Received: 29 March 2019 | Revised: 24 September 2019 | Accepted: 26 December 2019
DOI: 10.1002/nla.2286

RESEARCH ARTICLE

WILEY

Non-Hermitian perturbations of Hermitian matrix-sequences and applications to the spectral analysis of the numerical approximation of partial differential equations

Giovanni Barbarino[1] | Stefano Serra-Capizzano[2,3]

[1] Faculty of Sciences, Scuola Normale Superiore, Pisa, Italy

[2] Department of Humanities and Innovation, University of Insubria, Como, Italy

[3] Department of Information Technology, Uppsala University, Uppsala, Sweden

Correspondence
Giovanni Barbarino, Faculty of Sciences, Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy.
Email: giovanni.barbarino@sns.it

Summary

This article concerns the spectral analysis of matrix-sequences which can be written as a non-Hermitian perturbation of a given Hermitian matrix-sequence. The main result reads as follows. Suppose that for every n there is a Hermitian matrix X_n of size n and that $\{X_n\}_n \sim_\lambda f$, that is, the matrix-sequence $\{X_n\}_n$ enjoys an asymptotic spectral distribution, in the Weyl sense, described by a Lebesgue measurable function f; if $\|Y_n\|_2 = o(\sqrt{n})$ with $\|\cdot\|_2$ being the Schatten 2 norm, then $\{X_n + Y_n\}_n \sim_\lambda f$. In a previous article by Leonid Golinskii and the second author, a similar result was proved, but under the technical restrictive assumption that the involved matrix-sequences $\{X_n\}_n$ and $\{Y_n\}_n$ are uniformly bounded in spectral norm. Nevertheless, the result had a remarkable impact in the analysis of both spectral distribution and clustering of matrix-sequences arising from various applications, including the numerical approximation of partial differential equations (PDEs) and the preconditioning of PDE discretization matrices. The new result considerably extends the spectral analysis tools provided by the former one, and in fact we are now allowed to analyze linear PDEs with (unbounded) variable coefficients, preconditioned matrix-sequences, and so forth. A few selected applications are considered, extensive numerical experiments are discussed, and a further conjecture is illustrated at the end of the article.

KEYWORDS

approximation of PDEs, perturbation results, preconditioning, spectral distribution in the Weyl sense

1 | INTRODUCTION

A *matrix-sequence* $\{A_n\}_n$ is an ordered collection of complex matrices such that $A_n \in \mathbb{C}^{n\times n}$ and n belongs to $\mathbb{N}^+$ or to an infinite subset of $\mathbb{N}^+$, with $\mathbb{C}$ being the complex field and $\mathbb{N}^+$ being the set of positive integers. It is often observed in practice that matrix-sequences arising from the numerical discretization of linear differential equations possess a *spectral symbol*, that is, a measurable function $f : D \subseteq \mathbb{R}^q \to \mathbb{C}$, $q \geq 1$ describing the asymptotic distribution of the matrices eigenvalues in the Weyl sense,[1–3] meaning that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} F(\lambda_i(A_n)) = \frac{1}{\mu_q(D)} \int_D F(f(x))\, dx$$

holds for every continuous function $F : \mathbb{C} \to \mathbb{C}$ with compact support, where D is a measurable set with finite Lebesgue measure $\mu_q(D) > 0$ and $\lambda_i(A_n)$ are the eigenvalues of A_n. In this case we write

$$\{A_n\}_n \sim_\lambda f.$$

In this article, we prove new results regarding the spectral symbols of matrix-sequences which can be written as a non-Hermitian perturbation of a given Hermitian matrix-sequence. We remind that the knowledge of the spectral symbol has a practical impact in obtaining fine estimates on the convergence speed of Krylov methods,[4,5] when we face the problem of the efficient computation of the solution of large linear systems. Furthermore, especially in the context of generalized locally Toeplitz (GLT) matrix-sequences[2,6,7] arising in the approximation of PDEs, the computation and analysis of the spectral symbol[8–11] have been used for designing efficient solvers combining preconditioning and multigrid/multiiterative methods (see References 12–16 and references therein).

From the point of view of the applications the main result of this article is the following.

Theorem 1. *Let $\{X_n\}_n$ be a matrix-sequence such that each X_n is Hermitian and $\{X_n\}_n \sim_\lambda f$, where f is a measurable function defined on a subset of $\mathbb{R}^q$ for some q, with finite and positive Lebesgue measure. If $\|Y_n\|_2 = o(\sqrt{n})$, with $\|\cdot\|_2$ being the Frobenius norm, then $\{X_n + Y_n\}_n \sim_\lambda f$.*

In a previous article by Leonid Golinskii and the second author, a similar result was proved[17,theorem 3.4], but under the technical restrictive assumption that the involved matrix-sequences $\{X_n\}_n, \{Y_n\}_n$ are uniformly bounded in spectral norm. Nevertheless, the result had a remarkable impact in the analysis of both spectral distribution and clustering of matrix-sequences arising from various applications, including the study of the zeros of orthogonal polynomials[18] and the computation of the spectral symbol of matrix-sequences belonging to the algebra generated by Toeplitz sequences.[19] However, the main application remains the numerical approximation of PDEs (and fractional PDEs) along with the preconditioning of the related discretization matrices.[12–14,20–24] Indeed, when the approximation of a differential operator of order k is considered, the matrices coming from the discretization of the lower order operators are usually of negligible norm and hence, in general, $\{s_k A_n\}_n$ is exactly a sequence that can be written as a Hermitian dominant part plus a perturbation, where $s_k = 1$ if k is even and s_k is the imaginary unit if k is odd.

The new result (Theorem 1) extends in a substantial way the spectral analysis tools delivered by the former one,[17] and in fact we are now allowed to analyze linear PDEs with (unbounded) variable coefficients. Recent studies have also shown that this result is needed when studying the perturbations caused by nonregular domains and hence variation of the discretization that describes the PDEs near the boundary. Furthermore, Theorem 1 has an impact on the GLT theory, because it allows a generalization of property **GLT2** in References 2, p. 4 and 7, p. 6 (see also Section 3). We remind that the GLT matrix-sequences form a $*$-algebra of matrix-sequences including Toeplitz sequences generated by L^1 symbols,[1] their algebra, and virtually any kind of matrix-sequences arising from the approximation by local methods (finite differences, finite elements, finite volumes, isogeometric analysis,[25] etc.) of variable-coefficient differential and fractional operators.

The article is organized as follows. The theory is developed in Section 2. Section 3 contains the essentials of the GLT theory, for dealing with approximated differential operators. In Section 4, a few selected applications are considered. Extensive numerical experiments are discussed in Section 5 to show the correctness of the theory. A final Section 6 summarizes the findings of the article and a conjecture, supported by numerical experiments, is illustrated.

2 PERTURBATION RESULTS

In order to prove our main result, we need to introduce two distances on the space of matrix-sequences and cite some famous bounds. First, we recall a known theorem due to Hoffman and Wielandt.[26,theorem VI.4.1] It shows that the Schatten 2-norm (also called Frobenius norm) of the difference of normal matrices is bounded both from above and from below by the 2-norm of their eigenvalue differences in some order. Due to this result, we can prove a second lemma where the case A is Hermitian and B is any matrix that is taken into consideration. In what follows, we denote by S_n the collection of all permutations σ of the set $\{1, \dots, n\}$.

Theorem 2 (Hoffman-Wielandt). *Let $A, B \in \mathbb{C}^{n\times n}$ be normal matrices. If $\alpha_1, \alpha_2, \dots, \alpha_n$ and $\beta_1, \beta_2, \dots, \beta_n$ are the eigenvalues of A and B, respectively, then*

$$\min_{\sigma\in S_n} \sum_{i=1}^{n} |\alpha_i - \beta_{\sigma(i)}|^2 \le \|A-B\|_2^2 \le \max_{\sigma\in S_n} \sum_{i=1}^{n} |\alpha_i - \beta_{\sigma(i)}|^2.$$

Moreover, if A is Hermitian, $\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_n$ and $\Re(\beta_1) \ge \Re(\beta_2) \ge \cdots \ge \Re(\beta_n)$, then

$$\sum_{i=1}^{n} |\alpha_i - \beta_i|^2 \le \|A-B\|_2^2 \le \sum_{i=1}^{n} |\alpha_i - \beta_{n-i}|^2.$$

Lemma 1. *Let A be a Hermitian matrix and let B be any matrix with eigenvalues $\alpha_1, \alpha_2, \dots, \alpha_n$ and $\beta_1, \beta_2 \dots, \beta_n$, respectively. Suppose that $\alpha_1 \ge \alpha_2 \ge \cdots \ge \alpha_n$ and $\Re(\beta_1) \ge \Re(\beta_2) \ge \cdots \ge \Re(\beta_n)$. In this case,*

$$\left(\sum_{i=1}^{n} |\alpha_i - \beta_i|^2\right)^{1/2} \le \sqrt{2}\|A-B\|_2.$$

Proof. By the Schur normal form, we are allowed to perform a unitary base change in order to transform the matrix B into an upper triangular matrix. Indeed, such transformation does not change the eigenvalues, the Schatten 2-norm and the Hermitian nature of A. With an abuse of notations, we will continue to denote with A, B the matrices after the base change. We can thus decompose B into

$$B = D + \mathrm{i}N + R,$$

where D, N are real diagonal matrices, and R is a strictly upper triangular matrix. Moreover, any square matrix Z can be decomposed in terms of its real and imaginary parts as $Z = \Re(Z) + \mathrm{i}\Im(Z)$, where $\Re(Z) = \frac{1}{2}(Z+Z^*)$ and $\Im(Z) = \frac{1}{2\mathrm{i}}(Z-Z^*)$. Therefore we decompose $A - B$ in terms of its real and imaginary parts, namely,

$$X := \Re(A-B) = A - D - \frac{R+R^*}{2}, \quad Y := \Im(A-B) = -N - \frac{R-R^*}{2\mathrm{i}}.$$

Notice that N, R, and R^* are elementwise disjoint, since N is diagonal, R is strictly upper triangular, and R^* is strictly lower triangular. Moreover, if we decompose $\beta_i = \mu_i + \mathrm{i}\nu_i$, where μ_i and ν_i are real numbers, then we obtain

$$N = \mathrm{diag}(\nu_i)_{i=1,\dots,n}, \quad D = \mathrm{diag}(\mu_i)_{i=1,\dots,n},$$

and, as a consequence, we have

$$\|Y\|_2^2 = \|N\|_2^2 + \frac{1}{4}\left(\|R\|_2^2 + \|R^*\|_2^2\right) = \sum_{i=1}^{n} \nu_i^2 + \frac{1}{2}\|R\|_2^2. \tag{1}$$

Using the triangular property of the norm, we can state that

$$\|X\|_2 \ge \|A-D\|_2 - \frac{1}{2}\|R+R^*\|_2 = \|A-D\|_2 - \frac{1}{\sqrt{2}}\|R\|_2, \tag{2}$$

and due to Theorem 2, we know that

$$\|A-D\|_2^2 \ge \sum_{i=1}^{n} (\alpha_i - \mu_i)^2. \tag{3}$$

In addition, we can use the property of the trace $\mathrm{Tr}(XY) = \mathrm{Tr}(YX)$ in the following sequence of identities:

$$\begin{aligned}\|A - B\|_2^2 &= \|X + \mathrm{i}Y\|_2^2 \\ &= \mathrm{Tr}\left[(X + \mathrm{i}Y)^*(X + \mathrm{i}Y)\right] \\ &= \mathrm{Tr}\left[(X - \mathrm{i}Y)(X + \mathrm{i}Y)\right] \\ &= \mathrm{Tr}(X^2) + \mathrm{Tr}(Y^2) - \mathrm{i}\left[\mathrm{Tr}(XY) - \mathrm{Tr}(YX)\right] \\ &= \|X\|_2^2 + \|Y\|_2^2. \end{aligned} \tag{4}$$

Finally, by exploiting Equations (1)–(4), we prove the result as follows:

$$\begin{aligned}\sum_i |\alpha_i - \beta_i|^2 &= \sum_i (\alpha_i - \mu_i)^2 + \sum_i \nu_i^2 \\ &\le \|A - D\|_2^2 + \|N\|_2^2 \\ &\le \left(\|X\|_2 + \frac{1}{\sqrt{2}}\|R\|_2 \right)^2 + \|Y\|_2^2 - \frac{1}{2}\|R\|_2^2 \\ &= \|X\|_2^2 + \sqrt{2}\|R\|_2\|X\|_2 + \|Y\|_2^2 \\ &\le \|X\|_2^2 + 2\|Y\|_2\|X\|_2 + \|Y\|_2^2 \\ &\le 2\left(\|X\|_2^2 + \|Y\|_2^2\right) \\ &= 2\|A - B\|_2^2. \end{aligned}$$

■

2.1 | Distances on sequences

The next definition introduces a pseudometric on $\mathbb{C}^n$, which is known as the *optimal matching distance*[26,p. 153].

Definition 1. Given $v, w \in \mathbb{C}^n$, the optimal matching distance is defined as

$$d(v, w) := \min_{\sigma \in S_n} \max_{i=1,\dots,n} |v_i - w_{\sigma(i)}|.$$

We can modify the previous metric and introduce a new function d' called *modified optimal matching distance*.

Definition 2. Given $v, w \in \mathbb{C}^n$, the modified optimal matching distance is defined as

$$d'(v, w) := \min_{\sigma \in S_n} \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v - w_\sigma|_i^{\downarrow} \right\},$$

where

$$|v - w_\sigma| = [|v_1 - w_{\sigma(1)}|, |v_2 - w_{\sigma(2)}|, \dots, |v_n - w_{\sigma(n)}|],$$

and $|v - w_\sigma|_i^{\downarrow}$ is the ith largest element in $|v - w_\sigma|$, with the convention $|v - w_\sigma|_{n+1}^{\downarrow} := 0$.

Given $A \in \mathbb{C}^{n \times n}$, let $\Lambda(A) \in \mathbb{C}^n$ be the vector of the eigenvalues. We can extend the distances d, d' to matrices and sequences in the following ways.

Definition 3. Let $d(\cdot,\cdot)$ and $d'(\cdot,\cdot)$ be as in Definitions 1 and 2, respectively. Given $A, B \in \mathbb{C}^{n \times n}$, we define

$$d(A, B) := d(\Lambda(A), \Lambda(B)), \quad d'(A, B) := d'(\Lambda(A), \Lambda(B)).$$

Given two matrix-sequences $\{A_n\}_n, \{B_n\}_n$, we define

$$d(\{A_n\}_n, \{B_n\}_n) := \lim_{n\to\infty} \sup d(A_n, B_n), \quad d'(\{A_n\}_n, \{B_n\}_n) := \lim_{n\to\infty} \sup d'(A_n, B_n).$$

It was proved in References 27 that d' induces a complete pseudometric on the space of matrix-sequences, and one of the main result of the article is reported below.

Theorem 3. *If* $\{A_n\}_n \sim_\lambda f$, *then*

$$d'(\{A_n\}_n, \{B_n\}_n) = 0 \;\Leftrightarrow\; \{B_n\}_n \sim_\lambda f.$$

Notice that $d'(\{A_n\}_n, \{B_n\}_n) \le d(\{A_n\}_n, \{B_n\}_n)$ for every pair of sequences $\{A_n\}_n, \{B_n\}_n$, and hence we have the following corollary.

Corollary 1. *If* $\{A_n\}_n \sim_\lambda f$, *then*

$$d(\{A_n\}_n, \{B_n\}_n) = 0 \;\Rightarrow\; \{B_n\}_n \sim_\lambda f.$$

2.2 | Proof of the main result

We are now ready to prove our main result (Theorem 1).

