathscr{W}_n$,

$$\mathrm{a}(u_{j,n},v_n) = \lambda_{j,n}(b\,u_{j,n},v_n). \tag{2}$$

Assuming that both the exact and numerical eigenvalues are arranged in non-decreasing order, the pair $(\lambda_{j,n},u_{j,n})$ is taken as an approximation of the pair (λ_j,u_j) 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 $\{\varphi_1,\ldots,\varphi_{N_n}\}$, solving the Galerkin problem

(2) is equivalent to solving the generalized eigenvalue problem

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

where $\mathbf{u}_{j,n}$ is the coefficient vector of $u_{j,n}$ with respect to $\{\varphi_1,\dots,\varphi_{N_n}\}$ and

$$\begin{aligned}\mathbf{K}_n(a,G) := [\mathrm{a}(\psi_j,\psi_i)]_{i,j=1}^{N_n} &= \left[\int_\Omega a(x)\psi_j'(x)\psi_i'(x)\mathrm{d}x\right]_{i,j=1}^{N_n} \\ &= \left[\int_0^1 \frac{a(G(t))}{|G'(t)|}\varphi_j'(t)\varphi_i'(t)\mathrm{d}t\right]_{i,j=1}^{N_n}, \end{aligned}\tag{4}$$

$$\begin{aligned}\mathbf{M}_n(b,G) := [(b\,\psi_j,\psi_i)]_{i,j=1}^{N_n} &= \left[\int_\Omega b(x)\psi_j(x)\psi_i(x)\mathrm{d}x\right]_{i,j=1}^{N_n} \\ &= \left[\int_0^1 b(G(t))|G'(t)|\varphi_j(t)\varphi_i(t)\mathrm{d}t\right]_{i,j=1}^{N_n}. \end{aligned}\tag{5}$$

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 Ω, both $\mathbf{K}_n(a,G)$ and $\mathbf{M}_n(b,G)$ are always symmetric positive definite, regardless of the chosen basis functions $\varphi_1,\dots,\varphi_{N_n}$ and the map G. Moreover, it is clear from (3) that the numerical eigenvalues $\lambda_{j,n}$, $j=1,\dots,N_n$, are just the eigenvalues of the matrix

$$\mathbf{L}_n(a,b,G) := (\mathbf{M}_n(b,G))^{-1}\mathbf{K}_n(a,G). \tag{6}$$

Therefore, if $a_G, b_G \in L^1([0,1])$ and the basis functions $\varphi_1,\dots,\varphi_{N_n}$ are chosen as the p-degree C^k B-splines

$$B_{2,[p,k]},\dots,B_{n(p-k)+k,[p,k]}, \tag{7}$$

then

$$\mathbf{K}_n(a,G) = \left[\int_0^1 \frac{a(G(t))}{|G'(t)|}B'_{j+1,[p,k]}(t)B'_{i+1,[p,k]}(t)\mathrm{d}t\right]_{i,j=1}^{n(p-k)+k-1}, \tag{8}$$

$$\mathbf{M}_n(b,G) = \left[\int_0^1 b(G(t))|G'(t)|B_{j+1,[p,k]}(t)B_{i+1,[p,k]}(t)\mathrm{d}t\right]_{i,j=1}^{n(p-k)+k-1}. \tag{9}$$

From the *-algebra structure of GLT sequences, as proven in [17, 18], with the notions and notations in Definition 1 and Remark 1, we know that

$$\left\{\frac{1}{n}\mathbf{K}_n(a,G)\right\}_n \sim_\lambda \kappa_{[p,k]}, \tag{10}$$

$$\left\{n\mathbf{M}_n(b,G)\right\}_n \sim_\lambda \xi_{[p,k]}, \tag{11}$$

$$\left\{\frac{1}{n^2}\mathbf{L}_n(a,b,G)\right\}_n \sim_\lambda \zeta_{[p,k]}, \tag{12}$$

