# Analysis of the spectral symbol associated to discretization schemes of linear self-adjoint differential operators

**Davide Bianchi[1,2]** 📧

**Abstract**

Given a linear self-adjoint differential operator $\mathscr{L}$ along with a discretization scheme (like Finite Differences, Finite Elements, Galerkin Isogeometric Analysis, etc.), in many numerical applications it is crucial to understand how good the (relative) approximation of the whole spectrum of the discretized operator $\mathscr{L}^{(n)}$ is, compared to the spectrum of the continuous operator $\mathscr{L}$. The theory of Generalized Locally Toeplitz sequences allows to compute the spectral symbol function $\omega$ associated to the discrete matrix $\mathscr{L}^{(n)}$. Inspired by a recent work by T. J. R. Hughes and coauthors, we prove that the symbol $\omega$ can measure, asymptotically, the maximum spectral relative error $\mathscr{E} \geq 0$. It measures how the scheme is far from a good relative approximation of the whole spectrum of $\mathscr{L}$, and it suggests a suitable (possibly non-uniform) grid such that, if coupled to an increasing refinement of the order of accuracy of the scheme, guarantees $\mathscr{E} = 0$.

**Keywords** Spectral symbol · Discretization schemes · Linear self-adjoint differential operators · Maximum spectral relative error

**Mathematics Subject Classification** 47B35 · 35P05 · 65N99 · 65D25 · 65F99

## 1 Introduction

If $\mathscr{L}^{(n)}$ is a matrix that discretizes a linear self-adjoint differential operator $\mathscr{L}$, obtained by a discretization scheme like Finite Differences (FD), Finite Elements (FE), Isogeometric Galerkin Analysis (IgA), etc., several problems require that

📧 Davide Bianchi
d.bianchi9@uninsubria.it; bianchi@hit.edu.cn

1    University of Insubria, Como, Italy

2    Harbin Institute of Technology (Shenzhen), Shenzhen, China

the spectrum of $\mathscr{L}^{(n)}$ converges uniformly to the (point) spectrum of the operator $\mathscr{L}$. More precisely, let us set the notation: for any fixed $k, j \in \mathbb{N} \setminus \{0\}$, indicate with $\lambda_k$ and $\lambda_{-j}$ the $k$-th nonnegative and the $j$-th negative real eigenvalue of a given operator, respectively. Given a $d_n \times d_n$ matrix $\mathscr{L}^{(n)}$ with $d_n$ real eigenvalues, denote with $d_n^+$ the sum of the null and positive indices of inertia and with $d_n^-$ the negative index of inertia, i.e.,

$$d_n^+ := \left| \left\{ \lambda(\mathscr{L}^{(n)}) \ : \ \lambda(\mathscr{L}^{(n)}) \geq 0 \right\} \right|, \qquad d_n^- := \left| \left\{ \lambda(\mathscr{L}^{(n)}) \ : \ \lambda(\mathscr{L}^{(n)}) < 0 \right\} \right|.$$

Clearly, $d_n = d_n^- + d_n^+$. Finally, sort the eigenvalues in nondecreasing order, that is,

$$\ldots \leq \lambda_{-j}^{(n)} \leq \cdots \leq \lambda_{-1}^{(n)} < 0 \leq \lambda_1^{(n)} \leq \cdots \leq \lambda_k^{(n)} \leq \ldots$$

With this notation, the property of *uniform spectral convergence* we mentioned above translates into asking that

$$\mathscr{E} := \limsup_{n \to \infty} \mathscr{E}_n = 0, \qquad \mathscr{E}_n := \max_{\substack{k = -d_n^-, \ldots, d_n^+ \\ k \neq 0}} \left| \frac{\lambda_k^{(n)}}{\lambda_k} - 1 \right|, \qquad (1.1)$$

where $n$ is the mesh fineness parameter, and $\lambda_k^{(n)}$ and $\lambda_k$ are the $k$-th eigenvalues of the discretized and the continuous operator, respectively. Typically, $\mathscr{E} > 0$: the relative error estimates for the eigenvalues and eigenfunctions are good only in the *lowest modes*, that is, for $|k| = 1, \ldots, K_n$, where $K_n / d_n \to \sigma$ as $n \to \infty$ and such that $\sigma \ll 1$. In general, a large portion of the eigenvalues, the so-called *higher modes*, are not approximations of their continuous counterparts in any meaningful sense. This may negatively affect the solutions obtained by discrete approximations of elliptic boundary-value problems, or parabolic and hyperbolic initial value problems. In these cases, all modes may contribute in the solution to some extent and inaccuracies in the approximation of higher modes can not always be ignored. Regarding this, see for example the spectral-gap problem of the one dimensional wave equation for the uniform observability of the control waves [8, 20, 28, 35] or structural engineering problems [27, Sects. 3–6].

In this setting, the theory of Generalized Locally Toeplitz (GLT) sequences provides the necessary tools to understand whether the methods used to discretize the operator $\mathscr{L}$ are effective in approximating the whole spectrum. The GLT theory originated from the seminal work of P. Tilli on Locally Toeplitz sequences [48] and was later developed by S. Serra-Capizzano in [41, 42]. It was devised to compute and analyze the spectral distribution of matrices arising from the numerical discretization of integral equations and differential equations. Let $\hat{\mathscr{L}}^{(n)}$ be the discrete operator normalized by an appropriate power of $d_n$, which depends on the dimension of the underlying space and on the maximum order of derivatives involved. It usually happens that the sequence of matrices $\left\{ \hat{\mathscr{L}}^{(n)} \right\}_n$ enjoys an asymptotic spectral distribution as $n \to \infty$, i.e., as the mesh goes to zero. More precisely, for any test function $F(t) \in C_c(\mathbb{C})$ it holds that

$$\lim_{n \to \infty} \frac{1}{d_n} \sum_{\substack{k = -d_n^- \\ k \neq 0}}^{d_n^+} F\left(\lambda_k\left(\hat{\mathscr{L}}^{(n)}\right)\right) = \frac{1}{m(D)} \int_D F(\omega(\mathbf{y}))m(d\mathbf{y}),$$

where $\lambda_k\left(\hat{\mathscr{L}}^{(n)}\right), k = -d_n^-, \ldots, -1, 1, \ldots, d_n^+$ are the eigenvalues of the normalized operator $\hat{\mathscr{L}}^{(n)}$ and $\omega : D \subset \mathbb{R}^d \to \mathbb{C}$ is referred to as the *spectral symbol* of the sequence $\left\{ \hat{\mathscr{L}}^{(n)} \right\}_n$, see [49, Eq. (3.2)] in relation to Toeplitz matrices. The GLT theory allows to compute the spectral symbol $\omega$ related to $\mathscr{L}^{(n)}$, especially if the numerical method employed to produce $\mathscr{L}^{(n)}$ belongs to the family of the so-called *local methods*, such as FD methods, FE methods and collocation methods with locally supported basis functions.

In several recent papers the sampling of the spectral symbol was suggested to be used to approximate the spectrum of the discrete matrix operators $\hat{\mathscr{L}}^{(n)}$ and $\mathscr{L}^{(n)}$. Unfortunately, this approach is not always successful in general, as we will show with an example. The main reference is the paper by T. J. R. Hughes and coauthors [24], which reviews the state-of-the-art of the symbol-based analysis for the eigenvalues distribution carried on in the framework of the isogeometric Galerkin approximation (IgA). Inspired by the latter work [24], we show that the symbol $\omega$ can measure the maximum spectral relative error $\mathscr{E}$ defined in (1.1) and it suggests a suitable (non-uniform) grid such that, if coupled to an increasing refinement of the order of accuracy of the scheme, guarantees $\mathscr{E} = 0$.

The paper is organized as follows.

- In Sect. 2, the spectral symbol $\omega$ and the monotone rearrangement $\omega^*$ are introduced.
- In Sect. 3 we present an asymptotic result, Theorem 3.1, which connects the eigenvalue distribution of a matrix sequence and the monotone rearrangement of its spectral symbol. It is the main tool for the results in Sect. 4. In particular, under suitable regularity assumptions on $\omega$, thanks to Theorem 3.2 and Theorem 3.3 we prove that

$$\mathscr{E} = \sup_{\substack{x \in (0, 1) \\ x : \zeta^*(x) \neq 0}} \left| \frac{\omega^*(x)}{\zeta^*(x)} - 1 \right|, \tag{1.2}$$

where $\zeta$ is the (normalized) Weyl distribution function of the eigenvalues of $\mathscr{L}$ and $\omega$ is the spectral symbol associated to the numerical scheme applied to discretize $\mathscr{L}$.

- Section 4 is devoted entirely to numerical experiments and applications. The validity of (1.2) is shown in Table 2 and Fig. 2. Moreover, in Sect. 4.1.1 we provide an example about the unfeasibility to obtain an accurate approximation of the eigenvalues of a differential operator by just uniformly sampling the spectral symbol $\omega$. In Sects. 4.1.2 and 4.1.3 we generalize the aforementioned results to the case of central FD and IgA methods of higher order. By means of a suit-

able non-uniform grid, suggested by the spectral symbol and combined by an increasing order of the approximation, we obtain (1.1), namely $\mathscr{E} = 0$. Finally, in Sect. 4.2, we describe a simple application to the general $d$-dimensional Laplacian.

- Section 5 is devoted to draw conclusions and set the lines for future research.

Due to the technicalities that involve the results presented in Sects. 3 and 4, all the proofs are moved to the appendix to improving the fluency of the reading. The appendix is therefore organized as follows.

- Section A collects the proofs of the results presented in Sect. 3.
- Section B provides rigorous proofs for some numerical evidences observed in Sect. 4.
- In Sect. C is provided the FD numerical scheme used in Sect. 4.1.1 and Sect. 4.1.2.
- In Sect. D are listed some known results about the IgA numerical scheme used in Sect. 4.1.3.

## 2 Spectral symbol and monotone rearrangement

In this section we provide the definitions of spectral symbol of a sequence of matrices and its monotone rearrangement, which are the main tools we will use throughout this paper to study the asymptotic spectral distribution of sequences of discretization matrices. We will use the notation $X^{(n)}, Y^{(n)}$ to denote general $d_n \times d_n$ square matrices, while $T^{(n)}$ will be used for Toeplitz matrices, i.e., matrices with constant elements along their diagonals: $\left(T^{(n)}\right)_{i,j} = t_{i-j}$ for all $i, j = 1, \ldots, d_n$, and with

$$\mathbf{t} = \left[t_{-d_n+1}, \ldots, t_0, \ldots, t_{d_n-1}\right] \in \mathbb{C}^{2d_n-1}.$$

If $t_k$ is the $k$-th Fourier coefficient of a complex integrable function $f$ defined over the interval $[-\pi, \pi]$, then $T^{(n)} = T^{(n)}(f)$ is referred to be the Toeplitz matrix generated by $f$.

With the symbol $\left\{X^{(n)}\right\}_n$ we will denote a sequence of square matrices with increasing dimensions $d_n \times d_n$, i.e., such that $d_n \to \infty$ as $n \to \infty$. When we will confront the spectrum of two sequences of matrices, we will assume that they have the same dimension. The eigenvalues will be sorted in nondecreasing order, that is,

$$\lambda_{-d_n^-}^{(n)} \leq \cdots \leq \lambda_{-j}^{(n)} \leq \cdots \leq \lambda_{-1}^{(n)} < 0 \leq \lambda_1^{(n)} \leq \cdots \lambda_k^{(n)} \leq \cdots \leq \lambda_{d_n^+}^{(n)},$$

where $d_n^-$ is the negative index of inertia and $d_n^+$ is the sum of the null and positive indices of inertia. Clearly, $d_n = d_n^- + d_n^+$. If the eigenvalues do not depend on the parameter $n$, then we will omit the superscript $(n)$.

We will consider all the Euclidean spaces equipped with the usual Lebesgue measure $m(\cdot)$. We will not specify the dimension of $m(\cdot)$ by a subscript since it will be clear from the context.

### 2.1 Spectral symbol

The following definition of spectral symbol has been slightly modified in accordance to our notation and purposes.

**Definition 2.1** *(Spectral symbol)* Let $\left\{X^{(n)}\right\}_n$ be a sequence of matrices and let $\omega : D \subseteq \mathbb{R}^d \to \mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$) be a measurable function and $D$ a measurable set with $0 < m(D) < \infty$, such that the Lebesgue integral

$$\int_D \omega(\mathbf{y}) m(d\mathbf{y})$$

exists, finite or infinite. We say that $\{X^{(n)}\}_n$ is distributed like $\omega$ in the sense of the eigenvalues, in symbols $\{X^{(n)}\}_n \sim_\lambda \omega$, if

$$\lim_{n\to\infty} \frac{1}{d_n} \sum_{\substack{k = -d_n^- \\ k \neq 0}}^{d_n^+} F\left(\lambda_k\left(X^{(n)}\right)\right) = \frac{1}{m(D)} \int_D F(\omega(\mathbf{y})) m(d\mathbf{y}), \qquad \forall F \in C_c(\mathbb{K}).$$

(2.1)

We will call $\omega$ the (spectral) symbol of $\{X^{(n)}\}_n$.

Relation (2.1) is satisfied for example by Hermitian Toeplitz matrices generated by real-valued functions $\omega \in L^1([-\pi, \pi])$, i.e., $\left\{T^{(n)}(\omega)\right\}_n \sim_\lambda \omega$, see [26, 49, 50]. For a general overview on Toeplitz operators and spectral symbol, see [10, 11]. What is interesting to highlight is that matrices with a Toeplitz-like structure naturally arise when discretizing over a uniform grid problems which have a translation invariance property, such as linear differential operators with constant coefficients.

**Remark 1** If $D$ is compact, $\omega$ continuous, and $\lambda_k \in R_\omega = [\min \omega, \max \omega]$ for every $k \in \mathbb{N}$, then taking $F(t) = t\chi_{R_\omega}(t)$, with $\chi_{R_\omega}$ a $C^\infty$ cut-off such that $\chi_{R_\omega}(t) \equiv 1$ on $R_\omega$, gives

$$\lim_{n\to\infty} \frac{1}{d_n} \sum_{\substack{k = -d_n^- \\ k \neq 0}}^{d_n^+} \lambda_k\left(X^{(n)}\right) = \frac{1}{m(D)} \int_D \omega(\mathbf{y}) m(d\mathbf{y}).$$

(2.2)

Because the Riemannian sum over equispaced points converges to the integral of the right hand side of the above formula, (2.2) could suggest that the eigenvalues $\lambda_k\left(X^{(n)}\right)$ can be approximated by a pointwise evaluation of the symbol $\omega(\mathbf{y})$ over an equispaced grid of $D$, for $n \to \infty$, except for at most an $o(d_n)$ of outliers, see Definition 2.3. This is mostly the content of [22, Remark 3.2] and [24, Sect. 2.2].

We remark that in the special case of Toeplitz matrices generated by regular enough functions $\omega$,

$$\lambda_k\left(T^{(n)}(\omega)\right) = \omega\left(\frac{k\pi}{d_n+1}\right) + O(d_n^{-1}) \qquad \text{for every } k = 1, \dots, d_n,$$

see [51] and [9]. In some sense, this justifies the informal meaning given above to the spectral symbol.

That said, unfortunately the discretization of a linear differential operator does not always own a Toeplitz-like structure. Nevertheless, the GLT theory provides tools to prove the validity of (2.1) for more general matrix sequences. For a review of the GLT theory and its applications we mainly refer to [22, 23], and to [5, 6]. We conclude this subsection with two definitions.

**Definition 2.2** *(Essential range)* Let $\omega$ be a real valued measurable function and define the set $R_\omega \subset \mathbb{R}$ as

$$t \in R_\omega \quad \Longleftrightarrow \quad m(\{\mathbf{y} \in D : |\omega(\mathbf{y}) - t| < \epsilon\}) > 0 \quad \forall \epsilon > 0.$$

We call $R_\omega$ the *essential range* of $\omega$. Notice that $R_\omega$ is closed.

**Definition 2.3** *(Outliers)* Given a matrix sequence $\{X^{(n)}\}$ such that $\left\{X^{(n)}\right\}_n \sim_\lambda \omega$, if $\lambda_k\left(X^{(n)}\right) \notin R_\omega$ we call it an *outlier*.

## 2.2 Monotone rearrangement

Dealing with a univariate and monotone spectral symbol $\omega(\mathbf{y}) = \omega(\theta)$ has several advantages. Unfortunately, in general $\omega$ is multivariate or not monotone. Nevertheless, it is possible to consider a rearrangement $\omega^* : (0, 1) \to (\inf R_\omega, \sup R_\omega)$ such that $\omega^*$ is univariate, monotone nondecreasing and still satisfies the limit relation (2.1). This can be achieved in the following way.

**Definition 2.4** Let $\omega : D \subset \mathbb{R}^d \to \mathbb{R}$ be measurable. Define

$$\omega^* : (0, 1) \to R_\omega, \qquad \omega^*(x) = \inf\left\{t \in R_\omega : \phi_\omega(t) > x\right\} \tag{2.3}$$

where

$$\phi_\omega : \mathbb{R} \to [0, 1], \qquad \phi_\omega(t) := \frac{1}{m(D)} m(\{\mathbf{y} \in D : \omega(\mathbf{y}) \le t\}), \tag{2.4}$$

and where, in case of bounded $R_\omega$, we consider the extension $\omega^* : [0, 1] \to R_\omega$.

In analysis, $\omega^*$ is called *monotone rearrangement* (see [44, pg. 189]) while in probability theory it is called (generalized) inverse distribution function (see [47, pg. 260]). For "historical" reasons we will use the analysts' name; see [18, Definition 3.1 and Theorem 3.3] and [39], where the monotone rearrangement was first introduced in the context of spectral symbol functions. Clearly, $\omega^*$ is a.e. unique, univariate, and monotone increasing, which make it a good choice for an equispaced sampling. On the other

hand, $\omega^*$ could not have an analytical expression or it could be not feasible to compute, therefore it is often needed an approximation $\omega_r^*$. In Algorithm 1 we summarize the steps presented in [22, Example 10.2] and [24, Sect. 3] in order to approximate the eigenvalues $\lambda_k\left(X^{(n)}\right)$ by means of an equispaced sampling of the rearranged symbol $\omega^*$ (or its approximated version $\omega_r^*$). For the sake of clarity and without loss of generality, we suppose $D = [0, 1] \times [-\pi, \pi]$ and $\omega$ continuous. As standard result in approximations of monotone rearrangements, it holds that $\|\omega_{r(n)}^* - \omega^*\|_\infty \to 0$ as $n \to \infty$, see [13, 45]. For a detailed survey about monotone rearrangements we refer to the recent work by G. Talenti [46].