Proof of Theorem 1. Throughout this proof, the eigenvalues of X_n will be denoted by $\lambda_1, \dots, \lambda_n$ and the eigenvalues of $X_n + Y_n$ by $\mu_1, \dots, \mu_n$, where we suppose that $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_n$ and $\Re(\mu_1) \ge \Re(\mu_2) \ge \cdots \ge \Re(\mu_n)$. Due to Lemma 1, we know that

$$\left(\sum_{i=1}^{n} |\lambda_i - \mu_i|^2\right)^{1/2} \le \sqrt{2}\|X_n - (X_n + Y_n)\|_2 = \sqrt{2}\|Y_n\|_2.$$

If k_n is the number of indices $i \in \{1, \dots, n\}$ such that $|\lambda_i - \mu_i| > \varepsilon > 0$, then

$$\sqrt{k_n}\varepsilon \le \left(\sum_{i=1}^{n} |\lambda_i - \mu_i|^2\right)^{1/2} \le \sqrt{2}\|Y_n\|_2 \;\Rightarrow\; \frac{k_n}{n} \le 2\left(\frac{\|Y_n\|_2}{\sqrt{n}\varepsilon}\right)^2 \xrightarrow{n\to\infty} 0.$$

The last relation implies that

$$d'(\{X_n\}_n, \{X_n + Y_n\}_n) \le \lim_{n\to\infty} \sup \frac{k_n}{n} + \varepsilon = \varepsilon$$

for every $\varepsilon > 0$, and Theorem 3 allows one to conclude that $\{X_n + Y_n\}_n \sim_\lambda f$. ■

Corollary 2. *Let* $\{X_n\}_n$ *be a matrix-sequence such that each* X_n *is Hermitian and* $\{X_n\}_n \sim_\lambda f$, *where* f *is a measurable function defined on a subset of some* $\mathbb{R}^q$ *with finite and positive Lebesgue measure. Suppose that any of the following conditions is met.*

1. $\|Y_n\|_p = o(\sqrt{n})$ *with* $\|\cdot\|_p$ *being the Schatten p-norm for some* $1 \le p \le 2$.
2. $\|Y_n\| = o(1)$.

Then $\{X_n + Y_n\}_n \sim_\lambda f$.

Proof.

1. For any matrix A, we have $\|A\|_p \ge \|A\|_2$ whenever $1 \le p \le 2$. Hence,

$$\|Y_n\|_p = o(\sqrt{n}) \;\Rightarrow\; \|Y_n\|_2 = o(\sqrt{n}),$$

and the thesis follows from Theorem 1.

2. By definition of Schatten 2-norm,

$$\|Y_n\|_2 \leq \sqrt{n}\|Y_n\| = o(\sqrt{n}),$$

and the thesis follows from Theorem 1.

■

We remind the hypotheses on the sequences in Reference 17, theorem 3.4, that is, $\|X_n\|, \|Y_n\| \leq C$ and $\|Y_n\|_1 = o(n)$. Consequently, it is evident that Corollary 2 is not a direct generalization of Reference 17, theorem 3.4, since we changed the perturbation norm from $\|Y_n\|_1 = o(n)$ to $\|Y_n\|_1 = o(\sqrt{n})$, even if it permits to eliminate the assumption of boundedness for both sequences. Nevertheless, another corollary of Theorem 1 has the same order of perturbation for $\{Y_n\}_n$, but it reintroduces the upper bound condition for $\|Y_n\|$.

Corollary 3. *Let $\{X_n\}_n$ be a matrix-sequence such that each X_n is Hermitian and $\{X_n\}_n \sim_\lambda f$, where f is a measurable function defined on a subset of some $\mathbb{R}^q$ with finite and positive Lebesgue measure. Suppose that both the following conditions are met.*

1. *$\|Y_n\|_1 = o(n)$ with $\|\cdot\|_1$ being the Schatten 1-norm.*
2. *$\|Y_n\| = O(1)$.*

Then $\{X_n + Y_n\}_n \sim_\lambda f$.

Proof. The condition $\|Y_n\| = O(1)$ ensures the existence of a constant C such that $\|Y_n\| < C$ for every n. If we fix $\varepsilon > 0$, we can consider the number k_n of singular values of Y_n greater than ε. The first condition leads to

$$k_n\varepsilon \leq \|Y_n\|_1 = o(n) \ \Rightarrow \ k_n = o(n),$$

and therefore

$$\frac{\|Y_n\|_2}{\sqrt{n}} \leq \frac{\sqrt{C^2k_n + \varepsilon^2(n-k_n)}}{\sqrt{n}} = \sqrt{(C^2-\varepsilon^2)\frac{k_n}{n} + \varepsilon^2} = \varepsilon + o(1).$$

The last relation holds for every $\varepsilon > 0$, so we can conclude that $\|Y_n\|_2 = o(\sqrt{n})$ and the thesis follows from Theorem 1. ■

3 | GLT SEQUENCES

Along with the concept of spectral symbol already introduced, we need to recall the notion of *singular values symbol*, that is, a measurable function describing the asymptotic distribution of the singular values of a matrix-sequence. Given a matrix-sequence $\{A_n\}_n$, a singular value symbol associated with $\{A_n\}_n$ is a measurable function $f : D \subseteq \mathbb{R}^q \to \mathbb{C}$ satisfying

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} F(\sigma_i(A_n)) = \frac{1}{\mu_q(D)}\int_D F(|f(x)|)\, dx$$

for every continuous function $F : \mathbb{R} \to \mathbb{C}$ with compact support, where D is a measurable set with finite Lebesgue measure $\mu_q(D) > 0$ and $\sigma_i(A_n)$ are the singular values of A_n. In this case we write

$$\{A_n\}_n \sim_\sigma f.$$

In order to understand the applications that we will present below, we need to introduce a handy tool devised for solving the problem of computing/analyzing the spectral distribution of matrices arising from the numerical discretization of integrodifferential equations. It is often observed in practice that matrix-sequences $\{A_n\}_n$ arising from the discretization of such equations belong to the class of the so-called GLT sequences, and in particular they enjoy an asymptotic singular value and eigenvalue distribution as $n \to \infty$; we refer the reader to Reference 28 for a nice introduction to this subject and to References 2,3,7, and 29–33 for more advanced studies. Here we simply summarize the main properties of the theory of GLT sequences, both in the case where the considered integrodifferential equation is univariate and multivariate.

3.1 | Unidimensional case

A GLT sequence $\{A_n\}_n$ is a special matrix-sequence equipped with one of its singular values symbols κ, which is referred to as the *GLT symbol* of $\{A_n\}_n$ and is defined over the domain $D = [0,1] \times [-\pi, \pi]$. A point of D is often denoted by (x, θ), and we use the notation $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ to indicate that $\{A_n\}_n$ is a GLT sequence with symbol κ. The symbol of a GLT sequence is unique in the sense that if $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ and $\{A_n\}_n \sim_{\mathrm{GLT}} \xi$ then $\kappa = \xi$ almost everywhere (a.e.) in $[0,1] \times [-\pi, \pi]$. The main properties of GLT sequences are summarized below.

GLT 1. If $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ then $\{A_n\}_n \sim_\sigma \kappa$. If $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ and each A_n is Hermitian then $\{A_n\}_n \sim_\lambda \kappa$.

GLT 2. If $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ and $A_n = X_n + Y_n$, where

- every X_n is Hermitian,
- $\|X_n\|, \|Y_n\| \leq C$ for some constant C independent of n,
- $n^{-1}\|Y_n\|_1 \to 0$,

then $\{A_n\}_n \sim_\lambda \kappa$.

GLT 3. Here we list three fundamental examples of GLT sequences.

- Given a function f in $L^1([-\pi, \pi])$, its associated Toeplitz sequence is $\{T_n(f)\}_n$, where

$$T_n(f) = [f_{i-j}]_{i,j=1}^n, \quad f_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\theta) e^{-\mathrm{i}k\theta}\, d\theta.$$

  $\{T_n(f)\}_n$ is a GLT sequence with symbol $\kappa(x, \theta) = f(\theta)$.
- Given any a.e. continuous function $a : [0,1] \to \mathbb{C}$, its associated diagonal sampling sequence is $\{D_n(a)\}_n$, where

$$D_n(a) = \mathrm{diag}_{i=1,\dots,n}\, a\left(\frac{i}{n}\right).$$

  $\{D_n(a)\}_n$ is a GLT sequence with symbol $\kappa(x, \theta) = a(x)$.
- A zero-distributed sequence is a matrix-sequence such that $\{Z_n\}_n \sim_\sigma 0$, that is,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} F(\sigma_i(A_n)) = F(0),$$

  for every continuous function $F : \mathbb{R} \to \mathbb{C}$ with compact support. Any zero-distributed sequence is a GLT sequence with symbol $\kappa(x, \theta) = 0$.

GLT 4. If $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ and $\{B_n\}_n \sim_{\mathrm{GLT}} \xi$, then

- $\{A_n^*\}_n \sim_{\mathrm{GLT}} \overline{\kappa}$, where A_n^* is the conjugate transpose of A_n,
- $\{\alpha A_n + \beta B_n\}_n \sim_{\mathrm{GLT}} \alpha\kappa + \beta\xi$ for all $\alpha, \beta \in \mathbb{C}$,
- $\{A_n B_n\}_n \sim_{\mathrm{GLT}} \kappa\xi$.

GLT 5. If $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ and $\kappa \neq 0$ a.e., then $\{A_n^\dagger\}_n \sim_{\mathrm{GLT}} \kappa^{-1}$, where $A_n^\dagger$ is the Moore–Penrose pseudoinverse of A_n.

GLT 6. If $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ and each A_n is Hermitian, then $\{f(A_n)\}_n \sim_{\mathrm{GLT}} f(\kappa)$ for all continuous functions $f : \mathbb{C} \to \mathbb{C}$.

GLT 7. $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ if and only if there exist GLT sequences $\{B_{n,m}\}_n \sim_{\mathrm{GLT}} \kappa_m$ such that κ_m converges to κ in measure and $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ as $m \to \infty$.

GLT 8. Suppose $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ and $\{B_{n,m}\}_n \sim_{\mathrm{GLT}} \kappa_m$, where both A_n and $B_{n,m}$ have the same size. Then, $\{B_{n,m}\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ as $m \to \infty$ if and only if κ_m converges to κ in measure.

GLT 9. If $\{A_n\}_n \sim_{\mathrm{GLT}} \kappa$ then there exist functions $a_{i,m}, f_{i,m}$, $i = 1, \dots, N_m$, such that

- $a_{i,m} \in C^\infty([0,1])$ and $f_{i,m}$ is a trigonometric polynomial,
- $\sum_{i=1}^{N_m} a_{i,m}(x) f_{i,m}(\theta)$ converges to $\kappa(x,\theta)$ a.e.,
- $\left\{ \sum_{i=1}^{N_m} D_n(a_{i,m}) T_n(f_{i,m}) \right\}_n \xrightarrow{\text{a.c.s.}} \{A_n\}_n$ as $m \to \infty$.

It will not be necessary to introduce the a.c.s. convergence used in **GLT 7–GLT 9**, but we refer the reader to Reference 34 for the original definition and to References 2, chap 5, 33, and 35 for a detailed exploration of the topic.

3.2 | Multidimensional case

Similarly to the one-dimensional case, a multilevel GLT sequence $\{A_{\boldsymbol{n}}\}_n$ is a sequence of matrices with increasing size, equipped with one of its singular values symbols κ, which is referred to as the *GLT symbol* and is defined over a domain D of the form $[0,1]^q \times [-\pi,\pi]^q$, $q \geq 1$. A point of $D = [0,1]^q \times [-\pi,\pi]^q$ is usually denoted by $(\mathbf{x}, \boldsymbol{\theta})$, where $\mathbf{x} = (x_1, \dots, x_q)$ and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_q)$ are vectors of variables.

When dealing with multilevel sequences, matrices, and vectors, we will use the multiindex notation. A multiindex $\boldsymbol{i} \in \mathbb{Z}^q$, also called a q-index, is simply a vector in $\mathbb{Z}^q$; its components are denoted by $i_1, \dots, i_q$.

- $\mathbf{0}$, $\mathbf{1}$, $\mathbf{2}$, … are the vectors of all zeros, all ones, all twos, … (their size will be clear from the context).
- For any q-index $\boldsymbol{m}$, $N(\boldsymbol{m}) = \prod_{j=1}^q m_j$ and $\boldsymbol{m} \to \infty$ means that $\min(\boldsymbol{m}) = \min_{j=1,\dots,q} m_j \to \infty$.
- If $\boldsymbol{h}, \boldsymbol{k}$ are q-indices, $\boldsymbol{h} \leq \boldsymbol{k}$ means that $h_r \leq k_r$ for all $r = 1, \dots, q$, while $\boldsymbol{h} \nleq \boldsymbol{k}$ means that $h_r > k_r$ for at least one $r \in \{1, \dots, q\}$.
- If $\boldsymbol{h}, \boldsymbol{k}$ are q-indices such that $\boldsymbol{h} \leq \boldsymbol{k}$, the multiindex range $\boldsymbol{h}, \dots, \boldsymbol{k}$ is the set $\{\boldsymbol{j} \in \mathbb{Z}^q \ : \ \boldsymbol{h} \leq \boldsymbol{j} \leq \boldsymbol{k}\}$. We assume for the multiindex range $\boldsymbol{h}, \dots, \boldsymbol{k}$ the standard lexicographic ordering:

$$\left[\cdots \left[\left[(j_1, \dots, j_q) \right]_{j_q = h_q, \dots, k_q} \right]_{j_{q-1} = h_{q-1}, \dots, k_{q-1}} \cdots \right]_{j_1 = h_1, \dots, k_1}. \tag{5}$$

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

$$\begin{aligned} &(h_1,h_2),\ (h_1,h_2+1),\ \dots,\ (h_1,k_2),\ (h_1+1,h_2),\ (h_1+1,h_2+1),\ \dots,\ (h_1+1,k_2), \\ &\quad \dots,\ (k_1,h_2),\ (k_1,h_2+1),\ \dots,\ (k_1,k_2). \end{aligned}$$

- When a q-index $\boldsymbol{j}$ varies over a multiindex range $\boldsymbol{h}, \dots, \boldsymbol{k}$ (this is sometimes written as $\boldsymbol{j} = \boldsymbol{h}, \dots, \boldsymbol{k}$), it is understood that $\boldsymbol{j}$ varies from $\boldsymbol{h}$ to $\boldsymbol{k}$ following the specific ordering (5). For instance, if $\boldsymbol{m} \in \mathbb{N}^d$ and if we write $\mathbf{x} = [x_{\boldsymbol{i}}]_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{m}}$, then $\mathbf{x}$ is a vector of size $N(\boldsymbol{m})$ whose components $x_{\boldsymbol{i}}$, $\boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{m}$, are ordered in accordance with Equation (5): the first component is $x_{\mathbf{1}} = x_{(1,\dots,1,1)}$, the second component is $x_{(1,\dots,1,2)}$, and so on until the last component, which is $x_{\boldsymbol{m}} = x_{(m_1,\dots,m_q)}$. Similarly, if $X = [x_{\boldsymbol{ij}}]_{\boldsymbol{i},\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{m}}$, then X is a $N(\boldsymbol{m}) \times N(\boldsymbol{m})$ matrix whose components are indexed by two d-indices $\boldsymbol{i}, \boldsymbol{j}$, both varying from $\mathbf{1}$ to $\boldsymbol{m}$ according to the lexicographic ordering (5).
- Operations involving q-indices that have no meaning in the vector space $\mathbb{Z}^q$ must always be interpreted in the componentwise sense. For instance, $\boldsymbol{ij} = (i_1 j_1, \dots, i_q j_q)$, $\boldsymbol{i}/\boldsymbol{j} = (i_1/j_1, \dots, i_q/j_q)$, and so forth.

In this context, by a sequence of matrices (or matrix-sequence), we mean a sequence of the form $\{A_{\boldsymbol{n}}\}_n$, where $\boldsymbol{n} = (n_1, \dots, n_d)$ depends on n and $\boldsymbol{n} \to \infty$ as $n \to \infty$. In many cases, it is natural to assume that $\boldsymbol{n} = n\mathbf{c}$, where c is a vector of rational constants and n diverges to infinity. It is always understood that a matrix $A_{\boldsymbol{n}}$ parameterized by a q-index $\boldsymbol{n}$ has dimension $N(\boldsymbol{n}) = n_1 \cdot \ldots \cdot n_q$; its entries will be indexed by two q-indices $\boldsymbol{i}, \boldsymbol{j}$.

The main theoretical properties of one-dimensional GLT sequences **GLT 1–GLT 9** still hold in the multidimensional context, upon substituting the sequences $\{A_n\}_n$ with the multilevel sequences $\{A_{\boldsymbol{n}}\}_n$. The only exception is **GLT 3**, that has to be rewritten in order to include q-level Toeplitz matrices generated by an L^1 q-variate function and q-level diagonal sampling matrices associated with an a.e. continuous q-variate function.

GLT 3. Here we list three important examples of GLT sequences.

- Given a function f in $L^1([-\pi,\pi]^q)$, its associated Toeplitz sequence is $\{T_{\boldsymbol{n}}(f)\}_n$, where the elements are multidimensional Fourier coefficients of f:

$$T_{\boldsymbol{n}}(f) = [f_{\boldsymbol{i}-\boldsymbol{j}}]_{\boldsymbol{i},\boldsymbol{j}=\boldsymbol{1}}^{\boldsymbol{n}}, \quad f_{\boldsymbol{k}} = \frac{1}{(2\pi)^q}\int_{-\pi}^{\pi} f(\boldsymbol{\theta})e^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\theta}}\, d\boldsymbol{\theta}.$$

  $\{T_{\boldsymbol{n}}(f)\}_n$ is a GLT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta}) = f(\boldsymbol{\theta})$.

