SIAM J. NUMER. ANAL.
Vol. 57, No. 3, pp. 1369–1394

LAPLACIAN PRECONDITIONING OF ELLIPTIC PDEs: LOCALIZATION OF THE EIGENVALUES OF THE DISCRETIZED OPERATOR*

TOMÁŠ GERGELITS†, KENT-ANDRÉ MARDAL‡, BJØRN FREDRIK NIELSEN§, AND ZDENĚK STRAKOŠ¶

Abstract. In [*IMA J. Numer. Anal.*, 29 (2009), pp. 24–42], Nielsen, Tveito, and Hackbusch study the operator generated by using the inverse of the Laplacian as the preconditioner for second order elliptic PDEs $-\nabla \cdot (k(x)\nabla u) = f$. They prove that the range of $k(x)$ is contained in the spectrum of the preconditioned operator, provided that $k(x)$ is continuous. Their rigorous analysis only addresses mappings defined on infinite dimensional spaces, but the numerical experiments in the paper suggest that a similar property holds in the discrete case. Motivated by this investigation, we analyze the eigenvalues of the matrix $\mathbf{L^{-1}A}$, where $\mathbf{L}$ and $\mathbf{A}$ are the stiffness matrices associated with the Laplace operator and second order elliptic operators with a scalar coefficient function, respectively. Using only technical assumptions on $k(x)$, we prove the existence of a one-to-one pairing between the eigenvalues of $\mathbf{L^{-1}A}$ and the intervals determined by the images under $k(x)$ of the supports of the finite element nodal basis functions. As a consequence, we can show that the nodal values of $k(x)$ yield accurate approximations of the eigenvalues of $\mathbf{L^{-1}A}$. Our theoretical results, including their relevance for understanding how the convergence of the conjugate gradient method may depend on the whole spectrum of the preconditioned matrix, are illuminated by several numerical experiments.

Key words. second order elliptic PDEs, preconditioning by the inverse Laplacian, eigenvalues of the discretized preconditioned problem, nodal values of the coefficient function, Hall's theorem, convergence of the conjugate gradient method

AMS subject classifications. 65F08, 65F15, 65N12, 35J99

DOI. 10.1137/18M1212458

1. Introduction.

The classical analysis of Krylov subspace solvers for matrix problems with Hermitian matrices relies on their spectral properties; see, e.g., [1, 16]. Typically, one seeks a preconditioner which yields parameter independent bounds for the extreme eigenvalues; see, e.g., [9, 20, 27, 15, 26] for a discussion of this issue in terms of *operator preconditioning*. This approach consists of considering the mapping properties of the continuous operator between appropriate Sobolev spaces in order to derive a discrete preconditioner. This has the advantage that only the largest and smallest eigenvalues (in the absolute sense if an indefinite problem is solved) must

*Received by the editors September 11, 2018; accepted for publication (in revised form) April 3, 2019; published electronically June 13, 2019.

http://www.siam.org/journals/sinum/57-3/M121245.html

Funding: The first author's research was supported by Charles University project GA UK 172915. The research of the first and fourth authors was supported by the Grant Agency of the Czech Republic under contract 17-04150J. The third author's research was supported by Research Council of Norway project 239070.

†Institute of Computer Science of the Czech Academy of Sciences, 182 07 Prague 8, Czech Republic (gergelits@cs.cas.cz), and Faculty of Mathematics and Physics, Charles University, 186 75 Prague 8, Czech Republic (gergelits@karlin.mff.cuni.cz).

‡Department of Mathematics, University of Oslo, 0316 Oslo, Norway (kent-and@math.uio.no), and Department of Numerical Analysis and Scientific Computing, Simula Research Laboratory, 1325 Lysaker, Norway (kent-and@simula.uio.no).

§Faculty of Science and Technology, Norwegian University of Life Sciences, NO-1432 Ås, Norway (bjorn.f.nielsen@nmbu.no).

¶Faculty of Mathematics and Physics, Charles University, 186 75 Prague 8, Czech Republic (strakos@karlin.mff.cuni.cz).

be studied, and the bounds for the required number of Krylov subspace iterations can become independent of the mesh size and other important parameters. However, for problems with spatially variable coefficients, possibly varying by many orders of magnitudes, the condition number estimate provided by the operator preconditioning is of limited value when the variation of the coefficients is ignored, i.e., when the Sobolev spaces do not involve the variable coefficients. The corresponding condition number will then be huge. Furthermore, convergence bounds based on single number characteristics, such as the condition number, are too simple to capture the adaption of Krylov subspace methods to the data. In particular, Krylov subspace methods are strongly nonlinear in the input data (matrix and the initial residual), and therefore the whole spectral information is needed in order to capture the actual convergence behavior when these methods are applied to problems with self-adjoint operators with large condition numbers but structured spectra.[1]

The complete spectrum mattering has been known since the introduction of the conjugate gradient method (CG) in 1952. Chapters 14–19 of the seminal paper by Hestenes and Stiefel [19] link CG to orthogonal polynomials and continued fractions. They very clearly explain the link between CG and Gauss quadrature (approximating the distribution function determined via the spectral decomposition of the involved matrix, conveniently presented via the Riemann–Stieltjes integral); see, in particular, [19, Chapter 14, Theorems 14:1-14:3]. This classical view and the understanding which combines algorithmic development with approximation theory and functional analysis (see, e.g., [38, Chapter II, section 7]) have been further developed by several authors, including the beautiful and almost unknown monograph on the method of moments by Vorobyev [39, Chapter III]. For a recent description, we refer the reader to, e.g., the paper by Herzog and Sachs [18], section 5.2 in the monograph [26], and section 3.5 and Chapter 5 in the monograph [24].

The superlinear character of CG convergence has been observed and investigated in several studies [2, 3, 21, 22, 37, 31, 18], and the acceleration of convergence has been linked with the presence of large outlying eigenvalues and clustering of the eigenvalues. Since Krylov subspace methods for systems with Hermitian matrices use short recurrences, exact arithmetic considerations must be complemented with a thorough rounding error analysis; otherwise, it can in practice be misleading or even completely useless. This issue was again pointed out already by Hestenes and Stiefel in the paper [19], where Chapter 8 is devoted to propagation of rounding errors. The authors emphasized that the loss of orthogonality might increase so rapidly that the computed approximation would not be as good an estimate of the solution as desired. Simultaneously, the deterioration of convergence due to rounding errors in the presence of large outlying eigenvalues was reported, based on experiments, already in [23]; see also [8], [21, p. 72], the discussion in [37, p. 559], and the summary in [24, section 5.6.4, pp. 279–280].

In investigating the convergence behavior of Krylov subspace methods for Hermitian problems, we thus have to deal with two phenomena acting against each other. Large outlying eigenvalues (or well-separated clusters of large eigenvalues) can in theory, assuming exact arithmetic, be linked with acceleration of CG convergence. However, in practice, using finite precision computations, it can cause deterioration of the convergence rate. This intriguing situation has been fully understood thanks to the seminal work of Greenbaum [11] with the fundamental preceding analysis of

[1] The spectral information may not be descriptive for the convergence of Krylov subspace methods in general; see [12, 14] and [24, section 5.7].

the Lanczos method by Paige [33, 34]; see also [13, 36, 29, 28] and the recent paper [10] that addresses the question of validity of the CG composite convergence bounds based on the so-called effective condition number.

Summarizing, preconditioning that leads to favorable distributions of the eigenvalues of the preconditioned (Hermitian) matrix can lead to much faster convergence than preconditioning that only focuses on minimizing the condition number. As presented below, the term "favorable" is very subtle and its understanding requires knowledge of many associated results.

Motivated by these facts and the results in [32], the purpose of this paper is to show that approximations of *all* the eigenvalues of a classical generalized eigenvalue problem are readily available. More specifically, assuming that the function $k(x)$ is uniformly positive, bounded, and piecewise continuous, we will study finite element (FE) discretizations of

$$\begin{aligned} \nabla \cdot (k(x)\nabla u) &= \lambda \Delta u \quad \text{in } \Omega \subset \mathbb{R}^d, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{1.1}$$

$d = 1$, 2, or 3, which yields a system of linear equations in the form

$$\mathbf{A}\mathbf{v} = \lambda \mathbf{L}\mathbf{v}. \tag{1.2}$$

As mentioned above, mathematical properties of the continuous problem (1.1) are studied in [32]. In particular, the authors of that paper prove that[2]

$$k(x) \in \mathrm{sp}(\mathcal{L}^{-1}\mathcal{A})$$

for all $x \in \Omega$ at which $k(x)$ is *continuous*, where

$$\mathcal{A} : H_0^1(\Omega) \mapsto H^{-1}(\Omega), \quad \langle \mathcal{A}u, v\rangle = \int_\Omega k\nabla u \cdot \nabla v, \quad u, v \in H_0^1(\Omega), \tag{1.3}$$

$$\mathcal{L} : H_0^1(\Omega) \mapsto H^{-1}(\Omega), \quad \langle \mathcal{L}u, v\rangle = \int_\Omega \nabla u \cdot \nabla v, \quad u, v \in H_0^1(\Omega). \tag{1.4}$$

The authors also conjecture that the spectrum of the discretized preconditioned operator $\mathbf{L}^{-1}\mathbf{A}$ can be approximated by the nodal values of $k(x)$. In the present text, we show, without the continuity assumption on the coefficient function, how the function values of $k(x)$ are related to the generalized spectrum of the discretized operators (matrices) in (1.2). Our main results state the following:

- There exists a (potentially nonunique) pairing of the eigenvalues of $\mathbf{L}^{-1}\mathbf{A}$ and the intervals determined by the images under $k(x)$ of the supports of the FE basis functions; see Theorem 3.1 in section 3.
- The function values of $k(x)$ at the nodes of the FE grid can be paired with the individual eigenvalues of the discrete preconditioned operator $\mathbf{L}^{-1}\mathbf{A}$. Furthermore, these function values yield accurate approximations of the eigenvalues; see Corollary 3.2 in section 3.

The text is organized as follows. Notation, assumptions, a brief note on the CG convergence analysis, and a motivating example are presented in section 2. Section 3

[2] The spectrum of the operator $\mathcal{L}^{-1}\mathcal{A}$ on an infinite dimensional normed linear space is defined as

$$\mathrm{sp}(\mathcal{L}^{-1}\mathcal{A}) \equiv \left\{\lambda \in \mathbb{C};\ \mathcal{L}^{-1}\mathcal{A} - \lambda\mathcal{I} \text{ does not have a bounded inversion}\right\}.$$

contains theoretical results. The proof of the pairing in Theorem 3.1 uses the classical Hall's theorem from the theory of bipartite graphs. Corollary 3.2 then follows as a simple consequence. The numerical experiments in section 4 illustrate the results of our analysis. Moreover, using Theorem 3.1, the discussion at the end of section 4 explains the CG convergence behavior observed in the example presented in section 2. The text closes with concluding remarks in section 5.