---

**Algorithm 1** Monotone rearrangement

1: Fix $r \in \mathbb{N}$ such that $r = r(n) \geq d_n$, and fix the equispaced grid $\{(x_i, \theta_j)\}$ over $[0, 1] \times [-\pi, \pi]$, where $x_i = \frac{i}{r+1}$, $\theta_j = -\pi + \frac{2\pi j}{r+1}$ for $i, j = 1, \cdots, r$;

2: Get the set of samplings $\{\omega(x_i, \theta_j)\}_{i,j=1}^r$ and form a nondecreasing sequence $\{\omega_1 \leq \omega_2 \leq \cdots \leq \omega_{r^2}\}$;

3: Define $\omega_r^* : [0, 1] \to [\min \omega, \max \omega]$ as the piecewise linear nondecreasing function which interpolates the samples $\{\min \omega = \omega_0 \leq \omega_1 \leq \omega_2 \leq \cdots \leq \omega_{r^2} \leq \omega_{r^2+1} = \max \omega\}$ over the nodes $\{0, \frac{1}{r^2+1}, \frac{2}{r^2+1}, \cdots, \frac{r^2}{r^2+1}, 1\}$;

4: Sample $\omega_r^*$ over the set $\left\{\frac{k}{d_n+1}\right\}_{k=1}^{d_n}$ and define $\omega_{r,k}^{*,(n)} := \omega_r^*\left(\frac{k}{d_n+1}\right)$. Obviously, if $\omega^*$ is available then use it instead of $\omega_r^*$ and define $\omega_k^{*,(n)} := \omega^*\left(\frac{k}{d_n+1}\right)$.

---

## 3 Asymptotic spectral distribution

In this section we present the main results about the asymptotic spectral distribution of the eigenvalues of a matrix sequence with given spectral symbol. Since the proofs are rather technical, they can be found in Sect. A to not break the flow of the reading.

Theorem 3.1 connects the monotone rearrangement of the symbol function to the asymptotic eigenvalues distribution of a matrix sequence. Theorem 3.2 provides an analytic way to measure the maximum spectral relative error between two sequences of matrices, see Definition 3.1. Finally, Theorem 3.3 exploits the previous results to give an effective measure for the maximum spectral relative error that characterizes a discretization scheme, when approximating a linear self-adjoint differential operator.

### 3.1 Monotone rearrangement and spectral distribution

Let $\{X^{(n)}\}_n$ be a matrix sequence such that $\{X^{(n)}\}_n \sim_\lambda \omega$ and define the (discrete) *eigenvalues counting function* $N(X^{(n)}, \cdot) : \mathbb{R} \to \{0, 1, \ldots, d_n\}$,

$$N\left(X^{(n)}, t\right) := \left|\left\{k = -d_n^-, \ldots, -1, 1, \ldots, d_n^+ \ : \ \lambda_k\left(X^{(n)}\right) \leq t\right\}\right|.$$

Since the monotone rearrangement takes as argument values in $[0, 1]$, it is helpful to define a new indexing $\hat{k}$ associated to $X^{(n)}$ such that $\hat{k}/d_n \in [0, 1]$. Let us set

$$\vartheta_{X^{(n)}} : \left\{ -d_n^- \left( X^{(n)} \right), \dots, -1, 1, \dots, d_n^+ \left( X^{(n)} \right) \right\} \to \{1, \dots, d_n\},$$

$$\hat{k} = \vartheta_{X^{(n)}}(k) := \begin{cases} k + d_n^- \left( X^{(n)} \right) + 1 & \text{if } k < 0, \\ k + d_n^- \left( X^{(n)} \right) & \text{if } k > 0. \end{cases} \tag{3.1}$$

We have the following asymptotic relations.

**Theorem 3.1** (Discrete Weyl's law) *Let $\{X^{(n)}\}_n$ be a matrix sequence such that $\{X^{(n)}\}_n \sim_\lambda \omega$. Then*

$$\left\{ X^{(n)} \right\}_n \sim_\lambda \omega^*(x), \quad x \in (0, 1), \tag{3.2}$$

*and*

$$\lim_{n \to \infty} \frac{N\left( X^{(n)}, t \right)}{d_n} = \phi_\omega(t) \quad \text{for every point of continuity of } \phi_\omega. \tag{3.3}$$

*Suppose now that $\phi_\omega$ is continuous or, equivalently, that*

$$m(\{\mathbf{y} : \omega(\mathbf{y}) = t\}) = 0 \quad \forall t \in \mathbb{R}.$$

*Let $\hat{k} = \hat{k}(n)$ as in (3.1) be such that the sequence $\{\hat{k}(n)/d_n\}_n$ converges from the left to $x_0 \in (0, 1)$, as $n \to \infty$. Then, there exists the limit for the sequence $\left\{ \lambda_{\hat{k}(n)}\left( X^{(n)} \right) \right\}_n$ and it holds that*

$$\lim_{x \to x_0^-} \omega^*(x) = \sup \left\{ t \in R_\omega : t \leq \lim_{n \to \infty} \lambda_{\hat{k}(n)}\left( X^{(n)} \right) \right\}. \tag{3.4}$$

*In particular, if $\omega^*$ is continuous at $x_0$,*

$$\omega^*\left( x_0 \right) = \lim_{n \to \infty} \lambda_{\hat{k}(n)}\left( X^{(n)} \right). \tag{3.5}$$

*Finally, if $\lambda_{\hat{k}(n)}\left( X^{(n)} \right) \geq \inf \left( R_\omega \right) (\leq \sup(R_\omega))$ eventually, then (3.5) holds for $x_0 = 0$ ($x_0 = 1$) as well.*

**Proof** See Sect. A. □

In some sense, the limit relation (2.2) can be viewed as the strong law for large numbers for specially chosen sequences of dependent complex/real-valued random variables $\left\{ \lambda_{k(n)}\left( X^{(n)} \right) \right\}_n$. See [26] for the link between the spectrum of Hermitian matrices and equally distributed sequences (in the sense of Weyl) as defined in [31, Definition 7.1], and [34] as a recent survey about equidistributions from a probabilistic point of view.

**Corollary 3.1** *With the same notations of Theorem 3.1 and assuming $\phi_\omega$ continuous, it holds that*

$$\lim_{n \to \infty} \frac{\left|\left\{k = -d_n^-, \ldots, -1, 1, \ldots, d_n^+ \; : \; \lambda_k\left(X^{(n)}\right) \notin R_\omega\right\}\right|}{d_n} = 0,$$

*that is, the number of possible outliers is $o(d_n)$. Moreover,*

$$\lim_{n \to \infty} \frac{\left|\left\{k = -d_n^-, \ldots, -1, 1, \ldots, d_n^+ \; : \; \lambda_k\left(X^{(n)}\right) \le t, \; \lambda_k\left(X^{(n)}\right) \in R_\omega\right\}\right|}{d_n} = \phi_\omega(t).$$

**Proof** See Sect. A. □

**Corollary 3.2** *With the same notations of Theorem* 3.1, *assume that $\phi_\omega$ is continuous and $\omega^*$ is absolutely continuous. Let $\omega_n^*$ be the (pseudo) inverse function of $\frac{N(X^{(n)}, \cdot)}{d_n}$, then $\left\|\omega_n^* - \omega^*\right\|_\infty \to 0$ as $n \to \infty$. Let $h : \mathbb{N} \to \infty$ be such that*

$$\lim_{n \to \infty} h(d_n) = \infty, \qquad \lim_{n \to \infty} \frac{\left\|\omega_n^* - \omega^*\right\|_\infty}{\frac{1}{h(d_n)}} = 0.$$

*Let $g : (\inf R_\omega, \sup R_\omega) \to \mathbb{R}$ be a differentiable real function and let $\{\hat{k}(n)\}_n$ be a sequence of integers such that*

(i)  $\frac{\hat{k}(n)}{d_n} \to x_0 \in [0, 1]$;

(ii)  $\lambda_{\hat{k}(n)+1}\left(X^{(n)}\right) > \lambda_{\hat{k}(n)}\left(X^{(n)}\right) \in (\inf R_\omega, \sup R_\omega)$ *eventually for $n \to \infty$.*

*Then*

$$\lim_{n \to \infty} h(d_n) \left[ g\left( \lambda_{\hat{k}(n) + \left\lceil \frac{d_n}{h(d_n)} \right\rceil}\left(X^{(n)}\right) \right) - g\left( \lambda_{\hat{k}(n)}\left(X^{(n)}\right) \right) \right] = \lim_{x \to x_0} (g(\omega^*(x)))'$$

*almost everywhere.*

**Proof** See Sect. A. □

**Corollary 3.3** *Let $\phi_\omega$ and $\omega^*$ be continuous, and suppose that $R_\omega$ is bounded. Then, in presence of no outliers (eventually), the absolute error between a uniform sampling of $\omega^*$ and the eigenvalues of $X^{(n)}$ converges to zero, namely*

$$\max_{\substack{k = -d_n^-, \ldots, d_n^+ \\ k \ne 0}} \left\{ \left| \lambda_k\left(X^{(n)}\right) - \omega^*\left( \frac{\hat{k}}{d_n + 1} \right) \right| \right\} \to 0 \qquad \text{as } n \to \infty.$$

**Proof** See Sect. A. □

### 3.2 Local and maximum spectral relative errors

Exploiting the results obtained in the preceding subsection, we can now prescribe a way to measure the maximum relative error between two sequences of eigenvalues. Given two sequences of matrices $\{X^{(n)}\}_n, \{Y^{(n)}\}_n$ of the same dimension $d_n$, we will use the notation $\hat{k}$ as in (3.1) to have an index with a common set, $\hat{k} \in \{1, \ldots, d_n\}$, for both the eigenvalues associated to $X^{(n)}$ and $Y^{(n)}$. Moreover, we define

$$\vartheta := \vartheta_{X^{(n)}}^{-1} \circ \vartheta_{Y^{(n)}}, \tag{3.6}$$

where $\vartheta_{(\cdot)}$ was introduced in (3.1) as well. Let us observe that, if $\lambda_j(Y^{(n)})$ is the $j$-th eigenvalue associated to $Y^{(n)}$ and $\lambda_i(X^{(n)})$ is the $i$-th eigenvalue associated to $X^{(n)}$, then $\vartheta(j) = i$ if and only if $\hat{j} = \hat{i}$. See Fig. 1.

**Definition 3.1** *(Local and maximum spectral relative errors)* Given two sequences of matrices $\{X^{(n)}\}_n, \{Y^{(n)}\}_n$ of the same dimension $d_n$, define the sequence

$$\mathscr{E}_n(X^{(n)}, Y^{(n)}) := \max_{\hat{k}=1,\ldots,d_n} \left\{ \delta_{\hat{k}}^{(n)} \right\},$$
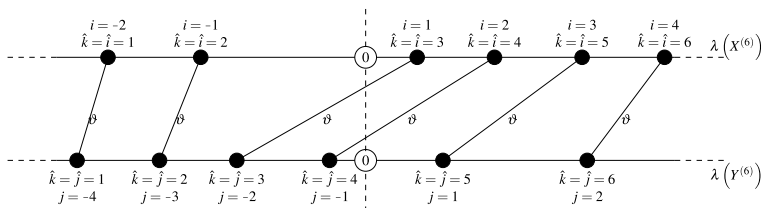
where

$$\delta_{\hat{k}}^{(n)} := \begin{cases} \left| \frac{\lambda_{\hat{k}}(X^{(n)})}{\lambda_{\hat{k}}(Y^{(n)})} - 1 \right| & \text{if } \lambda_{\hat{k}}(Y^{(n)}) \neq 0, \\ 0 & \text{if } \lambda_{\hat{k}}(Y^{(n)}) = \lambda_{\hat{k}}(X^{(n)}) = 0, \\ \infty & \text{if } \lambda_{\hat{k}}(X^{(n)}) \neq \lambda_{\hat{k}}(Y^{(n)}) = 0. \end{cases}$$

We call $\delta_{\hat{k}}^{(n)}$ *the local spectral relative error*, and we call

$$\mathscr{E}(X^{(n)}, Y^{(n)}) := \limsup_{n \to \infty} \mathscr{E}_n(X^{(n)}, Y^{(n)})$$

the *maximum spectral relative error* (MSRE). When both the matrix sequences $\{X^{(n)}\}_n, \{Y^{(n)}\}_n$ are well understood from the context, with abuse of notation we will just write $\mathscr{E}$, and this applies in particular when we will compare the eigenvalues



**Fig. 1** Example of the new index notation used to compare the eigenvalues of two matrices $X^{(n)}, Y^{(n)}$. In this case we set $n = 6$ and non-null eigenvalues. Observe that the new labeling is made necessary due to the fact that, in general, $X^{(n)}$ and $Y^{(n)}$ can have different indices of inertia. In this example, $d_6^-(X^{(6)}) = 2$ and $d_6^-(Y^{(6)}) = 4$. The function $\vartheta$ in (3.6) allows to pass from the original indices of $Y^{(n)}$ to the corresponding indices of $X^{(n)}$

of a discretized operator $\mathscr{L}^{(n)}$ with the eigenvalues of the continuous operator $\mathscr{L}$, whose first $d_n$ eigenvalues can be arranged as the diagonal entries of a diagonal matrix.

For brevity, let us set $\kappa^- := \inf_n \{d_n^-(Y^{(n)})\}, \kappa^+ := \sup_n \{d_n^+(Y^{(n)})\}$. Moreover, for any fixed $x_0 \in [0,1]$ let us define

$$F := \left\{ \{a_m\}_m \ : \ \exists \lim_{m \to \infty} a_m \text{ finite or infinite} \right\},$$

$$G_{x_0} := \left\{ \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} \ : \ \lim_{n_j \to \infty} \frac{\hat{k}(n_j)}{d_{n_j}} = x_0 \text{ eventually by passing to a subsequence} \right\},$$

$$H_{x_0} := \left\{ \lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} \ : \ \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} \in F \cap G_{x_0} \right\}, \quad \sigma_k = \limsup_{n \to \infty} \left| \frac{\lambda_{\vartheta(k)}(X^{(n)})}{\lambda_k(Y^{(n)})} - 1 \right|.$$

**Theorem 3.2** *Fix two sequences of matrices such that $\{X^{(n)}\}_n \sim_\lambda \omega_1$ and $\{Y^{(n)}\}_n \sim_\lambda \omega_2$. If*

$$m(\{y \ : \ \omega_1(y) = t\}) = m(\{z \ : \ \omega_2(z) = t\}) = 0 \quad \forall t \in \mathbb{R},$$

*then*

$$\mathscr{E}(X^{(n)}, Y^{(n)}) \geq \max \left\{ \sup_{\substack{x \in (0,1); \\ x \, : \, \omega_2^*(x) \neq 0}} \left| \frac{\omega_1^*(x)}{\omega_2^*(x)} - 1 \right|; \sigma \right\}, \tag{3.7}$$

*where*

$$\sigma := \begin{cases} 0 & \text{if } \omega_2^*(x) \neq 0 \, \forall x \in [0,1], \\ \sup H_{x_0} & \text{if } \exists x_0 \in [0,1] \text{ such that } \omega_2^*(x_0) = \omega_1^*(x_0) = 0, \\ +\infty & \text{if } \exists x_0 \in [0,1] \text{ such that } \omega_2^*(x_0) = 0 \neq \omega_1^*(x_0). \end{cases}$$

*If $\omega_2^*(x_0) = 0$ and*

$$\limsup_{n \to \infty} \left| \left| \frac{\lambda_{\vartheta(k(n))}(X^{(n)})}{\lambda_{k(n)}(Y^{(n)})} - 1 \right| - \sigma_{k(n)} \right| = 0 \quad \text{for every } k(n) \text{ such that } \lim_{n \to \infty} \frac{k(n)}{d_n} = 0,$$

*then*

$$\sup H_{x_0} = \sup_{\substack{k \in [\kappa^-, \kappa^+] \\ k \neq 0}} \{\sigma_k\}.$$

*Moreover, if $\omega_1^*, \omega_2^*$ are continuous, $R_{\omega_1}$ and $R_{\omega_2}$ are bounded and both the sequences do not have outliers, then equality holds in (3.7).*

**Proof** See Sect. A. □

### 3.3 Linear self-adjoint differential operators and eigenvalue distribution

The asymptotic distribution of the eigenvalues for partial differential operators on general manifolds has been widely studied and developed, see for example [32, 38] and all the references therein. The topic is too vast to cover it properly, therefore we will concentrate our examples only on a couple of cases: Sturm-Liouville operators for the one dimensional case and elliptic self-adjoint operators for the multi-dimensional case, see Sect. 4.1. Nevertheless, the tools presented in this section can be applied to study the quality of a discretization scheme to preserve the discrete spectrum of many classes of self-adjoint operators. The approach is the following: given an operator $\mathscr{L}$ and its discretized version $\mathscr{L}^{(n)}$, if $\{h(d_n)\mathscr{L}^{(n)}\}_n \sim_\lambda \omega$ for a function $h : \mathbb{N} \to \mathbb{R}$ (which depends on the dimension of the underlying space and the higher order of derivatives involved), then study the limit of

$$\zeta_n(t) = \frac{N^{(n)}(\mathscr{L}, t)}{d_n} := \frac{\left| k = -d_n^-, \ldots, -1, 1, \ldots, d_n^+ \ : \ h(d_n)\lambda_k(\mathscr{L}) \le t \right|}{d_n}. \quad (3.8)$$

We have the following result.

**Theorem 3.3** *Let $\mathscr{L}$ be a self-adjoint linear operator and let $\mathscr{L}^{(n)}$ be a $d_n \times d_n$ matrix obtained from $\mathscr{L}$ by a discretization scheme. Let $d_n^-, d_n^+$ be the negative and nonnegative indices of inertia of $\mathscr{L}^{(n)}$. Define*

$$X^{(n)} := \mathscr{L}^{(n)}, \qquad Y^{(n)} := \operatorname*{diag}_{\substack{k = -d_n^-, \ldots, d_n^+ \\ k \ne 0}} \left\{ \lambda_k(\mathscr{L}) \right\}.$$

*Suppose that:*

(i) $\left| \frac{\lambda_{\vartheta(k)}(\mathscr{L}^{(n)})}{\lambda_k(\mathscr{L})} - 1 \right| \to 0$ *as $n \to \infty$ for every $k = k(n)$ such that $\frac{k(n)}{d_n} \to 0$, where $\vartheta$ is defined in (3.6);*

(ii) $\{h(d_n)\mathscr{L}^{(n)}\}_n \sim_\lambda \omega$ *for some fixed $h : \mathbb{N} \to \mathbb{R}$;*

(iii) $m(\{\mathbf{y} \ : \ \omega(\mathbf{y}) = t\}) = 0$ *for every $t \in \mathbb{R}$, and $R_\omega$ bounded;*

(iv) $\lim_{n\to\infty} \zeta_n(t) = \zeta(t)$, *where $\zeta_n$ is defined in (3.8) and $\zeta : \mathbb{R} \to [0, 1]$ is continuous and surjective.*

*Then*

$$\mathscr{E} = \mathscr{E}\big(X^{(n)}, Y^{(n)}\big) \geq \sup_{\substack{x \in [0, 1] \\ x \,:\, \zeta^*(x) \neq 0}} \left| \frac{\omega^*(x)}{\zeta^*(x)} - 1 \right|.$$

*Moreover, if $\omega^*, \zeta^*$ are continuous and $h(d_n)\lambda_k\big(\mathscr{L}^{(n)}\big) \in R_\omega$ eventually, then equality holds.*

**Proof** See Sect. A. □

The (normalized) Weyl function $\zeta$ is known for many kind of differential operators. See the next section for a couple of examples.