- Given an a.e. continuous function, $a : [0,1]^q \to \mathbb{C}$, its associated diagonal sampling sequence $\{D_{\boldsymbol{n}}(a)\}_n$ is defined as

$$D_{\boldsymbol{n}}(a) = \mathrm{diag}\left(\left\{a\left(\frac{\boldsymbol{i}}{\boldsymbol{n}}\right)\right\}_{\boldsymbol{i}=1}^{\boldsymbol{n}}\right).$$

  $\{D_{\boldsymbol{n}}(a)\}_n$ is a GLT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta}) = a(\mathbf{x})$.

- Any zero-distributed sequence $\{Z_{\boldsymbol{n}}\}_n \sim_\sigma 0$ is a GLT sequence with symbol $\kappa(\mathbf{x},\boldsymbol{\theta}) = 0$.

In the next sections, we will see examples of matrix-sequences arising from relevant applications in which Theorem 1 is essential for deducing the eigenvalue distribution.

4 | A FEW APPLICATIONS

The section is divided in four subsections. Subsections 4.1 and 4.2 deal with matrix-sequences coming from the approximation of the same second-order one-dimensional differential equation, by using basic finite differences and linear finite elements, respectively. Subsections 4.3 and 4.4 are concerned with matrix-sequences coming from the approximation of a d dimensional PDE and a preconditioning problem, respectively. We stress that a similar analysis can be performed in several other approximation contexts, following the same steps and with the very same tools: higher order finite elements, higher order finite differences, isogeometric analysis, finite volumes, and so forth. With regard to the finite volumes approximation class, we recall that for convection dominated convection-diffusion-reaction problems, finite volumes represent the most appropriate choice.

Notice that the applications presented here have all been discussed in previous works, as Reference 2, but with different hypotheses on the coefficients of the PDEs. In particular, up till now, we had to suppose that the variable coefficients of the differential equations were bounded, since we needed a bound to the spectral norm of the matrices. Thanks to the powerful and more general hypotheses of Theorem 1, we are now able to find spectral properties of PDE discretizations even with unbounded variable coefficients.

4.1 | Finite differences: The one-dimensional setting

We consider the second-order differential equation with Dirichlet boundary conditions

$$\begin{cases} -(a(x)u'(x))' + b(x)u'(x) + c(x)u(x) = f(x), \; x \in (0,1), \\ u(0) = \alpha, \quad u(1) = \beta. \end{cases} \tag{6}$$

The well-posedness of the problem is guaranteed when $a(x) \in C^1(0,1)$, and the uniqueness and existence of the solution are guaranteed in the case $a(x) > 0$, $b(x) \geq 0$ and with continuous functions $b(x)$, $c(x)$ on $[0,1]$, and $f(x) \in L^2([0,1])$.[36] Both the elliptic and the parabolic equations with unbounded coefficients have been analyzed recently by analytical and probabilistic methods. For a discussion about the conditions of existence and uniqueness, even in the multidimensional case (18), we refer to References 37–40 and references therein. For the GLT analysis presented in this section, we only require the following assumptions.

- $a(x), c(x)$ are real-valued functions, continuous a.e., defined in $[0, 1]$,
- $b(x)$ is a real-valued function on $[0, 1]$, such that $|b(x)x^{\alpha}|$ is bounded for some $\alpha < 3/2$,

while $f(x)$ is a general function.

We employ central second-order finite differences for approximating the given equation. We define the stepsize $h = \frac{1}{n+1}$ and the points $x_k = kh$ for k belonging to the interval $[0, n+1]$. For every $j = 1, \dots, n$ we have

$$\begin{aligned}
-(a(x)u'(x))'|_{x=x_j} &\approx -\frac{a\left(x_{j+\frac{1}{2}}\right)u'\left(x_{j+\frac{1}{2}}\right) - a\left(x_{j-\frac{1}{2}}\right)u'\left(x_{j-\frac{1}{2}}\right)}{h} \\
&\approx -\frac{a\left(x_{j+\frac{1}{2}}\right)\frac{u(x_{j+1})-u(x_j)}{h} - a\left(x_{j-\frac{1}{2}}\right)\frac{u(x_j)-u(x_{j-1})}{h}}{h} \\
&= \frac{-a\left(x_{j+\frac{1}{2}}\right)u(x_{j+1}) + \left(a\left(x_{j+\frac{1}{2}}\right) + a\left(x_{j-\frac{1}{2}}\right)\right)u(x_j) - a\left(x_{j-\frac{1}{2}}\right)u(x_{j-1})}{h^2} \\
b(x)u'(x)|_{x=x_j} &\approx b(x_j)\frac{u(x_{j+1}) - u(x_{j-1})}{2h} \\
c(x)u(x)|_{x=x_j} &= c(x_j)u(x_j).
\end{aligned}$$

Let $a_k := a(x_{\frac{k}{2}})$ for any $k \in [0, 2n+2]$ and set $b_j := b(x_j)$, $c_j := c(x_j)$, $f_j := f(x_j)$ for every $j = 0, \dots, n+1$. We compute approximations u_j of the values $u(x_j)$ for $j = 1, \dots, n$ by solving the following linear system

$$A_n \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{pmatrix} + B_n \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{pmatrix} + C_n \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n-1} \\ u_n \end{pmatrix} = h^2 \begin{pmatrix} f_1 + \frac{1}{h^2}a_1\alpha + \frac{1}{2h}b_1\alpha \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n + \frac{1}{h^2}a_{2n+1}\beta - \frac{1}{2h}b_n\beta \end{pmatrix},$$

where

$$A_n = \begin{pmatrix} a_1 + a_3 & -a_3 & & & \\ -a_3 & a_3 + a_5 & -a_5 & & \\ & \ddots & \ddots & \ddots & \\ & & -a_{2n-3} & a_{2n-3} + a_{2n-1} & -a_{2n-1} \\ & & & -a_{2n-1} & a_{2n-1} + a_{2n+1} \end{pmatrix}, \tag{7}$$

$$B_n = \frac{h}{2}\begin{pmatrix} 0 & b_1 & & & \\ -b_2 & 0 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & -b_{n-1} & 0 & b_{n-1} \\ & & & -b_n & 0 \end{pmatrix}, \quad C_n = h^2\mathrm{diag}(c_1, \dots, c_n). \tag{8}$$

In the case where $a(x) \equiv 1$ and $b(x) \equiv 1$, in accordance with the notation of axiom **GLT 3** in Subsection 3.1, we obtain the basic Toeplitz structures

$$K_n = T_n(2 - 2\cos(\theta)) = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \tag{9}$$

$$H_n = T_n(\mathrm{i}\sin(\theta)) = \frac{1}{2}\begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}, \tag{10}$$

with $A_n = K_n$ and $B_n = hH_n$. In this case, it is also useful consider the first-order noncentral discretization operators, which, after a proper scaling, can be written as

$$K_n^+ = T_n(1 - e^{-\mathrm{i}\theta}) = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ & & & & 1 \end{pmatrix},$$

$$K_n^- = T_n(1 - e^{\mathrm{i}\theta}) = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix}. \tag{11}$$

By setting

$$D_n^+(a) := \mathrm{diag}(a_3, a_5, \dots, a_{2n+1}), \quad D_n^-(a) := \mathrm{diag}(a_1, a_3, \dots, a_{2n-1}),$$

the matrix-sequences $\{D_n^+(a)\}_n$ and $\{D_n^-(a)\}_n$ are GLT sequences with symbol $a(x)$ and the proof is virtually identical to that for showing $\{D_n(a)\}_n \sim_{\mathrm{GLT}} a(x)$.[2] As a consequence of the algebra structure of GLT sequences, that is, using axioms **GLT 3** and **GLT 4**, we obtain

$$A_n = D_n^+(a)K_n^+ + D_n^-(a)K_n^- \;\Rightarrow\; \{A_n\}_n \sim_{\mathrm{GLT}} a(x)(2 - 2\cos(\theta)).$$

Furthermore, regarding the matrices C_n, we have

$$\{D_n(c)\} \sim_{\mathrm{GLT}} c(x) \quad \{n^{-2}I_n\}_n \sim_{\mathrm{GLT}} 0 \;\Rightarrow\; \{C_n\}_n = \{n^{-2}D_n(c)\}_n \sim_{\mathrm{GLT}} 0,$$

which, again by axioms **GLT 3** and **GLT 4**, implies that $\{A_n\}_n + \{C_n\}_n \sim_{\mathrm{GLT}} a(x)(2 - 2\cos(\theta))$. By exploiting the real symmetry of all the considered matrices and in view of axiom **GLT 1**, we obtain

$$\{A_n\}_n + \{C_n\}_n \sim_\lambda a(x)(2 - 2\cos(\theta)).$$

For the matrices B_n, taking into account Theorem 1, we are interested in estimating their Schatten 2-norm (Frobenius norm in the numerical analysis community). Suppose that there exists a constant $C > 0$ such that $|b(x)x^\alpha| < C$. If $\alpha < 1$, then $|b(x)x| \leq |b(x)x^\alpha| < C$, so we analyze only the case $\alpha \geq 1$. We have

$$\|B_n\|_2^2 \leq \frac{h^2}{2}\sum_{i=1}^{n} b_i^2 = \frac{h^2}{2}\sum_{i=1}^{n} b(ih)^2 \leq \frac{C^2h^2}{2}\sum_{i=1}^{n}(ih)^{-2\alpha} = \frac{C^2h^{2-2\alpha}}{2}\sum_{i=1}^{n} i^{-2\alpha}. \tag{12}$$

Since $-2\alpha \leq -2$, we can estimate the last sum with the integral of $x^{2\alpha+1}$, in the following way

$$\sum_{i=1}^{n} i^{-2\alpha} \leq 1 + \int_1^n x^{-2\alpha}\, dx = \frac{2\alpha - n^{-2\alpha+1}}{2\alpha - 1}. \tag{13}$$

Substituting (13) into (12), we obtain

$$\|B_n\|_2^2 \leq \frac{C^2h^{2-2\alpha}}{2}\sum_{i=1}^{n} i^{-2\alpha} \leq \frac{C^2\alpha}{2\alpha-1}h^{2-2\alpha} - \frac{C^2}{4\alpha-2}h,$$

which implies

$$\|B_n\|_2 = O(n^{\alpha-1}). \tag{14}$$

Observe that $\frac{3}{2} > \alpha$ leads to $\|B_n\|_2 = o(\sqrt{n})$. By invoking Theorem 1, we simply conclude

$$\{A_n\}_n + \{B_n\}_n + \{C_n\}_n \sim_\lambda a(x)(2-2\cos(\theta)),$$

with $\{X_n\}_n = \{A_n\}_n + \{C_n\}_n$, $\{Y_n\}_n = \{B_n\}_n$, and where the matrix-sequence $\{B_n\}_n$ is nonsymmetric.

4.2 | Finite elements: The one-dimensional setting

Let us consider the same Equation (6) as in the previous subsection, but with slightly different conditions on the variable coefficients since we assume $a(x), b(x), c(x)$ simply Lebesgue integrable, while $f(x)$ is a generic function. We write it in weak form

$$\begin{cases} \int_0^1 a(x)u'(x)w'(x) + b(x)u'(x)w(x) + c(x)u(x)w(x)dx = \int_0^1 f(x)w(x)\,dx, & x \in (0,1), \\ u(0)=0, \quad u(1)=0, \end{cases}$$

where $w(x)$ is allowed to belong to the Sobolev space $H_0^1([0,1])$. We set $h = \frac{1}{n+1}$ and $x_k = kh$ for k integer in the interval $[0, n+1]$. We consider the so-called hat functions

$$\varphi_{i,n}(x) = \frac{1}{h}\left[(x-x_{i-1})\chi_{[x_{i-1},x_i)}(x) + (x_{i+1}-x)\chi_{[x_i,x_{i+1})}(x)\right], \quad i = 1,\dots,n,$$

and assume that the functions $u(x)$ and $w(x)$ belong to the linear space spanned by $\varphi_i(x)$, that is,

$$u(x) = \sum_{j=1}^{n} u_{j,n}\varphi_{j,n}(x), \quad w(x) = \sum_{i=1}^{n} w_{i,n}\varphi_{i,n}(x).$$

By substituting these expressions in the weak form, we obtain that the integral

$$\int_0^1 a(x)u'(x)w'(x) + b(x)u'(x)w(x) + c(x)u(x)w(x)\,dx$$

is approximated by

$$\sum_{i,j=1}^{n} u_{j,n}w_{i,n}\int_0^1 a(x)\varphi'_{i,n}(x)\varphi'_{j,n}(x)\,dx + \sum_{i,j=1}^{n} u_{j,n}w_{i,n}\int_0^1 b(x)\varphi_{i,n}(x)\varphi'_{j,n}(x)\,dx$$
$$+\sum_{i,j=1}^{n} u_{j,n}w_{i,n}\int_0^1 c(x)\varphi_{i,n}(x)\varphi_{j,n}(x)\,dx,$$

while

$$\int_0^1 f(x)w(x)\,dx = \sum_{j=1}^{n} w_{i,n}\int_0^1 f(x)\varphi_{i,n}(x)\,dx.$$

Therefore, if u^n denotes the vector of the unknowns $u_{i,n}$, w^n the vector of the values $w_{i,n}$, and

$$
\begin{aligned}
A_n &= \left(\int_0^1 a(x)\varphi'_{i,n}(x)\varphi'_{j,n}(x)dx \right)_{i,j}, \\
B_n &= \left(\int_0^1 b(x)\varphi_{i,n}(x)\varphi'_{j,n}(x)dx \right)_{i,j}, \\
C_n &= \left(\int_0^1 c(x)\varphi_{i,n}(x)\varphi_{j,n}(x)dx \right)_{i,j}, \\
f^n &= \left(\int_0^1 f(x)\varphi_{i,n}(x)dx \right)_{i},
\end{aligned}
$$

we deduce that the relationships

$$(w^n)^T(A_n + B_n + C_n)u^n = (w^n)^T f^n$$

have to be satisfied for every w^n. The latter is clearly equivalent to the linear system

$$(A_n + B_n + C_n)u^n = f^n.$$

We notice that

$$\int_0^1 a(x)\varphi'_{i,n}(x)\varphi'_{j,n}(x) = (n+1)^2 \begin{cases} 0 & |i-j| > 1 \\ -\int_{x_i}^{x_{i+1}} a(x)\ dx & j = i+1 \\ -\int_{x_{i-1}}^{x_i} a(x)\ dx & i = j+1 \\ \int_{x_{i-1}}^{x_{i+1}} a(x)\ dx & i = j \end{cases}$$

$$\int_0^1 b(x)\varphi_{i,n}(x)\varphi'_{j,n}(x) = (n+1) \begin{cases} 0 & |i-j| > 1 \\ -\int_{x_{i-1}}^{x_i} b(x)\varphi_{i,n}(x)\ dx & i = j+1 \\ \int_{x_i}^{x_{i+1}} b(x)\varphi_{i,n}(x)\ dx & j = i+1 \\ \int_{x_{i-1}}^{x_i} b(x)\varphi_{i,n}(x)\ dx - \int_{x_i}^{x_{i+1}} b(x)\varphi_{i,n}(x)\ dx & i = j \end{cases}$$

Let us compute the Schatten 2-norm of B_n:

$$
\begin{aligned}
\frac{1}{n+1}\|B_n\|_2 &= \frac{1}{n+1}\sqrt{\sum_{i,j=1}^{n}(B_n)_{i,j}^2} \le \frac{1}{n+1}\sum_{i,j=1}^{n}|(B_n)_{i,j}| \\
&= \sum_{i=1}^{n-1}\left|\int_{x_i}^{x_{i+1}} b(x)\varphi_{i,n}(x)\ dx\right| + \sum_{i=2}^{n}\left|\int_{x_{i-1}}^{x_i} b(x)\varphi_{i,n}(x)\ dx\right| \\
&\quad + \sum_{i=1}^{n}\left|\int_{x_{i-1}}^{x_i} b(x)\varphi_{i,n}(x)\ dx - \int_{x_i}^{x_{i+1}} b(x)\varphi_{i,n}(x)\ dx\right| \\
&\le \int_{x_0}^{x_1} |b(x)|\varphi_{1,n}(x)\ dx + \int_{x_n}^{x_{n+1}} |b(x)|\varphi_{n,n}(x)\ dx \\
&\quad + \sum_{i=1}^{n-1}\int_{x_i}^{x_{i+1}} 2|b(x)|(\varphi_{i+1,n}(x) + \varphi_{i,n}(x))\ dx \\
&\le 2\sum_{i=0}^{n}\int_{x_i}^{x_{i+1}} |b(x)|dx = 2\|b(x)\|_1.
\end{aligned}
\tag{15}
$$