2. Notation, a brief note on the CG convergence analysis, and an introductory example. We consider a self-adjoint second order elliptic PDE in the form

$$-\nabla \cdot (k(x)\nabla u) = f \quad \text{for } x \in \Omega, \qquad u = 0 \quad \text{for } x \in \partial\Omega \tag{2.1}$$

and the corresponding generalized eigenvalue problem (1.1) with the domain $\Omega \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, and the given function $f \in L^2(\Omega)$. We assume that the real valued scalar function $k(x) : \mathbb{R}^d \to \mathbb{R}$ is bounded and piecewise continuous and that it is uniformly positive, i.e.,

$$k(x) \geq \alpha > 0, \quad x \in \Omega.$$

Let $V \equiv H_0^1(\Omega)$ denote the Sobolev space of functions defined on Ω with zero trace at $\partial\Omega$ and with the standard inner product. The weak formulations of the problems (2.1) and (1.1) are to seek $u \in V$ (respectively, $u \in V$ and $\lambda \in \mathbb{R}$), such that

$$\mathcal{A}u = f \qquad (\text{respectively, } \mathcal{A}u = \lambda \mathcal{L}u), \tag{2.2}$$

where $\mathcal{A}$, $\mathcal{L} : V \to V^\#$, $f \in V^\#$ are defined in (1.3) and (1.4), and the function $f \in L^2(\Omega)$ is identified with the associated linear functional $f \in V^\#$ defined by

$$\langle f, v \rangle \equiv \int_\Omega f v \,. \tag{2.3}$$

(We use # to denote dual spaces.) Discretization via a conforming FE method, using, for simplicity of exposition, Lagrange elements, leads to the discrete operators

$$\mathcal{A}_h,\ \mathcal{L}_h : V_h \to V_h^\#,$$

where the finite dimensional subspace V_h is spanned by the piecewise polynomial basis functions $\phi_1, \ldots, \phi_N$ with the local supports

$$\mathcal{T}_i = \operatorname{supp}(\phi_i), \quad i = 1, \ldots, N.$$

The matrix representations $\mathbf{A}_h$ and $\mathbf{L}_h$ are defined as

$$[\mathbf{A}_h]_{ij} = \langle \mathcal{A}_h \phi_j, \phi_i \rangle = \int_\Omega \nabla \phi_i \cdot k \nabla \phi_j, \tag{2.4}$$

$$[\mathbf{L}_h]_{ij} = \langle \mathcal{L}_h \phi_j, \phi_i \rangle = \int_\Omega \nabla \phi_i \cdot \nabla \phi_j, \quad i, j = 1, \ldots, N. \tag{2.5}$$

In the text below, we will, for the sake of simple notation, omit the subscript h and write $\mathbf{A} \equiv \mathbf{A}_h$ and $\mathbf{L} \equiv \mathbf{L}_h$. Throughout this text, we assume that conforming elements are employed.

A brief note on the CG convergence analysis. In this short description, we will not include the effects of rounding errors; i.e., we make the unrealistic assumption of exact computations. This description is for the purpose of this paper sufficient. For the full account of the effects of rounding errors, we refer the reader to, e.g., [29] and [24, section 5.8].

The energy norm of the error in CG can be written as

$$\|\mathbf{x}-\mathbf{x}_j\|_{\mathbf{A}}^2 = \min_{\varphi\in\Pi_j} \|\varphi(\mathbf{A})(\mathbf{x}-\mathbf{x}_0)\|_{\mathbf{A}}^2, \quad j=1,2,\ldots, \tag{2.6}$$

where Π_j denotes the set of all polynomials of degree at most j having the value 1 at zero.[3] Using the spectral decomposition of the matrix $\mathbf{A}$, with $\lambda_1,\ldots,\lambda_N$ denoting its eigenvalues and $\mathbf{v}_1,\ldots,\mathbf{v}_N$ the associated orthonormal eigenvectors, formula (2.6) can be written as

$$\|\mathbf{x}-\mathbf{x}_j\|_{\mathbf{A}}^2 = \|\mathbf{r}_0\|^2 \sum_{l=1}^{N} \omega_l \frac{\left(\varphi_j^{CG}(\lambda_l)\right)^2}{\lambda_l}, \quad j=1,2,\ldots. \tag{2.7}$$

Here, $\varphi_j^{CG}(\lambda)$ represents the polynomial giving the minimum in (2.6) and

$$\mathbf{r}_0 = \mathbf{b}-\mathbf{A}\mathbf{x}_0,$$
$$\omega_l = (\mathbf{z},\mathbf{v}_l)^2,\ l=1,\ldots,N, \quad \mathbf{z}=\mathbf{r}_0/\|\mathbf{r}_0\|.$$

If we introduce the distribution function

$$\omega(\lambda) = \begin{cases} 0 & \text{for} \quad \lambda<\lambda_1, \\ \sum_{l=1}^{i}\omega_l & \text{for} \quad \lambda_i\le\lambda<\lambda_{i+1}, \\ 1 & \text{for} \quad \lambda_N\le\lambda, \end{cases} \tag{2.8}$$

then we can express the error in terms of a Riemann–Stieltjes integral:

$$\|\mathbf{x}-\mathbf{x}_j\|_{\mathbf{A}}^2 = \|\mathbf{r}_0\|^2 \int \frac{\left(\varphi_j^{CG}(\lambda)\right)^2}{\lambda}\, d\omega(\lambda), \quad j=1,2,\ldots. \tag{2.9}$$

In this way, it becomes clear that the distribution function $\omega(\lambda)$, defined by the points of increase $\lambda_1,\ldots,\lambda_N$ and the individual weights $\omega_1,\ldots,\omega_N$, "determines" the CG convergence behavior. Indeed, the CG polynomial $\varphi_j^{CG}(\lambda)$ is given by

$$\varphi_j^{CG}(\lambda) = \frac{(\lambda-\theta_1^{(j)})\ldots(\lambda-\theta_j^{(j)})}{(-1)^j\theta_1^{(j)}\ldots\theta_j^{(j)}}, \quad j=1,2,\ldots, \tag{2.10}$$

where $\theta_1^{(j)},\ldots,\theta_j^{(j)}$ are the Ritz values at the jth CG iteration, i.e., the approximations of the eigenvalues of $\mathbf{A}$ (implicitly) generated at the jth CG step, which are identical to the nodes of the associated Gauss quadrature that is accurate for all polynomials of degree at most $2j-1$. We can observe that the value of $\phi_j^{CG}(\lambda)$ at the eigenvalue λ_l of $\mathbf{A}$ vanishes whenever this eigenvalue is closely approximated by some Ritz value. (In finite precision computations, multiple Ritz values can approximate single eigenvalues,

[3] Here we consider CG applied to the linear algebraic system $\mathbf{A}\mathbf{x}=\mathbf{b}$ with the Hermitian positive definite matrix $\mathbf{A}$. An analogous formula holds for CG applied to preconditioned problems as well as to problems in infinite dimensional Hilbert spaces; see, e.g., [18, 26].

which causes a significant delay of convergence. This does not happen for our example below.)

Daniel [6] presented the relations (2.7) and (2.9) in his paper published in 1967. Subsequently, he (for the first time) presented the *bound*

$$\frac{\|\mathbf{x}-\mathbf{x}_j\|_{\mathbf{A}}}{\|\mathbf{x}-\mathbf{x}_0\|_{\mathbf{A}}} \le 2\left(\frac{\sqrt{\kappa(\mathbf{A})}-1}{\sqrt{\kappa(\mathbf{A})}+1}\right)^j, \quad \kappa(\mathbf{A})=\lambda_N/\lambda_1. \tag{2.11}$$

This bound is based on replacing $\varphi_j^{CG}(\lambda)$ by the jth (shifted and scaled) Chebyshev polynomial that uses, instead of the whole distribution function $\omega(\lambda)$, only the information about the smallest and largest eigenvalues λ_1 and λ_N, respectively.[4] Immediately after presenting the associated Theorem 1.2.2, Daniel issues a clear statement concerning the assumptions and the interpretation of the bound (we reformulate it using our notation):

> "... assuming only that the spectrum of $\mathbf{A}$ lies inside the interval $[\lambda_1, \lambda_N]$, we can do no better than Theorem 1.2.2."

Most unfortunately, the existing thorough insight present in the early papers was soon, to a large extent, overshadowed by the following narrow algorithmic simplification concerning the bound (2.11), which is here presented as a quote that can be found (with slight variations) in many papers:

> "This estimate shows that a smaller condition number results in faster convergence. Hence, the convergence may be improved by reducing the condition number."

Such a view identifies CG convergence behavior with the linear bound (2.11); i.e., it ignores the assumption of Theorem 1.2.2 emphasized by Daniel. This methodological misconception has been harmful for decades to theory as well as to practical computations. (Further details and references concerning this issue can be found in Chapter 11 of the monograph [26].) While mentioning this, we do not deny the use of the condition number bound (2.11) *where appropriate*. It can be very useful under particular circumstances, as mentioned at the beginning of this paper.

An example. The following example illustrates in detail the motivation outlined in section 1, i.e., that the condition number may be misleading in the characterization of the convergence behavior of the CG method. Consider the boundary value problem

$$-\nabla\cdot(k(x)\nabla u)=0 \quad \text{in } \Omega\,, \qquad u=u_D \quad \text{on } \partial\Omega\,, \tag{2.12}$$

where the domain $\Omega \equiv (-1,1)\times(-1,1)$ is divided into four subdomains Ω_i, $i=1, 2, 3, 4$, corresponding to the axis quadrants numbered counterclockwise. Let $k(x)$ be piecewise constant on the individual subdomains Ω_i, $k_1 = k_3 \approx 161.45$, $k_2 = k_4 = 1$. The Dirichlet boundary conditions are described in [30, section 5.3].

The numerical solution u of the problem (2.12) and the linear FE discretization, using the standard uniform triangulation, are shown in the left part of Figure 1. The resulting algebraic problem $\mathbf{Ax}=\mathbf{b}$ is preconditioned and solved by CG (algorithmically, $\mathbf{Ax}=\mathbf{b}$ is solved by the preconditioned conjugate gradient method (PCG)). In the right panel of Figure 1, we see the relative energy norm of the error as a function

[4] The early works relating iterations based on Chebyshev polynomials to CG are thoroughly recalled in the Historical note 5.5.3 of the monograph [24, pp. 254–256].

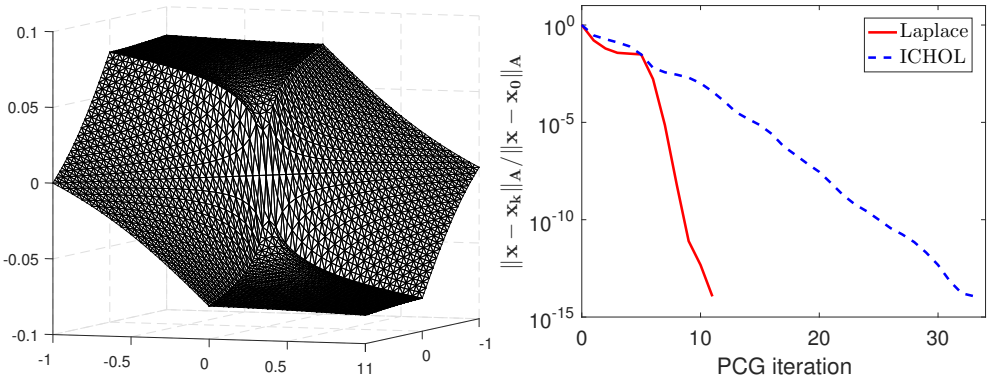

FIG. 1. *Left: The numerical solution of the problem* (2.12) *on the background of the linear FE triangulation. Right: The relative energy norm of the PCG error as a function of the iteration steps. The Laplace operator preconditioning (solid line) is much more efficient than the incomplete Cholesky preconditioning (dashed line), despite the fact that the condition numbers are* 161.54 *and close to* 16*, respectively. This can be explained by the differences in the associated distribution functions (see the end of section* 4 *below).*

of iteration steps for the Laplace operator preconditioning (solid line) and for the preconditioning using the algebraic incomplete Cholesky factorization of the matrix $\mathbf{A}$ (ICHOL) with the drop-off tolerance 10^{-2} (dashed line) where the problem has $N = 3969$ degrees of freedom. Despite the fact that the spectral condition number $\lambda_{max}/\lambda_{min}$ of the symmetrized preconditioned matrix for the Laplace operator preconditioning is an order of magnitude larger than for the ICHOL preconditioning, close to 161 and close to 16, respectively, PCG with the Laplace operator preconditioning clearly demonstrates much faster convergence.[5] This is due to the differences in the distribution of the eigenvalues with the nonnegligible components of the initial residuals in the direction of the associated eigenvectors and effects of rounding errors.