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

$$\frac{a(G(t))}{|G'(t)|}\mathbf{f}_{[p,k]}(\theta), \quad b(G(t))|G'(t)|\mathbf{h}_{[p,k]}(\theta), \quad \frac{a(G(t))}{b(G(t))|G'(t)|^2}\mathbf{e}_{[p,k]}(\theta),$$

respectively, with

$$\mathbf{f}_{[p,k]}(\theta) := \mathbf{K}_{[p,k]}^{[0]} + \sum_{\ell=1}^{\eta-1}\left(\mathbf{K}_{[p,k]}^{[\ell]}\mathrm{e}^{\mathrm{i}\ell\theta} + (\mathbf{K}_{[p,k]}^{[\ell]})^T\mathrm{e}^{-\mathrm{i}\ell\theta}\right), \tag{13}$$

$$\mathbf{h}_{[p,k]}(\theta) := \mathbf{M}_{[p,k]}^{[0]} + \sum_{\ell=1}^{\eta-1}\left(\mathbf{M}_{[p,k]}^{[\ell]}\mathrm{e}^{\mathrm{i}\ell\theta} + (\mathbf{M}_{[p,k]}^{[\ell]})^T\mathrm{e}^{-\mathrm{i}\ell\theta}\right), \tag{14}$$

$$\mathbf{e}_{[p,k]}(\theta) := (\mathbf{h}_{[p,k]}(\theta))^{-1}\mathbf{f}_{[p,k]}(\theta), \tag{15}$$

where the blocks $\mathbf{K}_{[p,k]}^{[\ell]}$ and $\mathbf{M}_{[p,k]}^{[\ell]}$, of size $p-k$, are defined in [18]. The analytical predictions of the eigenvalue errors are obtained through the sampling procedure described as follows

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

where $m := \min(n(p-k), n(p-k)+k-1)$.

However, such a procedure is not completely satisfactory especially in low frequencies (see [18]), when variable coefficients are considered. We address such a issue in the following section.

3 Global distribution results and extrapolation

As discussed in [18], 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 already illustrated in that paper. Drawing inspiration from [1, 10, 11, 14], we here describe — in the case of a isogeometric Galerkin method based on B-splines discretization of the eigenvalue problem (1) — an alternative interpolation–extrapolation 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 $c_p : [0,\pi] \to \mathbb{R}$, 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} B-spline discretization, the following property holds: for all n the eigenvalues of the matrix $\frac{1}{n^2}\mathbf{L}_n(a,b)$, sorted in increasing order, satisfy

$$\frac{\lambda_{j,n}}{n^2} \approx c_p(\theta_{j,n}) e_p(\theta_{j,n}), \qquad j = 1, \dots, \min(n+p-2, n), \tag{17}$$

where

$$e_p = \frac{f_p(\theta)}{h_p(\theta)} = \frac{(2-2\cos\theta) h_{p-1}(\theta)}{h_p(\theta)} = \frac{(2-2\cos\theta)\left(\phi_{[2p-1]}(p) + 2\sum_{\ell=1}^{p-1} \phi_{[2p-1]}(p-\ell)\cos(\ell\theta)\right)}{\phi_{[2p+1]}(p+1) + 2\sum_{\ell=1}^{p} \phi_{[2p+1]}(p+1-\ell)\cos(\ell\theta)}$$

is defined in [18], and

$$\theta_{j,n} = \frac{j\pi}{n}, \qquad j = 1, \dots, n.$$

2. Compute the eigenvalues of the matrix $\frac{1}{n_1^2} L_{n_1}(a,b)$ corresponding to a small value of n_1, i.e.,
$$\frac{\lambda_{j_1,n_1}}{n_1^2}, \qquad j_1 = 1, \dots, n_1 + p - 2.$$
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,
$$c_p(\theta_{j_1,n_1}) \approx \frac{\lambda_{j_1,n_1}}{n_1^2 e_p(\theta_{j_1,n_1})}, \qquad j_1 = 1, \dots, \min(n_1 + p - 2, n_1).$$
This means that we have an approximation of the unknown function c_p over the coarse uniform grid consisting of the points θ_{j_1,n_1}.
4. Interpolate the data
$$\left(\theta_{j_1,n_1}, \frac{\lambda_{j_1,n_1}}{n_1^2 e_p(\theta_{j_1,n_1})}\right), \qquad j_1 = 1, \dots, \min(n_1 + p - 2, n_1),$$
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,
$$\tilde{c}_p(\theta_{j_1,n_1}) = \frac{\lambda_{j_1,n_1}}{n_1^2 e_p(\theta_{j_1,n_1})}, \qquad j_1 = 1, \dots, \min(n_1 + p - 2, n_1).$$
5. Given a large n, compute approximations of the numerical eigenvalues $\lambda_{j,n}$ by replacing c_p with $\tilde{c}_p$ in (17):
$$\lambda_{j,n} \approx n^2 \tilde{c}_p(\theta_{j,n}) e_p(\theta_{j,n}), \qquad j = 1, \dots, \min(n+p-2, n).$$
6. Compute analytical predictions for the eigenvalue errors as follows:

$$\frac{\lambda_{j,n}}{\lambda_j} - 1 \approx \frac{n^2 \tilde{c}_p(\theta_{j,n}) e_p(\theta_{j,n})}{\lambda_j} - 1, \qquad j = 1, \ldots, \min(n+p-2, n).$$

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).

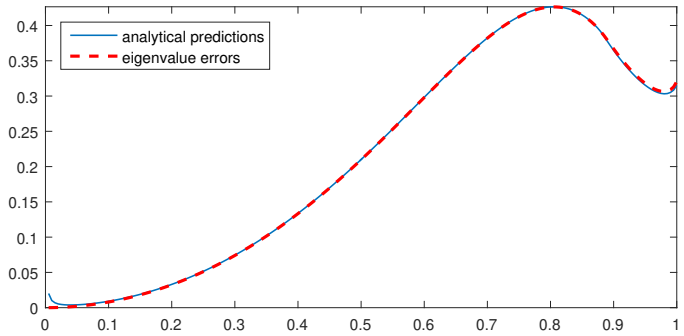

Fig. 1 Example 1 [Linear C^0 B-spline discretization with $a(x) = 2 + 0.5x$ and $b(x) = 1$]: analytical predictions $n^2 \zeta_r(\frac{j}{n})/\lambda_{j,n'} - 1$ and eigenvalue errors $\lambda_{j,n}/\lambda_{j,n'} - 1$ versus j/N_n, $j = 1, \ldots, N_n$ ($N_n = n-1$, $n = 200$, $n' = 1500$, $r = 10000$).

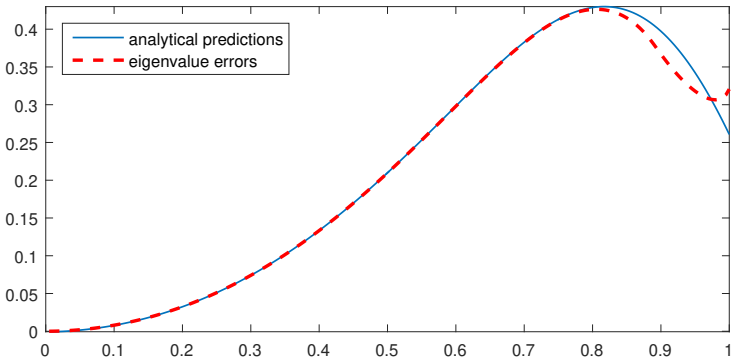

Fig. 2 Example 1 [Linear C^0 B-spline discretization with $a(x) = 2 + 0.5x$ and $b(x) = 1$]: analytical predictions $n^2 \tilde{c}_1(\theta_{j,n}) e_1(\theta_{j,n})/\lambda_{j,n'} - 1$ and eigenvalue errors $\lambda_{j,n}/\lambda_{j,n'} - 1$ versus j/N_n, $j = 1, \ldots, N_n$ ($N_n = n-1$, $n = 200$, $n' = 1500$, $n_1 = 10$).

Example 1. Let $p = 1$, $n = 200$, $a(x) = 2 + 0.5x$ and $b(x) = 1$. Let $n' = 1500 \gg n$ and take the first $n-1$ eigenvalues of $\mathbf{L}_{n'}(a,b)$, namely $\lambda_{1,n'}, \ldots, \lambda_{n-1,n'}$, as approximations of the unknown exact eigenvalues $\lambda_1, \ldots, \lambda_{n-1}$. In Figure 2 we plotted

the (approximate) analytical predictions $n^2\tilde{c}_1(\theta_{j,n})e_1(\theta_{j,n})/\lambda_{j,n'}-1$, obtained from the above interpolation–extrapolation algorithm for $n_1=10$, and the (approximate) eigenvalue errors $\lambda_{j,n}/\lambda_{j,n'}-1$ versus $j/(n-1)$, for $j=1,\ldots,n-1$. We clearly see that in Figure 2 the slight mismatch for small eigenvalues observed Figure 1 is not present.