# 4 Numerical experiments

We will write $\mathscr{L}^{(n,\eta)}$ to denote a $d_n \times d_n$ square matrix which is the discretization of a linear differential operator $\mathscr{L}$ by means of a numerical scheme of order of approximation $\eta$. In the case where the approximation order $\eta$ is clear by the context, then we will omit it. If the discretized operator $\mathscr{L}^{(n)}$ is normalized by a constant depending on the fineness mesh parameter $n$, then we will denote it with $\hat{\mathscr{L}}^{(n)}$. We will use the subscripts *dir* and *BCs* to denote a (discretized) linear differential operator characterized with Dirichlet or generic boundary conditions, respectively. When it will be necessary to highlight the dependency of the differential operator on a variable coefficient $p(x)$, we will write it as subscript. So, for example, the normalized discretization of a diffusion operator with Dirichlet BCs by means of the IgA scheme of order $\eta$ will be denoted by

$$\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir},p(x)}.$$

All the computations are performed on MATLAB R2018b running on a desktop-pc with an Intel i5-4460H @3.20 GHz CPU and 8GB of RAM.

## 4.1 Application to Euler-Cauchy differential operator

We begin our analysis with respect to a toy-model example. In this subsection, the main objectives are:

- to show numerical evidences of Theorem 3.3, i.e., that the monotone rearrangement $\omega^*(x)$ of the spectral symbol $\omega(y)$ measures the maximum spectral relative error $\mathscr{E}$. See Sects. 4.1.1, 4.1.2 and 4.1.3;
- to disprove that, in general, a uniform sampling of the spectral symbol $\omega(y)$ can provide an accurate approximation of the eigenvalues of the normalized and unnormalized discrete operators $\hat{\mathscr{L}}^{(n)}$ and $\mathscr{L}^{(n)}$, respectively. See Sect. 4.1.1;

- to show that a discretization scheme can lead to $\mathscr{E} = 0$ if coupled with a suitable (possibly non-uniform) grid and an increasing order of approximation. See Sects. 4.1.2 and 4.1.3.

Let us fix $\alpha > 0$ and let us consider the following self-adjoint operator with Dirichlet BCs,

$$\mathscr{L}_{\mathrm{dir},\alpha x^2}[u](x) := -\left(\alpha x^2 u'(x)\right)', \qquad \mathrm{dom}\left(\mathscr{L}_{\mathrm{dir},\alpha x^2}\right) = W_0^{1,2}(1, \mathrm{e}^{\sqrt{\alpha}}). \quad (4.1)$$

The formal equation is an Euler-Cauchy differential equation and by means of the Liouville transformation

$$y(x) = \int_1^x \left(\sqrt{\alpha t}\right)^{-1} dt, \quad (4.2)$$

the operator (4.1) has the same spectrum of

$$\Delta_{\mathrm{dir},\alpha}[v](y) := -v''(y) + \frac{\alpha}{4} v(y), \qquad \mathrm{dom}\left(\Delta_{\mathrm{dir},\alpha}\right) = W_0^{1,2}(0, 1),$$

which is a self-adjoint operator in Schroedinger form with constant potential $V(y) \equiv \alpha/4$. For a general review, see for example [17, 21, 52]. It is clear that

$$\lambda_k\left(\mathscr{L}_{\mathrm{dir},\alpha x^2}\right) = \lambda_k\left(\Delta_{\mathrm{dir},\alpha}\right) = \lambda_k\left(\Delta_{\mathrm{dir}}\right) + \frac{\alpha}{4} = k^2\pi^2 + \frac{\alpha}{4} \quad \text{for every } k \geq 1,$$

where

$$\Delta_{\mathrm{dir}}[v](y) := -v''(y), \qquad \mathrm{dom}\left(\Delta_{\mathrm{dir}}\right) = W_0^{1,2}(0, 1).$$

For later reference, notice that

$$\lim_{\alpha \to 0} \lambda_k\left(\mathscr{L}_{\mathrm{dir},\alpha x^2}\right) = \lambda_k\left(\Delta_{\mathrm{dir}}\right) = k^2\pi^2 \quad \text{for every } k \geq 1, \quad (4.3)$$

namely, the diffusion coefficient $p(x) = \alpha x^2$ produces a constant shift of $\alpha/4$ to the eigenvalues of the unperturbed Laplacian operator with Dirichlet BCs, i.e., $\Delta_{\mathrm{dir}}$.

We introduce the following definition.

**Definition 4.1** *(Numerical and analytic spectral relative errors)* Let $\mathscr{L}_{\mathrm{dir},\alpha x^2}^{(n)}$ be the discrete differential operator obtained from (4.1) by means of a generic numerical discretization method. If

$$\left\{(d_n + 1)^{-2} \mathscr{L}_{\mathrm{dir},\alpha x^2}^{(n)}\right\}_n \sim_\lambda \omega_\alpha(x, \theta), \qquad (x, \theta) \in [1, \mathrm{e}^{\sqrt{\alpha}}] \times [-\pi, \pi],$$

then fix $n, n', r \in \mathbb{N}$, with $n' \gg n$ and $r = r(n) \geq n$ and compute the following quantities

$$\mathbf{err}_{\alpha,k}^{(n)} = \left| \frac{\lambda_k\left(\mathcal{L}_{\mathrm{dir},\alpha x^2}^{(n)}\right)}{\lambda_k\left(\mathcal{L}_{\mathrm{dir},\alpha x^2}^{(n')}\right)} - 1 \right|, \qquad \mathbf{err}_{\alpha,r,k}^{*,(n)} = \left| \frac{(d_n+1)^2\omega_{\alpha,r,k}^{*,(n)}}{\lambda_k\left(\mathcal{L}_{\mathrm{dir},\alpha x^2}^{(n')}\right)} - 1 \right|$$

for $k = 1, \cdots, d_n$. Specifically, $\omega_{\alpha,r,k}^{*,(n)} = \omega_{\alpha,r}^*\left(\frac{k}{d_n+1}\right)$, where $\omega_{\alpha,r}^*$ is the (approximated) monotone rearrangement of the spectral symbol $\omega_\alpha$ obtained by the procedure described in Algorithm 1. We call $\mathbf{err}_{\alpha,k}^{(n)}$ the *numerical spectral relative error* and $\mathbf{err}_{\alpha,r,k}^{*,(n)}$ the *analytic spectral relative error*. The difference between these definitions and Definition 3.1 is that in this case we are using as a comparing sequence the eigenvalues of the same discrete operator on a finer mesh, since supposedly we do not know the exact eigenvalues of the continuous operator but we do know that $\lambda_k\left(\mathcal{L}_{\mathrm{dir},\alpha x^2}^{(n')}\right)$ converges to the exact eigenvalue $\lambda_k\left(\mathcal{L}_{\mathrm{dir},\alpha x^2}\right)$ as $n' \to \infty$. We say that $\omega_\alpha$ *spectrally approximates* the discrete differential operator $\mathcal{L}_{\mathrm{dir},\alpha x^2}^{(n)}$ if

$$\limsup_{n\to\infty} \left(\mathbf{err}_{\alpha,r,k}^{*,(n)}\right) = 0 \qquad \text{for every fixed } k.$$

### 4.1.1 Approximation by 3-points central FD method on uniform grid

In our example, if we apply the standard central 3-points FD scheme, see [43], then the sequence of the normalized discretization matrices $\hat{\mathcal{L}}_{\mathrm{dir},\alpha x^2}^{(n)} = (n+1)^{-2}\mathcal{L}_{\mathrm{dir},\alpha x^2}^{(n)}$ of the operator (4.1) has spectral symbol

$$\omega_\alpha(x,\theta) = \frac{\alpha x^2}{\left(\mathrm{e}^{\sqrt{\alpha}} - 1\right)^2}4\sin^2\left(\frac{\theta}{2}\right), \qquad \text{with} \quad D = [1, \mathrm{e}^{\sqrt{\alpha}}] \times [0, \pi],$$

see [22, Theorem 10.5]. Working with this toy-model problem using the 3-points central FD scheme provides us a further advantage, since we can analytically compute the monotone rearrangement $\omega_\alpha^*$, or at least a finer approximation than $\omega_{\alpha,r}^*$ which does not depend on the extra parameter $r$ and is less computationally expensive. Indeed, from Eq. (2.4) we have that

$$\phi_{\omega_\alpha} : \left[0, \frac{4\alpha\mathrm{e}^{2\sqrt{\alpha}}}{\left(\mathrm{e}^{\sqrt{\alpha}} - 1\right)^2}\right] \to [0, 1], \tag{4.4}$$

where

$$\phi_{\omega_\alpha}(t) = \frac{1}{\pi\left(\mathrm{e}^{\sqrt{\alpha}} - 1\right)}m\left(\left\{(x,\theta) \in [1, \mathrm{e}^{\sqrt{\alpha}}] \times [0, \pi] \ : \ \frac{4\alpha x^2\sin^2(\theta/2)}{\left(\mathrm{e}^{\sqrt{\alpha}} - 1\right)^2} \leq t\right\}\right).$$

After some analytical manipulation, $\phi_{\omega_\alpha}$ can be expressed explicitly

$$
\phi_{\omega_\alpha}(t) = \frac{1}{\pi\left(e^{\sqrt{\alpha}}-1\right)} \cdot
\begin{cases}
\Phi\left(t, e^{\sqrt{\alpha}}\right) - \Phi(t, 1) & \text{if } t \in \left[0, \frac{4\alpha}{\left(e^{\sqrt{\alpha}}-1\right)^2}\right], \\[2em]
\pi\left(\frac{\left(e^{\sqrt{\alpha}}-1\right)\sqrt{t}}{2\sqrt{\alpha}}-1\right) + \left[\Phi\left(t, e^{\sqrt{\alpha}}\right) - \Phi\left(t, \frac{\left(e^{\sqrt{\alpha}}-1\right)\sqrt{t}}{2\sqrt{\alpha}}\right)\right] & \text{if } t \in \left[\frac{4\alpha}{\left(e^{\sqrt{\alpha}}-1\right)^2}, \frac{4\alpha e^{2\sqrt{\alpha}}}{\left(e^{\sqrt{\alpha}}-1\right)^2}\right],
\end{cases}
$$

where

$$
\Phi(t, x) = \frac{\left(e^{\sqrt{\alpha}}-1\right)\sqrt{t}\, \log\left(2\alpha x\left(\sqrt{1 - \frac{\left(e^{\sqrt{\alpha}}-1\right)^2}{4\alpha x^2}}\right) + 1\right)}{\sqrt{\alpha}} + 2x \arcsin\left(\frac{\left(e^{\sqrt{\alpha}}-1\right)\sqrt{t}}{2\sqrt{\alpha}x}\right).
$$

Since we have an analytical expression for $\phi_{\omega_\alpha}(t)$, it is then possible to compute a numerical approximation of its generalized inverse $\omega_\alpha^*$ over the uniform grid $\left\{\frac{k}{n+1}\right\}_{k=1}^n$, for example by means of a Newton method. This approximation of the monotone rearrangement does not depend on the extra parameter $r$: therefore, when we will compute the analytical spectral relative error with respect to $\omega_\alpha^*$ we will write $\mathbf{err}_{\alpha,k}^{*,(n)}$ without the subscript $r$.

In Fig. 2 it is possible to check that an equispaced sampling of $(n+1)^2\omega_{\alpha,r}^*$ asymptotically distributes exactly as the eigenvalues of the unnormalized discrete operator $\mathscr{L}_{\mathrm{dir},\alpha x^2}^{(n)}$. Indeed, $\phi_{\omega_\alpha}$ is continuous and strictly monotone increasing which implies that $\omega_\alpha^*$ is continuous: then relation (3.5) applies. Moreover, according to Eq. (4.3), we observe that
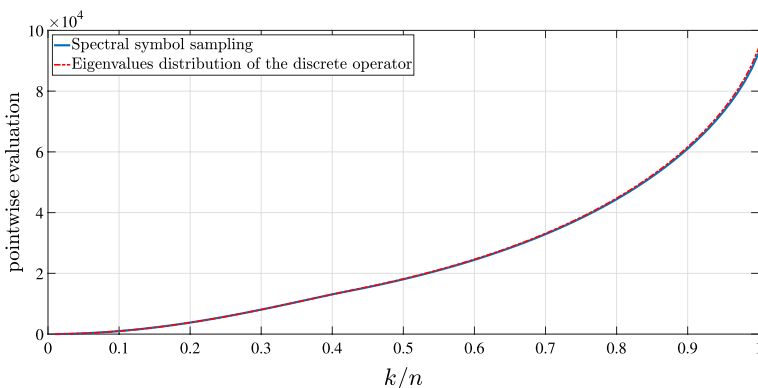


**Fig. 2** For $\alpha = 1$, comparison between the distribution of the first $n = 10^2$ eigenvalues of the discrete operator $\mathscr{L}_{\mathrm{dir},\alpha x^2}^{(n)}$ (red-dotted line) and the $n$-equispaced samples of $(n+1)^2\omega_{\alpha,r}^*$ with $r = 10^3$ (blue-continuous line). On the $x$-axis is reported the quotient $k/n$, for $k = 1, \ldots, n$. The sovrapposition of the graphs is explained by Theorem 3.1 and the limit (3.5)
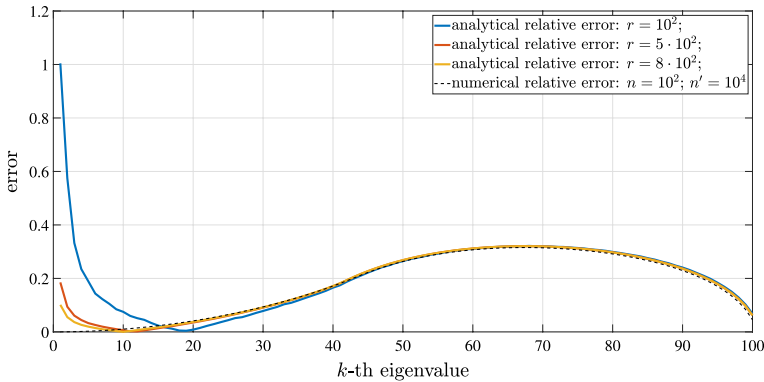
**Fig. 3** Comparison between the numerical spectral relative error $\mathbf{err}^{(n)}_{\alpha,k}$ and the analytic spectral relative error $\mathbf{err}^{*,(n)}_{\alpha,k,r}$ for increasing $r = 10^2, 5 \cdot 10^2, 8 \cdot 10^2$. The values of $n$ and $n'$ are fixed at $10^2$ and $10^4$, respectively, and $\alpha = 1$. The maximum discrepancy between the numerical relative error and the analytical relative errors is achieved for $k = 1$ in all the three cases, and apparently it decreases as $r$ increases

$$\lim_{\alpha \to 0} \phi_{\omega_\alpha}(t) = \frac{2}{\pi} \arcsin\left(\frac{\sqrt{t}}{2}\right) \qquad t \in [0, 4],$$

$$\lim_{\alpha \to 0} \omega^*_\alpha(x) = 4 \sin^2\left(\frac{\pi x}{2}\right) \qquad x \in [0, 1],$$

which means that the monotone rearrangement $\omega^*_\alpha$ converges to the spectral symbol $\omega$ as $\alpha \to 0$, that is, to the spectral symbol which characterizes the differential operator $\Delta_{\text{dir}}$ discretized by means of a 3-points FD scheme. The eigenvalues of $(n+1)^{-2}\left(\Delta^{(n)}_{\text{dir}}\right)$ are the exact sampling of $\omega(\theta) = 4\sin^2(\theta/2)$ over the uniform grid $\left\{\frac{k\pi}{n+1}\right\}^n_{k=1}$, see [43, p. 154]. This asymptotic behaviour reflects what we already observed in (4.3).

All these remarks would suggest that $\omega^*_\alpha$, or equivalently $\omega^*_{\alpha,r}$, spectrally approximates the normalized discrete operator $\hat{\mathscr{L}}^{(h)}_{\text{dir},\alpha x^2}$. Unfortunately, this conjecture looks to be partially proven wrong by Fig. 3: it shows the comparison between the graphs of the numerical spectral relative error $\mathbf{err}^{(n)}_{\alpha,k}$ and the analytic spectral relative error $\mathbf{err}^{*,(n)}_{\alpha,r,k}$, for several different increasing values of the parameter $r$. We observe a discrepancy in the analytical prediction of the eigenvalue error $\mathbf{err}^{*,(n)}_{\alpha,r,k}$, for small $k \ll n$, with respect to the numerical relative error $\mathbf{err}^{(n)}_{\alpha,k}$.

In particular, the maximum discrepancy is achieved at $k = 1$, for every $r$. The discrepancy apparently decreases as the number of grid points $r$ increases, as well observed in [24, Figure 48] for some test-problems in the setting of Galerkin discretization by linear $C^0$ B-splines. In that same paper, some plausible hypotheses and suggestions were advanced:

**Table 1** For every fixed $k$ and $\alpha$, the analytic relative error $\mathbf{err}_{\alpha,k}^{*,(n)}$ converges to the lower bound $c_{\alpha,k}$ as $n$ increases, where $c_{\alpha,k}$ is given in (4.5)

| | | $|\mathbf{err}_{\alpha,k}^{*,(n)}/c_{\alpha,k} - 1|$ | | |
| | | $n = 10^2$ | $n = 10^3$ | $n = 10^4$ |
|---|---|---|---|---|
| $\alpha = 0.1$ | $c_{\alpha,1} = 0.0025$ | 0.0326 | 3.3223e−04 | 3.3283e−06 |
| | $c_{\alpha,5} = 1.0131e - 04$ | 20.3811 | 0.2076 | 0.0021 |
| | $c_{\alpha,10} = 2.5530e - 05$ | 325.3811 | 3.3222 | 0.0333 |
| $\alpha = 1$ | $c_{\alpha,1} = 0.0247$ | 0.0041 | 4.1363e−05 | 4.1438e−07 |
| | $c_{\alpha,5} = 0.0010$ | 2.5395 | 0.0259 | 2.5899e−04 |
| | $c_{\alpha,10} = 2.5324e - 04$ | 40.6422 | 0.4136 | 0.0041 |
| $\alpha = 2$ | $c_{\alpha,1} = 0.0482$ | 0.0026 | 2.6120e−05 | 2.6167e−07 |
| | $c_{\alpha,5} = 0.0020$ | 1.6056 | 0.0163 | 1.6354e−04 |
| | $c_{\alpha,10} = 5.0635e - 04$ | 25.7979 | 0.2612 | 0.0026 |
| $\alpha = 5$ | $c_{\alpha,1} = 0.1124$ | 0.0020 | 2.0008e−05 | 2.0044e−07 |
| | $c_{\alpha,5} = 0.0050$ | 1.2389 | 0.0125 | 1.2528e−04 |
| | $c_{\alpha,10} = 0.0013$ | 20.4017 | 0.2002 | 1.2528e−04 |

Observe that $\mathbf{err}_{\alpha,k}^{*,(n)}$ seems to be monotone decreasing of order $O(n^{-2})$. The approximation of $\omega_\alpha^*$ is obtained evaluating $\phi_\alpha^{-1}$ from (4.4) by means of the `fzero()` function from MATLAB r2018b

- The discrepancy could depend on the fact that it was used $\omega_{\alpha,r}^*$ instead of $\omega_\alpha^*$, and then that discrepancy should tend to zero in the limit $r \to \infty$, since $\omega_{\alpha,r}^* \to \omega_\alpha^*$.
- Numerical instability of the analytic relative error $\mathbf{err}_{\alpha,r,k}^{*,(n)}$ for small eigenvalues, [24, Remark 3.1].
- Change the sampling grid into an "almost uniform" grid: see [24, Remark 3.2] for details.