The spectra and distribution functions associated with the discretized preconditioned problems are given in Figure 2 for $N = 49$ degrees of freedom and in Figure 3 for $N = 3969$ degrees of freedom. Here, $\mathbf{L} = \mathbf{L}^{1/2}\mathbf{L}^{1/2}$ is the matrix associated with the discretized Laplace operator and $\mathbf{CC}^* \approx \mathbf{A}$ is the matrix resulting from ICHOL using the drop-off tolerance 10^{-2}, with the eigenvalues and eigenvectors of the associated generalized eigenvalue problems (see (1.2))

$$\mathbf{A}\mathbf{v}_i^{\mathbf{L}} = \lambda_i^{\mathbf{L}}\mathbf{L}\mathbf{v}_i^{\mathbf{L}}, \quad i = 1, \ldots, N,$$
$$\mathbf{A}\mathbf{v}_i^{\mathbf{C}} = \lambda_i^{\mathbf{C}}\mathbf{C}\mathbf{C}^*\mathbf{v}_i^{\mathbf{C}}, \quad i = 1, \ldots, N.$$

The weights of the distribution function $\omega^{\mathbf{L}}(\lambda)$ (respectively, $\omega^{\mathbf{C}}(\lambda)$), associated with the eigenvalues $\lambda_i^{\mathbf{L}}$ (respectively, $\lambda_i^{\mathbf{C}}$, $i = 1 \ldots, N$), related to the preconditioned algebraic systems

$$\mathbf{A}_{\mathbf{L}}\,(\mathbf{L}^{1/2}\mathbf{x}) = \mathbf{L}^{-1/2}\mathbf{b}, \quad \mathbf{A}_{\mathbf{L}} = \mathbf{L}^{-1/2}\mathbf{A}\mathbf{L}^{-1/2},$$

[5] Here we do not compare the overall computational cost, which can be affected by the cost of the individual preconditioned iterations depending on the domain and the discretization as well as on the function $k(x)$.

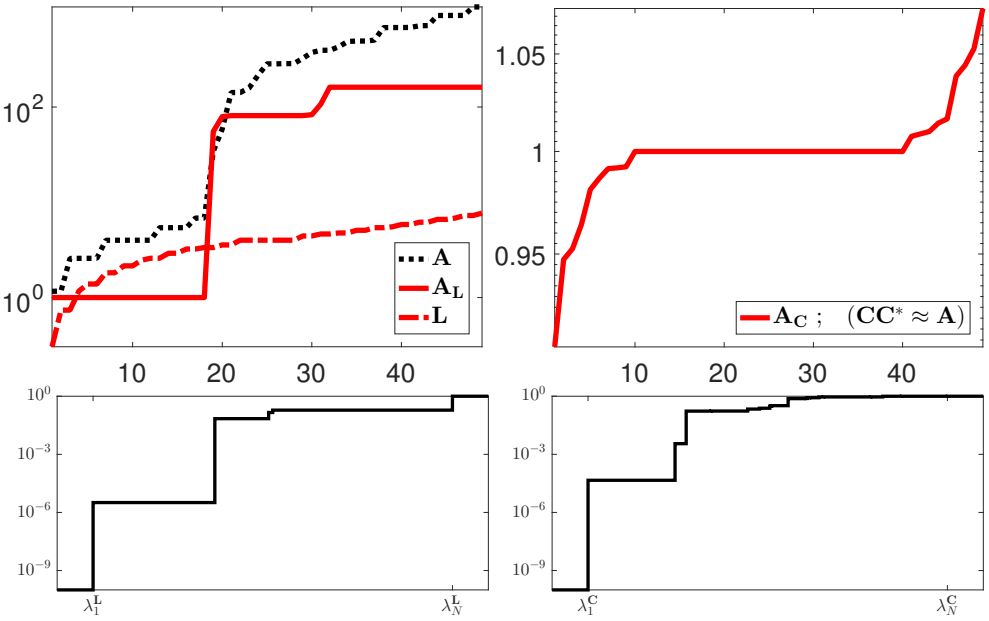

FIG. 2. *Top: Comparison of the spectra of the matrices* $\mathbf{A}$, $\mathbf{A_L}$, *and* $\mathbf{A_C}$, $N = 49$ *degrees of freedom. The dashed-dotted line in the upper left panel shows the eigenvalues of the discrete Laplace operator* $\mathbf{L}$ (2.5). *Due to a small drop-off tolerance, the eigenvalues of* $\mathbf{A}$ *and* $\mathbf{CC}^*$ *are graphically indistinguishable. Therefore, the right part only shows the eigenvalues of* $\mathbf{A_C}$ *(using a scale different from the left part of the figure). Bottom: Comparison of the distribution functions* $\omega^{\mathbf{L}}(\lambda)$ *(left) and* $\omega^{\mathbf{C}}(\lambda)$ *(right) associated with the preconditioned problems. The vertical axes are in the logarithmic scale, and* $\lambda_1^{\mathbf{L}} = 1$, $\lambda_N^{\mathbf{L}} = 161.45$, $\lambda_1^{\mathbf{C}} = 0.91$, $\lambda_N^{\mathbf{C}} = 1.07$.

respectively

$$\mathbf{A_C}\,(\mathbf{C}^*\mathbf{x}) = \mathbf{C}^{-1}\mathbf{b}, \quad \mathbf{A_C} = \mathbf{C}^{-1}\mathbf{A}\mathbf{C}^{-*},$$

are given by

$$\begin{aligned} \omega_i^{\mathbf{L}} &= |(\bar{\mathbf{v}}_i^{\mathbf{L}})^*\mathbf{q}^{\mathbf{L}}|^2, \quad i = 1, \ldots, N, \\ \omega_i^{\mathbf{C}} &= |(\bar{\mathbf{v}}_i^{\mathbf{C}})^*\mathbf{q}^{\mathbf{C}}|^2, \quad i = 1, \ldots, N. \end{aligned} \tag{2.13}$$

Here,

$$\bar{\mathbf{v}}_i^{\mathbf{L}} = \frac{\mathbf{L}^{1/2}\mathbf{v}_i^{\mathbf{L}}}{\|\mathbf{L}^{1/2}\mathbf{v}_i^{\mathbf{L}}\|} \quad \text{and} \quad \bar{\mathbf{v}}_i^{\mathbf{C}} = \frac{\mathbf{C}^*\mathbf{v}_i^{\mathbf{C}}}{\|\mathbf{C}^*\mathbf{v}_i^{\mathbf{C}}\|}$$

are the eigenvectors of the Hermitian and positive definite matrix $\mathbf{A_L}$ (respectively, $\mathbf{A_C}$), and

$$\mathbf{q}^{\mathbf{L}} = \frac{\mathbf{L}^{-1/2}\mathbf{b}}{\|\mathbf{L}^{-1/2}\mathbf{b}\|}, \quad \mathbf{q}^{\mathbf{C}} = \frac{\mathbf{C}^{-1}\mathbf{b}}{\|\mathbf{C}^{-1}\mathbf{b}\|}.$$

(We use the initial guess $\mathbf{x}_0 = 0$.) The distribution function $\omega^{\mathbf{C}}(\lambda)$ has its points of increase much more evenly distributed in the spectral interval $[\lambda_1(\mathbf{A_C}), \lambda_N(\mathbf{A_C})]$, which leads to a difference in the PCG convergence behavior. We will return to this issue, and offer a full explanation of the observed CG convergence behavior, after proving the main results and presenting their numerical illustrations.

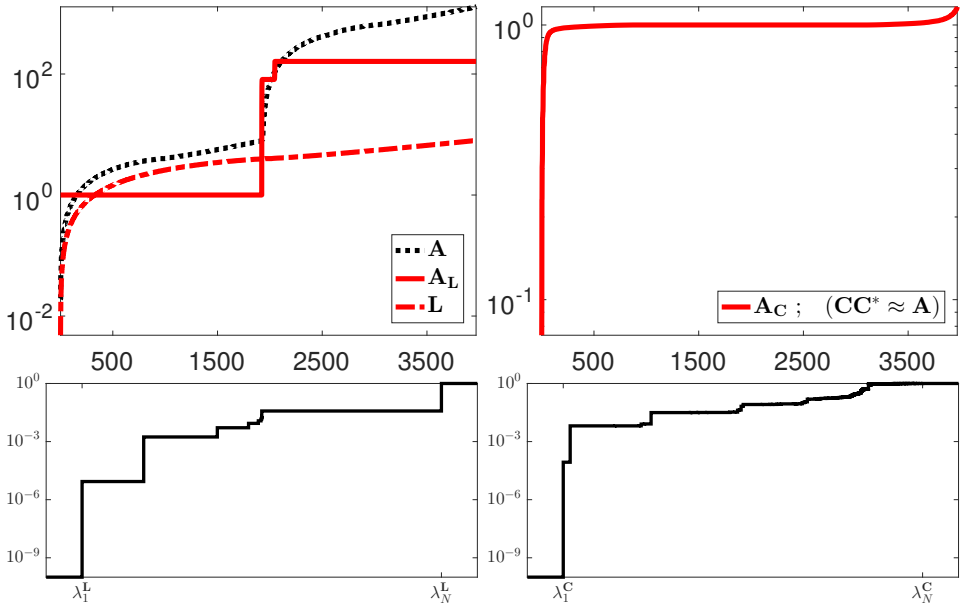

FIG. 3. *Top: Comparison of the spectra of the matrices* $\mathbf{A}$, $\mathbf{A_L}$, *and* $\mathbf{A_C}$, $N = 3969$ *degrees of freedom. The dashed-dotted line in the upper left panel shows the eigenvalues of the discrete Laplace operator* $\mathbf{L}$ (2.5). *Due to a small drop-off tolerance, the eigenvalues of* $\mathbf{A}$ *and* $\mathbf{CC}^*$ *are graphically indistinguishable. Therefore, the right part only shows the eigenvalues of* $\mathbf{A_C}$ *(using a scale different from the left part of the figure). Bottom: Comparison of the distribution functions* $\omega^{\mathbf{L}}(\lambda)$ *(left) and* $\omega^{\mathbf{C}}(\lambda)$ *(right) associated with the preconditioned problems. The vertical axes are in the logarithmic scale, and* $\lambda_1^{\mathbf{L}} = 1$, $\lambda_N^{\mathbf{L}} = 161.45$, $\lambda_1^{\mathbf{C}} = 7.4 \times 10^{-2}$, $\lambda_N^{\mathbf{C}} = 1.16$.

3. Analysis. As mentioned above, we will not only show that some function values of $k(x)$ are related to the spectrum of $\mathbf{L}^{-1}\mathbf{A}$, but also that there exists a one-to-one correspondence, i.e., a pairing, between the individual eigenvalues of $\mathbf{L}^{-1}\mathbf{A}$ and quantities given by the function values of $k(x)$ in relation to the supports of the FE basis functions. The proof does not require that $k(x)$ is continuous. If, moreover, $k(x)$ is constant on a part of the domain Ω that contains fully the supports of one or more basis functions, then the function value of $k(x)$ determines the associated eigenvalue *exactly* and the number of involved supports bounds from below the multiplicity of the associated eigenvalue. If $k(x)$ is slowly changing over the support of some basis function, then we get a very accurate localization of the associated eigenvalue.