From the inequalities above, we have

$$\left\| \frac{1}{n+1} B_n \right\|_2 = o(\sqrt{n}).$$

The matrices A_n and C_n are all real symmetric and, according to Reference 2, exercise 10.4, we already know that

$$\left\{ \frac{1}{n+1}(A_n + C_n) \right\}_n \sim_{\text{GLT}} a(x)(2 - 2\cos(\theta))$$

and

$$\left\{ \frac{1}{n+1}(A_n + C_n) \right\}_n \sim_{\lambda} a(x)(2 - 2\cos(\theta)).$$

In conclusion, by applying Theorem 1 with $\{X_n\}_n = \{\frac{1}{n+1}(A_n + C_n)\}_n$ and $\{Y_n\}_n = \{\frac{1}{n+1}B_n\}_n$, we deduce the spectral distribution of the complete (nonsymmetric) matrix-sequence, that is,

$$\left\{ \frac{1}{n+1} A_n \right\}_n + \left\{ \frac{1}{n+1} B_n \right\}_n + \left\{ \frac{1}{n+1} C_n \right\}_n \sim_{\lambda} a(x)(2 - 2\cos(\theta)).$$

4.2.1 | Minimal hypothesis

In the previous section, we actually proved that $\frac{1}{n+1}\|B_n\|_2 = O(1)$, so we still have to exploit the full power of Theorem 1. Suppose that $b(x)$ is a function on $[0, 1]$ with a single discontinuity point in zero of order α, in the sense that

$$|b(x)x^{-\alpha}| \leq C, \quad \forall x > 0,$$

where $C > 0$ is a constant. To ensure that the matrix B_n is well defined, we have to check at least that every element is finite. In the case $\alpha > -2$ we have

$$\left| \int_{x_0}^{x_1} \varphi_1(x)b(x) \right| \leq C \int_{x_0}^{x_1} \varphi_{1,n}(x)x^{\alpha}\, dx = \frac{C}{h} \int_0^h x^{\alpha+1}\, dx < \infty.$$

As a consequence, if $\alpha > -2$ then the matrices B_n are well defined. Furthermore, if $\alpha > -1$ then the function $|b(x)|$ belongs to $L^1(0, 1)$ and this case has already been addressed above. Here we want to explore the case $-2 < \alpha < -1$. With reference to the chain of equalities and inequalities in (15), we have

$$\begin{aligned} \frac{1}{n+1}\|B_n\|_2 &\leq \int_{x_0}^{x_1} |b(x)|\varphi_{1,n}(x)\, dx + \int_{x_n}^{x_{n+1}} |b(x)|\varphi_{n,n}(x)\, dx \\ &\quad + \sum_{i=1}^{n-1} \int_{x_i}^{x_{i+1}} 2|b(x)|(\varphi_{i+1,n}(x) + \varphi_{i,n}(x))\, dx \\ &\leq \int_{x_0}^{x_1} |b(x)|\varphi_{1,n}(x)\, dx + 2\int_{x_1}^{x_{n+1}} |b(x)|\, dx \\ &\leq \frac{C}{h} \int_0^h x^{\alpha+1}\, dx + 2C \int_h^1 x^{\alpha}\, dx = O(n^{-\alpha-1}). \end{aligned} \tag{16}$$

With the same computation, it is easy to show that if $\alpha = -1$ then $\frac{1}{n+1}\|B_n\|_2 = O(\log(n))$. Observe that $-\frac{3}{2} < \alpha$ leads to $\frac{1}{n+1}\|B_n\|_2 = o(\sqrt{n})$ and therefore, by Theorem 1, we conclude again that

$$\left\{ \frac{1}{n+1} A_n \right\}_n + \left\{ \frac{1}{n+1} B_n \right\}_n + \left\{ \frac{1}{n+1} C_n \right\}_n \sim_{\lambda} a(x)(2 - 2\cos(\theta)).$$

4.3 | Finite differences: A second-order approximation for a linear PDE in q dimensions

In this subsection, we extend to the q-dimensional setting the study carried out in Subsection 4.1 and we indicate a general framework for treating q-dimensional problems, including also a systematic recipe for extending the results of Subsection 4.2. One of the main contributions relies on the statement that no substantial difference is encountered when passing from 1 to q space dimensions. Of course, the q-dimensional GLT analysis involves several technicalities, but it is conceptually identical to the one-dimensional GLT analysis. The most important technicality which is encountered when passing from 1 to q space dimensions is the *multiindex language*, which allows one to maintain the one-dimensional notation by simply turning some letters (n, i, j, etc.) in boldface ($\boldsymbol{n}, \boldsymbol{i}, \boldsymbol{j}$, etc.). Regarding notation we remind that $\circ$ will indicate the componentwise product between matrices of the same size, while $\otimes$ will indicate the Kronecker product between matrices of arbitrary possibly different sizes.

Before starting, let us outline the main general ideas of a q-dimensional GLT analysis. Consider for example a linear second-order PDE such as

$$\begin{cases} -\nabla \cdot \mathrm{A}\nabla u + \mathbf{b} \cdot \nabla u + cu = f, & \text{in } (0,1)^q, \\ u = 0, & \text{on } \partial((0,1)^q), \end{cases} \tag{17}$$

where $\mathrm{A} : [0,1]^q \to \mathbb{R}^{q\times q}$ is a symmetric matrix of functions a_{hk}. As in the one-dimensional case, we should assume at least $a_{hk} \in C^1([0,1]^q)$ to ensure the well-posedness of problem (17). For the existence and uniqueness of the solution, we should assume all the coefficient to be bounded, the matrix $A(x)$ to be symmetric and positive definite with the least eigenvalue greater than a fixed value $\lambda_0 > 0$, and that 0 is not an eigenvalue of the corresponding differential operator.[41] For the case of unbounded coefficients, we again point to the references in Section 4.1. For the GLT analysis, we only need the conditions that a_{hk} are continuous a.e. on $[0,1]^q$. $\mathbf{b} : [0,1]^q \to \mathbb{R}^q$ is a vector of real-valued functions $b_k(\mathbf{x})$ on $[0,1]^q$ such that $|b_k(\mathbf{x})(x_1 \dots x_q)^\alpha|$ is bounded by the same constant for all indices k for some exponent $\frac{1}{q} + \frac{1}{2} > \alpha$. We also require that c is a real-valued a.e. continuous function on $[0,1]^q$. We now observe that (17) is equivalent to

$$\begin{cases} -\sum_{h,k=1}^{q} \frac{\partial}{\partial x_h}\left(a_{hk}\frac{\partial u}{\partial x_k}\right) + \sum_{k=1}^{d} b_k \frac{\partial u}{\partial x_k} + cu = f, & \text{in } (0,1)^q, \\ u = 0, & \text{on } \partial((0,1)^q). \end{cases} \tag{18}$$

Assume we discretize (18) by a local method; to fix the ideas, here we will assume that such method is a finite difference (FD) scheme. The resulting discretization matrices $A_{\boldsymbol{n}}$ are parametrized by a multiindex $\boldsymbol{n} = (n_1, \dots, n_q)$, where n_i is related to the discretization step h_i in the ith direction, and $n_i \to \infty$ if and only if $h_i \to 0$ (usually, $h_i \sim 1/n_i$). In order to simplify the notation, we choose $n_i = n$ for some $n \in \mathbb{N}$, that is, $\boldsymbol{n} = (n, \dots, n)$ and, consequently, $\{A_{\boldsymbol{n}}\}_n$ is a matrix-sequence. The matrix $A_{\boldsymbol{n}}$ can be decomposed according to the terms of the PDE as follows:

$$A_{\boldsymbol{n}} = \sum_{h,k=1}^{q} K_{\boldsymbol{n},hk}(a_{hk}) + \sum_{k=1}^{q} H_{\boldsymbol{n},k}(b_k) + I_{\boldsymbol{n}}(c),$$

where $K_{\boldsymbol{n},hk}(a)$, $H_{\boldsymbol{n},k}(b)$, and $I_{\boldsymbol{n}}(c)$ are the matrices resulting from the considered FD discretization of the differential operators

$$-\frac{\partial}{\partial x_h}\left(a\frac{\partial u}{\partial x_k}\right), \quad b\frac{\partial u}{\partial x_k}, \quad cu,$$

respectively. It usually turns out that, after a suitable normalization that we ignore in this discussion, the matrix-sequence $\{H_{\boldsymbol{n},k}(b) + I_{\boldsymbol{n}}(c)\}_n$ associated with the lower order differential operators of the PDE (18) is zero-distributed and the GLT analysis of $\{A_{\boldsymbol{n}}\}_n$ reduces to the GLT analysis of the matrix-sequence

$$\left\{\sum_{h,k=1}^{q} K_{\boldsymbol{n},hk}(a_{hk})\right\}_n$$

associated with the higher order differential operator of the PDE (18). Moreover, the sequences $\{K_{\boldsymbol{n},hk}(a_{hk})\}_{\boldsymbol{n}}$ often turns out to be GLT sequences of the form

$$K_{\boldsymbol{n},hk}(a_{hk}) = D_{\boldsymbol{n}}(a_{hk})T_{\boldsymbol{n}}(p_{hk}) + Z_{\boldsymbol{n},hk}, \quad \{Z_{\boldsymbol{n},hk}\}_{\boldsymbol{n}} \sim_{\sigma} 0,$$

where p_{hk} is the (separable) trigonometric polynomial that represents the FD formula used to discretize the derivative $-\frac{\partial^2 u}{\partial x_h \partial x_k}$. In conclusion, we have

$$\{K_{\boldsymbol{n},hk}\}_{\boldsymbol{n}} \sim_{\text{GLT}} a_{hk}(\mathbf{x})p_{hk}(\boldsymbol{\theta}) \tag{19}$$

and, consequently,

$$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\text{GLT}} \sum_{h,k=1}^{q} a_{hk}(\mathbf{x})p_{hk}(\boldsymbol{\theta}) = \mathbf{1}(\mathrm{A}(\mathbf{x}) \circ H(\boldsymbol{\theta}))\mathbf{1}^T, \tag{20}$$

where

$$H_{hk} = p_{hk}, \quad h, k = 1, \dots, q.$$

From (20) and Theorem 1, we arrive at the distribution relation

$$\{A_{\boldsymbol{n}}\}_{\boldsymbol{n}} \sim_{\lambda} \sum_{h,k=1}^{q} a_{hk}(\mathbf{x})p_{hk}(\boldsymbol{\theta}) = \mathbf{1}(\mathrm{A}(\mathbf{x}) \circ H(\boldsymbol{\theta}))\mathbf{1}^T.$$

In particular, with respect to the original result by Golinskii and the second author,[17,theorem 3.4] Theorem 1 allows us to relax the assumptions on $A(\mathbf{x})$ under which the previous relation holds. More specifically, the coefficients of $A(\mathbf{x})$ can now be supposed to be unbounded.

We note the formal analogy between the expression of the symbol $\mathbf{1}(\mathrm{A}(\mathbf{x}) \circ H(\boldsymbol{\theta}))\mathbf{1}^T$ and the expression of the higher order differential operator $\mathbf{1}(\mathrm{A} \circ Hu)\mathbf{1}^T$ in Equation (18). Because of this analogy, and especially because of (19), the matrix $H(\boldsymbol{\theta})$ in the Fourier variables $\boldsymbol{\theta}$ is usually referred to as the "symbol of the (negative) Hessian operator," although this terminology is clearly not rigorous from the mathematical viewpoint. If we change the FD formulas to discretize the derivative $-\frac{\partial^2 u}{\partial x_h \partial x_k}$, the symbol remains the same except for the matrix $H(\boldsymbol{\theta})$, which now collects the (separable) trigonometric polynomials associated with the new FD formulas—or with the "formulas" associated with the considered local method, which may be for example the finite element (FE) method or the isogeometric analysis (IgA). The only possible difference when passing from FDs to other local methods consists in the choice of the proper scaling, which depends only on the fact that either a Galerkin method (FEs, Galerkin IgA, etc.) or a collocation method (FDs, collocation IgA, etc.) is used: We remark that a slightly different situation occurs when considering finite volumes (see References 42 and 43 for a GLT analysis of matrix-sequences arising in the context of discretizations by finite volumes). This discussion motivates why we do not generalize explicitly the analysis in Subsection 4.2: It would be a plain combination of the analysis for finite elements in one dimension and the analysis in the present subsection.

4.3.1 | FD discretization of convection–diffusion–reaction equations

We consider the classical central FD discretizations of Equation (18). We choose $\boldsymbol{n} \in \mathbb{N}^q$ and we set $\boldsymbol{h} = \frac{\mathbf{1}}{\boldsymbol{n}+\mathbf{1}}$ and $x_{\boldsymbol{j}} = \boldsymbol{jh}$ for $\boldsymbol{j} = \mathbf{0}, \dots, \boldsymbol{n}+\mathbf{1}$.* Let $\mathbf{e}_k$ be the kth vector of the canonical basis of $\mathbb{R}^q$ and notice that $x_{\boldsymbol{j}} + sh_k\mathbf{e}_k = x_{\boldsymbol{j}+s\mathbf{e}_k}$. Then, for $\boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{n}$, we can approximate the terms appearing in (18) as follows:

*Operations involving q-indices in $\mathbb{Z}^q$ must be interpreted in the componentwise sense. In the present case, given $\boldsymbol{n} = (n_1, \dots, n_q)$, the vector of discretization steps $\boldsymbol{h} = \frac{\mathbf{1}}{\boldsymbol{n}+\mathbf{1}}$ and the grid points $x_{\boldsymbol{j}} = \boldsymbol{jh}$ are given by $\boldsymbol{h} = \left(\frac{1}{n_1+1}, \dots, \frac{1}{n_q+1}\right) = (h_1, \dots, h_q)$ and $x_{\boldsymbol{j}} = (j_1 h_1, \dots, j_q h_q)$.

$$\begin{aligned}\frac{\partial}{\partial x_k}\left(a_{kk}\frac{\partial u}{\partial x_k}\right)\bigg|_{x=x_{\mathbf{j}}} &\approx \frac{a_{kk}\frac{\partial u}{\partial x_k}(x_{\mathbf{j}+\mathbf{e}_k/2})-a_{kk}\frac{\partial u}{\partial x_k}(x_{\mathbf{j}-\mathbf{e}_k/2})}{h_k}\\ &\approx a_{kk}(x_{\mathbf{j}+\mathbf{e}_k/2})\frac{u(x_{\mathbf{j}+\mathbf{e}_k})-u(x_{\mathbf{j}})}{h_k^2}-a_{kk}(x_{\mathbf{j}-\mathbf{e}_k/2})\frac{u(x_{\mathbf{j}})-u(x_{\mathbf{j}-\mathbf{e}_k})}{h_k^2}\end{aligned} \tag{21}$$

$$\begin{aligned}\frac{\partial}{\partial x_h}\left(a_{hk}\frac{\partial u}{\partial x_k}\right)\bigg|_{x=x_{\mathbf{j}}} &\approx \frac{a_{hk}\frac{\partial u}{\partial x_k}(x_{\mathbf{j}+\mathbf{e}_h})-a_{hk}\frac{\partial u}{\partial x_k}(x_{\mathbf{j}-\mathbf{e}_h})}{2h_h}\\ &\approx a_{hk}(x_{\mathbf{j}+e_h})\frac{u(x_{\mathbf{j}+\mathbf{e}_h+\mathbf{e}_k})-u(x_{\mathbf{j}+\mathbf{e}_h-\mathbf{e}_k})}{4h_hh_k}-a_{hk}(x_{\mathbf{j}-e_h})\frac{u(x_{\mathbf{j}-\mathbf{e}_h+\mathbf{e}_k})-u(x_{\mathbf{j}-\mathbf{e}_h-\mathbf{e}_k})}{4h_hh_k}\end{aligned} \tag{22}$$

$$b_k\frac{\partial u}{\partial x_k}\bigg|_{x=x_{\mathbf{j}}}\approx b_k(x_{\mathbf{j}})\frac{u(x_{\mathbf{j}+\mathbf{e}_k})-u(x_{\mathbf{j}-\mathbf{e}_k})}{2h_k}, \tag{23}$$

$$cu|_{x=x_{\mathbf{j}}}=c(x_{\mathbf{j}})u(x_{\mathbf{j}}), \tag{24}$$

for $h,k=1,\dots,q,h\neq k$. The evaluations $u(x_{\mathbf{j}})$ of the solution of Equation (18) at the grid points $x_{\mathbf{j}}$ are approximated by the values $u_{\mathbf{j}}$, where $u_{\mathbf{j}}=0$ for $\boldsymbol{j}\in\{\mathbf{0},\dots,\boldsymbol{n}+\mathbf{1}\}\backslash\{\mathbf{1},\dots,\boldsymbol{n}\}$, and the vector $\mathbf{u}=(u_{\mathbf{1}},\dots,u_{\boldsymbol{n}})^T$ is the solution of the linear system

$$\begin{aligned}&-\sum_{k=1}^{q}a_{kk}(x_{\mathbf{j}+\mathbf{e}_k/2})\frac{u_{\mathbf{j}+\mathbf{e}_k}-u_{\mathbf{j}}}{h_k^2}-a_{kk}(x_{\mathbf{j}-\mathbf{e}_k/2})\frac{u_{\mathbf{j}}-u_{\mathbf{j}-\mathbf{e}_k}}{h_k^2}\\ &-\sum_{\substack{h,k=1\\h\neq k}}^{q}a_{hk}(x_{\mathbf{j}+e_h})\frac{u_{\mathbf{j}+\mathbf{e}_h+\mathbf{e}_k}-u_{\mathbf{j}+\mathbf{e}_h-\mathbf{e}_k}}{4h_hh_k}-a_{hk}(x_{\mathbf{j}-e_h})\frac{u_{\mathbf{j}-\mathbf{e}_h+\mathbf{e}_k}-u_{\mathbf{j}-\mathbf{e}_h-\mathbf{e}_k}}{4h_hh_k}\\ &+\sum_{k=1}^{q}b_k(x_{\mathbf{j}})\frac{u_{\mathbf{j}+\mathbf{e}_k}-u_{\mathbf{j}-\mathbf{e}_k}}{2h_k}+c(x_{\mathbf{j}})u_{\mathbf{j}}=f(x_{\mathbf{j}}),\quad \boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}.\end{aligned} \tag{25}$$