The problem is that these hypotheses, which stem from numerical observations, cannot be validated: the descent to zero of the observed discrepancy as $r$ increases is only apparent. Indeed, what happens is that, for every fixed $k$ it holds

$$\left| \mathbf{err}_{\alpha,r,k}^{*,(n)} \right| \to c_{\alpha,k} = \frac{\frac{\alpha}{4}}{k^2 \pi^2 + \frac{\alpha}{4}} > 0 \quad \text{as } r, n \to \infty, \tag{4.5}$$

with $c_{\alpha,k}$ independent of $n$ and $n'$. This is the content of Proposition B.1 and Remark 2 in Sect. B, with $p(x) = \alpha x^2$, $w(x) \equiv 1$, $q(x) \equiv 0$. We then have a lower bound for the analytic spectral relative error which can not be avoided by refining the grid points. Of course, as $n \to \infty$, then $c_{\alpha,k} \to 0$ as $k$ increases. Those remarks are summarized in Table 1.

The problem lies on the wrong informal interpretation given to the limit relation in Definition 2.1, and suggested by Remark 1. Indeed, the asymptotic equality (2.1) tells us that

$$\left(\frac{k(n)}{n}, \lambda_{k(n)}\left(\hat{\mathscr{L}}^{(n)}_{\mathrm{dir},\alpha x^2}\right)\right) \to \left(x, \omega^*_\alpha(x)\right) \qquad \text{as } n \to \infty$$

for every $k(n)$ such that $k(n)/n \to x \in [0,1]$, see Theorem 3.1. Therefore, since $\omega^*_\alpha$ is continuous and $R_\omega$ is bounded, it follows that

$$\left\| \omega^{*,(n)}_{\alpha,r,k} - \lambda_k\left(\hat{\mathscr{L}}^{(n)}_{\mathrm{dir},\alpha x^2}\right) \right\|_\infty \to 0 \qquad \text{as } n \to \infty,$$

by Corollary 3.3 and as observed for example in [22, Example 10.2 p. 198]. On the contrary, a uniform sampling of the symbol $\omega$ does not necessarily provide an accurate approximation of the eigenvalues of the operator $\hat{\mathscr{L}}^{(n)}_{\mathrm{dir},\alpha x^2}$, in the sense of the relative error. The uniform sampling of the symbol works perfectly only for specific subclasses of discretization schemes and operators, but it fails in general.

As a last remark, there does not exist an "almost" uniform grid as well, nor in an asymptotic sense as described in [24, Remark 3.2]. Knowing the exact sampling grid which guarantees $\omega^*_\alpha$ to spectrally approximate the discrete differential operator is equivalent to know the eigenvalue distribution of the original differential operator.

What we can do instead is to apply Theorem 3.3. With reference to (3.8), it is easy to prove that

**Table 2** In this table we check numerically the validity of Theorem 3.3 for different values of $\alpha$ and $n$

| | | $n = 10^2$ | $n = 10^3$ | $n = 5 \cdot 10^3$ |
|---|---|---|---|---|
| $\alpha = 0.5$ | $\left|\frac{\max |\lambda_k^{(n)}/\lambda_k|}{\max |\omega_\alpha^*(x)/x^2\pi^2|} - 1\right|$ | 0.0104 | 0.0010 | 2.0853e-04 |
| | $k_0/n$ | 0.7900 | 0.7880 | 0.7878 |
| $\alpha = 1$ | $\left|\frac{\max |\lambda_k^{(n)}/\lambda_k|}{\max |\omega_\alpha^*(x)/x^2\pi^2|} - 1\right|$ | 0.0158 | 0.0016 | 3.1754e-04 |
| | $k_0/n$ | 0.6700 | 0.6680 | 0.6676 |
| $\alpha = 1.2$ | $\left|\frac{\max |\lambda_k^{(n)}/\lambda_k|}{\max |\omega_\alpha^*(x)/x^2\pi^2|} - 1\right|$ | 0.0180 | 0.0018 | 3.6226e-04 |
| | $k_0/n$ | 0.64 | 0.6310 | 0.6302 |
| $\alpha = 3$ | $\left|\frac{\max |\lambda_k^{(n)}/\lambda_k|}{\max |\omega_\alpha^*(x)/x^2\pi^2|} - 1\right|$ | 0.0518 | 0.0097 | 0.0032 |
| | $k_0/n$ | 1 | 1 | 1 |

It can be seen that for every $\alpha$, as $n$ increases then the relative error between $\max_{k=1,\ldots,n}\left|\lambda_k\left(\mathscr{L}^{(n)}_{\mathrm{dir},\alpha x^2}\right)/\lambda_k\left(\mathscr{L}_{\mathrm{dir},\alpha x^2}\right)\right|$ and $\max_{x \in [0,1]}\left|\omega_\alpha^*(x)/x^2\pi^2\right|$ decreases, validating (4.7). In the table it is reported as well the ratio $k_0/n$, where $k_0$ is the $k_0$-th eigenvalue which achieves the maximum relative error between $\lambda_k\left(\mathscr{L}^{(n)}_{\mathrm{dir},\alpha x^2}\right)$ and $\lambda_k\left(\mathscr{L}_{\mathrm{dir},\alpha x^2}\right)$. We can notice that $k_0/n$ tends to a fixed value in $(0, 1]$ as $n$ increases. The approximation of $\omega_\alpha^*$ is obtained evaluating $\phi_\alpha^{-1}$ from (4.4) by means of the `fzero()` function from MATLAB r2018b
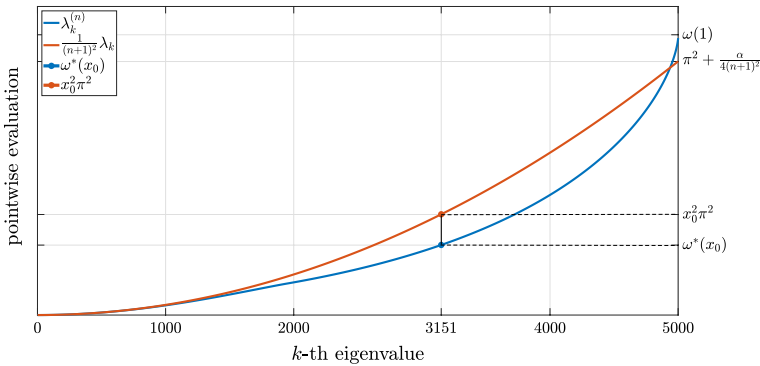
**Fig. 4** For $\alpha = 1.2$ and $n = 5 \cdot 10^3$, comparison between the eigenvalues distribution of the normalized discrete differential operator $\hat{\mathscr{L}}^{(n)}_{\text{dir},\alpha x^2}$ and the exact eigenvalues of the differential operator $\mathscr{L}_{\text{dir},\alpha x^2}$, normalized by $(n+1)^2$. The maximum spectral relative error $\mathscr{E}$ is obtained for $k_0 \approx 3151$, which corresponds to the maximum of $|\omega_\alpha^*(x)/x^2\pi^2 - 1|$, achieved at $\bar{x} \approx 0.6301 \approx k_0/n$. See Table 2

$$
\zeta_n(t) = \frac{\left| \left\{ k = 1, \ldots, n \; : \; \frac{\lambda_k(\mathscr{L}_{\text{dir},\alpha x^2})}{n^2} \leq t \right\} \right|}{n} \to \zeta(t) = \frac{\sqrt{t}}{\pi},
$$

that is,

$$
\zeta^*(x) = \pi^2 x^2. \tag{4.6}
$$

For every fixed $k$ it holds that $\lambda_k\left(\hat{\mathscr{L}}^{(n)}_{\text{dir},\alpha x^2}\right)$ converges to $\lambda_k\left(\hat{\mathscr{L}}_{\text{dir},\alpha x^2}\right)$, see [43], and by the continuity of $\phi_{\omega_\alpha}$ and by Eq. (4.6), then items (i)–(iv) of Theorem 3.3 are satisfied. Finally, since it holds that $\hat{\mathscr{L}}^{(n)}_{\text{dir},\alpha\tau(x)^2}$ does not have outliers by [40, Theorem 2.2], we conclude that

$$
\mathscr{E} = \max_{x \in [0,1]} \left| \frac{\omega_\alpha^*(x)}{x^2\pi^2} - 1 \right| > 0. \tag{4.7}
$$

In Fig. 4 and Table 2 it is numerically checked the validity of (4.7).

### 4.1.2 Discretization by $(2\eta + 1)$-points central FD method on non-uniform grid

Clearly, everything said in the preceding Sect. 4.1.1 remains valid even if we increase the order of accuracy of the FD method. Within this subsection we will make use of the FD scheme presented in Appendix C. The theoretical analysis of the scheme requires a proper standalone treatment, therefore we leave it to an upcoming future work and some of its properties are listed as conjectures in Conjecture C.1. Nevertheless, we introduce it now since it is helpful to the flow of our discussion and to provide (numerical) evidence to our conclusions.

First of all, we can say that the spectral symbol $\omega_{\alpha,\eta}$ in (C.2) does not spectrally approximate the discrete differential operator $\mathscr{L}_{\text{dir},\alpha x^2}^{(n,\eta)}$, in the sense of the relative error, for any $\eta \geq 1$, as we have already seen in the previous subsection for $\eta = 1$.

What is interesting instead is to change the sampling grid and to increase the order of accuracy $\eta$ of the FD discretization method. Indeed, as it was observed in (4.7), it is not possible to achieve $\mathscr{E} = 0$ if $\omega_{\alpha,\eta}^*(x) \neq x^2\pi^2$. From (C.2), for every $\eta \geq 1$ it is easy to check that $\max_{(x,\theta)\in[a,b]\times[0,\pi]} \omega_{\alpha,\eta}(x,\theta) = \omega_{\alpha,\eta}^*(1) \neq \pi^2$, and so we do not have any improvement by just increasing the order of accuracy $\eta$. On the other hand, observe that if we fix a new sampling grid $\{\bar{x}_j\}_{j=1}^n = \{\tau(x_j)\}_{j=1}^n$, with $\tau : [1, e^{\sqrt{\alpha}}] \to [1, e^{\sqrt{\alpha}}]$ a diffeomorphism, then

- from Corollary C.1 and (C.2),

$$\lim_{\eta\to\infty} \omega_{\alpha,\eta}(x,\theta) = \frac{\alpha\tau(x)^2}{(\tau'(x))^2\left(e^{\sqrt{\alpha}} - 1\right)^2}\theta^2 \qquad \text{for every } (x,\theta) \in [1, e^{\sqrt{\alpha}}] \times [0,\pi];$$

- defining $\tau_2(y) = e^{\sqrt{\alpha}y}$, then $\frac{\alpha\tau_2(y)^2}{(\tau_2'(y))^2} \equiv 1$.

In some sense, the spectral symbol $\omega_{\alpha,\eta}$ suggests us to change the uniform grid $\{x_j\}_{j=1}^n \subset [1, e^{\sqrt{\alpha}}]$ by means of the diffeomorphism induced by the Liouville transformation. Indeed, from (4.2) we have that

$$y(x) = \frac{\log(x)}{\sqrt{\alpha}} \quad \text{for } x \in [1, e^{\sqrt{\alpha}}], \qquad x(y) = e^{\sqrt{\alpha}y} \quad \text{for } y \in [0,1],$$

and therefore we can compose a $C^\infty$-diffeomorphism $\tau : [1, e^{\sqrt{\alpha}}] \to [1, e^{\sqrt{\alpha}}]$ such that

$$\tau : [1, e^{\sqrt{\alpha}}] \xrightarrow{\tau_1} [0,1] \xrightarrow{\tau_2} [1, e^{\sqrt{\alpha}}], \qquad \tau_1(x) = \frac{1}{e^{\sqrt{\alpha}} - 1}(x - 1), \qquad \tau_2(y) = e^{\sqrt{\alpha}y}.$$
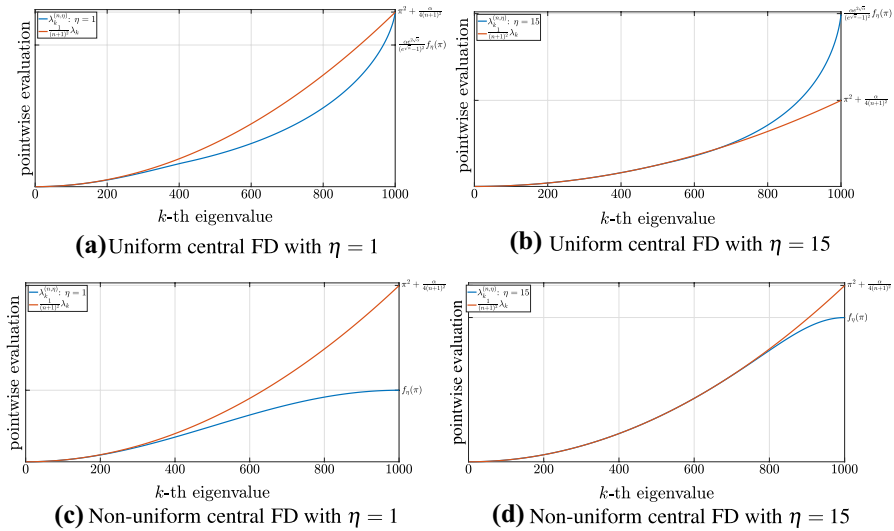
$$(4.8)$$

The new non-uniform grid is then given by

$$\{\bar{x}_j\}_{j=1}^n = \{\tau(x_j)\}_{j=1}^n, \qquad (4.9)$$

and it holds that

$$\lim_{\eta\to\infty} \omega_{\alpha,\eta}^*(x) = \pi^2 x^2 \qquad x \in [0,1],$$

that is exactly (4.6). Since the discretization scheme is convergent, for every fixed $k$ the local spectral relative error $\delta_k^{(n)}$ (see Definition 3.1) converges to zero as $n \to \infty$, and by [40, Theorem 2.2] it would hold that $\lambda_k\left(\mathscr{L}_{\text{dir},\alpha\tau(x)^2}^{(n,\eta)}\right) \in R_{\omega_{\alpha,\eta}}$ for every $k, \ldots, n$, and for every fixed $\eta \geq 1$. From these remarks, if the validity of Conjecture C.1 would be proved right, then we would be able to apply Theorem 3.3 and it would follow that

**(a)** Uniform central FD with $\eta = 1$

**(b)** Uniform central FD with $\eta = 15$

**(c)** Non-uniform central FD with $\eta = 1$

**(d)** Non-uniform central FD with $\eta = 15$

**Fig. 5** Graphic comparison between the eigenvalues distribution of the normalized discrete differential operators $\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir},\alpha\tau(x)^2}$ obtained by means of $(2\eta + 1)$-points central FD discretization on uniform ($\tau(x) = x$) and non-uniform grids ($\tau$ as in (4.8)). The parameters $\alpha$ and $n$ are fixed, with $\alpha = 1$ and $n = 10^3$, while $\eta$ changes. Let us observe that in Fig. 5c,d, i.e., in the case of central FD discretization on the non-uniform grid given by (4.9), the graph of the eigenvalue distribution seems to converge uniformly to the graph of the exact eigenvalues $(n + 1)^{-2}\lambda_k$, as $\eta$ increases. The same phenomenon does not happen in the case of central FD discretization on uniform grid, as it is clear from Fig. 5a,b. See Table 3 for a numerical comparison of the maximum spectral relative errors

**Table 3** Comparison between the maximum spectral relative errors of the discrete differential operators $\mathscr{L}^{(n,\eta)}_{\mathrm{dir},\alpha\tau(x)^2}$ obtained by means of $(2\eta + 1)$-points central FD discretization on uniform ($\tau(x) = x$) and non-uniform grids ($\tau$ as in (4.8)), for increasing values of $n$ and $\eta$

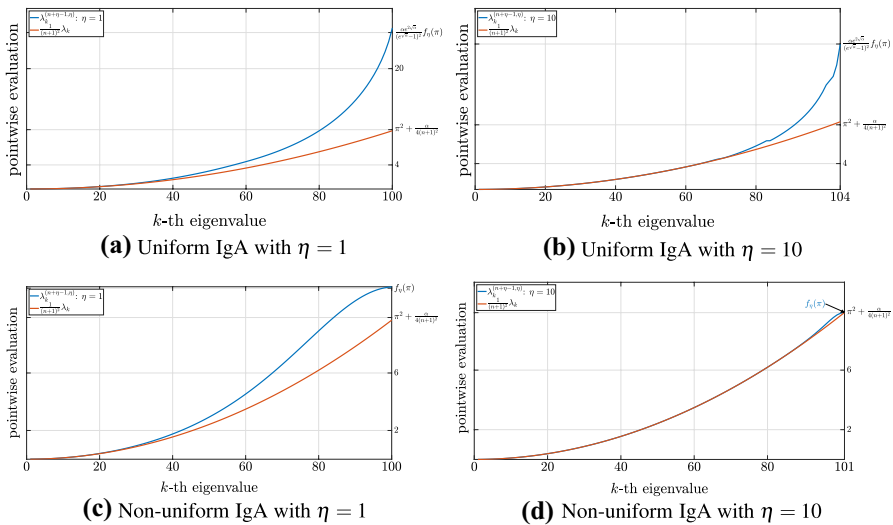|  | $\mathscr{E}_n = \mathscr{E}_n(\eta)$ | | |
| --- | --- | --- | --- |
|  | $\eta = 1$ | $\eta = 10$ | $\eta = 15$ |
|  | $n = 10^2$ | $n = 10^3$ | $n = 15 \cdot 10^2$ |
| Uniform grid | 0.3155 | 0.9057 | 1.0101 |
| Non-uniform grid | 0.5867 | 0.2210 | 0.1819 |

The parameter $\alpha$ is fixed, with $\alpha = 1$. We observe that in the uniform grid case, as $n$ and $\eta$ increase the maximum increases as well. On the contrary, in the non-uniform grid case given by (4.9), the maximum decreases as both $n$ and $\eta$ increase

$$\mathscr{E} = \mathscr{E}(\eta) = \left| \frac{\omega^*_{\alpha,\eta}(x)}{\pi^2 x^2} - 1 \right| \to 0 \qquad \text{as } \eta \to \infty.$$

**Table 4** Comparison between the maximum spectral relative errors of the discrete differential operators $\mathscr{L}_{\mathrm{dir},\alpha\tau(x)^2}^{(n,\eta)}$ obtained by means of IgA discretization of order $\eta$ on uniform ($\tau(x) = x$) and non-uniform grids ($\tau$ as in (4.8)), for increasing values of $\eta$ and $n$

| | $\mathscr{E}_n = \mathscr{E}_n(\eta)$ | | |
| --- | --- | --- | --- |
| | $\eta = 1$ | $\eta = 5$ | $\eta = 10$ |
| | $n = 10^2$ | $n = 5 \cdot 10^2$ | $n = 10^3$ |
| Uniform grid | 1.7653 | 1.1971 | 1.2005 |
| Non-uniform grid | 0.4433 | 0.0513 | 0.0268 |

The parameter $\alpha$ is fixed, with $\alpha = 1$. We observe that in the non-uniform grid case given by (4.9), the maximum spectral relative error decreases as $n$ and $\eta$ increase. Let us notice that we did not take in consideration the outliers, see Definition 2.3



**(a)** Uniform IgA with $\eta = 1$

**(b)** Uniform IgA with $\eta = 10$

**(c)** Non-uniform IgA with $\eta = 1$

**(d)** Non-uniform IgA with $\eta = 10$

**Fig. 6** Graphic comparison between the eigenvalues distribution of the normalized discrete differential operators $\mathscr{L}_{\mathrm{dir},\alpha\tau(x)^2}^{(n,\eta)}$ obtained by means of IgA discretization of order $\eta$ on uniform ($\tau(x) = x$) and non-uniform grids ($\tau$ as in (4.8)). The parameters $\alpha$ and $n$ are fixed, with $\alpha = 1$ and $n = 10^2$, while $\eta$ changes. Let us observe that in Fig. 6c, d, i.e., in the case of IgA discretization on the non-uniform grid given by (4.9), the graph of the eigenvalue distribution seems to converge uniformly to the graph of the exact eigenvalues $(n + 1)^{-2}\lambda_k$, as $\eta$ increases. The same phenomenon does not happen in the case of IgA discretization on uniform grid, as it is clear from Fig. 6a, b. Let us notice that we did not take in consideration the outliers, see Definition 2.3

See Fig. 5 and Table 3.