Our approach is based upon the intervals

$$k(\mathcal{T}_j) \equiv \left[\inf_{x \in \mathcal{T}_j} k(x), \sup_{x \in \mathcal{T}_j} k(x) \right], \quad j = 1, \ldots, N, \tag{3.1}$$

where $\mathcal{T}_j = \text{supp}(\phi_j)$.[6] We will first formulate two main results. Theorem 3.1 localizes the positions of *all* the individual eigenvalues of the matrix $\mathbf{L}^{-1}\mathbf{A}$ by pairing them with the intervals $k(\mathcal{T}_j)$ given in (3.1). Using the given pairing, Corollary 3.2 describes the closeness of the eigenvalues to the nodal function values of the scalar function $k(x)$.

[6] If $k(x)$ is assumed only to be measurable and bounded (i.e., $k(x) \in L^\infty(\Omega)$), then the intervals can be defined as $k(\mathcal{T}_j) \equiv [\text{ess inf}_{x \in \mathcal{T}_j} k(x), \text{ess sup}_{x \in \mathcal{T}_j} k(x)]$. On the other hand, if $k(x)$ is continuous on $\mathcal{T}_j$, then $k(\mathcal{T}_j)$ coincides with the closure of the range of $k(x)$ over $\mathcal{T}_j$.

The proof of Theorem 3.1 combines perturbation theory for matrices with a classical result from the theory of bipartite graphs. For clarity of exposition, the proof will be presented after stating the corollaries of Theorem 3.1.

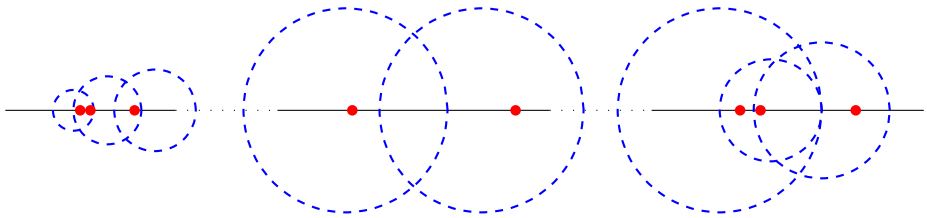

FIG. 4. *Illustration of Theorem* 3.1. *The diameters of the dashed circles indicate the size of the intervals* $k(\mathcal{T}_j)$, $j = 1, \ldots, N$. *The dots represent the eigenvalues* λ_j, $j = 1, \ldots, N$, *of the matrix* $\mathbf{L}^{-1}\mathbf{A}$. *We can find a pairing between the intervals* $k(\mathcal{T}_j)$ *and the eigenvalues* λ_i, *but the pairing may not be uniquely determined.*

THEOREM 3.1 (pairing the eigenvalues and the intervals $k(\mathcal{T}_j)$, $j = 1, \ldots, N$). *Using the previous notation, let* $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$ *be the eigenvalues of* $\mathbf{L}^{-1}\mathbf{A}$, *where* $\mathbf{A}$ *and* $\mathbf{L}$ *are defined by* (2.4) *and* (2.5), *respectively (with the subscript* h *dropped). As in* (1.1), *let* $k(x)$ *be bounded and piecewise continuous. Then there exists a (possibly nonunique) permutation* π *such that the eigenvalues of the matrix* $\mathbf{L}^{-1}\mathbf{A}$ *satisfy*

$$\lambda_{\pi(j)} \in k(\mathcal{T}_j), \quad j = 1, \ldots, N, \tag{3.2}$$

where the intervals $k(\mathcal{T}_j)$ *are as defined in* (3.1).

The statement is illustrated in Figure 4. The proof of the following corollary uses the one-to-one pairing of the intervals $k(\mathcal{T}_j)$ defined in (3.1), and therefore also the values of $k(\hat{x}_j)$ at any associated representatives $\hat{x}_j \in \mathcal{T}_j$, with the eigenvalues $\lambda_{\pi(j)}$.

COROLLARY 3.2 (pairing the eigenvalues and the nodal values; see Figure 5). *Using the notation and assumption of Theorem* 3.1, *consider any point* $\hat{x}_j$ *such that* $\hat{x}_j \in \mathcal{T}_j$. *Then the associated eigenvalue* $\lambda_{\pi(j)}$ *of the matrix* $\mathbf{L}^{-1}\mathbf{A}$ *satisfies*

$$|\lambda_{\pi(j)} - k(\hat{x}_j)| \le \sup_{x \in \mathcal{T}_j} |k(x) - k(\hat{x}_j)|, \quad j = 1, \ldots, N. \tag{3.3}$$

If, in addition, $k(x) \in \mathcal{C}^2(\mathcal{T}_j)$, *then*

$$\begin{aligned} |\lambda_{\pi(j)} - k(\hat{x}_j)| &\le \sup_{x \in \mathcal{T}_j} |k(x) - k(\hat{x}_j)| \\ &\le \hat{h}\|\nabla k(\hat{x}_j)\| + \tfrac{1}{2}\hat{h}^2 \sup_{x \in \mathcal{T}_j} \|D^2 k(x)\|, \quad j = 1, \ldots, N, \end{aligned} \tag{3.4}$$

where $\hat{h} = \operatorname{diam}(\mathcal{T}_j)$ *and* $D^2 k(x)$ *is the second order derivative of the function* $k(x)$.[7] *In particular,* (3.3) *and* (3.4) *hold for any discretization mesh node* $\hat{x}_j$ *such that* $\hat{x}_j \in \mathcal{T}_j$.

Proof. Since both $\lambda_{\pi(j)} \in k(\mathcal{T}_j)$ and $k(\hat{x}_j) \in k(\mathcal{T}_j)$, it trivially follows that

$$|\lambda_{\pi(j)} - k(\hat{x}_j)| \le \sup_{x \in \mathcal{T}_j} |k(x) - k(\hat{x}_j)|.$$

[7] See [5, section 1.2] for the definition of the second order derivative.

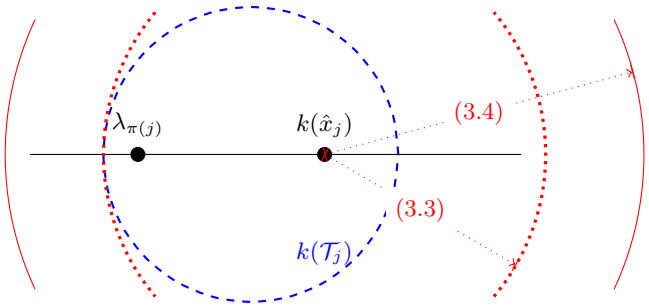

FIG. 5. *Illustration of Corollary* 3.2. *The relation* (3.2) *(indicated by the dashed blue circle) can give significantly better localization of the position of the individual eigenvalues than the bounds* (3.3) *(indicated by the dotted red circle) and* (3.4) *(indicated by the solid red circle). When* $k(x)$ *is constant over* $\mathcal{T}_j$*, then* $k(\mathcal{T}_j)$ *reduces to one point* $\lambda_{\pi(j)}$*; see also* (3.3). *The bound* (3.4) *is weaker than* (3.2) *and* (3.3)*, but the evaluation of its first term might be easier in practice.*

Moreover, for any $x \in \mathcal{T}_j$, the multidimensional Taylor expansion (see, e.g., [5, p. 11, section 1.2]) gives for $k(x) \in \mathcal{C}^2(\mathcal{T}_j)$ that

$$k(x) - k(\hat{x}_j) = \nabla k(\hat{x}_j)(x - \hat{x}_j) + \tfrac{1}{2} D^2 k(\hat{x}_j + \alpha(x - \hat{x}_j))(x - \hat{x}_j, x - \hat{x}_j),$$

where $\alpha \in [0, 1]$, with the absolute value obeying

$$|k(x) - k(\hat{x}_j)| \le \|\nabla k(\hat{x}_j)\| \|(x - \hat{x}_j)\| + \tfrac{1}{2} \|x - \hat{x}_j\|^2 \|D^2 k(\hat{x}_j + \alpha(x - \hat{x}_j))\|,$$

giving the statement. ∎

We now give the proof of Theorem 3.1. Lemma 3.3 below and its Corollary 3.4, identify the groups of eigenvalues in any union of intervals

$$\bar{k}(\mathcal{T}_{\mathcal{J}}) \equiv \bigcup_{j \in \mathcal{J}} k(\mathcal{T}_j), \quad \mathcal{J} \subset \{1, \ldots, N\}. \tag{3.5}$$

This enables us to apply Hall's theorem (see [4, Theorem 5.2] or, e.g., [17, Theorem 1]) to prove Theorem 3.1. (For the sake of completeness, we have also formulated Hall's result below in Theorem 3.5.)

LEMMA 3.3. *Using the notation introduced above and the assumption of Theorem* 3.1*, let* $\mathcal{J} \subset \{1, \ldots, N\}$ *and* $\mathcal{T}_{\mathcal{J}} = \bigcup_{j \in \mathcal{J}} \mathcal{T}_j$*. Then there exist at least* $p = |\mathcal{J}|$ *eigenvalues* $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_p$ *of* $\mathbf{L}^{-1}\mathbf{A}$ *such that*

$$\tilde{\lambda}_\ell \in \left[\inf_{x \in \mathcal{T}_{\mathcal{J}}} k(x), \sup_{x \in \mathcal{T}_{\mathcal{J}}} k(x) \right], \quad \ell = 1, \ldots, p. \tag{3.6}$$

Proof. In brief, the proof is based on the theory of eigenvalue perturbations of matrices. We locally modify the scalar function $k(x)$ by setting it equal to a positive constant K in the union $\mathcal{T}_{\mathcal{J}}$ of the supports $\mathcal{T}_j$, $j \in \mathcal{J}$. This will result, after discretization, in a modified matrix $\tilde{\mathbf{A}}_{\mathcal{J}}$ such that K is an eigenvalue of $\mathbf{L}^{-1}\tilde{\mathbf{A}}_{\mathcal{J}}$ of at least p multiplicity. An easy bound for the eigenvalues of

$$\mathbf{L}^{-1}\mathbf{E}_{\mathcal{J}}, \quad \text{where} \quad \mathbf{E}_{\mathcal{J}} = \mathbf{A} - \tilde{\mathbf{A}}_{\mathcal{J}}, \tag{3.7}$$

combined with a standard perturbation theorem for matrices, then provides a bound for the associated p eigenvalues of $\mathbf{L}^{-1}\mathbf{A}$. A particular choice of the positive constant K will finish the proof.