We now want to understand the structure of the matrix $A_{\boldsymbol{n}}$ associated with the linear system (25). This is clearly important for the GLT analysis of the next paragraph. Luckily, the multiindex language allows us to provide a compact and easy-to-manage expression of this matrix. First, we note that $A_{\boldsymbol{n}}$ admits the following natural decomposition:

$$\begin{aligned}A_{\boldsymbol{n}}=&\sum_{k=1}^{q}\frac{1}{h_k^2}\left(\operatorname{diag}_{\mathbf{j}=\mathbf{1},\dots,\boldsymbol{n}}a_{kk}(x_{\mathbf{j}+\mathbf{e}_k/2})\right)K_{\boldsymbol{n},kk}^{+}\\ &+\sum_{k=1}^{q}\frac{1}{h_k^2}\left(\operatorname{diag}_{\mathbf{j}=\mathbf{1},\dots,\boldsymbol{n}}a_{kk}(x_{\mathbf{j}-\mathbf{e}_k/2})\right)K_{\boldsymbol{n},kk}^{-}\\ &+\sum_{\substack{h,k=1\\h\neq k}}^{q}\frac{1}{h_hh_k}\left(\operatorname{diag}_{\mathbf{j}=\mathbf{1},\dots,\boldsymbol{n}}a_{hk}(x_{\mathbf{j}+\mathbf{e}_h})\right)K_{\boldsymbol{n},hk}^{+}\\ &+\sum_{\substack{h,k=1\\h\neq k}}^{q}\frac{1}{h_hh_k}\left(\operatorname{diag}_{\mathbf{j}=\mathbf{1},\dots,\boldsymbol{n}}a_{hk}(x_{\mathbf{j}-\mathbf{e}_h})\right)K_{\boldsymbol{n},hk}^{-}\\ &+\sum_{k=1}^{q}\frac{1}{h_k}\left(\operatorname{diag}_{\mathbf{j}=\mathbf{1},\dots,\boldsymbol{n}}b_k(x_{\mathbf{j}})\right)H_{\boldsymbol{n},k}+\left(\operatorname{diag}_{\mathbf{j}=\mathbf{1},\dots,\boldsymbol{n}}c(x_{\mathbf{j}})\right),\end{aligned} \tag{26}$$

where the matrices $K_{\boldsymbol{n},hk}^{\pm}$ and $H_{\boldsymbol{n},k}$ are defined by their action on a generic vector $\mathbf{u}\in\mathbb{R}^{N(\boldsymbol{n})}$, as follows:

$$(K^{\pm}_{\boldsymbol{n},kk}\mathbf{u})_{\boldsymbol{j}} = u_{\boldsymbol{j}} - u_{\boldsymbol{j}\pm\mathbf{e}_k}, \tag{27}$$

$$(K^{\pm}_{\boldsymbol{n},hk}\mathbf{u})_{\boldsymbol{j}} = \frac{u_{\boldsymbol{j}\pm(\mathbf{e}_h-\mathbf{e}_k)} - u_{\boldsymbol{j}\pm(\mathbf{e}_h+\mathbf{e}_k)}}{4}, \tag{28}$$

$$(H_{\boldsymbol{n},k}\mathbf{u})_{\boldsymbol{j}} = \frac{1}{2}(-u_{\boldsymbol{j}-\mathbf{e}_k} + u_{\boldsymbol{j}+\mathbf{e}_k}), \tag{29}$$

$\boldsymbol{j} = \mathbf{1}, \dots, \boldsymbol{n}$, $k, h = 1, \dots, q$, $h \neq k$ (it is understood that $u_{\boldsymbol{i}} = 0$ whenever $\boldsymbol{i} \notin \{\mathbf{1}, \dots, \boldsymbol{n}\}$). Using the multiindex language, it is not difficult to see that

$$K^{\pm}_{\boldsymbol{n},kk} = \left(\bigotimes_{r=1}^{k-1} I_{n_r}\right) \otimes K^{\pm}_{n_k} \otimes \left(\bigotimes_{r=k+1}^{q} I_{n_r}\right), \quad k = 1, \dots, q, \tag{30}$$

$$K^{\pm}_{\boldsymbol{n},hk} = \mp\frac{1}{2}\left(\bigotimes_{r=1}^{h-1} I_{n_r}\right) \otimes J^{\pm}_{n_h} \otimes \left(\bigotimes_{r=h+1}^{k-1} I_{n_r}\right) \otimes H_{n_k} \otimes \left(\bigotimes_{r=k+1}^{q} I_{n_r}\right), \quad 1 \le h \neq k \le q, \tag{31}$$

$$H_{\boldsymbol{n},k} = \left(\bigotimes_{r=1}^{k-1} I_{n_r}\right) \otimes H_{n_k} \otimes \left(\bigotimes_{r=k+1}^{q} I_{n_r}\right), \quad k = 1, \dots, q, \tag{32}$$

where $K^{\pm}_n$, H_n (see Equations (10) and (11)), I_n are the matrices associated with the first-order FD discretizations of constant-coefficient one-dimensional diffusion equations and $J^{\pm}_n$ are the $n \times n$ Jordan nilpotent matrices

$$J^{\pm}_n = I_n - K^{\pm}_n.$$

Since $K^{\pm}_n = T_n(1 - e^{\pm \mathrm{i}\theta})$, $H_n = \mathrm{i}\; T_n(\sin\theta)$, and $J^{\pm}_n = T_n(e^{\pm\mathrm{i}\theta})$, from (30)–(32) and the relation

$$T_{n_1}(f_1) \otimes \dots \otimes T_{n_q}(f_q) = T_{\boldsymbol{n}}(f_1 \otimes \dots \otimes f_q),$$

we deduce that

$$K^{\pm}_{\boldsymbol{n},kk} = T_{\boldsymbol{n}}(1 - e^{\pm\mathrm{i}\theta_k}), \quad k = 1, \dots, q, \tag{33}$$

$$K^{\pm}_{\boldsymbol{n},hk} = \mp\frac{\mathrm{i}}{2}\; T_{\boldsymbol{n}}(e^{\pm\mathrm{i}\theta_h}\sin\theta_k), \quad 1 \le h \neq k \le q, \tag{34}$$

$$H_{\boldsymbol{n},k} = \mathrm{i}\; T_{\boldsymbol{n}}(\sin\theta_k), \quad k = 1, \dots, q, \tag{35}$$

and in particular,

$$K^{+}_{\boldsymbol{n},kk} + K^{-}_{\boldsymbol{n},kk} = T_{\boldsymbol{n}}(2 - 2\cos\theta_k), \quad k = 1, \dots, q, \tag{36}$$

$$K^{+}_{\boldsymbol{n},hk} + K^{-}_{\boldsymbol{n},hk} = T_{\boldsymbol{n}}(\sin\theta_h \sin\theta_k), \quad 1 \le h \neq k \le q. \tag{37}$$

If $H : [0,1]^q \to \mathbb{R}^{q\times q}$ is the symmetric matrix of continuous functions defined by

$$(H(\boldsymbol{\theta}))_{kk} = 2 - 2\cos\theta_k, \quad k = 1, \dots, q, \tag{38}$$

$$(H(\boldsymbol{\theta}))_{hk} = \sin\theta_h \sin\theta_k, \quad 1 \le h \ne k \le q, \tag{39}$$

then we prove

$$\{n^{-2}A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} f^{(\boldsymbol{\nu})}, \tag{40}$$

$$\{n^{-2}A_{\boldsymbol{n}}\}_n \sim_{\sigma,\ \lambda} f^{(\boldsymbol{\nu})}, \tag{41}$$

where

$$f^{(\boldsymbol{\nu})}(\mathbf{x},\boldsymbol{\theta}) = \boldsymbol{\nu}(\mathrm{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{\nu}^T = \sum_{h,k=1}^{q} \nu_h \nu_k a_{hk}(\mathbf{x})(H(\boldsymbol{\theta}))_{hk}. \tag{42}$$

Despite the technicalities intrinsic to any q-dimensional analysis, the proof of Equations (40)–(41) is essentially the same as in the one-dimensional case. It consists of the following steps. Throughout this proof, C denotes a generic constant independent of n. From now on, we assume to have a single discretization parameter n that varies in the infinite set of indices such that $\boldsymbol{n}+\mathbf{1} = \boldsymbol{\nu}n \in \mathbb{N}^q$, where $\boldsymbol{\nu} = (\nu_1, \dots, \nu_q) \in \mathbb{Q}^q$ is an a priori fixed vector with positive components. The relation $\boldsymbol{n}+\mathbf{1} = \boldsymbol{\nu}n$ should be kept in mind while reading the proof.

Step 1. Decompose $A_{\boldsymbol{n}}$ as follows:

$$A_{\boldsymbol{n}} = K_{\boldsymbol{n}} + B_{\boldsymbol{n}} + C_{\boldsymbol{n}}, \tag{43}$$

where

$$B_{\boldsymbol{n}} = \sum_{k=1}^{q} \frac{1}{h_k}\left(\ \mathrm{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} b_k(x_{\boldsymbol{j}})\right) H_{\boldsymbol{n},k} \tag{44}$$

is the FD convection matrix, resulting from the FD discretization of the first-order term in Equation (18), while

$$C_{\boldsymbol{n}} = \left(\ \mathrm{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} c(x_{\boldsymbol{j}})\right) \tag{45}$$

is the matrix resulting from the FD discretization of the lower order term (the reaction term). We show that

$$\|n^{-2}B_{\boldsymbol{n}}\|_2 = o(\sqrt{N(\boldsymbol{n})}) \tag{46}$$

and

$$\{n^{-2}C_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} 0. \tag{47}$$

To prove Equation (46), we can notice that the matrices $H_{\boldsymbol{n},k}$ are elementwise disjoint, meaning that

$$[H_{\boldsymbol{n},k}]_{\boldsymbol{i},\boldsymbol{j}} \ne 0 \ \Rightarrow\ [H_{\boldsymbol{n},h}]_{\boldsymbol{i},\boldsymbol{j}} = 0 \quad \forall h \ne k.$$

Consequently, for every q-uple of diagonal matrices D_k of size $N(\boldsymbol{n})$, we have

$$\left\|\sum_{k=1}^{q} D_k H_{\boldsymbol{n},k}\right\|_2^2 = \sum_{k=1}^{q} \|D_k H_{\boldsymbol{n},k}\|_2^2 .$$

Let us keep in mind that the functions $b_k(\mathbf{x})$ satisfy

$$|b_k(\mathbf{x})(x_1 \dots x_d)^{\alpha}| \le C, \quad \forall x > 0,\ \forall k,$$

where $C > 0$ is a constant.

$$\begin{aligned}
\|n^{-2}B_{\boldsymbol{n}}\|_2^2 &= \left\| \sum_{k=1}^{q} \frac{1}{n^2 h_k} \left(\operatorname{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} b_k(x_{\boldsymbol{j}}) \right) H_{\boldsymbol{n},k} \right\|_2^2 \\
&= \frac{1}{n^2} \sum_{k=1}^{q} v_k^2 \left\| \left(\operatorname{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} b_k(x_{\boldsymbol{j}}) \right) H_{\boldsymbol{n},k} \right\|_2^2 \\
&\le \frac{2}{n^2} \sum_{k=1}^{q} v_k^2 \sum_{\boldsymbol{j}=\mathbf{1}}^{\mathbf{n}} b_k^2(x_{\boldsymbol{j}})
\end{aligned} \tag{48}$$

$$\begin{aligned}
&\le \frac{2C^2}{n^2} \sum_{k=1}^{q} v_k^2 \sum_{\boldsymbol{j}=1}^{\boldsymbol{n}} (j_1 h_1 \cdot \ldots \cdot j_q h_q)^{-2\alpha} \\
&= \left(\frac{2C^2}{n^{2-2q\alpha} N(\boldsymbol{v})^{-2\alpha}} \sum_{k=1}^{q} v_k^2 \right) \prod_{k=1}^{q} \sum_{j=1}^{n_k} j^{-2\alpha}.
\end{aligned} \tag{49}$$

If $\alpha < 1/2$, then $|b_k(\mathbf{x})(x_1 \dots x_d)^{1/2}| \le |b_k(\mathbf{x})(x_1 \dots x_q)^{\alpha}| \le C$, so we explore the case $\alpha \ge 1/2$.

- If $\alpha = 1/2$, then, due to Equation (48), we obtain

$$\begin{aligned}
\|n^{-2}B_{\boldsymbol{n}}\|_2^2 &\le \left(\frac{2C^2}{n^{2-q} N(\boldsymbol{v})^{-1}} \sum_{k=1}^{q} v_k^2 \right) \prod_{k=1}^{q} \sum_{j=1}^{n_k} j^{-1} \\
&\le \left(\frac{2C^2}{n^{2-q} N(\boldsymbol{v})^{-1}} \sum_{k=1}^{q} v_k^2 \right) \prod_{k=1}^{q} \left(1 + \log(n_k) \right)
\end{aligned} \tag{50}$$

$$= O(\log^{q/2}(n) n^{q/2-1}). \tag{51}$$

- If $\alpha > 1/2$, then referring to Equation (48) and due to (13), we obtain

$$\begin{aligned}
\|n^{-2}B_{\boldsymbol{n}}\|_2^2 &\le \left(\frac{2C^2}{n^{2-2q\alpha} N(\boldsymbol{v})^{-2\alpha}} \sum_{k=1}^{q} v_k^2 \right) \prod_{k=1}^{q} \sum_{j=1}^{n_k} j^{-2\alpha} \\
&\le \left(\frac{2C^2}{n^{2-2q\alpha} N(\boldsymbol{v})^{-2\alpha}} \sum_{k=1}^{q} v_k^2 \right) \prod_{k=1}^{q} \frac{-n_k^{-2\alpha+1} + 2\alpha}{2\alpha - 1} \\
&\le \frac{2C^2}{N(\boldsymbol{v})^{-2\alpha}} \left(\sum_{k=1}^{q} v_k^2 \right) \left(\frac{2\alpha}{2\alpha - 1} \right)^{q} n^{-2+2q\alpha}.
\end{aligned} \tag{52}$$

We conclude that

$$\alpha \ge 1/2 \ \Rightarrow \ \|n^{-2}B_{\boldsymbol{n}}\|_2 = O(n^{-1+q\alpha}). \tag{53}$$

Observe that $\frac{1}{q} + \frac{1}{2} > \alpha$ implies $\|n^{-2}B_{\boldsymbol{n}}\|_2^2 = o(n^{q/2}) = o(\sqrt{N(\boldsymbol{n})})$. Noticing that

$$\{C_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} c(\mathbf{x}), \quad \{n^{-1} I_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} 0$$

and using the structure of algebra of GLT sequences, we infer that property Equation (47) is met.