Example 2. Figure 4 shows the comparison between the (approximate) analytical predictions

$$\frac{n^2\tilde{c}_p(\theta_{j,n})e_p(\theta_{j,n})}{\lambda_{j,n'}}-1, \qquad j=1,\ldots,n,$$

and the (approximate) eigenvalue errors

$$\frac{\lambda_{j,n}}{\lambda_{j,n'}}-1, \qquad j=1,\ldots,n+p-2,$$

for $p=2,3,4,5$, $n=200$, $a(x)=2.1\cdot 10^9+1.05\cdot 10^9 x$ and $b(x)=8000$. The approximation parameters n' and n_1 have been chosen as $n'=1500$ and $n_1=10$, respectively, and the eigenvalues $\lambda_{j,n'}$ have been chosen correspondingly. We see from Figure 4 that the mismatch for small eigenvalues observed in Figure 3 has disappeared.

4 Dyadic decomposition argument and extreme eigenvalues

While the distribution results are available both in the Galerkin [16] and in the collocation setting [8], the use of extrapolation methods, as those described in the previous section and in [11, 12, 18], 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, 21] and references therein).

We point out that the inequality we are going to prove is new, and therefore it may certainly be considered as a (not so irrelevant) novelty of this paper. Consider the one-dimensional Poisson problem

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

where $\mathrm{f}\in C([0,1])$. Suppose we approximate (18) by using the isogeometric collocation method based on uniform B-splines (see [8, Section 4] for the details on this method). Then, the resulting discretization matrix is