### 4.1.3  IgA discretization by B-spline of degree $\eta$ and smoothness $C^{\eta-1}$

In this subsection we continue our analysis in the IgA framework, see Appendix D where we present some known results that are useful for our analysis. Here we just collect all the numerical results of the tests, which confirm again what observed in Sects. 4.1.1 and 4.1.2. The only difference relies on the fact that we took out the largest eigenvalues of the discrete operator $\mathscr{L}^{(n,\eta)}_{\mathrm{dir},\alpha\tau(x)^2}$. This is due to the fact that the IgA discretization suffers of a fixed number of outliers which depends on the degree $\eta$ and it is independent of $n$, see [15, Chapter 5.1.2 p. 153]. So, we consider only the eigenvalues $\lambda_k\left(\mathscr{L}^{(n,\eta)}_{\mathrm{dir},\alpha\tau(x)^2}\right)$ which belong to $R_{\omega_{\alpha,\eta}}$.

In Fig. 6 we compare the graphs of the eigenvalue distributions between the discrete eigenvalues $\lambda_k\left(\mathscr{L}^{(n,\eta)}_{\mathrm{dir},\alpha\tau(x)^2}\right)$ and the exact eigenvalues $\lambda_k\left(\mathscr{L}_{\mathrm{dir},\alpha x^2}\right)$, for different values of $\eta$ on uniform ($\tau(x) = x$) and non-uniform grids ($\tau$ as in (4.8)). They line up with the numerics of Table 4: if the sampling grid is given by (4.9), the maximum spectral relative error $\mathscr{E}$ decreases as the order of approximation increases.

### 4.2  *d*-dimensional Dirichlet Laplacian

The results of Sect. 3 can be generalized to study the maximum spectral relative error between more general linear self-adjoint differential operators and the numerical scheme implemented for their discretization. Let us sketch the ideas with a plain example.

Let $\Delta : W^{1,2}_0\left([0,1]^d\right) \to L^2\left([0,1]^d\right)$ be the second order elliptic operator defined by the formal equation

$$\Delta[u](\boldsymbol{x}) := -\sum_{j=1}^{d} \frac{\partial^2 u}{\partial_{x_j^2}}(\boldsymbol{x}), \qquad [0,1]^d \subset \mathbb{R}^d.$$

It is well-known that it has empty essential spectrum and that

$$\lambda_k(\Delta) = \lambda_{k_1\cdots k_d}(\Delta) = \sum_{j=1}^{d} \pi^2 k_j^2 \quad \text{for } k_1,\ldots,k_d \geq 1,$$

$$\zeta(t) = \lim_{n\to\infty} \frac{\left| k = 1,\ldots,n^d \ : \ \frac{\lambda_k(\Delta)}{(n+1)^2} \leq t \right|}{n^d} = c_d\left(\frac{t}{4\pi^2}\right)^{\frac{d}{2}}, \quad t \in [0, 4\pi^2/(c_d)^{2/d}],$$

where $c_d$ is the volume of the unit ball in $\mathbb{R}^d$, see [32]. Define

$$Y^{(n)} := (n+1)^{-2} \operatorname*{diag}_{k=1,\ldots,n^d} \left\{ \lambda_k(\Delta) \right\}.$$

Then it holds that $\left\{ Y^{(n)} \right\}_n \sim_\lambda \zeta^*(x) = 4\pi^2\left(\frac{x}{c_d}\right)^{2/d}$, $x \in [0,1]$.

On the other hand, a discretization of the operator $\Delta$ by means of separation of variables and the classic equispaced $(2d+1)$-points FD scheme, leads to

$$\Delta^{(n)} = (n+1)^2 \sum_{k=1}^{d} \underbrace{I^{(n)} \otimes \cdots \otimes I^{(n)}}_{k-1} \otimes T^{(n)}\big(2 - 2\cos(\theta_k)\big) \otimes \underbrace{I^{(n)} \otimes \cdots \otimes I^{(n)}}_{d-k},$$

for $\theta_k \in [0, \pi]$, where $I^{(n)}$ is the identity matrix and $\otimes$ is the Kronecker product, and all the matrices are of dimension $d_n = n^d$. It holds that

$$\left\{(n+1)^{-2}\Delta^{(n)}\right\}_n = \left\{\hat{\Delta}^{(n)}\right\}_n \sim_\lambda \omega(\boldsymbol{\theta}) = \sum_{k=1}^{d}(2 - 2\cos(\theta_k))$$

where

$$\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d) \in [0, \pi]^d,$$

see [23, Chapter 7.3]. Since the discretization scheme is convergent, for every fixed $k$ the local spectral relative error $\delta_k^{(n)}$ (see Definition 3.1) converges to zero as $n \to \infty$, and by [40, Theorem 2.2] it holds that $\lambda_k\big(\Delta^{(n)}\big) \in R_\omega$ for every $k$, $n$. Moreover, it is immediate to check that both $\omega^*, \zeta^*$ are continuous, therefore we can apply Theorem 3.3 to get the maximum spectral relative error:

$$\mathcal{E} = \sup_{x \in (0,1]} \left| \frac{\omega^*(x)}{\zeta^*(x)} - 1 \right| \geq \left| \frac{\omega^*(1)}{\zeta^*(1)} - 1 \right| = \left| \frac{4d}{\frac{4\pi^2}{(c_d)^{2/d}}} - 1 \right| > 0.$$

## 5 Conclusions

Given a differential operator $\mathcal{L}$ discretized by means of a numerical scheme, the knowledge of the spectral symbol $\omega$ provides a way to measure how far the spectrum of the discretized operator is from a uniform approximation of all the spectrum of the original differential operator $\mathcal{L}$. Moreover, the symbol itself gives insights about how to build a discretization scheme such that the discrete operator could benefit of a spectrum that converges uniformly to the spectrum of the continuous operator, i.e., such that $\mathcal{E} = 0$, see (1.1). This is crucial in some engineering applications and our results improve those in the recent review paper by T. J. R. Hughes and coauthors.

In light of the proposed analysis, it becomes a priority to devise new specific discretization schemes with mesh-dependent order of approximation which guarantee a good balance between convergence to zero of the maximum spectral relative error and computational costs.

Finally, other applications of the spectral symbol related to the computation of the maximum spectral relative error have been recently addressed in the graphs setting, for PDEs approximation, see [1].

## Proofs of Section 3

In this section we provide the rigorous proofs of the results that we presented in Sect. 3. Before going further in the details, we need some preliminary definitions and notations from measure theory, see [29].

Given a probability space $(S, \mathscr{S}, v)$ and a measurable function $g : S \to \mathbb{R}$, the *distribution function* $\phi_g : \mathbb{R} \to [0, 1]$ associated to $g$ is defined by

$$\phi_g(t) := v(\{s \in S \ : \ g(s) \le t\}).$$

It holds that $\phi_g$ is right-continuous, proper and with nonnegative increments. Moreover, any function $\phi : \mathbb{R} \to [0, 1]$ satisfying those stated properties determines a probability measure $\mu$ on $\mathbb{R}$ such that

$$\mu((-\infty, t]) = \phi(t),$$

and vice-versa, see [29, Theorem 4.25]. In this case, we call $\phi$ the *distribution* associated to $\mu$.

**Definition A.1** *(Vague convergence)* A sequence $\{\mu_n\}_n$ of locally-finite measures on $\mathbb{R}$ is said to *converge vaguely* to a measure $\mu$ if and only if

$$\lim_{n \to \infty} \int F(t)\mu_n(dt) = \int F(t)\mu(dt), \quad \forall F \in C_c(\mathbb{R}). \tag{A.1}$$

In this case, we write $\mu_n \xrightarrow{v} \mu$.

If $\{\mu_n\}_n$ and $\mu$ are probability measures, that is, if $\mu_n(\mathbb{R}) = \mu(\mathbb{R}) = 1$, and $\{\phi_n\}_n$ and $\phi$ are their distributions, respectively, then

$$\mu_n \xrightarrow{v} \mu \quad \text{if and only if} \quad \phi_n(t) \to \phi(t) \tag{A.2}$$

for every $t \in \mathbb{R}$ which is a point of continuity of $\phi$, see [14, Theorem 4.4.1].

Definition A.1 and Definition 2.1 are connected. Let us set the atomic measure $\delta_x : \mathscr{B} \to \{0, 1\}$ from the Borel sets $\mathscr{B}$ of $\mathbb{R}$ such that

$$\delta_x(B) := \begin{cases} 1 & \text{if } x \in B, \\ 0 & \text{if } x \notin B, \end{cases}$$

and define

$$\mu_n(\cdot) := \frac{1}{d_n} \sum_{\substack{k = -d_n^- \\ k \ne 0}}^{d_n^+} \delta_{\lambda_k(X^{(n)})}(\cdot). \tag{A.3}$$

Clearly, $\mu_n(\cdot)$ is a probability measure and for every measurable function $F : \mathbb{R} \to \mathbb{R}$ it holds that

$$\int F(t)\mu_n(dt) = \frac{1}{d_n} \sum_{\substack{k=-d_n^- \\ k \neq 0}}^{d_n^+} F\big(\lambda_k\big(X^{(n)}\big)\big).$$

**Lemma A.1** *Let $\{X^{(n)}\}_n$ be a matrix sequence and $\omega : D \subset \mathbb{R}^M \to \mathbb{R}$ be a measurable function, $D$ a measurable set with $0 < m(D) < \infty$, and such that the Lebesgue integral $\int_D \omega(\mathbf{y})m(d\mathbf{y})$ exists, finite or infinite. Then $\{X^{(n)}\}_n \sim_\lambda \omega$ if and only if $\mu_n \xrightarrow{v} \mu$, where $\mu$ is the measure on $\mathbb{R}$ determined by the distribution function $\phi_\omega$ associated to $\omega$ and defined in (2.4), i.e., such that $\mu((-\infty, t]) = \phi_\omega(t)$.*

**Proof** It is an immediate consequence of (A.1) and (2.1), and the fact that

$$\int F(t)\mu(dt) = \frac{1}{m(D)} \int_D F(\omega(\mathbf{y}))m(d\mathbf{y})$$

for every measurable function $F$, see (2.4) and [29, Lemma 1.24]. □

The next lemma is a slight generalization of a well know result by Pólya, [36, Satz I] (see also [47, Exercise 6.4.1]).

**Lemma A.2** *Let $\phi_n, \phi : (a, b) \subseteq \mathbb{R} \to \mathbb{R}$, where $-\infty \leq a < b \leq +\infty$. Suppose that $\phi_n$ and $\phi$ are nonnegative, monotone nondecreasing and that*

(i) $\lim_{t \to a} \phi_n(t) = a_n$ and $\lim_{n \to \infty} a_n = \lim_{t \to a} \phi(t) = 0$;
(ii) $\lim_{t \to b} \phi_n(t) = b_n$ and $\lim_{n \to \infty} b_n = \lim_{t \to b} \phi(t) = c \in (0, +\infty)$;
(iii) $\lim_{n \to \infty} \phi_n(t) = \phi(t)$ for every $t$ in a dense subset $I$ of $[a, b]$.

*If $\phi$ is continuous then $\phi_n$ converges uniformly to $\phi$.*

**Proof** For the convenience of the reader we report here the proof. Fix $\epsilon > 0$. Clearly, there exists $\delta_1 = \delta_1(\epsilon) > 0$ such that $0 \leq \phi(t) < \epsilon$ and $c - \epsilon < \phi(t) \leq c$ for every $t \in (a, a + \delta_1)$ and $(b - \delta_1, b)$, respectively, with $a + \delta_1 < b - \delta_1$. Choose $t' \in (a, a + \delta_1) \cap I, t'' \in (b - \delta_1, b) \cap I$. For the continuity of $\phi$ and the compactness of $[t', t'']$, there exists $\delta_2 = \delta_2(\epsilon) > 0$ such that $0 \leq \phi(t_i) - \phi(t_j) < \epsilon$ for every $t_i, t_j$ in $[t', t'']$ such that $0 < t_i - t_j < \delta_2$. Choose $k = k(\epsilon)$ and a set of points $\{t_i\}_{i=0}^{k+1}$ such that $t_0 = t', t_{k+1} = t'', t_i \in (t', t'') \cap I$ for $i = 1, \ldots, k$ and

$$0 < t_{i+1} - t_i < \delta_2 \quad \text{for } i = 0, \ldots, k.$$

By hypothesis, $\lim_{n \to \infty} a_n = 0$, $\lim_{n \to \infty} b_n = c$ and $\lim_{n \to \infty} \phi_n(t_i) = \phi(t_i)$ for every $i = 0, \ldots, k + 1$, and therefore there exists $N = N(\epsilon)$ such that

$$0 \leq a_n < \epsilon, \qquad 0 \leq |b_n - c| < \epsilon, \qquad 0 \leq |\phi_n(t_i) - \phi(t_i)| < \epsilon$$

for every $n > N$ and for every $i = 0, \ldots, k + 1$. By monotonicity, for every $t \in (a, b)$ it holds that

$$
\phi_n(t) - \phi(t)
$$
$$
\leq \begin{cases} (\phi_n(t_0) - \phi(t_0)) + \phi(t_0) < 2\epsilon & \text{if } t \in (a, t_0), \\ \big(\phi_n(t_{i+1}) - \phi(t_{i+1})\big) - \big(\phi(t_{i+1}) - \phi(t_i)\big) < 2\epsilon & \text{if } t \in [t_i, t_{i+1}], \ i = 0, \ldots, k, \\ b_n - \phi(t_{k+1}) < c + \epsilon - \phi(t_{k+1}) < 2\epsilon & \text{if } t \in (t_{k+1}, b), \end{cases}
$$

for every $n > N$. By a symmetric argument it can be shown that $\phi(t) - \phi_n(t) < 2\epsilon$, and therefore

$$
\sup_{t \in [a,b]} |\phi_n(t) - \phi(t)| < 2\epsilon \qquad \forall\, n > N.
$$

$\square$

We are ready to provide the proofs of Theorem 3.1, Corollary 3.1 and Corollary 3.2.

**Theorem** 3.1.

*Proof* Let us observe that $\omega^*$ is monotone increasing, right continuous and the possibly non-empty set of its point of discontinuity (jumps) is at most countable. Moreover, the function $\phi_\omega$ defined in (2.4) is the distribution function associated to both $\omega$ and $\omega^*$, that is,

$$
\frac{1}{m(D)} m(\{y \in D \ : \ \omega(y) \leq t\}) = \phi_\omega(t) = \phi_{\omega^*}(t) = m(\{x \in (0, 1) \ : \ \omega^*(x) \leq t\}).
$$

The proof of the last statement follows immediately by adapting [47, Lemma 6.3.10] to our definition of nondecreasing monotone rearrangement. Therefore, if we indicate with $\mu$ the measure determined by $\phi_\omega$, by [29, Lemma 1.24] it holds that

$$
\frac{1}{m(D)} \iint_D F(\omega(y)) m(dy) = \int F(t) \mu(dt) = \int_0^1 F(\omega^*(x)) m(dx)
$$

for every measurable function $F$. The above identity plus (2.1) gives (3.2). Define now

$$
\phi_n(t) := \frac{N\big(X^{(n)}, t\big)}{d_n}.
$$

Clearly, $\phi_n(t) = \mu_n((-\infty, t])$ for every $t \in \mathbb{R}$, where $\mu_n$ is the probability measure defined in (A.3). By Lemma A.1 we have that $\mu_n \xrightarrow{v} \mu$, and by the equivalence (A.2) we get (3.3). Suppose now that $\phi_\omega$ is continuous. Without loss of generality and for the sake of simplicity, let now $d_n^-(X^{(n)}) = 0$, such that $\hat{k} = k$, and define $\lambda_{k(n)} := \lambda_{k(n)}(X^{(n)})$. Applying Lemma A.2 it holds that $\phi_n \to \phi_\omega$ uniformly. On the other hand, it holds that

$$x_0 = \lim_{n\to\infty} \frac{k(n)}{d_n} = \lim_{n\to\infty} \frac{N\big(X^{(n)}, \lambda_{k(n)}\big)}{d_n} = \lim_{n\to\infty} \phi_n\big(\lambda_{k(n)}\big).$$

Therefore, for every $\epsilon > 0$ there exist $N_1 = N_1(\epsilon), N_2 = N_2(\epsilon) \in \mathbb{N}$ such that

$$\sup_{t\in\mathbb{R}} \big|\phi_n(t) - \phi_\omega(t)\big| < \epsilon \quad \forall\, n > N_1,$$

$$\big|\phi_n\big(\lambda_{k(n)}\big) - x_0\big| < \epsilon \quad \forall\, n > N_2.$$

It follows easily that

$$\lim_{n\to\infty} \phi_\omega\big(\lambda_{k(n)}\big) = \phi_\omega\Big(\lim_{n\to\infty} \lambda_{k(n)}\Big) = x_0.$$

Since $x_0 \in (0,1)$, then by (2.4) it holds that

$$\lim_{n\to\infty} \lambda_{k(n)} = t_0 \in (\inf R_\omega, \sup R_\omega).$$

Finally, by the above relation and by (2.3), we can conclude that

$$\lim_{x\to x_0^-} \omega^*(x) = \lim_{n\to\infty} \omega^*\left(\frac{k(n)}{d_n}\right) = \sup_{t\in R_\omega} \{t \le t_0\},$$

which is (3.4). Let us observe now that $x_0$ is a jump discontinuity point for $\omega^*$ if and only if there exist $t_1 < t_2 \in R_\omega$ such that $R_\omega \subseteq (\inf R_\omega, t_1] \cup [t_2, \sup R_\omega)$, and $\phi_\omega(t) = x_0$ if and only if $t \in [t_1, t_2]$. Therefore, if $\omega^*$ is continuous in $x_0$, then $t_0 = t_1 = t_2 \in R_\omega$ and we have (3.5). $\qquad\square$