Step 2. Consider the matrix $K_{\boldsymbol{n}}$. By Equations (27) and (43)–(45), we know that

$$\begin{aligned} K_{\boldsymbol{n}} = & \sum_{k=1}^{q} \frac{1}{h_k^2} \left(\operatorname{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} a_{kk}(x_{\boldsymbol{j}+\mathbf{e}_k/2}) \right) K_{\boldsymbol{n},kk}^{+} \\ & + \sum_{k=1}^{q} \frac{1}{h_k^2} \left(\operatorname{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} a_{kk}(x_{\boldsymbol{j}-\mathbf{e}_k/2}) \right) K_{\boldsymbol{n},kk}^{-} \\ & + \sum_{\substack{h,k=1 \\ h\neq k}}^{q} \frac{1}{h_h h_k} \left(\operatorname{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} a_{hk}(x_{\boldsymbol{j}+\mathbf{e}_h}) \right) K_{\boldsymbol{n},hk}^{+} \\ & + \sum_{\substack{h,k=1 \\ h\neq k}}^{q} \frac{1}{h_h h_k} \left(\operatorname{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} a_{hk}(x_{\boldsymbol{j}-\mathbf{e}_h}) \right) K_{\boldsymbol{n},hk}^{-}. \end{aligned} \tag{54}$$

In view of axiom **GLT 3**, all the diagonal matrix-sequences that appear in Equation (54) belong to the class of GLT sequences. More precisely, we observe

$$\{\operatorname{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} a_{kk}(x_{\boldsymbol{j}\pm\mathbf{e}_k/2})\}_n \sim_{\mathrm{GLT}} a_{kk}(\mathbf{x}), \tag{55}$$

$$\{\operatorname{diag}_{\boldsymbol{j}=\mathbf{1},\dots,\boldsymbol{n}} a_{hk}(x_{\boldsymbol{j}\pm\mathbf{e}_h})\}_n \sim_{\mathrm{GLT}} a_{hk}(\mathbf{x}). \tag{56}$$

Using the identities Equations (36)–(37) and axiom **GLT 4**, we thus obtain

$$\{n^{-2}K_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \sum_{k=1}^{q} v_k^2 a_{kk}(\mathbf{x})(2-2\cos\theta_k) + \sum_{\substack{h,k=1 \\ h\neq k}}^{q} v_h v_k a_{hk}(\mathbf{x}) \sin\theta_h \sin\theta_k, \tag{57}$$

that can be rewritten as

$$\{n^{-2}K_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT}} \boldsymbol{v}(\mathrm{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{v}^T.$$

Now $n^{-2}(K_{\boldsymbol{n}}+C_{\boldsymbol{n}})$ are Hermitian matrices, and since $\{n^{-2}C_{\boldsymbol{n}}\}_n$ is zero-distributed, by axioms **GLT 4** and **GLT 1** we infer

$$\{n^{-2}K_{\boldsymbol{n}} + n^{-2}C_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT},\sigma,\lambda} \boldsymbol{v}(\mathrm{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{v}^T.$$

Due to (46), even the sequence $\{n^{-2}B_{\boldsymbol{n}}\}_n$ is zero-distributed, and hence again by axioms **GLT 4** and **GLT 1** we obtain

$$\{n^{-2}A_{\boldsymbol{n}}\}_n = \{n^{-2}K_{\boldsymbol{n}} + n^{-2}B_{\boldsymbol{n}} + n^{-2}C_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT},\sigma} \boldsymbol{v}(\mathrm{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{v}^T.$$

Finally, using Theorem 1 we conclude that

$$\{n^{-2}A_{\boldsymbol{n}}\}_n = \{n^{-2}K_{\boldsymbol{n}} + n^{-2}B_{\boldsymbol{n}} + n^{-2}C_{\boldsymbol{n}}\}_n \sim_{\lambda} \boldsymbol{v}(\mathrm{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{v}^T.$$

4.4 | A basic application in a preconditioning context

We start this subsection by coming back to the one-dimensional finite difference discretization for the problem (6) with the same assumptions on the functions $a(x), b(x), c(x), f(x)$, as indicated in Subsection 4.1. After making the usual scaling by h^2, the matrix arising from the discretization of the second-order operator is $A_n = A_n(a)$ as in Equation (7), while the matrices arising from the discretization of the lower order operators are $B_n = B_n(b)$, $C_n = C_n(c)$ as in Equation (8). Now let

$$Z_n = Z_n(a, b, c, h) = A_n(a) + B_n(b) + C_n(c)$$

be the global discretization matrix and let us consider the standard real-imaginary part representation.[26] Then

$$\begin{aligned}&\Re(Z_n) = A_n(a) + C_n(c) + \Re(B_n(b)), \quad \Im(Z_n) = \Im(B_n(b)),\\ &B_n(b) = -h\mathrm{diag}(b_1, \dots, b_n)\ T_n(\mathrm{i}\sin(\theta)).\end{aligned}$$

From Subsection 4.1, we recall that

$$\{A_n(a)\}_n \sim_{\mathrm{GLT}} a(x)(2 - 2\cos(\theta)), \quad \{B_n(b)\}_n \sim_{\mathrm{GLT}} 0, \quad \{C_n(c)\}_n \sim_{\mathrm{GLT}} 0,$$

and $\{Z_n(a, b, c, h)\}_n \sim_{\lambda} a(x)(2 - 2\cos(\theta))$.

Now we make the problem even less Hermitian by considering the preconditioning by the positive definite Toeplitz matrices $A_n(1) = K_n = T_n(2 - 2\cos(\theta))$ (see Equation (9)) of the matrices $Z_n(a, b, c, h)$. It is evident that $\{K_n\}_n \sim_{\mathrm{GLT}} 2 - 2\cos(\theta)$ and hence, by the algebra structure of GLT sequences,

$$\left\{K_n^{-1} Z_n(a, b, c, h)\right\}_n \sim_{\mathrm{GLT}} a(x).$$

In particular, we have $\left\{K_n^{-1} Z_n(a, b, c, h)\right\}_n \sim_{\sigma} a(x)$. However, for the convergence of a preconditioned Krylov method, it would be useful to have information on the eigenvalues and Theorem 1 is a possible tool. In the present context, the problem in using Theorem 1 relies on the fact the K_n^{-1} is dense, and consequently it is difficult to evaluate any Schatten norm of the imaginary part of the matrix $K_n^{-1} Z_n(a, b, c, h)$. To work around this problem, we consider a symmetrization trick very popular for handling proofs in a preconditioning setting. In fact, $K_n^{-1} Z_n(a, b, c, h)$ is similar to

$$K_n^{-1/2} Z_n(a, b, c, h) K_n^{-1/2}$$

so that for proving the eigenvalue distribution of $\{K_n^{-1} Z_n(a, b, c, h)\}_n$ we can focus on the sequence

$$\{K_n^{-1/2} Z_n(a, b, c, h) K_n^{-1/2}\}_n.$$

Now K_n is positive definite and $\{K_n\}_n$ is a GLT sequence with symbol $2 - 2\cos(\theta)$. Therefore, the inverse of K_n and its square root are well defined and, again by the powerful properties of the GLT sequences (specifically, axioms **GLT 3**, **GLT 4**, **GLT 5**, **GLT 6**) the matrix-sequences $\{K_n^{-1}\}_n$, $\{K_n^{-1/2}\}_n$, and

$$\{K_n^{-1/2} Z_n(a, b, c, h) K_n^{-1/2}\}_n$$

are GLT sequences with symbols $1/(2 - 2\cos(\theta))$, $1/\sqrt{2 - 2\cos(\theta)}$, and $a(x)$, respectively. As a consequence, the singular value distribution results for the sequences $\{K_n^{-1}\}_n$, $\{K_n^{-1/2}\}_n$, and

$$\{K_n^{-1/2} Z_n(a, b, c, h) K_n^{-1/2}\}_n$$

are obvious. The difficulty is that we are interested in the eigenvalue distribution of $\{K_n^{-1/2} Z_n(a, b, c, h) K_n^{-1/2}\}_n$ and the involved matrices are non-Hermitian. Hence, in order to apply Theorem 1 in the most convenient way, we consider the real-imaginary part representation of $K_n^{-1/2} Z_n(a, b, c, h) K_n^{-1/2}$, that is,

$$\begin{aligned}K_n^{-1/2} Z_n(a, b, c, h) K_n^{-1/2} = {} & K_n^{-1/2} \left(A_n(a) + C_n(c) + \Re(B_n(b))\right) K_n^{-1/2}\\ & - \mathrm{i}h K_n^{-1/2} \Im\left(\mathrm{diag}(b_1, \dots, b_n)\ T_n(\mathrm{i}\sin(\theta))\right) K_n^{-1/2}.\end{aligned}$$

It is evident that $R_n = K_n^{-1/2}\left(A_n(a) + C_n(c) + \Re(B_n(b))\right) K_n^{-1/2}$ is symmetric and that $\{R_n\}_n \sim_{\text{GLT}} a(x)$. Therefore it remains to study a proper Schatten norm of

$$S_n = -hK_n^{-1/2}\Im\left(\text{diag}(b_1, \dots, b_n)T_n(\text{i}\sin(\theta))\right) K_n^{-1/2},$$

which is a simplified task since the perturbation matrix $\text{i}\, S_n$ is skew-Hermitian (and hence normal) so that the singular values are the moduli of the eigenvalues. Indeed the key point here is that the matrix S_n is not easy to evaluate in terms of entries, but the spectrum can be evaluated and in fact this has been already done in References 22, theorems 3.1 and 3.7 (see also Reference 21). In these theorems, under the assumption that b is bounded, it is proved that the sequence $\{S_n\}_n$ is properly clustered at zero in the eigenvalue sense and is spectral bounded. This means that for every $\epsilon > 0$, there exists a nonnegative integer number N_ϵ such that the number of eigenvalues exceeding in modulus ϵ are bounded by N_ϵ, and the spectral norm of S_n is bounded by a constant C independent of n. We are now in the position of applying Theorem 1 and in fact, for every $\epsilon > 0$, we obtain

$$\|S_n\|_2^2 \le \|S_n\|^2 N_\epsilon + (n - N_\epsilon)\epsilon^2 \le C^2 N_\epsilon + (n - N_\epsilon)\epsilon^2,$$

so that $\|S_n\|_2 = o(\sqrt{n})$ by the arbitrariness of ϵ. In conclusion, setting $X_n = R_n$, $Y_n = \text{i}\, S_n$, and recalling that $K_n^{-1/2} Z_n(a,b,c,h) K_n^{-1/2} = R_n + \text{i}\ S_n$, Theorem 1 implies that $\{K_n^{-1/2} Z_n(a,b,c,h) K_n^{-1/2}\}_n \sim_\lambda a(x)$, that is, by similarity,

$$\left\{K_n^{-1} Z_n(a,b,c,h)\right\}_n \sim_\lambda a(x). \tag{58}$$

Before briefly giving the results in the multidimensional setting, we introduce a pair of general and useful tools for generalizing relation (58). The considered tools do not depend on the dimensionality of the underlying differential operator and hence they can be applied verbatim in the context of discretized PDEs.

Theorem 4. *Let S_n be either a Hermitian matrix or a skew-Hermitian matrix of size n and consider P_n^+, P_n^- two positive definite matrices of size n. Take the preconditioned matrices $F_n^+ = [P_n^+]^{-1}S_n$, $F_n^- = [P_n^-]^{-1}S_n$ and sort the eigenvalues of $\nu F_n^{\pm}$ in nondecreasing order, with $\nu = 1$ if S_n is Hermitian and $\nu = -\text{i}$ if S_n is skew-Hermitian, that is,*

$$\lambda_1(\nu F_n^{\pm}) \le \cdots \le \lambda_n(\nu F_n^{\pm}).$$

Under the assumption that $P_n^+ \ge P_n^-$ (i.e., $P_n^+ - P_n^-$ is positive semidefinite) we have

$$|\lambda_j(\nu F_n^+)| \le |\lambda_j(\nu F_n^-)|, \quad j = 1, \dots, n. \tag{59}$$

Proof. The proof consists in applying the minimax characterization

$$\lambda_j(\nu F_n^{\pm}) = \min_{\dim(V)=j} \ \max_{\mathbf{x}\in V,\ \mathbf{x}\neq 0} \frac{\mathbf{x}^*\nu S_n \mathbf{x}}{\mathbf{x}^* P_n^{\pm}\mathbf{x}}$$

and in observing that $\mathbf{x}^* P_n^+ \mathbf{x} \ge \mathbf{x}^* P_n^- \mathbf{x} > 0$, because $P_n^+ - P_n^-$ is positive semidefinite and P_n^+, P_n^- are positive definite. ■

Corollary 4. *Under the very same assumptions and notations as in Theorem 4, we have*

$$\|F_{n,\text{S}}^+\|_p \le \|F_{n,\text{S}}^-\|_p$$

for any $p \in [1, \infty]$, where

$$F_{n,\text{S}}^{\pm} = [P_n^{\pm}]^{-1/2} S_n [P_n^{\pm}]^{-1/2}.$$

Proof. The Schatten p-norm of a matrix is the l_p-norm of the vector of its singular values. Since $F_{n,\text{S}}^{\pm}$ is Hermitian if and only if S_n is Hermitian and $F_{n,\text{S}}^{\pm}$ is skew-Hermitian if and only if S_n is skew-Hermitian, it follows that the matrices $F_{n,\text{S}}^{\pm}$ are normal and hence

$$\sigma_j(F_{n,\mathrm{S}}^{\pm}) = |\lambda_j(F_n^{\pm})|, \quad j = 1, \dots, n,$$

because $F_{n,\mathrm{S}}^{\pm}$ is similar to $F_n^{\pm}$. Given the above relations, the claimed thesis follows from Equation (59). ■

The consequences of the corollary above are quite strong. Take any $\{P_n\}_n$ asymptotically equivalent to $\{K_n\}_n$, that is take any $\{P_n\}_n$ such that each P_n is positive definite and there exist positive constants c and C independent of n for which

$$cK_n \leq P_n \leq CK_n.$$

Assume that $\{P_n\}_n$ is a GLT sequence with symbol $\phi(x, \theta)$ and consider the sequence $\{Z_n(a, b, c, h)\}_n$. We have the following:

- $P_n^{-1} Z_n(a, b, c, h)$ is similar to $P_n^{-1/2} Z_n(a, b, c, h) P_n^{-1/2}$;
- $\{P_n^{-1} Z_n(a, b, c, h)\}_n$ is a GLT sequence with symbol

  $$a(x)(2 - 2\cos(\theta))/\phi(x, \theta)$$

  and therefore this function is also the singular value symbol;
- $\{P_n^{-1} \mathfrak{R}(Z_n(a, b, c, h))\}_n$ is a GLT sequence with symbol

  $$a(x)(2 - 2\cos(\theta))/\phi(x, \theta)$$

  and therefore this function is both the singular value and eigenvalue symbol because

  $$P_n^{-1} \mathfrak{R}(Z_n(a, b, c, h))$$

  is similar to the Hermitian matrix

  $$P_n^{-1/2} \mathfrak{R}(Z_n(a, b, c, h)) P_n^{-1/2};$$

- finally, thanks to the spectral equivalence and to the crucial Corollary 4, $\|P_n^{-1/2} \mathfrak{R}(Z_n(a, b, c, h)) P_n^{-1/2}\|_p$ and $\|K_n^{-1/2} \mathfrak{R}(Z_n(a, b, c, h)) K_n^{-1/2}\|_p$ are asymptotically equivalent sequences with equivalence constants $1/C$ and $1/c$.

As a consequence of the previous items, Theorem 1 can be applied with the same conclusions as in the basic case of $P_n = K_n$ and hence

$$\left\{P_n^{-1} Z_n(a, b, c, h)\right\}_n \sim_\lambda \frac{a(x)(2 - 2\cos(\theta))}{\phi(x, \theta)}. \tag{60}$$

The interesting fact is that P_n can be chosen as $D_n(a^{1/2}) K_n D_n(a^{1/2})$, which is a positive definite matrix (the diffusion coefficient $a(x)$ is positive). In this way, due to the structure of algebra of GLT sequences, the resulting sequence $\{P_n\}_n$ is a GLT sequence with symbol $\phi(x, \theta) = a(x)(2 - 2\cos(\theta))$. Furthermore, the two sequences $\{P_n\}_n$ and $\{K_n\}_n$ are asymptotically equivalent under the assumption that $a(x)$ is positive. Consequently, looking at Equation (60), the spectral symbol of $\{P_n^{-1} Z_n(a, b, c, h)\}_n$ is exactly 1, which means that the eigenvalues are clustered at 1 in a weak sense (see the analysis of the strong clustering in Reference 44 under the additional assumptions that $a(x)$ is two times differentiable in $[0, 1]$ and $b(x) = c(x)$ identically zero). Of course, in a practical preconditioning scheme the inversion of K_n can be replaced by one step (or few steps) of a multigrid method in order to reduce the related computational cost.