Let

$$\tilde{k}_{\mathcal{J}}(x) = \begin{cases} K & \text{for} \quad x \in \mathcal{T}_{\mathcal{J}}, \\ k(x) & \text{otherwise}, \end{cases}$$

with

$$\left\langle \tilde{\mathcal{A}}_{\mathcal{J},h} u, v \right\rangle \equiv \int_\Omega \nabla u \cdot \tilde{k}_{\mathcal{J}} \nabla v \,, \quad u, v \in V_h,$$

where, analogously to (2.4),

$$[\tilde{\mathbf{A}}_{\mathcal{J}}]_{ij} = \left\langle \tilde{\mathcal{A}}_{\mathcal{J},h} \phi_j, \phi_i \right\rangle = \int_\Omega \nabla \phi_i \cdot \tilde{k}_{\mathcal{J}} \nabla \phi_j, \quad i,j = 1, \ldots, N.$$

Since $\tilde{k}_{\mathcal{J}}$ is constant on each $\mathcal{T}_j$, $j \in \mathcal{J}$, and the support of the basis function ϕ_j is $\mathcal{T}_j$, it holds for any $v \in V_h$ that

$$\left\langle \tilde{\mathcal{A}}_{\mathcal{J},h} \phi_j, v \right\rangle = \int_\Omega \nabla \phi_j \cdot \tilde{k}_{\mathcal{J}} \nabla v \, = \int_{\mathcal{T}_j} \nabla \phi_j \cdot \tilde{k}_{\mathcal{J}} \nabla v \tag{3.8}$$

$$= K \int_{\mathcal{T}_j} \nabla \phi_j \cdot \nabla v \, = K \left\langle \mathcal{L}_h \phi_j, v \right\rangle, \quad j \in \mathcal{J}. \tag{3.9}$$

Thus, K is an eigenvalue of the operator $\mathcal{L}_h^{-1} \tilde{\mathcal{A}}_{\mathcal{J},h}$ associated with the eigenfunctions ϕ_j, $j \in \mathcal{J}$, and therefore K is the eigenvalue of the matrix $\mathbf{L}^{-1}\tilde{\mathbf{A}}_{\mathcal{J}}$ with the multiplicity at least p. This can also be verified by construction by observing that

$$\tilde{\mathbf{A}}_{\mathcal{J}} \mathbf{e}_j = K \, \mathbf{L} \mathbf{e}_j, \quad j \in \mathcal{J}.$$

Consider now the eigenvalues of $\mathbf{L}^{-1}\mathbf{E}_{\mathcal{J}}$; see (3.7). The Rayleigh quotient for an eigenpair $(\theta, \mathbf{q})$, $\|\mathbf{q}\| = 1$, and the associated eigenfunction $q = \sum_{j=1}^N \nu_j \phi_j$, where $\mathbf{q}^T = [\nu_1, \ldots, \nu_N]$, satisfies

$$\theta = \frac{\mathbf{q}^T \mathbf{E}_{\mathcal{J}} \mathbf{q}}{\mathbf{q}^T \mathbf{L} \mathbf{q}} = \frac{\mathbf{q}^T (\mathbf{A} - \tilde{\mathbf{A}}_{\mathcal{J}}) \mathbf{q}}{\mathbf{q}^T \mathbf{L} \mathbf{q}} = \frac{\left\langle (\mathcal{A}_h - \tilde{\mathcal{A}}_{\mathcal{J},h}) q, q \right\rangle}{\langle \mathcal{L}_h q, q \rangle}$$

$$= \frac{\int_\Omega \nabla q \cdot (k(x) - \tilde{k}_{\mathcal{J}}(x)) \nabla q \, dx}{\int_\Omega \|\nabla q\|^2 \, dx} = \frac{\int_{\mathcal{T}_{\mathcal{J}}} (k(x) - K) \|\nabla q\|^2 \, dx}{\int_\Omega \|\nabla q\|^2 \, dx},$$

giving

$$|\theta| \leq \sup_{x \in \mathcal{T}_{\mathcal{J}}} |k(x) - K|. \tag{3.10}$$

Next, consider the symmetric matrices

$$\mathbf{A}_{\mathbf{L}} = \mathbf{L}^{-1/2} \mathbf{A} \mathbf{L}^{-1/2}, \quad \mathbf{E}_{\mathbf{L}} = \mathbf{L}^{-1/2} \mathbf{E}_{\mathcal{J}} \mathbf{L}^{-1/2}, \quad \tilde{\mathbf{A}}_{\mathbf{L}} = \mathbf{L}^{-1/2} \tilde{\mathbf{A}}_{\mathcal{J}} \mathbf{L}^{-1/2}.$$

According to a standard result from the perturbation theory of symmetric matrices (see, e.g., [35, Corollary 4.9, p. 203]), we find that

$$\lambda_s(\mathbf{A}_{\mathbf{L}}) = \lambda_s(\tilde{\mathbf{A}}_{\mathbf{L}} + \mathbf{E}_{\mathbf{L}}) \in [\lambda_s(\tilde{\mathbf{A}}_{\mathbf{L}}) + \theta_{min}, \lambda_s(\tilde{\mathbf{A}}_{\mathbf{L}}) + \theta_{max}], \quad s = 1, \ldots, N,$$

where θ_{min} and θ_{max} are the smallest and largest eigenvalues of $\mathbf{E_L}$, respectively. Since the matrices $\mathbf{L}^{-1}\mathbf{A}$, $\mathbf{L}^{-1}\mathbf{E}_{\mathcal{J}}$, and $\mathbf{L}^{-1}\tilde{\mathbf{A}}_{\mathcal{J}}$ have the same spectrum as the matrices $\mathbf{A_L}$, $\mathbf{E_L}$, and $\tilde{\mathbf{A}}_{\mathbf{L}}$, respectively, it follows that

$$\lambda_s(\mathbf{L}^{-1}\mathbf{A}) = \lambda_s(\mathbf{L}^{-1}\tilde{\mathbf{A}}_{\mathcal{J}} + \mathbf{L}^{-1}\mathbf{E}_{\mathcal{J}}) \in [\lambda_s(\mathbf{L}^{-1}\tilde{\mathbf{A}}_{\mathcal{J}}) + \theta_{min}, \lambda_s(\mathbf{L}^{-1}\tilde{\mathbf{A}}_{\mathcal{J}}) + \theta_{max}].$$

Due to (3.10),

$$\theta_{min} \geq -\sup_{x\in\mathcal{T}_{\mathcal{J}}} |k(x) - K|,$$
$$\theta_{max} \leq \sup_{x\in\mathcal{T}_{\mathcal{J}}} |k(x) - K|,$$

and thus, since K is at least a p-multiple eigenvalue of $\mathbf{L}^{-1}\tilde{\mathbf{A}}_{\mathcal{J}}$, there exist p eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_p$ of $\mathbf{L}^{-1}\mathbf{A}$ such that

$$\tilde{\lambda}_\ell \in \left[K - \sup_{x\in\mathcal{T}_{\mathcal{J}}} |k(x) - K|,\ K + \sup_{x\in\mathcal{T}_{\mathcal{J}}} |k(x) - K| \right], \quad \ell = 1, \ldots, p. \tag{3.11}$$

Setting

$$K = \tfrac{1}{2}\left(\inf_{x\in\mathcal{T}_{\mathcal{J}}} k(x) + \sup_{x\in\mathcal{T}_{\mathcal{J}}} k(x) \right)$$

gives

$$\tilde{\lambda}_\ell \in \left[\inf_{x\in\mathcal{T}_{\mathcal{J}}} k(x),\ \sup_{x\in\mathcal{T}_{\mathcal{J}}} k(x) \right], \quad \ell = 1, \ldots, p. \qquad \square$$

Applying Lemma 3.3 N times with $\mathcal{J} = \{1\}$, $\mathcal{J} = \{2\}, \ldots, \mathcal{J} = \{N\}$, we see that, for the support of any basis function ϕ_j, there is an eigenvalue $\tilde{\lambda}$ of $\mathbf{L}^{-1}\mathbf{A}$ such that $\tilde{\lambda} \in k(\mathcal{T}_j)$. Moreover, as an additional important consequence, for any subset $\mathcal{J} \subset \{1, \ldots, N\}$ the associated union of intervals $\bar{k}(\mathcal{T}_{\mathcal{J}})$ (see (3.5)) contains at least $p = |\mathcal{J}|$ eigenvalues of $\mathbf{L}^{-1}\mathbf{A}$; see the following corollary.

COROLLARY 3.4. *Let, as above, $\mathcal{J} \subset \{1, \ldots, N\}$ and $\mathcal{T}_{\mathcal{J}} = \cup_{j\in\mathcal{J}}\mathcal{T}_j$. Then there exist at least $p = |\mathcal{J}|$ eigenvalues $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_p$ of $\mathbf{L}^{-1}\mathbf{A}$ such that*

$$\tilde{\lambda}_\ell \in \bar{k}(\mathcal{T}_{\mathcal{J}}) \equiv \bigcup_{j\in\mathcal{J}} k(\mathcal{T}_j), \quad \ell = 1, \ldots, p. \tag{3.12}$$

Moreover, taking $\mathcal{J} = \{1, \ldots, N\}$, (3.12) immediately implies that any eigenvalue $\tilde{\lambda}$ of $\mathbf{L}^{-1}\mathbf{A}$ belongs to (at least one) interval $k(\mathcal{T}_j)$, $j \in \{1, \ldots, N\}$.

Proof. Since $\bar{k}(\mathcal{T}_j) = k(\mathcal{T}_j)$, for any $j \in \mathcal{J}$, is an interval (3.1), the set $\bar{k}(\mathcal{T}_{\mathcal{J}})$ consists of at most p intervals. We decompose $\bar{k}(\mathcal{T}_{\mathcal{J}})$ into $\tilde{p}$ mutually disjoint intervals, $\tilde{p} \leq p$, and denote

$$\bar{k}(\mathcal{T}_{\mathcal{J}_i}) \equiv \bigcup_{j\in\mathcal{J}_i} k(\mathcal{T}_j), \quad i = 1, \ldots, \tilde{p}.$$

Lemma 3.3 then ensures that each interval $\bar{k}(\mathcal{T}_{\mathcal{J}_i})$ contains at least $|\mathcal{J}_i|$ eigenvalues of $\mathbf{L}^{-1}\mathbf{A}$. Summing up, at least $\sum_{i=1,\ldots,\tilde{p}} |\mathcal{J}_i| = |\mathcal{J}|$ eigenvalues of $\mathbf{L}^{-1}\mathbf{A}$ must be contained in the union $\bar{k}(\mathcal{T}_{\mathcal{J}})$. $\square$

In order to finalize the proof of Theorem 3.1, we still need to show the existence of a one-to-one pairing between the individual eigenvalues and the individual intervals $k(\mathcal{T}_j),\ j = 1, \ldots, N$. The relationship between the intervals $k(\mathcal{T}_j),\ j = 1, \ldots, N$, and the eigenvalues of $\mathbf{L}^{-1}\mathbf{A}$ described in Lemma 3.3 and Corollary 3.4 can be represented by the following bipartite graph. Let, as above, $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$ be the eigenvalues of $\mathbf{L}^{-1}\mathbf{A}$. Consider the bipartite graph

$$(\mathcal{S}, \mathcal{I}, E) \tag{3.13}$$

with the sets of nodes $\mathcal{S} = \mathcal{I} = \{1, \ldots, N\}$ and the set of edges E, where

$$\{s, i\} \in E \quad \text{if and only if} \quad \lambda_s \in k(\mathcal{T}_i), \quad s \in \mathcal{S},\ i \in \mathcal{I}.$$

A subset of edges $M \subset E$ is called matching if no edges from M share a common node; see [4, section 5.1]. We will use the following famous theorem.

THEOREM 3.5 (Hall's theorem). *Let $(\mathcal{S}, \mathcal{I}, E)$ be a bipartite graph. Given $\mathcal{J} \subset \mathcal{I}$, let $G(\mathcal{J}) \subset \mathcal{S}$ denote the set of all nodes adjacent to any node from $\mathcal{J}$, i.e.,*

$$G(\mathcal{J}) = \{s \in \mathcal{S};\ \exists i \in \mathcal{J} \text{ such that } \{s, i\} \in E\}.$$

Then there exists a matching $M \subset E$ that covers $\mathcal{I}$ if and only if

$$|G(\mathcal{J})| \ge |\mathcal{J}| \quad \text{for any} \quad \mathcal{J} \subset \mathcal{I}; \tag{3.14}$$

see, e.g., [4, Theorem 5.2] *and the original formulation* [17, Theorem 1].