**Corollary** 3.1.

***Proof*** It is immediate from (3.3). Let us observe that

$$0 \le \lim_{n\to\infty} \frac{\big|\{k\ :\ \lambda_k\big(X^{(n)}\big) < \inf R_\omega\}\big|}{d_n} \le \lim_{n\to\infty} \frac{N(X^{(n)}, \inf R_\omega)}{d_n} = \phi_\omega\big(\inf R_\omega\big) = 0.$$

Moreover, since

$$\frac{\big|\{k\ :\ \lambda_k\big(X^{(n)}\big) \notin R_\omega\}\big|}{d_n} = \frac{\big|\{k\ :\ \lambda_k\big(X^{(n)}\big) \in \mathbb{R}\}\big|}{d_n} - \frac{\big|\{k\ :\ \lambda_k\big(X^{(n)}\big) \in R_\omega\}\big|}{d_n}$$

$$= 1 - \frac{\big|\{k\ :\ \lambda_k\big(X^{(n)}\big) \in R_\omega\}\big|}{d_n}$$

$$= 1 - \frac{N(X^{(n)}, \sup R_\omega)}{d_n} + \frac{\big|\{k\ :\ \lambda_k\big(X^{(n)}\big) < \inf R_\omega\}\big|}{d_n},$$

then passing to the limit we get

$$\lim_{n\to\infty} \frac{\left|\left\{k \ : \ \lambda_k(X^{(n)}) \notin R_\omega\right\}\right|}{d_n} = 1 - \phi_\omega(\sup R_\omega) = 0.$$

The second part of the thesis follows instead by (3.3) and the easy fact that

$$\left|\left\{k \ : \ \lambda_k(X^{(n)}) \leq t, \ \lambda_k(X^{(n)}) \in R_\omega\right\}\right| = \left|\left\{k \ : \ \lambda_k(X^{(n)}) \leq t\right\}\right|$$
$$- \left|\left\{k \ : \ \lambda_k(X^{(n)}) \leq t, \ \lambda_k(X^{(n)}) \notin R_\omega\right\}\right|,$$

where

$$\left|\left\{k \ : \ \lambda_k(X^{(n)}) \leq t, \ \lambda_k(X^{(n)}) \notin R_\omega\right\}\right| \leq \left|\left\{k \ : \ \lambda_k(X^{(n)}) \notin R_\omega\right\}\right| = o(d_n).$$

$\square$

**Corollary** 3.2.

**Proof** Since $\omega^*$ is absolutely continuous then it is differentiable almost everywhere. Let $x_0 \in [0,1]$ such that $(\omega^*)'|_{x=x_0}$ exists finite. It is clear that $\omega_n^*$ is monotone (piecewise constant) and right-continuous for any $n$, that $\omega_n^*(\hat{m}/d_n) = \lambda_{\hat{m}}(X^{(n)})$ for any $\hat{m} \in \{1,\dots,d_n\}$, and that $\omega_n^* \to \omega^*$ uniformly as $n \to \infty$, because of (3.5) and Lemma A.2. Therefore, by hypothesis we have that

$$\lim_{n\to\infty} \frac{g\left(\lambda_{\hat{k}(n)+\left\lceil \frac{d_n}{h(d_n)}\right\rceil}(X^{(n)})\right) - g\left(\lambda_{\hat{k}(n)}(X^{(n)})\right)}{\frac{1}{h(d_n)}} = \lim_{n\to\infty} \frac{g\left(\omega_n^*\left(\frac{\hat{k}(n)}{d_n} + \frac{1}{h(d_n)}\right)\right) - g\left(\omega_n^*\left(\frac{\hat{k}(n)}{d_n}\right)\right)}{\frac{1}{h(d_n)}}$$
$$= g'(\omega^*(x_0)) \lim_{n\to\infty} \frac{\omega_n^*\left(\frac{\hat{k}(n)}{d_n} + \frac{1}{h(d_n)}\right) - \omega_n^*\left(\frac{\hat{k}(n)}{d_n}\right)}{\frac{1}{h(d_n)}}$$
$$= g'(\omega^*(x_0)) \lim_{n\to\infty} \frac{\omega^*\left(\frac{\hat{k}(n)}{d_n} + \frac{1}{h(d_n)}\right) - \omega^*\left(\frac{\hat{k}(n)}{d_n}\right)}{\frac{1}{h(d_n)}}$$
$$= g'(\omega^*(x_0)) \lim_{n\to\infty} \omega'(\xi_n), \ \xi_n \in \left[\frac{\hat{k}(n)}{d_n}, \frac{\hat{k}(n)}{d_n} + \frac{1}{h(d_n)}\right]$$
$$= g'(\omega^*(x_0))(\omega^*)'(x_0).$$

Since $m\left(\left\{x \in [0,1] \ : \ \nexists(\omega^*)'(x) \text{ or } \omega^*(x) = 0\right\}\right) = 0$, we conclude.

**Corollary** 3.3.

**Proof** Let us observe that since there are no outliers, then eventually $\min(R_\omega) \leq \lambda_k(X^{(n)}) \leq \max(R_\omega)$ for every $k$. Without loss of generality and for the sake of simplicity, let $d_n^- = 0$. Suppose the thesis is false. Then there exists

a sequence $k = k(n)$ such that $|\lambda_{k(n)}(X^{(n)}) - \omega^*(k(n)/(d_n + 1))| > \epsilon$ for some $\epsilon > 0$. On the other hand, by the boundedness of $k(n)/(d_n + 1)$ and by Theorem 3.1, passing to a subsequence we can assume that $k(n)/(d_n + 1)$ converges to a point $x_0$ in [0, 1] and that (3.5) holds. If $\omega^*(k(n)/(d_n + 1)) = 0$ eventually, then $|\lambda_{k(n)}(X^{(n)}) - \omega^*(k(n)/(d_n + 1))| \to 0$, which is a contradiction. At the same time, if $\omega^*(k(n)/(d_n + 1))$ is not eventually identical to zero, then by passing again to a subsequence we can assume that $\omega^*(k(n)/(d_n + 1)) \neq 0$, and by the boundedness of $R_\omega$

$$\left| \lambda_{k(n)}(X^{(n)}) - \omega^*\left( \frac{k(n)}{d_n + 1} \right) \right| \leq c \left| \frac{\lambda_{k(n)}(X^{(n)})}{\omega^*\left( \frac{k(n)}{d_n+1} \right)} - 1 \right| \to 0,$$

which is again a contradiction. $\qquad\square$

Before introducing the proofs of Theorem 3.2 and Theorem 3.3, we need a couple of lemmas.

**Lemma A.3** *With the same notations of Theorem 3.2, if $\phi_{\omega_2}$ is continuous and $\omega_2^*(x_0) = 0$ then*

$$\frac{\hat{k}(n)}{d_n} \to x_0 \quad \text{if and only if} \quad \frac{k(n)}{d_n} \to 0.$$

**Proof** Notice that

$$\frac{d_n^-(Y^{(n)})}{d_n} = \mu_n((-\infty, 0))$$

where $\mu_n$ is the probability measure defined in (A.3). Since $\phi_{\omega_2}$ is continuous, applying [14, Theorem 4.3.2]

$$\lim_{n \to \infty} \frac{d_n^-(Y^{(n)})}{d_n} = \phi_{\omega_2}(0).$$

By definition, $\phi_{\omega_2}$ is the inverse of $\omega_2^*$ and therefore $\omega_2^*(x_0) = 0$ implies $\phi_{\omega_2}(0) = x_0$. From this result and (3.1) we get that

$$\lim_{n \to \infty} \frac{\hat{k}(n)}{d_n} = \lim_{n \to \infty} \frac{k(n)}{d_n} + \lim_{n \to \infty} \frac{d_n^-(Y^{(n)})}{d_n} + \lim_{n \to \infty} \frac{\mathbb{1}_{(-\infty,0)}(k(n))}{d_n}$$

$$= \lim_{n \to \infty} \frac{k(n)}{d_n} + \phi_{\omega_2}(0) = \lim_{n \to \infty} \frac{k(n)}{d_n} + x_0,$$

and we conclude. $\qquad\square$

**Lemma A.4** *With the same hypotheses and notations of Theorem 3.2, it holds that*

$$\mathscr{E} = \sup \left\{ \lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} \ : \ \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} \in F \right\}, \tag{A.4}$$

*or equivalently*

$$\mathscr{E} = \max \left\{ \sup \left\{ \lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} \ : \ \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} \in F \cap G_{x_0}^c \right\}; \right.$$

$$\sup \left\{ \lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} \ : \ \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} \in F \cap G_{x_0} \right\} \right\} \tag{A.5}$$

$$= \max \left\{ \sup \left\{ \lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} \ : \ \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} \in F \cap G_{x_0}^c \right\}; \sup H_{x_0} \right\}.$$

*In particular, if $\omega_2^*(x_0) = 0$, then we have that*

$$\sup H_{x_0} \geq \sup_{\substack{k \in [\kappa^-, \kappa^+] \\ k \neq 0}} \{\sigma_k\},$$

*and if*

$$\limsup_{n \to \infty} \left| \left\| \frac{\lambda_{\vartheta(k(n))}(X^{(n)})}{\lambda_{k(n)}(Y^{(n)})} - 1 \right\| - \sigma_{k(n)} \right| = 0 \quad \text{for every } k(n) \text{ such that } \lim_{n \to \infty} \frac{k(n)}{d_n} = 0,$$

*then*

$$\sup H_{x_0} = \sup_{\substack{k \in [\kappa^-, \kappa^+] \\ k \neq 0}} \{\sigma_k\}. \tag{A.6}$$

**Proof** The first part follows from the definition. Indeed,

$$\mathscr{E} = \limsup_n \mathscr{E}_n = \limsup_n \max_{\hat{k}=1,\dots,d_n} \left\{ \delta_{\hat{k}}^{(n)} \right\}$$

$$= \sup \left\{ \lim_{n_j \to \infty} \max_{\hat{k}=1,\dots,d_{n_j}} \left\{ \delta_{\hat{k}}^{(n_j)} \right\} \ : \ \left\{ \max_{\hat{k}=1,\dots,d_{n_j}} \left\{ \delta_{\hat{k}}^{(n_j)} \right\} \right\}_{n_j} \in F \right\}$$

$$= \sup(A).$$

If we call $B := \left\{ \lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} \ : \ \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} \in F \right\}$, then it is easy to check that $A \subset B$, and then $\sup(A) \leq \sup(B)$. On the other hand, for every element $b \in B$ there exists an increasing sequence of indices $\{n_{j_m}\} \subset \{n_j\}$ such that

$$b = \lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} \le \limsup_{n_j \to \infty} \max_{\hat{k}=1,\ldots,d_{n_j}} \left\{ \delta_{\hat{k}}^{(n_j)} \right\} = \lim_{n_{j_m} \to \infty} \max_{\hat{k}=1,\ldots,d_{n_{j_m}}} \left\{ \delta_{\hat{k}}^{(n_{j_m})} \right\}.$$

Therefore, for every $b \in B$ there exists $a \in A$ such that $b \le a$, and then it follows that $\sup(B) \le \sup(A)$, that is, $\sup(B) = \sup(A)$ which is exactly (A.4). Identity (A.5) follows trivially by the definition of $H_{x_0}$. By Lemma A.3, if $\omega_2^*(x_0) = 0$ we have that

$$G_{x_0} = \left\{ \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} : \lim_{n_j \to \infty} \frac{k(n_j)}{d_{n_j}} = 0 \text{ eventually by passing to a subsequence} \right\},$$

where $k(n)$ is the sub-index associated to $\lambda_{\hat{k}(n)}(Y^{(n)})$, that is, $k(n) = \vartheta_{Y^{(n)}}^{-1}(\hat{k}(n))$. Without loosing the generality of the proof, we will suppose hereafter that $\lim_{n_j \to \infty} \frac{k(n_j)}{d_{n_j}} = 0$ without passing to a subsequence, and that $d_n^-(Y^{(n)}) = d_n^-(X^{(n)})$, that is, $\vartheta(k) = k$ and then we can avoid to pass to the index $\hat{k}$. Because of this last assumption, we have that

$$\left| \frac{\lambda_{\vartheta_{Y \to X}(k)}(X^{(n)})}{\lambda_k(Y^{(n)})} - 1 \right| = \left| \frac{\lambda_k(X^{(n)})}{\lambda_k(Y^{(n)})} - 1 \right| = \delta_k^{(n)}, \quad k \in \left\{ d_n^-(Y^{(n)}), \ldots, d_n^+(Y^{(n)}) \right\}.$$

Let us define now

$$\limsup_{n \to \infty} \delta_k^{(n)} = \sigma_k \in [0, \infty].$$

Then there exists a subsequence $\{n_j\}$ such that

$$\sigma_k = \lim_{n_j \to \infty} \delta_k^{(n_j)}.$$

So, $\left\{ \delta_k^{(n_j)} \right\} \in F$ and clearly $\frac{k}{d_{n_j}} \to 0$, consequently $\sigma_k \in H_{x_0}$. Therefore $\sup_{k \in [\kappa^-, \kappa^+]} \{\sigma_k\} \le \sup H_{x_0}$. On the other hand, assuming that $k \ne 0$

$$\limsup_{n \to \infty} \left| \delta_{k(n)}^{(n)} - \sigma_{k(n)} \right| = 0 \quad \text{for every } k(n) \text{ such that } \lim_{n \to \infty} \frac{k(n)}{d_n} = 0,$$

then, given an element $\delta_{k(n)}^{(n)} \in H_{x_0}$, there exists a subsequence such that

$$\lim_{j \to \infty} \left| \delta_{k(n_j)}^{(n_j)} - \sigma_{k(n_j)} \right| = 0.$$

Therefore,

$$\lim_{n \to \infty} \delta_{k(n)}^{(n)} \leq \lim_{j \to \infty} \left| \delta_{k(n_j)}^{(n_j)} - \sigma_{k(n_j)} \right| + \sigma_{k(n_j)}$$

$$\leq \lim_{j \to \infty} \left| \delta_{k(n_j)}^{(n_j)} - \sigma_{k(n_j)} \right| + \sup_{\substack{k \in [\kappa^-, \kappa^+] \\ k \neq 0}} \{\sigma_k\}$$

$$= \sup_{\substack{k \in [\kappa^-, \kappa^+] \\ k \neq 0}} \{\sigma_k\},$$

and it follows that $\sup H_{x_0} \leq \sup_{\substack{k \in [\kappa^-, \kappa^+] \\ k \neq 0}} \{\sigma_k\}$.

**Theorem** 3.2.

***Proof*** Let us begin observing that, by (A.4),

$$\mathscr{E} = \sup \left\{ \lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} : \left\{ \delta_{\hat{k}(n_j)}^{(n_j)} \right\}_{n_j} \in F \right\},$$

and that the set $J$ of points such that $\omega_1^*$ or $\omega_2^*$ are discontinuous is at most countable, and the possible discontinuity points can be only jumps. Moreover, due to the strict monotonicity of $\omega_2^*$, there can exists only one point $x_0 \in (0, 1)$ such that $\omega_2^*(x_0) = 0$. Fix $x_0 \in (0, 1)$ such that both $\omega_1^*$ and $\omega_2^*$ are continuous in $x_0$, and such that $\omega_2^*(x_0) \neq 0$. Let $\{\hat{k}(n_j)\}$ be a sequence such that $\hat{k}(n_j)/d_{n_j} \to x_0^-$ as $n_j \to \infty$. Then

$$\lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} = \left| \frac{\omega_1^*(x_0)}{\omega_2^*(x_0)} - 1 \right| < \infty$$

due to (3.4). This proves that

$$\mathscr{E} \geq \sup_{\substack{x \in (0,1) \setminus J; \\ x : \omega_2^*(x) \neq 0}} \left| \frac{\omega_1^*(x)}{\omega_2^*(x)} - 1 \right| = \sup_{\substack{x \in (0,1); \\ x : \omega_2^*(x) \neq 0}} \left| \frac{\omega_1^*(x)}{\omega_2^*(x)} - 1 \right| = \max \left\{ \sup_{\substack{x \in (0,1); \\ x : \omega_2^*(x) \neq 0}} \left| \frac{\omega_1^*(x)}{\omega_2^*(x)} - 1 \right| ; 0 \right\}.$$

Let now assume that there exists $x_0 \in [0, 1]$ such that $\omega_2(x_0) = 0$. If there exists $x_0 \in [0, 1]$ such that $\omega_2(x_0) = 0 \neq \omega_1(x_0)$, then by the right-continuity of $\omega_2^*$,

$$\sup_{\substack{x \in (0,1); \\ x : \omega_2^*(x) \neq 0}} \left| \frac{\omega_1^*(x)}{\omega_2^*(x)} - 1 \right| = +\infty = \sigma.$$

On the contrary, if there exists $x_0 \in [0, 1]$ such that $\omega_2(x_0) = 0 = \omega_1(x_0)$, then by (A.5) it holds that

$$\mathscr{E} \geq \sup \left\{ \lim_{n_j \to \infty} \delta^{(n_j)}_{\hat{k}(n_j)} \; : \; \left\{ \delta^{(n_j)}_{\hat{k}(n_j)} \right\}_{n_j} \in F \cap G_{x_0} \right\} = \sup H_{x_0} = \sigma.$$

Therefore, we can conclude that

$$\mathscr{E} \geq \max \left\{ \sup_{\substack{x \in (0,1); \\ x \, : \, \omega_2^*(x) \neq 0}} \left| \frac{\omega_1^*(x)}{\omega_2^*(x)} - 1 \right|; \sigma \right\},$$

which is exactly (3.7).