The same steps can be applied verbatim in the multidimensional case. We consider the FD scheme for problem (18). As preconditioner, we consider, for example, the positive definite matrix coming from the same discretization to the same problem in Equation (18) with the advection–reaction coefficients set to zero. As diffusion matrix, instead of $A(\mathbf{x})$ we consider $\hat{A}(\mathbf{x})$, where $\hat{A}(\mathbf{x})$ can be formally any matrix-valued function, which is positive definite a.e. The conclusions are that

$$\{P_{\boldsymbol{n}}^{-1}A_{\boldsymbol{n}}\}_n \sim_{\mathrm{GLT},\sigma} \frac{\boldsymbol{v}(\mathrm{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{v}^T}{\boldsymbol{v}(\hat{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{v}^T} \tag{61}$$

and

$$\{P_{\boldsymbol{n}}^{-1}A_{\boldsymbol{n}}\}_n \sim_{\lambda} \frac{\boldsymbol{v}(\mathrm{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{v}^T}{\boldsymbol{v}(\hat{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\boldsymbol{v}^T} \tag{62}$$

using Theorem 1, by following exactly the same steps performed in the one-dimensional setting. Of course, the choice of $\hat{A}(\mathbf{x})$ is important from a computational viewpoint. Indeed, $\hat{A}(\mathbf{x})$ has to be selected in such a way that the related linear system is easier to solve (e.g., $\hat{A}(\mathbf{x})$ with diagonal structure) and the range of the spectrum of $[\hat{A}(\mathbf{x})]^{-1}A(\mathbf{x})$ has minimal convex hull, in order to have a fast convergence of the related (preconditioned) Krylov method.

5 | NUMERICAL EXPERIMENTS

We now perform some numerical experiments for confirming the conclusions of Theorem 1 and for showing that the same conclusions are observed in practice under weaker assumptions (see Conjecture 1). In particular, we consider the following settings.

1. We consider the differential Equation (6) with unbounded coefficients $a(x) = c(x) = -\log(1-x)$, $b(x) = 1/\sqrt[4]{x^5}$. We numerically show that for the matrices arising from the FD discretization described in Section 4.1, it holds that

$$\{A_n\}_n + \{B_n\}_n + \{C_n\}_n \sim_\lambda a(x)(2-2\cos(\theta)).$$

2. We numerically show that for the matrices arising from the FE discretization described in Section 4.2, it holds that

$$\left\{\frac{1}{n+1}A_n\right\}_n + \left\{\frac{1}{n+1}B_n\right\}_n + \left\{\frac{1}{n+1}C_n\right\}_n \sim_\lambda a(x)(2-2\cos(\theta)).$$

3. We consider the differential Equation (6) with unbounded coefficients $a(x) = c(x) = -\log(1-x)$, $b(x) = 1/\sqrt[4]{x^3}$. We numerically show that for the preconditioned matrices arising from the FD discretization as described in Section 4.4, it holds that

$$\left\{K_n^{-1}Z_n(a,b,c,h)\right\}_n \sim_\lambda a(x).$$

4. In the case when $P_n = D_n(a^{1/2})K_nD_n(a^{1/2})$, we also show that (60) holds, that is,

$$\left\{P_n^{-1}Z_n(a,b,c,h)\right\}_n \sim_\lambda 1.$$

5. We consider the differential Equation (17) in two space dimensions, with unbounded coefficients $a_{1,1}(x,y) = c(x,y) = 1/xy$, $a_{2,2}(x,y) = -xy$, $a_{1,2}(x,y) = x+y$, $b_1(x,y) = b_2(x,y) = 1/\sqrt[4]{(xy)^3}$. We numerically show that for the matrices arising from the FD discretization described in Section 4.3, it holds that

$$\{A_{\boldsymbol{n}}\}_n \sim_\lambda \mathbf{1}(\mathrm{A}(\mathbf{x})\circ H(\boldsymbol{\theta}))\mathbf{1}^T.$$

6. We consider the differential Equation (17) in two space dimensions, with unbounded coefficients $a_{1,1}(x,y) = a_{2,2}(x,y) = 1 + 1/\sqrt{x} + 1/\sqrt{y}$, $c(x,y) = 1/xy$, $a_{1,2}(x,y) = 0$, $b_1(x,y) = b_2(x,y) = 1/\sqrt[4]{(xy)^3}$. We numerically show that for the preconditioned matrices arising from the FD discretization as described in Section 4.4, it holds that

$$\{P_{\boldsymbol{n}}^{-1}A_{\boldsymbol{n}}\}_n \sim_\lambda 1.$$

The matrices B_n are the source of non-Hermitianess in every setting, so we can choose the functions $b(x)$, $b_1(x, y)$, $b_2(x, y)$ with values as large as possible, in order to test the validity of Theorem 1. In fact, we notice that in the experiments the functions b always satisfy the minimal hypothesis given in the respective sections. In the bidimensional FD case, we have also chosen a system that is not coercive or elliptic at every point, so that the related problem will be foreign to most real applications; yet, the experiments will demonstrate that Theorem 1 holds true even when considering artificially difficult examples. Actually, we want to test if Theorem 1 still holds when the perturbation has an $o(n)$ Schatten 1-norm instead of an $o(\sqrt{n})$ Schatten 2-norm.

1. We consider again the differential Equation (6). If $b(x)$ has a singularity of order $\alpha > -2$ at $x = 0$, then we obtain a perturbation whose Schatten 1-norm is of order $o(n)$. Therefore, we consider again the coefficients $a(x) = c(x) = -\log(1 - x)$ and we modify the perturbation $b(x) = 1/\sqrt[4]{x^7}$. We check by numerical experiments that for the matrices arising from the FD discretization described in Section 4.1, it holds that

$$\{A_n\}_n + \{B_n\}_n + \{C_n\}_n \sim_\lambda a(x)(2 - 2\cos(\theta)).$$

2. We numerically show that for the matrices arising from the FE discretization as described in Section 4.2, it holds that

$$\left\{\frac{1}{n+1}A_n\right\}_n + \left\{\frac{1}{n+1}B_n\right\}_n + \left\{\frac{1}{n+1}C_n\right\}_n \sim_\lambda a(x)(2 - 2\cos(\theta)).$$

3. We consider the differential Equation (17), in two space dimensions. If $b_k(\mathbf{x})$ have singularities of order $\alpha > -\frac{3}{2}$ at zero, then we obtain a perturbation whose Schatten 1-norm is of order $o(N(\boldsymbol{n}))$. Therefore, we consider the coefficients $a_{1,1}(x, y) = a_{2,2}(x, y) = 1/\sqrt{xy}$, $c(x, y) = 1/xy$, $a_{1,2}(x, y) = x + y$, $b_1(x, y) = b_2(x, y) = 1/\sqrt[4]{(xy)^5}$. We show that for the matrices arising from the FD discretization described in Section 4.3, it holds that

$$\{A_{\boldsymbol{n}}\}_n \sim_\lambda \mathbf{1}(\mathrm{A}(\mathbf{x}) \circ H(\boldsymbol{\theta}))\mathbf{1}^T.$$

The aim of the experiments is to show that the chosen sequences of matrices possess the declared spectral symbols. We will show that the portion of eigenvalues with nonnegligible imaginary part tends to zero, and that the distributions of eigenvalues converge to the corresponding symbols.

5.1 | Tables and graphs

Since all the spectral symbols in our experiments are real functions, we are interested in evaluating the number of eigenvalues with a nonnegligible imaginary part. In Table 1 we report the number of eigenvalues with imaginary part greater than a fixed threshold and their percentage with respect to the dimension of the linear system. In particular, we fix the thresholds $\epsilon = 10^{-1}, 10^{-2}$ and consider each experiment from 1. to 6. with dimensions $N = 50, 100, 200, 400, 800$. In the bidimensional FD context, we take $\boldsymbol{n} = (n, n)$ with $n = 7, 10, 14, 20, 28$, in order to obtain $N(\boldsymbol{n}) = n^2$ the closest possible to 50,100,200,400,800. We clearly observe how the percentage of outliers tends to zero when the perturbation has order $o(\sqrt{n})$, in accordance with the theory. For the case of two dimensions, since the fineness parameter is n^{-1} and the dimension is n^2, the choice $\epsilon = 10^{-2}$ is not appropriate to see the convergence with $n = 7, 10, 14, 20, 28$; however also in this case, the trend is quite clear.

In Table 2 we report the number and rate of eigenvalues with imaginary part greater than the same thresholds for experiments from 7. to 9., and we consider the same dimensions as Table 1. We observe how the percentage of outliers tends to zero also when the perturbation has Schatten 1-norm of order $o(n)$ but Schatten 2-norm which is not $o(\sqrt{n})$, thus violating the assumptions of Theorem 1. As expected, the convergence to zero is slower when compared with the previous experiments. In any case, the numerical results support the generalization of Theorem 1 stated in Conjecture 1 below.

After having assessed that the imaginary parts of the eigenvalues tend to zero, let us look at the real parts of the eigenvalues. In order to show that they converge to the spectral symbol, we plot the symbol and the eigenvalues of a chosen set of matrices in the sequence. In Figures 1 to 4 the red line is always the increasing sampling of the spectral symbol performed on 10,000 points. The blue lines are linear plots that connect the real parts of the eigenvalues referred to the

TABLE 1 Number and percentage of eigenvalues with imaginary part greater than ε

N		50	100	200	400	800
FD-1-dim	$\varepsilon = 10^{-1}$	4/8%	6/6%	10/5%	12/3%	14/1.75%
	$\varepsilon = 10^{-2}$	6/12%	8/8%	14/7%	20/5%	28/3.5%
FE	$\varepsilon = 10^{-1}$	4/8%	6/6%	8/4%	12/3%	14/1.75%
	$\varepsilon = 10^{-2}$	4/8%	8/8%	12/6%	14/4.5%	26/3.25%
Prec-K	$\varepsilon = 10^{-1}$	2/4%	2/2%	2/1%	2/0.5%	4/0.5%
	$\varepsilon = 10^{-2}$	6/12%	8/8%	10/5%	12/3%	16/2%
Prec-P	$\varepsilon = 10^{-1}$	14/28%	20/20%	30/15%	44/11%	62/7.75%
	$\varepsilon = 10^{-2}$	32/64%	54/54%	86/43%	130/32.5%	194/24.25%
FD-2-dim	$\varepsilon = 10^{-1}$	4/8.16%	2/2%	2/1.02%	2/0.5%	2/0.26%
	$\varepsilon = 10^{-2}$	8/16.33%	12/12%	30/15.31%	52/13%	92/11.74%
Prec-P-2-dim	$\varepsilon = 10^{-1}$	12/24.49%	18/18%	20/10.2%	22/5.5%	28/3.57%
	$\varepsilon = 10^{-2}$	42/85.71%	82/82%	142/72.45%	256/64%	430/54.85%

Note: The coefficients used are $a(x) = c(x) = -\log(1-x)$, $b(x) = 1/\sqrt[4]{x^5}$ in FD-1-dim, FE settings, and $a(x) = c(x) = -\log(1-x)$, $b(x) = 1/\sqrt[4]{x^3}$ in both the monodimensional preconditioned cases. The bidimensional coefficients are $a_{1,1}(x,y) = a_{2,2}(x,y) = 1 + 1/\sqrt{x} + 1/\sqrt{y}$, $c(x,y) = 1/xy$, $a_{1,2}(x,y) = 0$, $b_1(x,y) = b_2(x,y) = 1/\sqrt[4]{(xy)^3}$ in the preconditioned setting, and $a_{1,1}(x,y) = c(x,y) = 1/xy$, $a_{2,2}(x,y) = -xy$, $a_{1,2}(x,y) = x + y$, $b_1(x,y) = b_2(x,y) = 1/\sqrt[4]{(xy)^3}$ in the FD-2-dim setting.

TABLE 2 Number and percentage of eigenvalues with imaginary part greater than ε

N		50	100	200	400	800
FD-1-dim	$\varepsilon = 10^{-1}$	8/16%	12/12%	18/9%	28/7%	40/5%
	$\varepsilon = 10^{-2}$	8/16%	14/14%	22/11%	34/8.5%	74/9.25%
FE	$\varepsilon = 10^{-1}$	6/12%	12/12%	18/9%	26/6.5%	40/5%
	$\varepsilon = 10^{-2}$	4/8%	8/8%	12/6%	14/4.5%	26/3.25%
FD-2-dim	$\varepsilon = 10^{-1}$	12/24.29%	20/20%	30/15.31%	44/11%	58/7.4%
	$\varepsilon = 10^{-2}$	16/32.653%	24/24%	42/21.43%	64/16%	114/14.54%

Note: The coefficients used are $a(x) = c(x) = -\log(1-x)$, $b(x) = 1/\sqrt[4]{x^7}$ in FD-1-dim, FE and Prec settings and $a_{1,1}(x,y) = a_{2,2}(x,y) = 1/\sqrt{xy}$, $c(x,y) = 1/xy$, $a_{1,2}(x,y) = x + y$, $b_1(x,y) = b_2(x,y) = 1/\sqrt[4]{(xy)^5}$ in FD-2-dim settings.

specified matrix. In particular, given a $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that $\Re(\lambda_1) \leq \Re(\lambda_2) \leq \cdots \leq \Re(\lambda_n)$, the blue line is a piecewise linear function connecting the points $\left(\Re(\lambda_i), \frac{i}{n-1}\right)$ for $i = 1, 2, \ldots, n$.

- In Figure 1, are reported the symbol and the eigenvalues of matrices referred to Experiment 1. We consider linear systems of dimension $N = 4, 8, 16, 32, 64, 128$.
- In Figure 2, are reported the symbol and the eigenvalues of matrices referred to Experiment 2. We consider linear systems of dimension $N = 4, 8, 16, 32, 64, 128$.
- In Figure 3, are reported the symbol and the eigenvalues of matrices referred to Experiment 9. We consider linear systems of dimension $\boldsymbol{n} = (n, n)$ with $n = 7, 10, 14, 20, 28$, in order to obtain $N(\boldsymbol{n}) = n^2$ the closest possible to 50, 100, 200, 400, 800.
- In Figure 4, are reported the symbol and the eigenvalues of matrices referred to Experiment 3. We consider linear systems of dimension $N = 4, 8, 16, 32, 64, 128$.