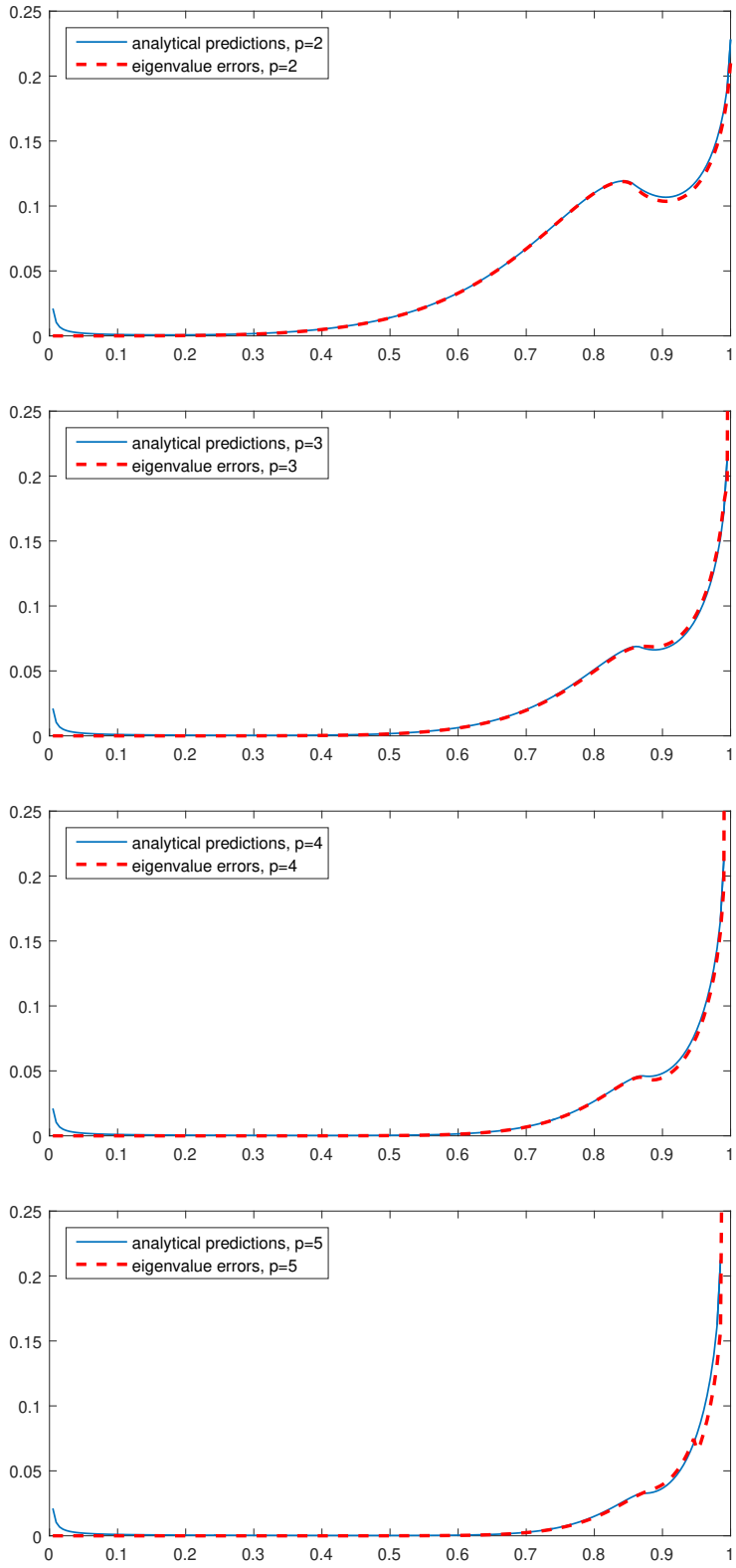

Fig. 3 Example 2 [p-degree C^{p-1} B-spline discretization for $p=2,3,4,5$ with $a(x)=2.1\cdot 10^9+1.05\cdot 10^9\,x$ and $b(x)=8000$]: analytical predictions $n^2\zeta_{r,[p,p-1]}(\frac{j}{n})/\lambda_{j,n'}-1$ and eigenvalue errors $\lambda_{j,n}/\lambda_{j,n'}-1$ versus j/N_n, $j=1,\ldots,N_n$ ($N_n=n+p-2\approx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