Let us now suppose that both $\omega_1^*, \omega_2^*$ are continuous, that $\{X^{(n)}\}_n, \{Y^{(n)}\}_n$ do not have outliers and that $R_{\omega_1}, R_{\omega_2}$ are bounded, that is, compact. If we prove the reverse inequality in (3.7), then we can conclude. By our assumptions it follows easily that

$$\lambda_{\hat{k}}\left(x^{(n)}\right) \in R_{\omega_1} = \omega_1^*([0,1]), \qquad \lambda_{\hat{k}}\left(Y^{(n)}\right) \in R_{\omega_2} = \omega_2^*([0,1])$$

for every $n$ and $\hat{k} = 1, \ldots, d_n$, that is, for every $n$ and for every $\hat{k}$ there exist $x_1, x_2 \in [0,1]$ such that

$$\lambda_{\hat{k}}\left(X^{(n)}\right) = \omega_1^*(x_1), \qquad \lambda_{\hat{k}}\left(Y^{(n)}\right) = \omega_2^*(x_2).$$

Notice that if there exists $x_0 \in [0,1]$ such that $\omega_2^*(x_0) = 0 \neq \omega_1^*(x_0)$ then

$$\mathscr{E} = +\infty = \sup_{\substack{x \in (0,1); \\ x \, : \, \omega_2^*(x) \neq 0}} \left| \frac{\omega_1^*(x)}{\omega_2^*(x)} - 1 \right| = \sigma.$$

Therefore, hereafter we will suppose that either $\omega_2^*(x) \neq 0$ for every $x \in [0,1]$ or that there exists $x_0 \in [0,1]$ such that $\omega_2^*(x_0) = 0 = \omega_1^*(x_0)$. Fix a sequence $\left\{ \delta^{(n_j)}_{\hat{k}(n_j)} \right\}_{n_j} \in F$. By the boundedness of $\hat{k}(n_j)/d_{n_j}$, there exists a subsequence $\{n_{j_m}\}$ such that

$$\lim_{n_{j_m} \to \infty} \frac{\hat{k}(n_{j_m})}{d_{n_{j_m}}} = x_0 \in [0,1].$$

Then by Theorem 3.1 it happens that

$$\begin{cases} \lim_{n_{j_m} \to \infty} \lambda_{\hat{k}(n_{j_m})}\left(X^{(n_{j_m})}\right) = \omega_1^*(x_0) \in (-\infty, +\infty), \\ \lim_{n_{j_m} \to \infty} \lambda_{\hat{k}(n_{j_m})}\left(Y^{(n_{j_m})}\right) = \omega_2^*(x_0) \in (-\infty, +\infty). \end{cases}$$

If $\omega_2^*(x_0) \neq 0$, then $\lim_{n_j \to \infty} \delta^{(n_j)}_{\hat{k}(n_j)}$ is finite, in particular

$$\lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} = \lim_{n_{j_m} \to \infty} \left| \frac{\lambda_{\hat{k}(n_{j_m})}\left(X^{(n_{j_m})}\right)}{\lambda_{\hat{k}(n_{j_m})}\left(Y^{(n_{j_m})}\right)} - 1 \right| = \left| \frac{\omega_1^*(x_0)}{\omega_2^*(x_0)} - 1 \right|.$$

If, instead, $x_0$ is such that $\omega_2^*(x_0) = 0 = \omega_1^*(x_0)$, then by (A.6) it holds that

$$\lim_{n_j \to \infty} \delta_{\hat{k}(n_j)}^{(n_j)} \leq \sup_{\substack{k \in [\kappa^-, \kappa^+] \\ k \neq 0}} \left\{ \limsup_{n \to \infty} \left| \frac{\lambda_{\vartheta(k)}\left(X^{(n)}\right)}{\lambda_k\left(Y^{(n)}\right)} - 1 \right| \right\} = \sigma.$$

Combining all the results obtained till now, we can conclude that

$$\mathscr{E} \leq \max \left\{ \sup_{\substack{x \in (0,1); \\ x \,:\, \omega_2^*(x) \neq 0}} \left| \frac{\omega_1^*(x)}{\omega_2^*(x)} - 1 \right|; \sigma \right\}.$$

$\square$

**Theorem** 3.3.

*Proof* Clearly, if we set

$$\hat{X}^{(n)} := h(d_n)X^{(n)}, \qquad \hat{Y}^{(n)} := h(d_n)Y^{(n)},$$

then

$$\mathscr{E}\left(X^{(n)}, Y^{(n)}\right) = \mathscr{E}\left(\hat{X}^{(n)}, \hat{Y}^{(n)}\right).$$

By Item (iv), Lemma A.1 and Theorem 3.1, it follows immediately that $\{\hat{Y}^{(n)}\} \sim_\lambda \zeta^*$, that $\lambda_k\left(\hat{Y}^{(n)}\right) \in R_{\zeta^*} = \overline{\zeta^*([0,1])}$ for every $k = -d_n^-, \ldots, -1, 1, \ldots, d_n^+$, and that therefore the sequence $\{\hat{Y}^{(n)}\}$ satisfies the hypotheses of Theorem 3.2. By items (ii)-(iii), we have that $\{\hat{X}^{(n)}\} \sim_\lambda \omega$ and it also satisfies the hypotheses of Theorem 3.2. Therefore,

$$\mathscr{E}\left(\hat{X}^{(n)}, \hat{Y}^{(n)}\right) \geq \max \left\{ \sup_{\substack{x \in [0,1] \\ x \,:\, \zeta^*(x) \neq 0}} \left| \frac{\omega^*(x)}{\zeta^*(x)} - 1 \right|; \sigma \right\}.$$

By Item (i), it follows that $\sigma = 0$ and we conclude.

$\square$

## Proofs of Section 4

**Proposition B.1** *Let us consider a Sturm-Liouville operator defined by the formal equation*

$$\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}[u](x) := \frac{1}{w(x)}\left[-\partial_x\big(p(x)\partial_x u(x)\big) + q(x)u(x)\right], \qquad x \in (a,b) \subset \mathbb{R},$$

*such that*

(i)   $p, p', w, w', q, (pw)', (pw)'' \in C([a,b])$;
(ii)   $p, w > 0$;
(iii)   *regular BCs.*

*Discretize the above self-adjoint linear differential operator $\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}$ by means of a numerical scheme, and let $\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}^{(n,\eta)}$ be the correspondent discrete operator, where n is the mesh fineness parameter and $\eta$ is the order of approximation of the numerical scheme. Define $B = \int_a^b \sqrt{\frac{w(x)}{p(x)}}m(dx)$. If:*

(a)   *for every fixed $k \in \mathbb{N}$,*

$$\lim_{n\to\infty} \lambda_k\left(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}^{(n,\eta)}\right) = \lambda_k\left(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}\right);$$

(b)   *there exists $\omega : [a,b] \times [-\pi,\pi] \to \mathbb{R}$, $\omega \in \mathrm{L}^1([a,b] \times [-\pi,\pi])$ such that*

$$\left\{(d_n+1)^{-2}\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}^{(n,\eta)}\right\}_n$$
$$= \left\{\hat{\mathscr{L}}_{\mathrm{BCs},p(x),q(x),w(x)}^{(n,\eta)}\right\}_n \sim_\lambda \omega(x,\theta) \quad (x,\theta) \in [a,b] \times [-\pi,\pi];$$

(c)   *for every $\eta$, the monotone rearrangement $\omega_\eta^*$, as defined in (2.3), is such that*

$$\omega_\eta^*(x) \sim \frac{x^2\pi^2}{B^2}, \qquad \text{as } x \to 0.$$

*Then, for every fixed $k \in \mathbb{N}$*

$$\lim_{n\to\infty} \left| \frac{(d_n+1)^2\omega_\eta^*\left(\frac{k}{d_n+1}\right)}{\lambda_k\left(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}^{(n,\eta)}\right)} - 1 \right|$$

$$= \lim_{n\to\infty} \left| \frac{(d_n+1)^2\omega_\eta^*\left(\frac{k}{d_n+1}\right)}{\lambda_k\left(\mathscr{L}_{\mathrm{BCs},V}\right)} - 1 \right| = c_k \geq 0, \qquad \lim_{k\to\infty} c_k = 0,$$

with $c_k$ independent of $\eta$, and where $\mathscr{L}_{\mathrm{BCs},V}$ is the differential operator associated to the normal form of $\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}$ by the Liouville transform $y(x) := \int_a^x \sqrt{\frac{w(t)}{p(t)}}\, m(dt)$. In particular,

$$\lim_{n\to\infty} \mathbf{err}^{*,(n)}_{\alpha,r,k} = c_k,$$

where $\mathbf{err}^{*,(n)}_{\alpha,r,k}$ is the analytic spectral relative error of Definition 4.1.

**Proof** By standard theory it holds that

$$\lambda_k\big(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}\big) = \lambda_k\big(\mathscr{L}_{\mathrm{BCs},V(y)}\big) \qquad \text{and} \qquad \lambda_k\big(\mathscr{L}_{\mathrm{dir},V\equiv 0}\big) = \lambda_k\big(\varDelta_{\mathrm{dir}}\big) = \frac{k^2\pi^2}{B^2}.$$

Therefore, from item (c)

$$\lim_{n\to\infty}(d_n+1)^2\omega^*_\eta\!\left(\frac{k}{d_n+1}\right) = \frac{k^2\pi^2}{B^2} = \lambda_k\big(\varDelta_{\mathrm{dir}}\big). \tag{B.1}$$

Then it is immediate to prove that if

$$\lambda_k\big(\mathscr{L}_{\mathrm{BCs},V(y)}\big) \neq \frac{k^2\pi^2}{B^2} = \lambda_k\big(\varDelta_{\mathrm{dir}}\big),$$

then

$$\lim_{n\to\infty}\left|\frac{(d_n+1)^2\omega^*_\eta\!\left(\frac{k}{d_n+1}\right)}{\lambda_k\big(\mathscr{L}^{(n,\eta)}_{\mathrm{BCs},p(x),q(x),w(x)}\big)} - 1\right| = \lim_{n\to\infty}\left|\frac{(d_n+1)^2\omega^*_\eta\!\left(\frac{k}{d_n+1}\right)}{\lambda_k\big(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}\big)} - 1\right|$$

$$= \lim_{n\to\infty}\left|\frac{\frac{k^2\pi^2}{B^2}}{\lambda_k\big(\mathscr{L}_{\mathrm{BCs},V(y)}\big)} - 1\right|$$

$$= c_k > 0.$$

Moreover,

$$\lim_{n\to\infty}\lambda_k\big(\mathscr{L}^{(n,\eta)}_{\mathrm{BCs},p(x),q(x),w(x)}\big) = \lambda_k\big(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}\big)$$

$$= \lambda_k\big(\mathscr{L}_{\mathrm{BCs},V(y)}\big) \sim \frac{k^2\pi^2}{B^2} \qquad \text{for } k\to\infty,$$

and then $c_k \to 0$ as $k \to \infty$. $\qquad\square$

**Corollary B.1** *If*

$$\omega_\eta(x,\theta) = \frac{p(x)}{w(x)(b-a)^2}f_\eta(\theta), \qquad (x,\theta)\in[a,b]\times[0,\pi],$$

with $f_\eta(\theta)$ nonnegative, nondecreasing and such that $f_\eta(\theta) \sim \theta^2$ as $\theta \to 0$, then item
(c) of Proposition B.1 is satisfied.

**Proof** From (2.4) and (2.3), for all $t \in [0, t_0]$, with

$$t_0 = (b-a)^{-2} \max_{[a,b]} (p(y)/w(y)) \sup_{[0,\pi]} (f_\eta(\theta)),$$

we have that

$$\phi_{\omega_\eta}(t) = \frac{1}{\pi(b-a)} \int_a^b \left( \int_0^\pi \mathbb{1}_{\Omega_y(t)}(\theta)d\theta \right) dy,$$

where

$$\Omega_y(t) := \left\{ \theta \in [0, \pi] \ : \ 0 \le f_\eta(\theta) \le \frac{(b-a)^2 w(y)}{p(y)} t \right\}.$$

By the monotonicity of $f_\eta$, it holds that $\Omega_y(t) = [0, \theta_y(t)]$,
$\theta_y(t) := \sup\{\theta \in [0, \pi] \ : \ f_\eta(\theta) \le \frac{(b-a)^2 w(y)}{p(y)} t\}$, and then $\theta_y(t) \to 0$ as $t \to 0$. So, by
the boundedness of $w(y)/p(y)$, for every $\epsilon > 0$ there exists $\delta_\epsilon > 0$ independent of $y$
such that for every $t \in [0, \min\{t_0; \delta_\epsilon\}]$ then $(1-\epsilon)\theta^2 < f_\eta(\theta) < (1+\epsilon)\theta^2$, and

$$\frac{1}{\pi(b-a)} \int_a^b m\left( \Omega_y^+(t) \right) m(dy) \le \phi_{\omega_\eta}(t) \le \frac{1}{\pi(b-a)} \int_a^b m\left( \Omega_y^-(t) \right) m(dy),$$

where

$$\Omega_y^+(t) := \left\{ \theta \in [0, \pi] \ : \ \theta^2 \le \frac{(b-a)^2}{1+\epsilon} \frac{w(y)}{p(y)} t \right\},$$

$$\Omega_y^-(t) := \left\{ \theta \in [0, \pi] \ : \ \theta^2 \le \frac{(b-a)^2}{1-\epsilon} \frac{w(y)}{p(y)} t \right\}.$$

Then it holds that

$$\frac{B}{\pi\sqrt{1+\epsilon}}\sqrt{t} \le \phi_{\omega_\eta}(t) \le \frac{B}{\pi\sqrt{1-\epsilon}}\sqrt{t}, \qquad B = \int_a^b \sqrt{\frac{w(y)}{p(y)}} dy.$$

By definition (2.3), $t \to 0$ as $x \to 0$ and then

$$(1-\epsilon)\frac{x^2\pi^2}{B^2} \le \omega_\eta^*(x) \le (1+\epsilon)\frac{x^2\pi^2}{B^2} \qquad \text{for } x \text{ small enough},$$

and the thesis follows. $\qquad\qquad\square$

**Remark 2** The matrix methods of Sects. C, D satisfy the hypotheses of Proposition
B.1, see Conjecture C.1 and Theorem D.1, and Corollary C.1 and Corollary D.1.
Therefore, in general, a uniform sampling of their spectral symbols does not provide

an accurate approximation of the eigenvalues $\lambda_k\left(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}^{(n,\eta)}\right)$ and $\lambda_k\left(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}\right)$, in the sense of the relative error. See Sects. 4.1.1, 4.1.2 and 4.1.3 for numerical examples. On the other hand,

$$\left\|\omega_\eta^*\left(\frac{k}{n+1}\right) - \lambda_k\left(\mathscr{L}_{\mathrm{BCs},p(x),q(x),w(x)}^{(n,\eta)}\right)\right\|_\infty \to 0 \qquad \text{as } n \to \infty,$$

if we exclude the outliers, see Corollary 3.3 and Fig. 2.

## $(2\eta + 1)$-points central FD discretization

In this section we present an high-order central FD discretization scheme which is a slight modification of the standard ones available in the literature. Since it is used only in the numerical Sect. 4.1.2 and since the main focus of this paper is to analytically calculate the MSRE, see Theorem 3.2 and Theorem 3.3, we will not provide proofs of the main properties of this scheme, in order to keep the length of the paper contained. Even if the main properties will be stated as *conjectures* and the rigorous theoretic analysis will be left for a possible future work, we will give some hints and comment about how to prove them. Nevertheless, we wish to point out that in Sect. 4.1.2 there are consistent numerical evidences of the validity of Conjecture C.1.

Let us consider the following one dimensional self-adjoint operator

$$\mathscr{L}_{\mathrm{dir},p(x)}[u](x) := -\big(p(x)u'(x)\big)', \qquad \mathrm{dom}\big(\mathscr{L}_{\mathrm{dir},p(x)}\big) = W_0^{1,2}((a,b)), \quad \text{(C.1)}$$

with $p \in C^2([a,b])$ and $p > 0$ (see [17, 21]). For a general review of FD methods we refer to [43]. Fix $\eta, n \in \mathbb{N}$ and $\eta \geq 1$. Given a standard equispaced grid $\mathbf{x} = \{x_j\}_{j=1-\eta}^{n+\eta} \subseteq [\bar{a}, \bar{b}]$, with

$$\bar{a} = a - (b-a)\frac{\eta - 1}{n+1} < a, \quad \bar{b} = b + (b-a)\frac{n+\eta}{n+1} > b, \quad x_j = a + (b-a)\frac{j}{n+1},$$

let us consider a $C^1$-diffeomorphism $\tau : [a,b] \to [a,b]$ such that $\tau'(x) \neq 0$, $\tau(a) = a, \tau(b) = b$ and let us consider its piecewise $C^1$-extension $\bar{\tau} : [\bar{a}, \bar{b}] \to [\bar{a}, \bar{b}]$ such that

$$\bar{\tau}(x) = \begin{cases} x & \text{if } x \leq a, \\ \tau(x) & \text{if } x \in (a,b) \\ x & \text{if } x \geq b. \end{cases}$$

By means of the piecewise $C^1$-diffeomorphism $\bar{\tau}$ we have a new (extended) grid $\bar{\mathbf{x}} = \{\bar{x}_j\}_{j=1-\eta}^{n+\eta} \subset [\bar{a}, \bar{b}]$, non necessarily uniformly equispaced, with $\bar{x}_j = \bar{\tau}(x_j)$. Combining together the high-order central FD schemes presented in [2, 3, 33], we obtain the following matrix operator as approximation of (C.1):

$$
\mathscr{L}_{\mathrm{dir},p(\bar{x})}^{(n,\eta)} =
\begin{bmatrix}
l_{1,1} & \cdots & l_{1,1+\eta} & 0 & \cdots & & & 0 \\
l_{2,1} & l_{2,2} & \cdots & l_{2,2+\eta} & 0 & \cdots & & 0 \\
& & \ddots & & & \ddots & & \\
& & & \ddots & & \ddots & \ddots & \\
0 & \cdots & 0 & l_{k,k-\eta} & \cdots & l_{k,k} & \cdots & l_{k,k+\eta} & 0 & \cdots & 0 \\
& & & & \ddots & & \ddots & \ddots & \ddots \\
& & & & & & & & \\
& & & & & & \cdots & l_{n,n}
\end{bmatrix}
\in \mathbb{R}^{n\times n},
$$

where, if we define the $C^0$-extension of $p(x)$ to $[\bar{a},\bar{b}]$ as

$$
\bar{p}(x) = \begin{cases}
p(a) & \text{for } x \le a, \\
p(x) & \text{for } x \in (a,b) \\
p(b) & \text{for } x \ge b,
\end{cases}
$$

and the element $l_{i,j}$ as

$$
l_{i,j} = \frac{2\bar{p}\left(\frac{\bar{x}_j + \bar{x}_i}{2}\right) \sum\limits_{\substack{m = i-\eta \\ m \ne i,j}}^{i+\eta} \prod\limits_{\substack{k = i-\eta \\ k \ne i,j,m}}^{i+\eta} (\bar{x}_k - \bar{x}_i)}{(\bar{x}_j - \bar{x}_i) \prod\limits_{\substack{k = i-\eta \\ k \ne i,j}}^{i+\eta} (\bar{x}_k - \bar{x}_j)}, \qquad \text{for } i \ne j, \ \begin{cases} i = 1,\dots,n, \\ j = i-\eta,\dots,i+\eta, \end{cases}
$$

then the generic matrix element of $\mathscr{L}_{\mathrm{dir},p(\bar{x})}^{(n,\eta)}$ is given by

$$
\left(\mathscr{L}_{\mathrm{dir},p(\bar{x})}^{(n,\eta)}\right)_{i,j} =
\begin{cases}
l_{i,j} & \text{for } i \ne j, \ |i-j| \le \eta, \ i,j = 1,\dots,n, \\
-\sum\limits_{\substack{k = i-\eta \\ k \ne i}}^{i+\eta} l_{i,k} & \text{for } i = j, \ i = 1,\dots,n, \\
0 & \text{for } i \ne j, \ |i-j| > \eta, \ i,j = 1,\dots,n.
\end{cases}
$$

The extended functions $\bar{\tau}, \bar{p}$ on $[\bar{a},\bar{b}] \supset [a,b]$ serve to implement correctly the BCs. With abuse of notation, we will call $\eta$ the *order of approximation* of the central FD method. We conjecture now some of the properties of this scheme, providing then a sketch of the proofs.