Now we are ready to finalize our argument.

Proof of Theorem 3.1. Consider the bipartite graph defined by (3.13), and let $G(\mathcal{J}) \subset \mathcal{S}$ be the set of all nodes (representing the eigenvalues) adjacent to any node from $\mathcal{J}$, $\mathcal{J} \subset \mathcal{I}$ (representing the intervals). In other words, $G(\mathcal{J})$ represents the indices of all eigenvalues $\{\lambda_s;\ s \in G(\mathcal{J})\}$ located in $\bar{k}(\mathcal{T}_\mathcal{J}) = \cup_{j \in \mathcal{J}} k(\mathcal{T}_j)$. Corollary 3.4 of Lemma 3.3 ensures that assumption (3.14) in Theorem 3.5 is satisfied, i.e.,

$$|G(\mathcal{J})| \ge |\mathcal{J}|. \tag{3.15}$$

Thus, according to Theorem 3.5, there exists a matching $M \subset E$ that covers $\mathcal{I}$. Since $|\mathcal{I}| = |\mathcal{S}|$, this matching defines the permutation $\pi(i),\ i = 1, 2, \ldots, N$, such that

$$\lambda_{\pi(i)} \in k(\mathcal{T}_i), \quad i = 1, \ldots, N,$$

which finishes the proof.

4. Numerical experiments. In this section, we will illustrate the theoretical results by a series of numerical experiments. We will investigate how well the nodal values of k correspond to the eigenvalues and assess the sharpness of the estimates in Corollary 3.2 in a few examples, including both uniform and local mesh refinement. Furthermore, we will compute the corresponding intervals $k(\mathcal{T}_i),\ i = 1, \ldots, N$, and consider the pairing in Theorem 3.1.

Test problems. We will consider four test problems defined on the domain $\Omega \equiv (0, 1) \times (0, 1)$, where we slightly abuse notation above and let $k = k(x, y)$. The first three problems use a continuous coefficient function $k(x, y)$:

(P1) $k(x, y) = \sin(x + y)$,
(P2) $k(x, y) = 1 + 50 \exp(-5(x^2 + y^2))$,
(P3) $k(x, y) = 2^7(x^7 + y^7)$.

The fourth problem uses a discontinuous function $k(x, y)$,

$$\text{(P4)} \quad k(x,y) = \begin{cases} \text{(P1)} & \text{for } (x,y) \in (0,1) \times (\tfrac{1}{2}, 1), \\ \text{(P2)} & \text{otherwise.} \end{cases}$$

Numerical experiments were computed using FEniCS [25] and MATLAB.[8] If not specified otherwise, we consider a triangular uniform mesh with piecewise linear discretization basis functions.

4.1. Illustration of Theorem 3.1 and Corollary 3.2. Throughout this section, we consider the increasingly sorted eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N$ and the increasingly sorted values $k_1 \leq k_2 \leq \cdots \leq k_N$ of the coefficient function k at the nodes (x_j, y_j), $\phi_i(x_j, y_j) = \delta_{ij}$, which determines the reordering R,

$$k_{R(i)} = k(x_i, y_i), \quad i = 1, 2, \ldots, N. \tag{4.1}$$

In Figure 6, we show the nodal values $k_1, \ldots, k_N$ and the eigenvalues $\lambda_1, \ldots, \lambda_N$ for the problem with $N = 81$ degrees of freedom. Clearly, there is a close correspondence between the nodal values and the eigenvalues even at this relatively coarse resolution, but there are some notable differences for (P3) and (P4) that are clearly visible: The nodal values of k provide much better approximations of the eigenvalues in the cases (P1) and (P2) than in the examples (P3) and (P4). In (P3), even though the function $k(x, y)$ is smooth, the norm of the gradient $\|\nabla k\|$ is large in significant subregions, which results, in accordance with Corollary 3.2, in less accurate correspondence between the nodal values and the eigenvalues at the coarse grid. In (P4), a discontinuous coefficient function is used.

Theorem 3.1 states that there exists a permutation π such that $\lambda_{\pi(i)} \in k(\mathcal{T}_i)$ for every $i = 1, \ldots, N$. The proof is not constructive, and it is therefore interesting to consider potential pairings. In Figure 7, we show the results for the pairing defined by sorting the eigenvalues and the nodal values of k increasingly (which gives the ordering of the associated intervals). The pairing appears to work quite well, except for the case (P4), where in particular the eigenvalues between 30–40 are outside the intervals provided by this pairing.

In order to ensure that we employ a proper pairing, i.e., to guarantee that $\lambda_{\pi(i)} \in k(\mathcal{T}_i)$, $i = 1, \ldots, N$, we construct the adjacency matrix $\mathbf{G}$ such that

$$\mathbf{G}_{si} = \begin{cases} 1, & \lambda_s \in k(\mathcal{T}_i), \\ 0, & \lambda_s \notin k(\mathcal{T}_i). \end{cases} \tag{4.2}$$

By using the Dulmage–Mendelsohn decomposition[9] of this adjacency matrix $\mathbf{G}$ (provided by the MATLAB command `dmperm`) we get a pairing π satisfying $\lambda_{\pi(i)} \in k(\mathcal{T}_i)$ for every $i = 1, \ldots, N$. Figure 8 illustrates the pairing π from Theorem 3.1 for (P4) and the comparison of the eigenvalues $\lambda_{\pi(i)}$ with the nodal values $k(x_i, y_i)$ associated with the basis function ϕ_i and the interval $k(\mathcal{T}_i)$ (the plots in Figure 8 should be compared with the lower right panels of Figures 6 and 7). To summarize, the pairing π from the Theorem 3.1 can be for problems (P1)–(P3) identified with the reordering R in (4.1). For problem (P4), the pairing can be given via the Dulmage–Mendelsohn decomposition of the adjacency matrix.

[8] FEniCS version 2017.2.0 and MATLAB Version: 8.0.0.783 (R2012b).

[9] See, e.g., the original paper [7].

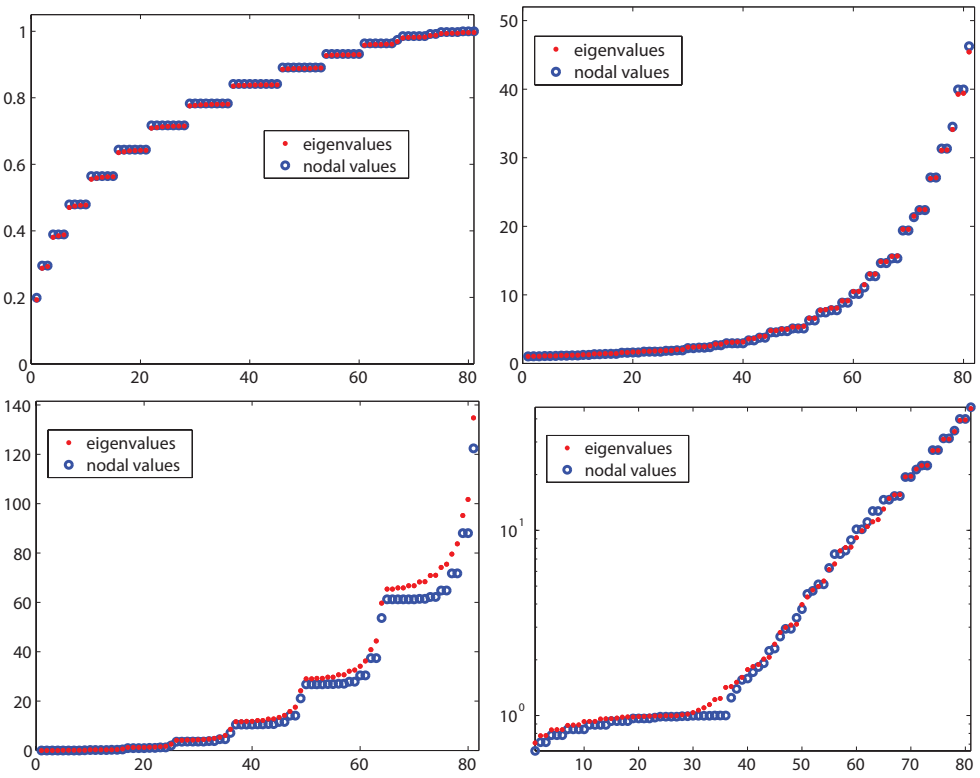

FIG. 6. *Comparison of the eigenvalues λ_s, $s = 1, \ldots, N$ (red dots), and the increasingly sorted nodal values of k (blue circles). Top left:* (P1)*, top right:* (P2)*, bottom left:* (P3)*, bottom right:* (P4)*. As in Figure* 7*, we use the semilogarithmic scale in the lower right panel* (P4).

The difference between the nodal values and the corresponding eigenvalues is estimated in (3.4), and to assess the sharpness of this estimate, Figure 9 compares the quantities $|\lambda_{\pi(i)} - k(x_i, y_i)|$ (red dots) with the first term on the right-hand side of (3.4) (black stars). We observe that the first term of (3.4) in general overestimates the differences at this coarse resolution.

4.2. Effects of *h*-adaptivity. Corollary 3.2 states that the estimated difference $|\lambda_{\pi(i)} - k(x_i, y_i)|$ improves at least linearly as the mesh is refined. Figure 10 shows the improvement of both the nodal value estimates of k and the associated intervals $k(\mathcal{T}_i)$ for problems (P1) and (P3) with $N = 59^2 = 3481$ degrees of freedom.

Corollary 3.2 is a local estimate which allows local mesh refinement for improving accuracy of the eigenvalue estimate. To see the effect of locally refined mesh on the spectrum of the preconditioned problem, we consider the test problem (P2), where we refine the mesh in the subdomain $[0, 0.2] \times [0, 0.2]$, i.e., in the area with large gradient of the function $k(x, y)$. Figure 11 shows the discretization mesh (top), the eigenvalues with the associated intervals (middle), and the associated nodal values (bottom). As expected, we observe more eigenvalues in the upper part of the spectrum as well as their better localization; see also, for comparison, the top right panels of Figures 6 and 7.

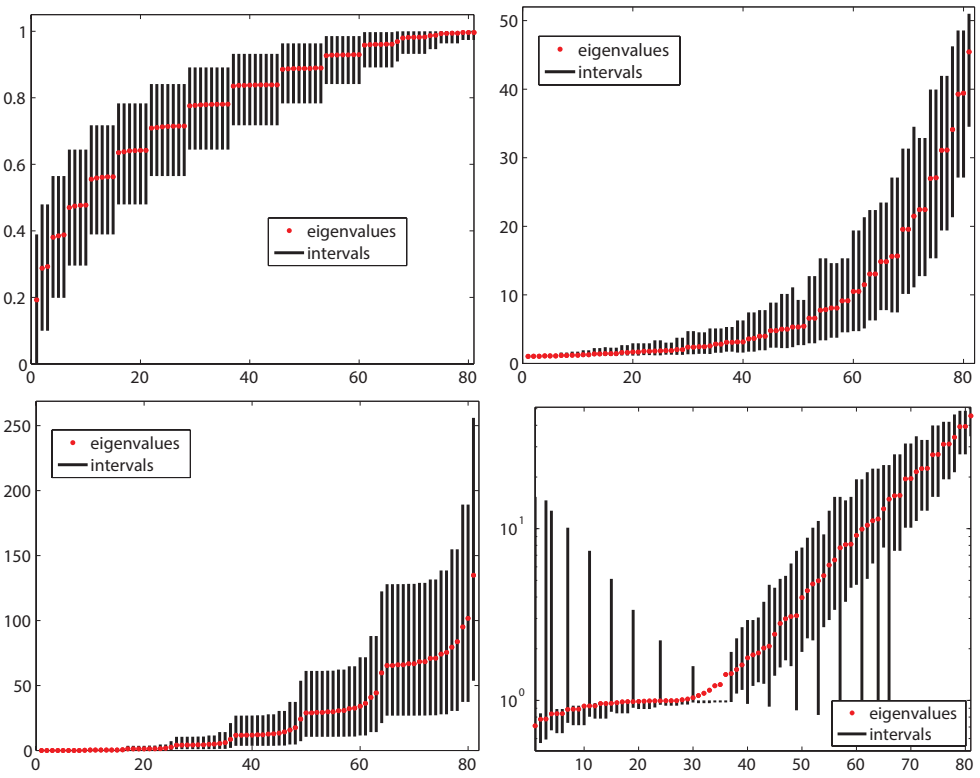

FIG. 7. *The eigenvalues $\lambda_1 \leq \cdots \leq \lambda_N$ (red dots) and the associated intervals (black vertical lines), where the corresponding pairing is defined by sorting of the nodal values of k increasingly; see* (4.1) *and Figure* 6. *Top left:* (P1), *top right:* (P2), *bottom left:* (P3), *bottom right:* (P4). *We observe that for* (P4) *some of the eigenvalues are not inside the associated intervals, and therefore the ordering based on pairing the eigenvalues with increasing nodal values of k does not in this case conform to π from Theorem* 3.1.