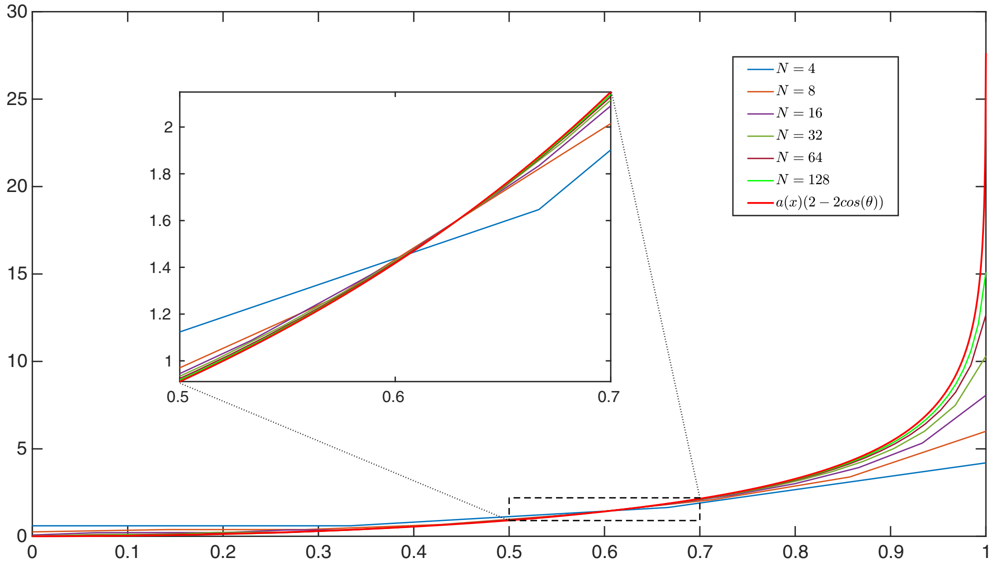

FIGURE 1 Plot of the real part of eigenvalues in FD-1-dim case against the rearranged symbol. The used coefficients are $a(x) = c(x) = -\log(1-x)$, $b(x) = 1/\sqrt[4]{x^5}$

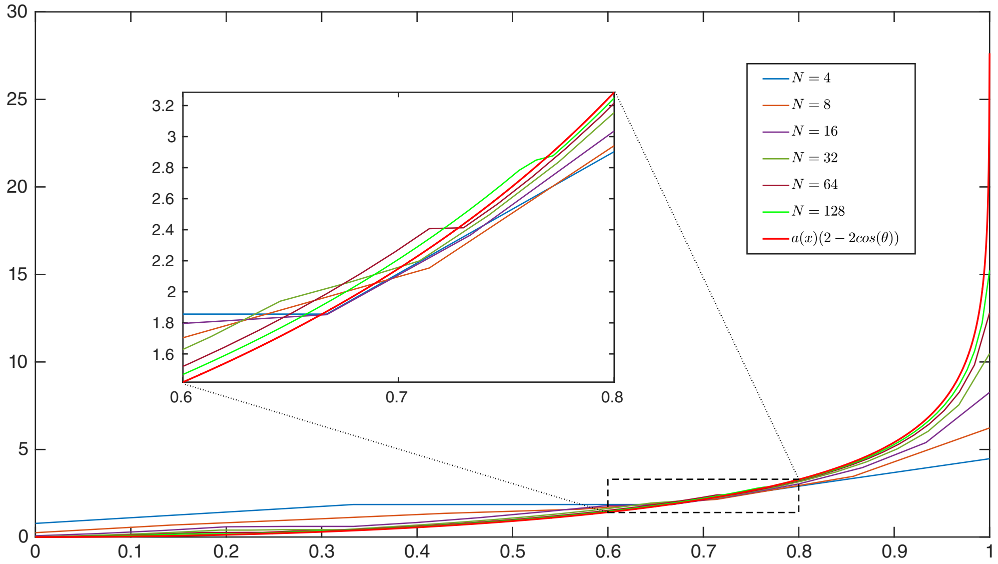

FIGURE 2 Plot of the real part of eigenvalues in FE case against the rearranged symbol. The used coefficients are $a(x) = c(x) = -\log(1-x)$, $b(x) = 1/\sqrt[4]{x^5}$

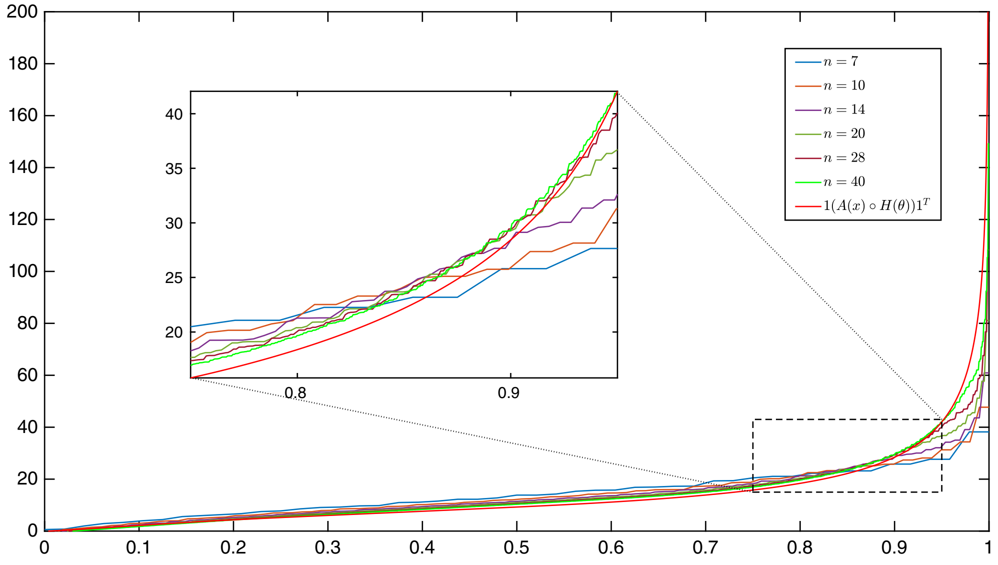

FIGURE 3 Plot of the real part of eigenvalues in FD-2-dim case against the rearranged symbol. The used coefficients are $a_{1,1}(x,y) = a_{2,2}(x,y) = 1/\sqrt{xy}$, $a_{1,2}(x,y) = x+y$, $b_1(x,y) = b_2(x,y) = 1/\sqrt[4]{(xy)^5}$, $c(x,y) = 1/xy$

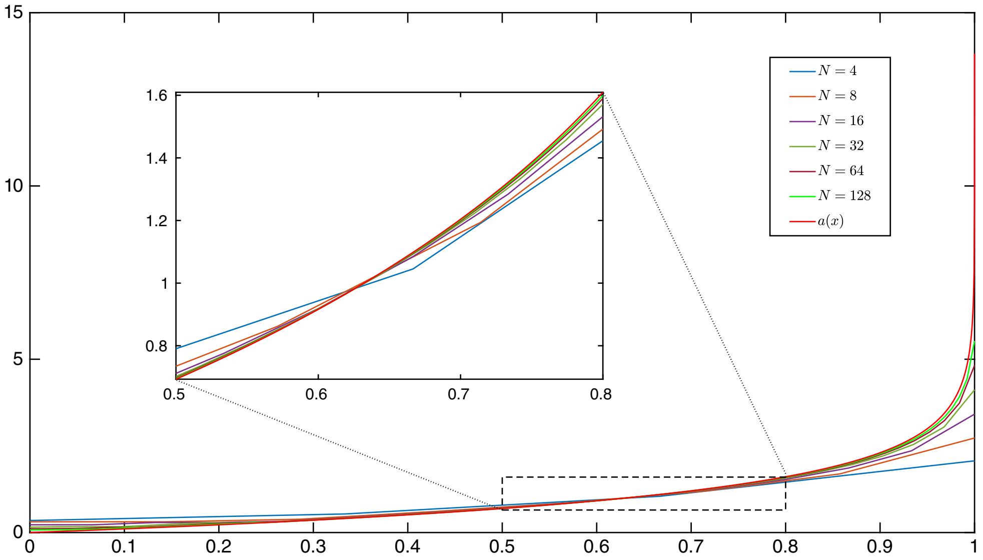

FIGURE 4 Plot of the real part of eigenvalues in Prec case against the rearranged symbol. The used coefficients are $a(x) = c(x) = -\log(1-x)$, $b(x) = 1/\sqrt[4]{x^3}$

In Figure 1, 2, and 4 we have chosen the sizes $N = 4, 8, 16, 32, 64, 128$ because it is difficult to distinguish the plot of eigenvalues and the symbol plot for $N \geq 200$. In all the figures, we have added a focus on particular points to observe the behavior closer. In all the figures, we see that the eigenvalue plots converge to the respective symbols as the dimension increases.

6 | CONCLUSIONS

We have proved a result (Theorem 1) which allows one to compute the asymptotic spectral distribution of matrix-sequences that can be written as a non-Hermitian perturbation of a given Hermitian matrix-sequence. As shown in Corollary 3, this result provides a noteworthy extension of a previous theorem due to Leonid Golinskii and the second author[17, theorem 3.4] (see also Reference 2, corollary 4.1). In particular, as illustrated in this article, we are now able to compute the asymptotic spectral distribution of PDE discretization matrices even in the case where the PDE coefficients possess minimal regularity properties (only L^1 in the case of finite elements!). It is also worth noting that the proof of the new result has not involved any functional analysis argument, which means that the proof of Reference 17, theorem 3.4 can be performed through purely matrix analysis arguments as in this article; in particular, there is no need to resort to Mergelyan's theorem.

Extensive numerical experiments have been discussed. What we have observed seems to indicate that our new result can be made even stronger, by considering the weaker condition $\|Y_n\|_1 = o(n)$ used in Reference 17 and simultaneously dropping the assumption of boundedness of the involved matrix-sequences as done in this note in Theorem 1. We therefore state the following conjecture, which can be viewed as a generalization of Theorem 1.

Conjecture 1. *Let X_n be a Hermitian matrix of size n, with $\{X_n\}_n \sim_\lambda f$. If $\|Y_n\|_1 = o(n)$ then*

$$\{X_n + Y_n\}_n \sim_\lambda f.$$

We refer to future articles for showing different and various applications of the result in more general contexts. For example, discretization of systems of PDEs, multidimensional FE methods (IgA, collocation, etc.), and when the equations are defined only on irregular domains, and consequentially the usage of nonregular grids adapted to the problem geometry is prescribed.

CONFLICTS OF INTEREST

This work does not have any conflicts of interest.

ORCID

Giovanni Barbarino https://orcid.org/0000-0003-0504-7361

REFERENCES

1. Böttcher A, Silbermann B. Introduction to large truncated Toeplitz matrices. New York, NY: Springer, 1999.
2. Garoni C, Serra-Capizzano S. Generalized locally Toeplitz sequences: Theory and applications (volume I). Cham, Switzerland: Springer, 2017.
3. Tilli P. Locally Toeplitz sequences: Spectral properties and applications. Linear Algebra Appl. 1998;278:91–120.
4. Beckermann B, Kuijlaars ABJ. Superlinear convergence of conjugate gradients. SIAM J Numer Anal. 2001;39:300–329.
5. Kuijlaars ABJ. Convergence analysis of Krylov subspace iterations with methods from potential theory. SIAM Rev. 2006;48:3–40.
6. Garoni C, Serra-Capizzano S. Generalized locally Toeplitz sequences: A spectral analysis tool for discretized differential equations. Lecture notes in mathematics, CIME foundation subseries 2219. Cham, Switzerland: Springer, 2018; p. 161–236.
7. Garoni C, Serra-Capizzano S. Generalized locally Toeplitz sequences: Theory and applications (Volume II). Cham, Switzerland: Springer, 2018.
8. Dorostkar A, Neytcheva M, Serra-Capizzano S. Spectral analysis of coupled PDEs and of their Schur complements via generalized locally Toeplitz sequences in 2D. Comput Methods Appl Mech Eng. 2016;309:74–105.
9. Garoni C, Manni C, Pelosi F, Serra-Capizzano S, Speleers H. On the spectrum of stiffness matrices arising from isogeometric analysis. Numer Math. 2014;127:751–799.
10. Garoni C, Manni C, Serra-Capizzano S, Sesana D, Speleers H. Spectral analysis and spectral symbol of matrices in isogeometric Galerkin methods. Math Comp. 2017;86:1343–1373.
11. Garoni C, Manni C, Serra-Capizzano S, Sesana D, Speleers H. Lusin theorem, GLT sequences and matrix computations: An application to the spectral analysis of PDE discretization matrices. J Math Anal Appl. 2017;446:365–382.
12. Beckermann B, Serra-Capizzano S. On the asymptotic spectrum of finite element matrix sequences. SIAM J Numer Anal. 2007;45:746–769.
13. Donatelli M, Garoni C, Manni C, Serra-Capizzano S, Speleers H. Robust and optimal multi-iterative techniques for IgA Galerkin linear systems. Comput Methods Appl Mech Eng. 2015;284:230–264.
14. Donatelli M, Garoni C, Manni C, Serra-Capizzano S, Speleers H. Robust and optimal multi-iterative techniques for IgA collocation linear systems. Comput Methods Appl Mech Eng. 2015;284:1120–1146.
15. Hofreither C, Jüttler B, Kiss G, Zulehner W. Multigrid methods for isogeometric analysis with THB-splines. Comput Methods Appl Mech Eng. 2016;308:96–112.
16. Hofreither C, Takacs S, Zulehner W. A robust multigrid method for isogeometric analysis in two dimensions using boundary correction. Comput Methods Appl Mech Eng. 2017;316:22–42.
17. Golinskii L, Serra-Capizzano S. The asymptotic properties of the spectrum of nonsymmetrically perturbed Jacobi matrix sequences. J Approx Theory. 2007;144:84–102.
18. Kuijlaars ABJ, Serra-Capizzano S. Asymptotic zero distribution of orthogonal polynomials with discontinuously varying recurrence coefficients. J Approx Theory. 2001;113:142–155.
19. Serra-Capizzano S, Sesana D, Strouse E. The eigenvalue distribution of products of Toeplitz matrices—Clustering and attraction. Linear Algebra Appl. 2010;432:2658–2678.
20. Bertaccini D, Durastante F. Limited memory block preconditioners for fast solution of fractional partial differential equations. J Sci Comput. 2018;77-2:950–970.
21. Bertaccini D, Golub GH, Serra-Capizzano S. Spectral analysis and superlinear convergence of a preconditioned iterative method for the convection-diffusion equation. SIAM J Matrix Anal Appl. 2007;29-1:260–278.
22. Bertaccini D, Golub GH, Serra-Capizzano S, Tablino-Possio C. Preconditioned HSS methods for the solution of non-Hermitian positive definite linear systems and applications to the discrete convection-diffusion equation. Numer Math. 2005;99:441–484.
23. Donatelli M, Mazza M, Serra-Capizzano S. Spectral analysis and structure preserving preconditioners for fractional diffusion equations. J Comput Phys. 2016;307:262–279.
24. Ng MK. Iterative methods for Toeplitz systems. Numerical mathematics and scientific computation. New York, NY: Oxford University Press, 2004.
25. Cottrell JA, Hughes TJR, Bazilevs Y. Isogeometric analysis: Toward integration of CAD and FEA. Chichester, UK: John Wiley & Sons, 2009.
26. Bhatia R. Matrix analysis. New York: Springer, 1997.
27. Barbarino G. Spectral measures. Structured matrices in numerical linear algebra. Cham, Switzerland; Springer; 2019; p. 1–24. Springer INdAM Ser., volume 30.
28. Böttcher A, Garoni C, Serra-Capizzano S. Exploration of Toeplitz-like matrices with unbounded symbols is not a purely academic journey. Sb Math. 2017;208:1602–1627.
29. Garoni C, Serra-Capizzano S. The theory of locally Toeplitz sequences: A review, an extension, and a few representative applications. Bol Soc Mat Mex. 2016;22:529–565.
30. Garoni C, Serra-Capizzano S. The theory of generalized locally Toeplitz sequences: A review, an extension, and a few representative applications. Oper Theory Adv Appl. 2017;259:353–394.

31. Serra-Capizzano S. Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations. Linear Algebra Appl. 2003;366:371–402.
32. Serra-Capizzano S. The GLT class as a generalized Fourier analysis and applications. Linear Algebra Appl. 2006;419:180–233.
33. Barbarino G. Equivalence between GLT sequences and measurable functions. Linear Algebra Appl. 2017;529:397–412.
34. Serra-Capizzano S. Distribution results on the algebra generated by Toeplitz sequences: A finite dimensional approach. Linear Algebra Appl. 2001;328:121–130.
35. Barbarino G, Garoni C. From convergence in measure to convergence of matrix-sequences through concave functions and singular values. Electr J Linear Algebra. 2017;32:500–513.
36. Brezis H. Functional analysis, Sobolev spaces and partial differential equations. New York, NY: Springer, 2011.
37. Grigoryan A. Heat kernels on weighted manifolds and applications. Contemp Math. 2006;398:93–191.
38. Tesei A. On uniqueness of the positive Cauchy problem for a class of parabolic equations. Problemi Attuali dell'Analisi e della Fisica Matematica (Taormina 1998), Rome: Aracne, 2000; p. 145–160.
39. Pozio MA, Tesei A. On the uniqueness of bounded solutions to singular parabolic problems. Discrete Contin Dyn Syst. 2005;13:117–137.
40. Punzo F, Tesei A. Uniqueness of solutions to degenerate elliptic problems with unbounded coefficients. Ann I H Poincaré. 2009;26:2001–2024.
41. Evans LC. Partial differential equations. Graduate studies in mathematics. Providence, RI: AMS, 2010.
42. Bertaccini D, Donatelli M, Durastante F, Serra-Capizzano S. Optimizing a multigrid Runge-Kutta smoother for variable-coefficient convection-diffusion equations. Linear Algebra Appl. 2017;533:507–535.
43. Donatelli M, Mazza M, Serra-Capizzano S. Spectral analysis and multigrid methods for finite volume approximations of space-fractional diffusion equations. SIAM J Sci Comput. 2018;40-6:A4007–A4039.
44. Serra-Capizzano S. The rate of convergence of Toeplitz based PCG methods for second order nonlinear boundary value problems. Numer Math. 1999;81:461–495.

How to cite this article: Barbarino G, Serra-Capizzano S. Non-Hermitian perturbations of Hermitian matrix-sequences and applications to the spectral analysis of the numerical approximation of partial differential equations. *Numer Linear Algebra Appl*. 2020;e2286. https://doi.org/10.1002/nla.2286