**Conjecture C.1** *In the above assumptions, for $\eta \ge 1$ it holds that*

(i) *the eigenvalues $\lambda_k\left(\mathscr{L}_{\mathrm{dir},p(\bar{x})}^{(n,\eta)}\right)$ are real for every $k$ and*

$$\lim_{n\to\infty} \lambda_k\left(\mathscr{L}^{(n,\eta)}_{\mathrm{dir},p(\bar{x})}\right) = \lambda_k\left(\mathscr{L}_{\mathrm{dir},p(x)}\right) \qquad \text{for every fixed } k.$$

$$\left\{(n+1)^{-2}\mathscr{L}^{(n,\eta)}_{\mathrm{dir},p(\bar{x})}\right\}_n = \left\{\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir},p(\bar{x})}\right\}_n \sim_\lambda \omega_\eta(x,\theta) \qquad (x,\theta)\in[a,b]\times[0,\pi], \text{ (C.2)}$$

(ii)

where

$$\omega_\eta(x,\theta) = \frac{p(\tau(x))}{(\tau'(x))^2(b-a)^2}f_\eta(\theta),$$

and

$$f_\eta(\theta) = d_{\eta,0} + 2\sum_{k=1}^{\eta} d_{\eta,k}\cos(k\theta), \qquad d_{\eta,k} = \begin{cases} (-1)^k\frac{\eta!\eta!}{(\eta-k)!(\eta+k)!}\frac{2}{k^2} & \text{for } k=1,\cdots\eta, \\ -2\sum_{j=1}^{\eta} d_{\eta,j} & \text{for } k=0. \end{cases}$$
(C.3)

(iii)   $\lambda_k\left(\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir},p(\bar{x})}\right) \in R_{\omega_\eta}$ for every $k, n$.

Item (i) can be proved by a straightforward generalization of standard techniques, see [12, Theorem 1] and [16, 25]. About item (ii), we recall that the "hat" means that the matrix operator has been normalized by $(n+1)^2$. Let us preliminarily observe that in the case of $\tau(x) = x$ and $p(x) \equiv 1$ then $\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir}}$ is a symmetric Toeplitz matrix defined by

$$\left(\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir}}\right)_{i,j} = \frac{d_{\eta,|i-j|}}{(b-a)^2} \quad \text{or equivalently} \quad \hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir}} = (b-a)^{-2}T_n(f_\eta),$$

where $d_{\eta,k}$ are the coefficients of the trigonometric polynomial $f_\eta$ in (C.3), see [33, Corollary 2.2] and [30, Eq. (27)]. By standard results on the eigenvalues distribution of Toeplitz matrices (see for example [22, item **T 4** p. 168]) it holds that

$$\left\{\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir}}\right\}_n \sim_\lambda \frac{f_\eta(\theta)}{(b-a)^2}, \qquad \theta\in[0,\pi]. \tag{C.4}$$

The strategy to prove (ii) is the following:

- show that $\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir},p(\bar{x})} = X_n + Y_n$, with $X_n$ symmetric and such that $\|X_n\|, \|Y_n\| \le C$, $n^{-1}\|Y_n\|_1 \to 0$ as $n\to\infty$, where $\|\cdot\|$ and $\|\cdot\|_1$ are the spectral norm and the Schatten 1-norm, respectively;
- show that $\{X_n\}_n \sim_{\mathrm{GLT}} \omega_\eta(x,\theta)$, see [22, Definition 8.1];
- Conclude that $\{\hat{\mathscr{L}}^{(n,\eta)}_{\mathrm{dir},p(\bar{x})}\}_n \sim_\lambda \omega_\eta(x,\theta)$ by [22, item **GLT 2** p. 170].

Finally, item (iii) can be recovered from [40, Theorem 2.2].

**Remark 3** Notice that the spectral symbol $\omega_\eta(x, \theta)$ consists of the product of two functions:

- the function $\frac{p(\tau(x))}{(\tau'(x))^2(b-a)^2}$ which consists of the diffusion coefficient $p(x)$ and the diffeomorphism $\tau(x)$, which are all intrinsic to the differential operator (C.1) itself and depend on the spatial variable $x$;
- $f_\eta(\theta)$, which is intrinsic to the discretization method and depends on the spectral variable $\theta$.

Let us now state some analytical properties of the symbol function $f_\eta$ in (C.3). We want to highlight that it shares the same properties of the symbol function associated to the IgA scheme, see Corollary D.1. This is pretty interesting, since it seems to suggest that, whatever is the (consistent) numerical approximating scheme, the limit symbol converges to the inverse Weyl distribution function of the continuous operator, as the order of approximation increases.

**Corollary C.1** *For every fixed $\eta$, the function $f_\eta$ is differentiable, nonnegative, monotone increasing on $[0, \pi]$ and it holds that*

$$f_\eta(\theta) \sim \theta^2 \quad \text{as } \theta \to 0, \qquad \lim_{\eta \to \infty} \sup_{\theta \in [0,\pi]} \left| f_\eta(\theta) - \theta^2 \right| = 0.$$

**Proof** $f_\eta(\theta)$ is obviously $C^\infty([0, \pi])$. Let us begin to prove that $f_\eta(\theta) \sim \theta^2$ as $\theta \to 0$ for every fixed $\eta$, and that it is monotone nonnegative on $[0, \pi]$. By the Taylor expansion at $\theta = 0$ we get

$$f_\eta(\theta) = d_{\eta,0} + \sum_{\substack{k=-\eta \\ k \neq 0}}^{\eta} d_{\eta,k} \cos(k\theta)$$

$$= d_{\eta,0} + \sum_{\substack{k=-\eta \\ k \neq 0}}^{\eta} d_{\eta,k} \left( 1 - \frac{(k\theta)^2}{2} \right) + o(\theta^3)$$

$$= -\theta^2 \sum_{\substack{k=-\eta \\ k \neq 0}}^{\eta} (-1)^k \frac{\eta!\eta!}{(\eta-k)!(\eta+k)!} + o(\theta^3)$$

$$= -\theta^2 \left[ (-1)^\eta \frac{\eta!\eta!}{(2\eta)!} \sum_{\substack{m=0 \\ m \neq \eta}}^{2\eta} (-1)^m \binom{2\eta}{m} \right] + o(\theta^3)$$

$$= \theta^2 + o(\theta^3).$$

Moreover, let us observe that

$$f_\eta(\theta) = d_{\eta,0} + 2 \sum_{k=1}^{\eta} \left| d_{\eta,k} \right| \cos(k(-\theta + \pi)), \quad f_\eta'(\theta) = 2 \sum_{k=1}^{\eta} k \left| d_{\eta,k} \right| \sin(k(-\theta + \pi)).$$

Define then

$$g(\psi) = \sum_{k=1}^{\eta} a_k \sin(k\psi), \qquad \text{with } \psi = -\theta + \pi \in [0, \pi], \qquad \text{and } a_k = 2k \left| d_{\eta,k} \right|.$$

It is immediate to check that $a_1 \geq a_2 \geq \ldots \geq a_n > 0$ and that

$$(2k)a_{2k} \leq (2k - 1)a_{2k-1} \quad \forall k \geq 1.$$

By [4, Theorem 1] we can conclude that $g(\psi) > 0$ on $(0, \pi)$ and then $f_\eta'(\theta) > 0$ on $(0, \pi)$. Since $f_\eta(0) = 0$, we deduce that $f_\eta(\theta) \geq 0$ on $[0, \pi]$. The second part of the thesis is an immediate consequence of identities (C.3). Indeed,

$$|d_{\eta,k}| \leq \frac{2}{k^2}, \qquad \sum_{k=1}^{\infty} \frac{2}{k^2} < \infty,$$

and then, by dominate convergence, it is easy to prove that for every fixed $\theta \in [0, \pi]$ it holds that

$$\lim_{\eta \to \infty} f_\eta(\theta) = \lim_{\eta \to \infty} \left( -2 \sum_{k=1}^{\eta} d_{\eta,k} \right) + 2 \lim_{\eta \to \infty} \left( \sum_{k=1}^{\eta} d_{\eta,k} \cos(k\theta) \right)$$

$$= \frac{\pi^2}{3} + 4 \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2} \cos(k\theta)$$

$$= \theta^2,$$

since $\{\pi^2/3\} \cup \{(-1)^k 4/k^2\}_{k\geq1}$ are the Fourier coefficients of $\theta^2$ on $[0, \pi]$. The convergence is uniform by Lemma A.2.                                                                                                 $\square$

## Isogeometric Galerkin discretization by B-splines of degree $\eta$ and smoothness $C^{\eta-1}$

For a general review of the IgA discretization method we refer the reader to [15, 27], so we skip all the introductions and we present directly some known results.

**Theorem D.1** *Let* $(a, b)$ *be discretized by a uniform mesh* $\{x_j\}_{j=1}^{n}$ *of step-size* $(n + 1)^{-1}$ *and let* $\tau : [a, b] \to [a, b]$ *a* $C^1$-*diffeomorphism such that* $\tau(a) = a$, $\tau(b) = b$ *and* $\tau'(x) \neq 0$ *for every* $x \in [a, b]$. *Let* $\bar{x} := \tau(x)$, $\eta \geq 1$, *and let us indicate with* $\mathscr{L}_{\text{dir},p(\bar{x})}^{(n,\eta)}$ *the discrete operator obtained from* (C.1) *by an IgA discretization scheme with B-splines of degree* $\eta$ *and smoothness* $C^{\eta-1}$. *Then*

(i)  *the eigenvalues $\lambda_k\left(\mathcal{L}_{\mathrm{dir},p(\bar{x})}^{(n,\eta)}\right)$ are real for every fixed k and*

$$\lim_{n\to\infty} \lambda_k\left(\mathcal{L}_{\mathrm{dir},p(\bar{x})}^{(n,\eta)}\right) = \lambda_k\left(\mathcal{L}_{\mathrm{dir},p(x)}\right);$$

(ii)  *it holds that*

$$\left\{(n+1)^{-2}\mathcal{L}_{\mathrm{dir},p(\bar{x})}^{(n,\eta)}\right\}_n = \left\{\hat{\mathcal{L}}_{\mathrm{dir},p(\bar{x})}^{(n,\eta)}\right\}_n \sim_\lambda \omega_\eta(x,\theta), \qquad (x,\theta) \in [a,b] \times [0,\pi],$$

*where*

$$\omega_\eta(x,\theta) = \frac{p(\tau(x))}{(\tau'(x))^2(b-a)^2} f_\eta(\theta),$$

*For example, for $\eta = 1, 2$, $f_\eta(\theta)$ has the following analytical expressions*

$$f_1(\theta) = \frac{6(1-\cos\theta)}{2+\cos\theta}, \qquad f_2(\theta) = \frac{20(3 - 2\cos\theta - \cos(2\theta))}{33 + 26\cos\theta + \cos(2\theta)}.$$

**Proof** For item (i) see [7, 37]. For item (ii), see [22, Theorem 10.15]. □

About the spectral symbol $\omega_\eta$ see again Remark 3. We have an analogue of Corollary C.1.

**Corollary D.1** *For every fixed $\eta$, the function $f_\eta$ is differentiable, nonnegative, monotone increasing on $[0,\pi]$ and it holds that*

$$f_\eta(\theta) \sim \theta^2 \quad \text{as } \theta \to 0, \qquad \lim_{\eta\to\infty}\sup_{\theta\in[0,\pi]} \left| f_\eta(\theta) - \theta^2 \right| = 0.$$

**Proof** See [21, Theorem 1, Theorem 2 and Lemma 1]. □

# References

1. Adriani, A., Bianchi, D., Serra-Capizzano, S.: Asymptotic spectra of large (Grid) graphs with a uniform local structure (Part I). Milan J. Math. **88**, 409–454 (2020)
2. Amodio, P., Sgura, I.: High-order finite difference schemes for the solution of second-order BVPs. J. Comput. Appl. Math. **176**(1), 59–76 (2005)
3. Amodio, P., Settani, G.: A matrix method for the solution of Sturm-Liouville problems. JNAIAM **6**(1–2), 1–13 (2011)
4. Askey, R., Steinig, J.: Some positive trigonometric sums. Trans. Amer. Math. Soc. **187**, 295–307 (1974)
5. Barbarino, G.: Equivalence between GLT sequences and measurable functions. Linear Algebra Appl. **529**, 397–412 (2017)
6. Barbarino, G.: Spectral measures. In: Bini, D., Di Benedetto, F., Tyrtyshnikov, E., Van Barel, M. (eds.) Structured Matrices in Numerical Linear Algebra. Springer INdAM Series, Springer, Cham (2019)
7. Bazilevs, Y., Beirao da Veiga, L., Cottrell, J.A., Hughes, T.J.R., Sangalli, G.: Isogeometric analysis: approximation, stability and error estimates for h-refined meshes. Math. Models Methods Appl. Sci. **16**(07), 1031–1090 (2006)
8. Bianchi, D., Serra-Capizzano, S.: Spectral analysis of finite-dimensional approximations of 1d waves in non-uniform grids. Calcolo **55**(47), 1–28 (2018)
9. Bogoya, J.M., Böttcher, A., Grudsky, S.M., Maximenko, E.A.: Eigenvalues of Hermitian Toeplitz matrices with smooth simple-loop symbols. J. Math. Anal. Appl. **442**(2), 1308–1334 (2015)
10. Böttcher, A., Grudsky, S.M.: Toeplitz Matrices, Asymptotic Linear Algebra and Functional Analysis, vol. 67. Springer, Berlin (2000)
11. Böttcher, A., Silbermann, B.: Analysis of Toeplitz Operators. Springer, Berlin (2006)
12. Carasso, A.: Finite-difference methods and the eigenvalue problem for nonselfadjoint Sturm-Liouville operators. Math. Comp. **23**(108), 717–729 (1969)
13. Chiti, G., Pucci, C.: Rearrangements of functions and convergence in Orlicz spaces. Appl. Anal. **9**(1), 23–27 (1979)
14. Chung, K.L.: A Course in Probability Theory. Academic Press, Cambridge (2001)
15. Cottrell, J.A., Hughes, T.J.R., Bazilevs, Y.: Isogeometric Analysis: Toward Integration of CAD and FEA. John Wiley, Hoboken (2009)
16. Courant, R., Friedrichs, K., Lewy, H.: Über die partiellen Differenzengleichungen der mathematischen Physik. Math. Ann. **100**(1), 32–74 (1928)
17. Davies, E.B.: Spectral Theory and Differential Operators. Cambridge University Press, Cambridge (1996)
18. Di Benedetto, F., Fiorentino, G., Serra-Capizzano, S.: CG preconditioning for Toeplitz matrices. Comput. Math. Appl. **25**(6), 35–45 (1993)
19. 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. **25**(5), e2198 (2018)
20. Ervedoza, S., Marica, A., Zuazua, E.: Numerical meshes ensuring uniform observability of onedimensional waves: construction and analysis. IMA J. Numer. Anal. **36**, 503–542 (2016)
21. Everitt, W.N., Markus, L.: Boundary Value Problems and Symplectic Algebra for Ordinary Differential and Quasi-differential Operators. American Mathematical Society, Providence (1999)
22. Garoni, C., Serra-Capizzano, S.: Generalized Locally Toeplitz Sequences: Theory and Applications. Springer, Cham (2017)
23. Garoni, C., Serra-Capizzano, S.: Generalized Locally Toeplitz Sequences: Theory and Applications, vol. II. Springer, Cham (2018)
24. Garoni, C., Speleers, H., Ekström, S.-E., Reali, A., Serra-Capizzano, S., Hughes, T.J.R.: Symbol-based analysis of finite element and isogeometric B-spline discretizations of eigenvalue problems: exposition and review. Arch. Computat. Methods Eng. **26**(5), 1639–1690 (2019)
25. Gary, J.: Computing eigenvalues of ordinary differential equations by finite differences. Math. Comp. **19**(91), 365–379 (1965)
26. Grenander, U., Szegö, G.: Toeplitz Forms and their Applications. University of California Press, Berkeley (1958)

27. Hughes, T.J.R., Evans, J.A., Reali, A.: Finite element and NURBS approximations of eigenvalue, boundary-value, and initial-value problems. Comput. Methods Appl. Mech. Eng. **272**, 290–320 (2014)
28. Infante, J.A., Zuazua, E.: Boundary observability for the space semi discretizations of the 1-d wave equation. Math. Model. Num. Ann. **33**, 407–438 (1999)
29. Kallenberg, O.: Foundations of Modern Probability, 3rd edn. Springer Nature, Switzerland AG (2021)
30. Khan, I.R., Ohba, R.: Closed-form expressions for the finite difference approximations of first and higher derivatives based on Taylor series. J. Comput. Appl. Math. **107**(2), 179–193 (1999)
31. Kuipers, L., Niederreiter, H.: Uniform Distribution of Sequences. John Wiley & Sons Inc, New York (1974)
32. Levendorskii, S.: Asymptotic Distribution of Eigenvalues of Differential Operators. Springer Science & Business Media, Berlin (1990)
33. Li, J.: General explicit difference formulas for numerical differentiation. J. Comput. Appl. Math. **183**(1), 29–52 (2005)
34. Limic, V., Limić, N.: Equidistribution, uniform distribution: a probabilist's perspective. Probab. Surv. **15**, 131–155 (2018)
35. Marica, A., Zuazua, E.: Propagation of 1D waves in regular discrete heterogeneous media: a Wigner measure approach. Found. Comput. Math. **15**(6), 1571–1636 (2015)
36. Pólya, G.: Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem. Math. Z. **8**(3–4), 171–181 (1920)
37. Puzyrev, V., Deng, Q., Calo, V.: Spectral approximation properties of isogeometric analysis with variable continuity. Comput. Methods Appl. Mech. Eng. **334**, 22–39 (2018)
38. Safarov, Yu., Vassiliev, D.: The Asymptotic Distribution of Eigenvalues of Partial Differential Operators. American Mathematical Society, Providence (1997)
39. Serra-Capizzano, S.: An ergodic theorem for classes of preconditioned matrices. Linear Algebra Appl. **282**(1–3), 161–183 (1998)
40. Serra-Capizzano, S.: A note on the asymptotic spectra of finite difference discretizations of second order elliptic partial differential equations. Asian J. Math. **4**(3), 499–514 (2000)
41. Serra-Capizzano, S.: Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations. Linear Algebra Appl. **366**, 371–402 (2003)
42. Serra-Capizzano, S.: The GLT class as a generalized Fourier analysis and applications. Linear Algebra Appl. **419**, 180–233 (2006)
43. Smith, G.D.: Numerical Solution of Partial Differential Equations: Finite Difference Methods, 3rd edn. Clarendon Press, Oxford (1985)
44. Stein, E.M., Weiss, G.: Introduction to Fourier Analysis on Euclidean Spaces. Princeton University Press, Princeton (1971)
45. Talenti, G.: Rearrangements of functions and partial differential equations. In: Fasano, A., Primicerio, M. (eds.) Nonlinear Diffusion Problems. Lecture Notes in Mathematics, pp. 153–178. Springer, Berlin (1986)
46. Talenti, G.: The art of rearranging. Milan J. Math. **84**, 105–157 (2016)
47. Taylor, J.C.: An Introduction to Measure and Probability. Springer, Berlin (1997)
48. Tilli, P.: Locally Toeplitz sequences: spectral properties and applications. Linear Algebra Appl. **278**(1–3), 91–120 (1998)
49. Tyrtyshnikov, E.E.: A unifying approach to some old and new theorems on distribution and clustering. Linear Algebra Appl. **232**, 1–43 (1996)
50. Tyrtyshnikov, E.E., Zamarashkin, N.: Spectra of multilevel Toeplitz matrices: advanced theory via simple matrix relationships. Linear Algebra Appl. **270**, 15–27 (1998)
51. Widom, H.: On the eigenvalues of certain Hermitian operators. Trans. Amer. Math. Soc. **88**(2), 491–522 (1958)
52. Zettl, A.: Sturm-Liouville Theory. American Mathematical Society, Providence (2005)