4.3. Re-entrant corner domain. The local considerations of Corollary 3.2 do not require additional regularity for the solutions of the associated PDEs, and our theoretical results are valid for domains of any shape. To illuminate that no additional regularity is needed, we conduct experiments with function $k(x, y)$ from test problem (P3) on a domain with a re-entrant corner, i.e.,

$$\Omega = [0,1] \times [0,1] \setminus \{(x,y) : x > 0.8y + 0.1 \quad \text{and} \quad y < 0.8x + 0.1\}.$$

The domain is shown in the left panel in Figure 12. The right panel shows the eigenvalues $\lambda_1, \ldots, \lambda_N$ (red dots) with the increasingly sorted nodal values $k_1, \ldots, k_N$ (green circles) and the associated intervals.

4.4. Convergence of the introductory example explained. We will now finish our exposition by returning back to the motivation example presented in section 2 and by explaining the difference in the behavior of PCG with the Laplace operator preconditioning and with the ICHOL preconditioning; see the right part of Figure 1.

First we present Figure 13, a modification of Figure 1, showing that at the fifth

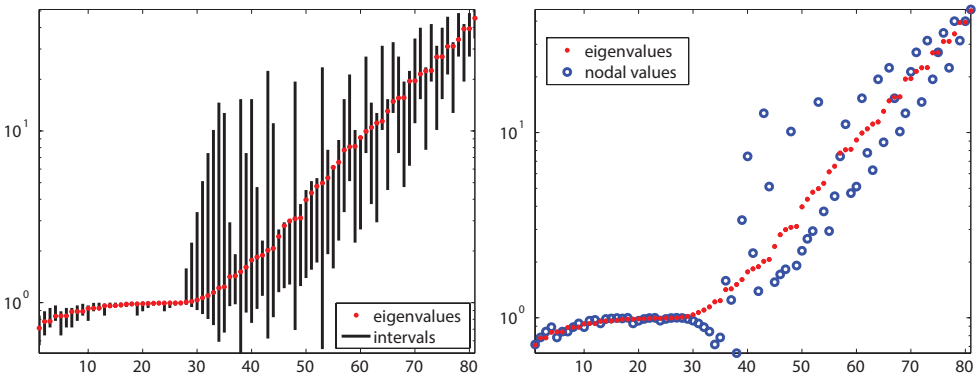

FIG. 8. *Illustration of the pairing π computed by the Dulmage–Mendelsohn decomposition of the corresponding adjacency matrix* $\mathbf{G}$ *(see* (4.2)*) for problem* (P4). *Left: The intervals $k(\mathcal{T}_i)$ (black vertical lines) are paired with the associated eigenvalues $\lambda_{\pi(i)}$ (red dots). The pairs are then plotted (for nicer graphical appearance) in increasing order of the size of the eigenvalues (not in the order $\pi(i)$, $i = 1, \ldots, N$). Right: The comparison of the eigenvalues (using the same ordering as in the left part) and the associated nodal values $k(x_i, y_i)$ (blue circles).*

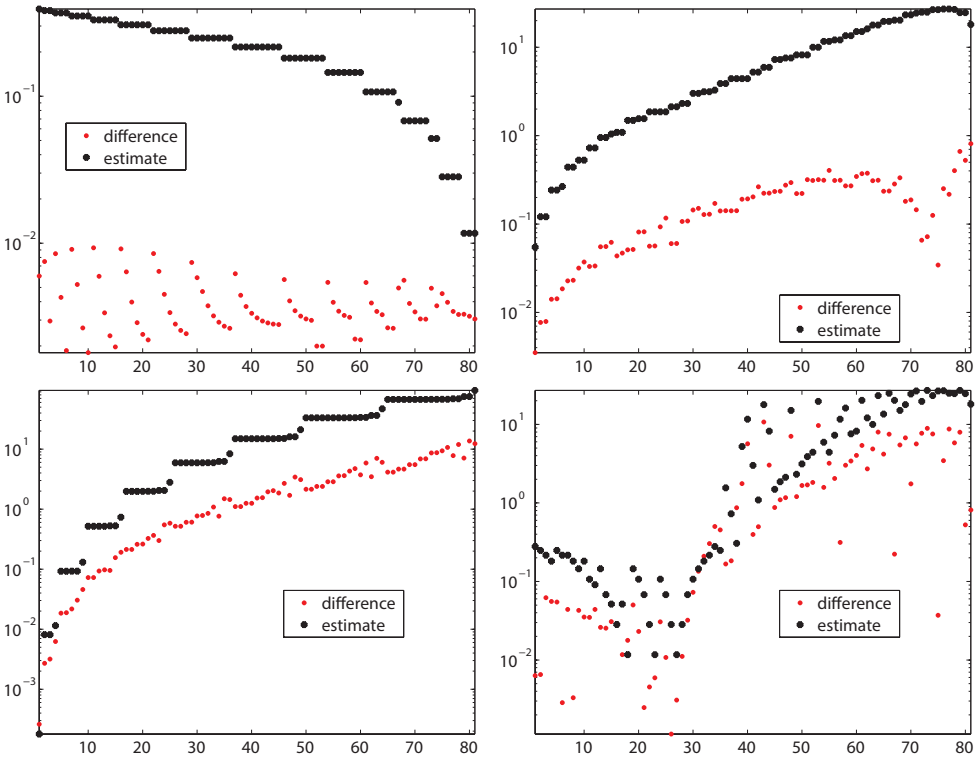

FIG. 9. *Illustration of Corollary* 3.2. *Comparison of the absolute difference $|\lambda_{\pi(i)} - k(x_i, y_i)|$ (red dots) and its estimate by the first term on the right-hand side of* (3.4) *(black stars). Top left:* (P1), *top right:* (P2), *bottom left:* (P3), *bottom right:* (P4).

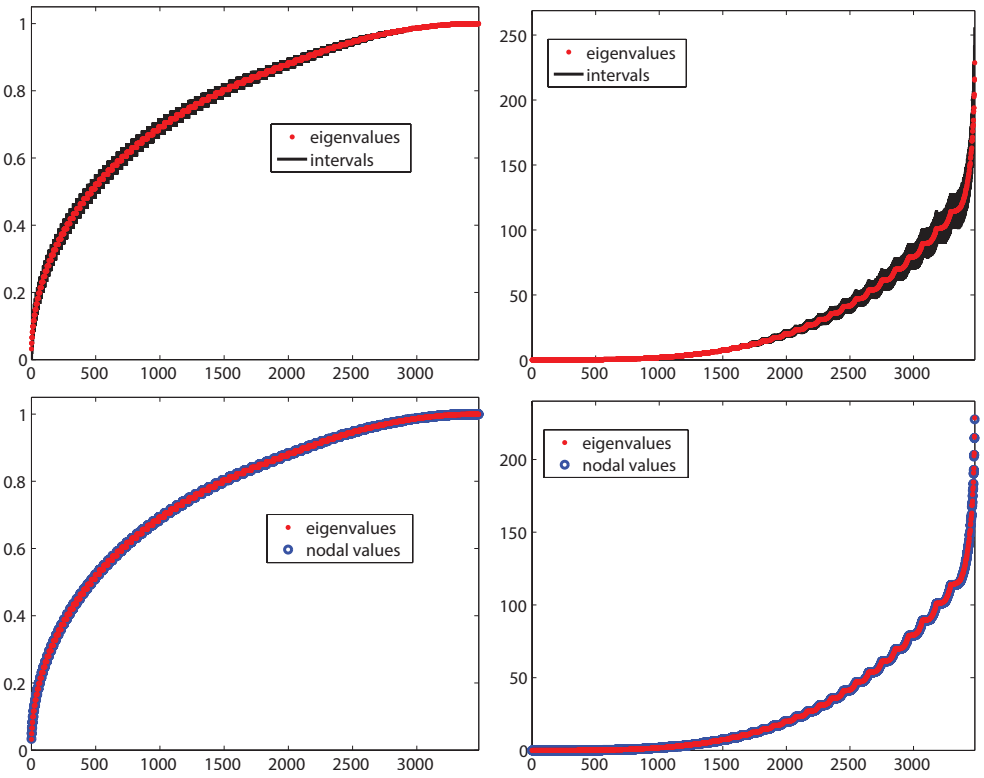

FIG. 10. *Top: The intervals $k(\mathcal{T}_i)$ (black vertical lines) and the associated eigenvalues $\lambda_{\pi(i)}$ (red dots) (plotted in increasing order of their size). Bottom: Comparison of the eigenvalues $\lambda_{\pi(i)}$ (plotted in increasing order of their size) and the associated nodal values $k(x_i, y_i)$ (blue circles). Here we use uniform mesh with $N = 3481$ degrees of freedom. Left:* (P1)*, right:* (P3)*. We can observe a dramatic improvement of the approximation accuracy; cf. Figures* 6 *and* 7.

iteration we can identify with remarkable accuracy the slope of the PCG convergence curves for most of the subsequent iterations, with the convergence being almost linear without a substantial acceleration. The rate of convergence is for the Laplace operator preconditioning remarkably faster than for the ICHOL preconditioning.

The convergence of the PCG method with the Laplace operator preconditioning can be completely explained using Theorem 3.1 and the results about the CG convergence behavior from the literature. Since $k(x)$ is in the given experiment constant for most of the supports of the basis functions (being equal to one (respectively, to 161.45)), according to Theorem 3.1 the preconditioned system matrix must have many multiple eigenvalues equal to one (respectively, to 161.45). This is illustrated by the computed quantities presented in Table 1. We see that 1922 eigenvalues are equal to one, 1922 are equal to 161.45, and the rest are spread between ≈ 28 and ≈ 134 (with the eigenvalues between 81.226 and 134 of weight so negligible (see (2.13)) that they do not contribute within the small number of iterations to the computations; they are for CG computations within the given number of iterations practically not visible; see [24, section 5.6.4]).