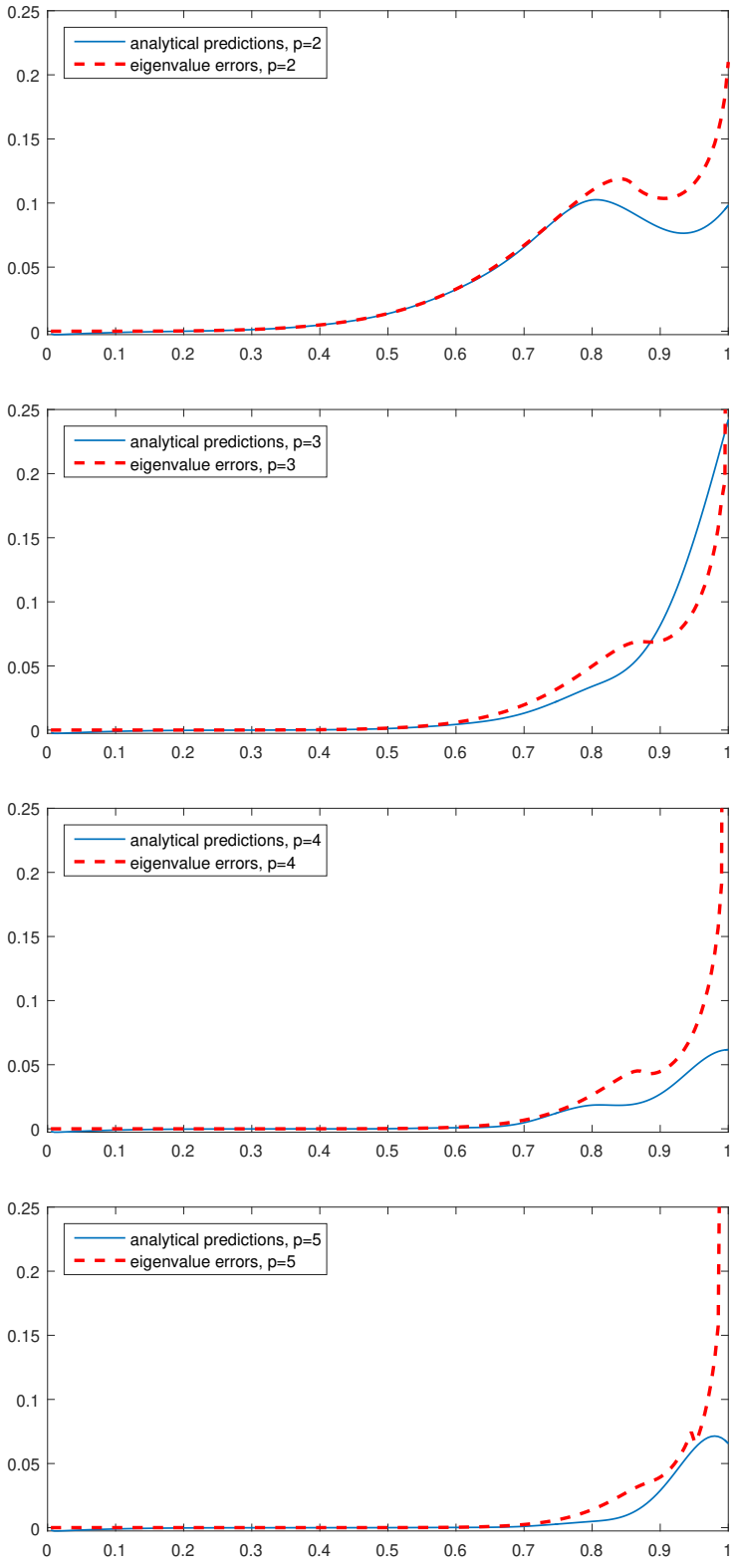

Fig. 4 Example 2 [p-degree C^{p-1} B-spline discretization for $p = 2,3,4,5$ with $a(x) = 2.1 \cdot 10^9 + 1.05 \cdot 10^9\, x$ and $b(x) = 8000$]: analytical predictions $n^2 \tilde{c}_p(\theta_{j,n}) e_p(\theta_{j,n}) / \lambda_{j,n'} - 1$ and eigenvalue errors $\lambda_{j,n}/\lambda_{j,n'} - 1$ versus j/N_n, $j = 1,\ldots,N_n$ ($N_n = n + p - 2 \approx 200$, $n = 200$, $n' = 1500$, $n_1 = 10$).

$$n^2\tilde{\mathbf{K}}_n^{[p]} := \left[-N''_{j,[p]}(\xi_{i,[p]})\right]_{i,j=2}^{n+p-1}, \tag{19}$$

where $\xi_{i,[p]},\ i=2,\dots,n+p-1$, are the Greville abscissae defined in [8, Eq. (4.6)], while $N_{j,[p]},\ j=2,\dots,n+p-1$, are the usual B-spline basis functions of maximal regularity $k=p-1$. In other words, with reference to (7), we have

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

We denote by $\Re(\tilde{\mathbf{K}}_n^{[p]})$ the real part of $\tilde{\mathbf{K}}_n^{[p]}$. In Theorem 1 we prove for $p=3$ the following result:

$$\exists c_p>0:\ \Re(\tilde{\mathbf{K}}_n^{[p]}) \geq c_p\,\tau_{n+p-2}(2-2\cos\theta), \quad \forall n\geq 2, \tag{20}$$

where

$$\Re(\tilde{\mathbf{K}}_n^{[3]}) = \begin{bmatrix} \frac{11}{2} & -\frac{4}{3} & -\frac{1}{6} & & & & & & \\ -\frac{4}{3} & \frac{5}{2} & -1 & & & & & & \\ -\frac{1}{6} & -1 & 2 & -1 & & & & & \\ & & -1 & 2 & -1 & & & & \\ & & & \ddots & \ddots & \ddots & & & \\ & & & & -1 & 2 & -1 & & \\ & & & & & -1 & 2 & -1 & -\frac{1}{6} \\ & & & & & & -1 & \frac{5}{2} & -\frac{4}{3} \\ & & & & & & -\frac{1}{6} & -\frac{4}{3} & \frac{11}{2} \end{bmatrix}$$

and

$$\tau_{n+1}(2-2\cos\theta) = \begin{bmatrix} 2 & -1 & & & \\ -1 & \ddots & \ddots & & \\ & \ddots & & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{bmatrix}.$$

Theorem 1. *There exists a constant $c_3>0$ such that*

$$\Re(\tilde{\mathbf{K}}_n^{[3]}) \geq c_3\,\tau_{n+1}(2-2\cos\theta), \quad \forall n\geq 2.$$

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

$$\Re(\tilde{\mathbf{K}}_n^{[3]}) \geq c\,\tau_{n+1}(2-2\cos\theta), \qquad \forall n\geq 6. \tag{21}$$

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

$$\begin{bmatrix}
\frac{11}{2}-2c & -\frac{4}{3}+c & -\frac{1}{6} & & & & & & \\
-\frac{4}{3}+c & \frac{5}{2}-2c & -1+c & & & & & & \\
-\frac{1}{6} & -1+c & 2-2c & -1+c & & & & & \\
 & & -1+c & 2-2c & -1+c & & & & \\
 & & & \ddots & \ddots & \ddots & & & \\
 & & & & -1+c & 2-2c & -1+c & & \\
 & & & & & -1+c & 2-2c & -1+c & -\frac{1}{6} \\
 & & & & & & -1+c & \frac{5}{2}-2c & -\frac{4}{3}+c \\
 & & & & & & & -\frac{1}{6} & -\frac{4}{3}+c & \frac{11}{2}-2c
\end{bmatrix}$$

and can be decomposed as follows:

$$\begin{bmatrix} r & -\frac{4}{3}+c \\ -\frac{4}{3}+c & s \end{bmatrix}_2^1 + \begin{bmatrix} \frac{11}{2}-2c-r & 0 & -\frac{1}{6} \\ 0 & 0 & 0 \\ -\frac{1}{6} & 0 & t \end{bmatrix}_3^1 + \begin{bmatrix} \frac{5}{2}-2c-s & -1+c \\ -1+c & 1-c-t \end{bmatrix}_3^2$$
$$+ \begin{bmatrix} 1-c & -1+c & & & \\ -1+c & 2-2c & -1+c & & \\ & \ddots & \ddots & \ddots & \\ & & -1+c & 2-2c & -1+c \\ & & & -1+c & 1-c \end{bmatrix}_{n-1}^3$$
$$+ \begin{bmatrix} 1-c-t & -1+c \\ -1+c & \frac{5}{2}-2c-s \end{bmatrix}_n^{n-1} + \begin{bmatrix} t & 0 & -\frac{1}{6} \\ 0 & 0 & 0 \\ -\frac{1}{6} & 0 & \frac{11}{2}-2c-r \end{bmatrix}_{n+1}^{n-1} + \begin{bmatrix} s & -\frac{4}{3}+c \\ -\frac{4}{3}+c & r \end{bmatrix}_{n+1}^{n}.$$

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,

$$\begin{bmatrix} \frac{11}{2}-2c-r & 0 & -\frac{1}{6} \\ 0 & 0 & 0 \\ -\frac{1}{6} & 0 & t \end{bmatrix}_3^1 = \begin{bmatrix} \frac{11}{2}-2c-r & 0 & -\frac{1}{6} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \vdots & & \vdots \\ -\frac{1}{6} & 0 & t & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & & \vdots \\ \vdots & & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{bmatrix}.$$

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

that each term of the summation is nonnegative definite:

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

These conditions are satisfied, for instance, with

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

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 $p\geq 2$. We also remark that (20) can be used in connection with the Fan-Hoffman theorem [4] in order to show that the spectral condition number of $\tilde{\mathbf{K}}_n^{[p]}$ is bounded from above by $C_p n^2$, where C_p is a constant depending only on p. Moreover, 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, it follows that the spectral condition number of $\tilde{\mathbf{K}}_n^{[p]}$ is bounded from below by $c_p n^2$, with c_p being a constant depending only on p.

In conclusion the spectral conditioning of the collocation stiffness matrices grows asymptotically as n^2 exactly as in the Galerkin setting [15].

We finally stress (20) is also important in a multigrid context for proving the optimality of the related two-grid and multigrid techniques (see [2, 7, 9]), by following the theory reported by Ruge and Stüben in [19].

References

1. F. Ahmad, E.S. Al-Aidarous, D. Abdullah Alrehaili, S.-E. Ekström, I. Furci, S. Serra-Capizzano, *Are the eigenvalues of preconditioned banded symmetric Toeplitz matrices known in almost closed form?*, Numer. Alg. 73-3 (2018) 867–893.
2. A. Aricò, M. Donatelli, S. Serra-Capizzano, *V-cycle optimal convergence for certain (multilevel) structured linear systems*, SIAM J. Matrix Anal. Appl. 26 (2004) 186–214.
3. B. Beckermann, S. Serra-Capizzano, *On the asymptotic spectrum of Finite Elements matrices*, SIAM J. Numer. Anal. 45 (2007) 746–769.
4. R. Bhatia, *Matrix analysis*, Springer-Verlag, New York, 1997.
5. J.A. Cottrell, T.J.R. Hughes, Y. Bazilevs, *Isogeometric analysis: toward integration of CAD and FEA*, John Wiley & Sons, Chichester, 2009.

6. M. DONATELLI, C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, H. SPELEERS, *Robust and optimal multi-iterative techniques for IgA Galerkin linear systems*, Comput. Meth. Appl. Mech. Engrg. 284 (2015) 230–264.
7. M. DONATELLI, C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, H. SPELEERS, *Two-grid optimality for Galerkin linear systems based on B-splines*, Comput. Vis. Sci. 17-3 (2015) 119–133.
8. M. DONATELLI, C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, H. SPELEERS, *Spectral analysis of matrices in isogeometric collocation methods*, Math. Comp. 85-300 (2016) 1639–1680.
9. M. DONATELLI, C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, H. SPELEERS, *Symbol-based construction and analysis of optimal/robust multigrid methods for B-splines Isogeometric Analysis*, SIAM J. Numer. Anal. 55-1 (2017) 31–62.
10. S.-E. EKSTRÖM, *Matrix-less Methods for Computing Eigenvalues of Large Structured Matrices*, Ph.D. Thesis, Uppsala University, 2018.
11. S.-E. EKSTRÖM, I. FURCI, C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, H. SPELEERS, *Are the eigenvalues of the B-spline IgA approximation of $-\Delta u = \lambda u$ known in almost closed form?*, Numer. Linear Algebra Appl. (2018), DOI: 10.1002/nla.2151, in press.
12. S.-E. EKSTRÖM, I. FURCI, S. SERRA-CAPIZZANO, *Exact formulae and matrix-less eigensolvers for block banded symmetric Toeplitz matrices*, BIT Numer. Math. (2018), DOI: 10.1007/s10543-018-0715-z, in press.
13. S.-E. EKSTRÖM, C. GARONI, *A matrix-less and parallel interpolation–extrapolation algorithm for computing the eigenvalues of preconditioned banded symmetric Toeplitz matrices* Numer. Alg. (2018), DOI: 10.1007/s11075-018-0508-0, in press.
14. S.-E. EKSTRÖM, C. GARONI, S. SERRA-CAPIZZANO, *Are the eigenvalues of banded symmetric Toeplitz matrices known in close form?*, Exp. Math. (2017), DOI: 10.1080/10586458.2017.1320241, in press.
15. C. GARONI, C. MANNI, F. PELOSI, S. SERRA-CAPIZZANO, H. SPELEERS, *On the spectrum of stiffness matrices arising from isogeometric analysis*, Numer. Math. 127 (2014) 751–799.
16. C. GARONI, C. MANNI, S. SERRA-CAPIZZANO, D. SESANA, H. SPELEERS, SPECTRAL ANALYSIS AND SPECTRAL SYMBOL OF MATRICES IN ISOGEOMETRIC GALERKIN METHODS, Math. Comp. 86-305 (2017) 1343–1373.
17. C. GARONI, S. SERRA-CAPIZZANO, *The theory of Generalized Locally Toeplitz sequences: theory and applications - Vol I*, Springer Monographs in Mathematics, Cham, 2017.
18. C. GARONI, H. SPELEERS, S.-E. EKSTRÖM, T.J.R. HUGHES, A. REALI, S. SERRA-CAPIZZANO, *Symbol-based analysis of finite element and isogeometric B-spline discretizations of eigenvalue problems: Exposition and review*, submitted.
19. J.W. RUGE, K. STÜBEN, *Algebraic multigrid*, Chapter 4 of the book *Multigrid methods* by S. McCormick, SIAM publications (1987) 73–130.
20. S. SERRA-CAPIZZANO, *Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences*, Numer. Math. 92 (2002) 433–465.
21. S. SERRA-CAPIZZANO, C. TABLINO-POSSIO, *Spectral and structural analysis of high precision Finite Difference matrices for Elliptic Operators*, Linear Algebra Appl. 293 (1999) 85–131.