Assuming exact arithmetic, van der Sluis and van der Vorst prove in the seminal

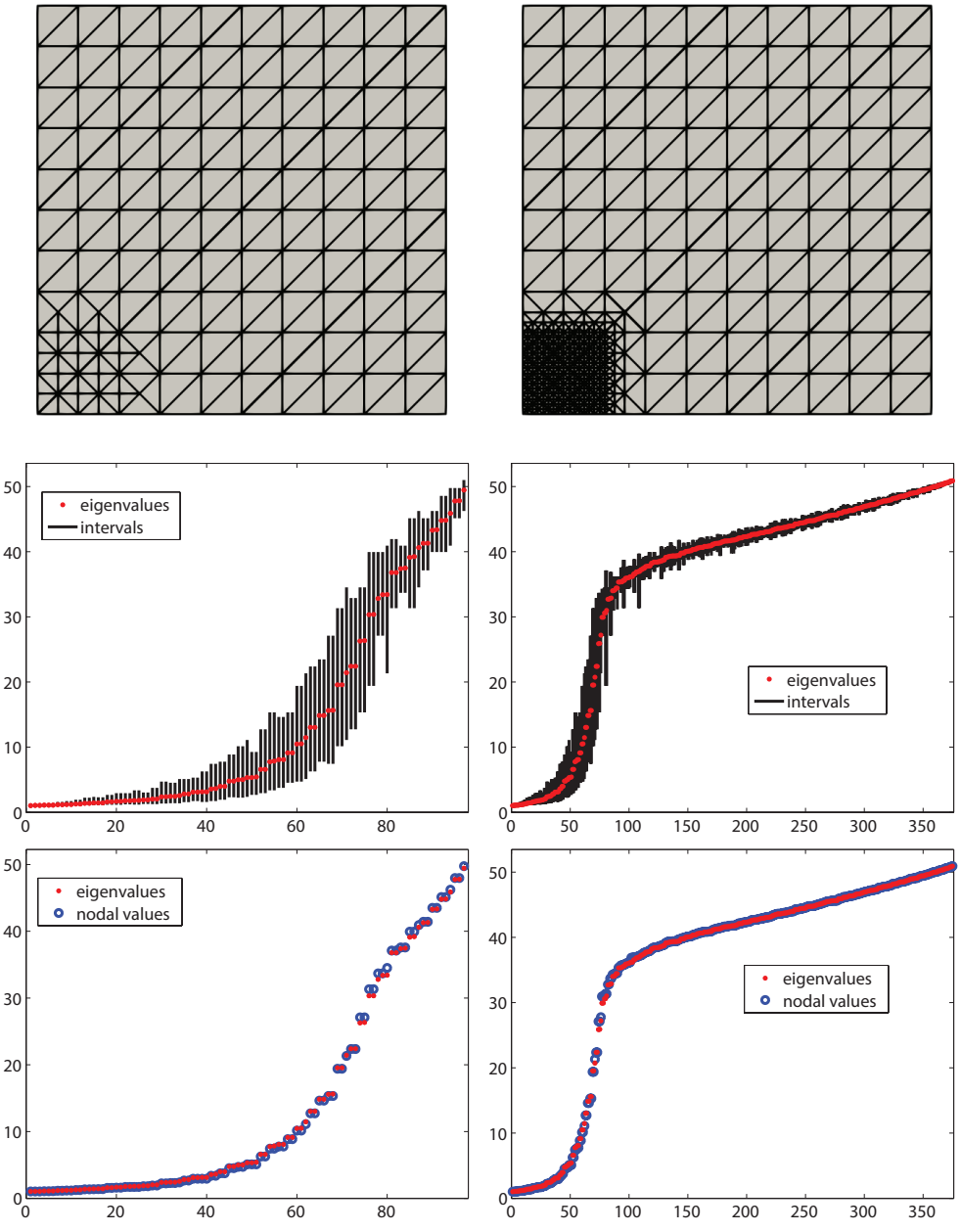

FIG. 11. *The influence of the locally refined mesh in the subdomain* $(0, 0.2) \times (0, 0.2)$ *for the test problem* (P2). *Left: One refinement step. Right: Three refinement steps. We use the same symbols as in Figures* 6 *and* 7.

paper [37] that if the Ritz values approximate (in a rather moderate way) the eigenvalues at the *lower end of the spectrum*, the computations further proceed with a rate as if the approximated eigenvalues are not present. Analysis of rounding errors in CG

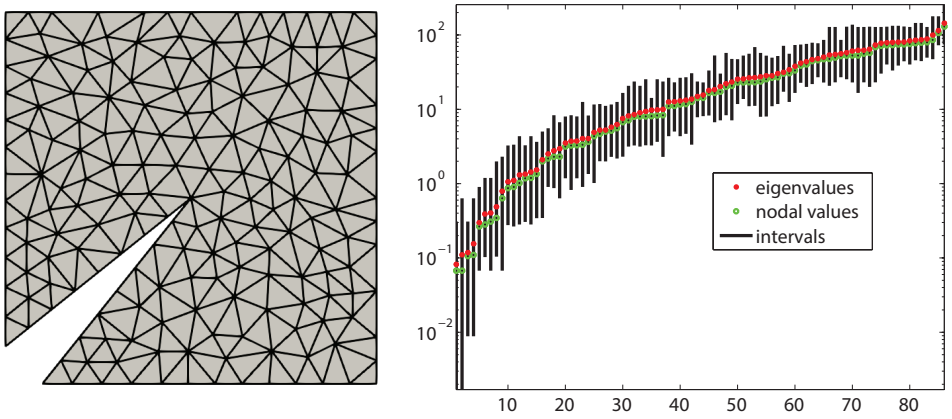

FIG. 12. *Comparison of the eigenvalues $\lambda_1 \le \cdots \le \lambda_N$ (red dots) with the increasingly sorted nodal values $k_1 \le \cdots \le k_N$ (green circles) and the associated intervals (black vertical lines) for the test problem* (P3) *on the re-entrant corner domain.*

and Lanczos by Paige, Greenbaum, and others, mentioned above in section 1, then proves that this argumentation concerning the lower end of the spectrum remains valid also in finite precision arithmetic computations. At the fifth iteration, the eigenvalues 1, 28.5, 61.4, 75.3 at the lower end of the spectrum and also the largest eigenvalue 161.45 are approximated by the Ritz values; see Figure 14. Therefore, from then on PCG converges, using the effective condition number upper bound

$$\frac{\|\mathbf{x}-\mathbf{x}_k\|_{\mathbf{A}}}{\|\mathbf{x}-\mathbf{x}_0\|_{\mathbf{A}}} \le 2\left(\frac{\sqrt{\kappa_e^{\mathbf{L}}}-1}{\sqrt{\kappa_e^{\mathbf{L}}}+1}\right)^{k-5}, \quad \kappa_e^{\mathbf{L}} = \frac{\lambda_{2039}^{\mathbf{L}}}{\lambda_{1926}^{\mathbf{L}}} = 1.02, \quad k > 5, \tag{4.3}$$

at least as fast as the right-hand side in (4.3) suggests. The convergence is in the iterations 6–9 very fast, and there are no further well separated eigenvalues that can be approximated within these iterations by the Ritz values. Therefore, we do not practically observe any further acceleration. At iteration 10, the convergence slows down. This is due to the effect of rounding errors that cause the forming of a second Ritz value that approximates the largest eigenvalue 161.45 (as mentioned above, the appearance of large outlying eigenvalues can cause deterioration of convergence due to roundoff; the detailed explanation is given, e.g., in [13], [24, section 5.9.1; see, in particular, Figures 5.14 and 5.15], and [10]).

Also, for the incomplete Cholesky preconditioning an analogous argumentation holds with the difference that the approximation of the five leftmost eigenvalues by the Ritz values slightly accelerate convergence. The bound (4.3) is valid with replacing $\kappa_e^{\mathbf{L}}$ by

$$\kappa_e^{\mathbf{C}} = \frac{\lambda_{3969}^{\mathbf{C}}}{\lambda_6^{\mathbf{C}}} = 3.75; \tag{4.4}$$

see the computed quantities in Table 2. We can see from Figure 14 that at the fifth iteration the five smallest eigenvalues are not yet approximated by the Ritz values. This needs about five additional iterations. From the tenth iteration the convergence remains very close to linear and slow because no further acceleration can take place

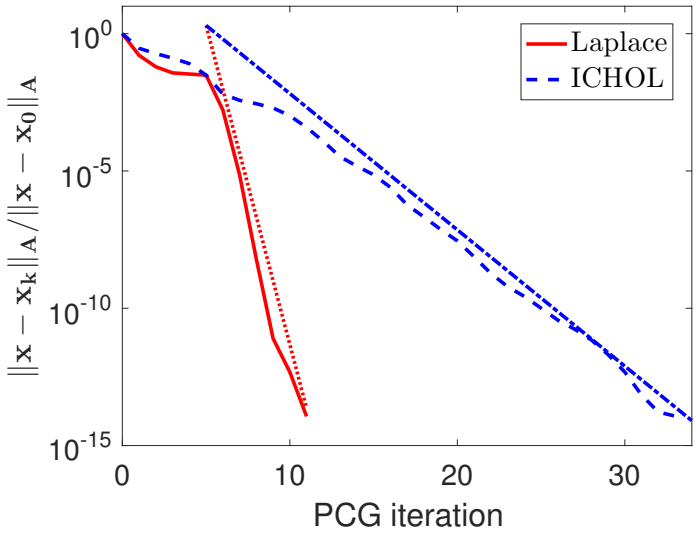

FIG. 13. *Explanation of the PCG behavior from Figure* 1. *The dotted and dash-dotted lines show the estimates of the PCG error based on the so-called effective condition number, which* here *(see the discussion in the text) fully describes the PCG behavior starting from the sixth iteration.*

TABLE 1

Detail of the points of increase (Ritz values) and the weights (see (2.13)*) of the distribution function* $\omega^{\mathbf{L}}(\lambda)$ *associated with the problem preconditioned by the Laplace operator. The effective condition number is for the given example determined by* $\lambda^{\mathbf{L}}_{1926}$ *and* $\lambda^{\mathbf{L}}_{2039}$*; see the top part of Figure* 15.

Index	1–1922	1923	1924	1925	1926
Eigenvalues	1	28.508	61.384	75.324	$\lambda^{\mathbf{L}}_{1926} = 79.699$
Total weight	9×10^{-6}	$\approx 10^{-3}$	$\approx 10^{-3}$	$\approx 10^{-3}$	$\approx 10^{-3}$
Index	1927–1930	1931–2039	2040–2047		2048–3969
Eigenvalues	80.875 – 81.222	$\lambda^{\mathbf{L}}_{2039} = 81.224$	81.226 – 133.94		161.45
Total weight	$\approx 10^{-3}$	1.8×10^{-2}	8×10^{-10}		0.96

due to the widespread eigenvalues and the effects of roundoff (no further eigenvalue approximation by the Ritz values can significantly affect the convergence behavior). The part of the spectra that practically determines the convergence rates after the fifth iteration of the Laplace operator PCG (respectively, after the tenth iteration of the ICHOL PCG) is illustrated in Figure 15.

Remark 4.1. Although we do not present in this section any formal statement with a formal proof, the presented explanation is much more than a discussion of the experimental results. Knowing the spectral information presented in Tables 1 and 2 and illustrated in Figure 15, the referenced theoretical results *prove* that starting from a *certain small iteration number greater than or equal to five*, the convergence behavior of CG will be very close to linear. We do know *a priori* the subsequent linear rates of convergence for both cases, which are determined by the effective condition numbers presented in (4.3) (respectively, (4.4)). A formal proof of the fact that it happens for the Laplace preconditioning case precisely at the fifth iteration (and not at the sixth or the seventh) cannot be done due to technically complicated terms (see the related results in the referenced literature). However, if we look at the Ritz values at the fifth iteration (see Figure 14), we can say with certainty that the acceleration

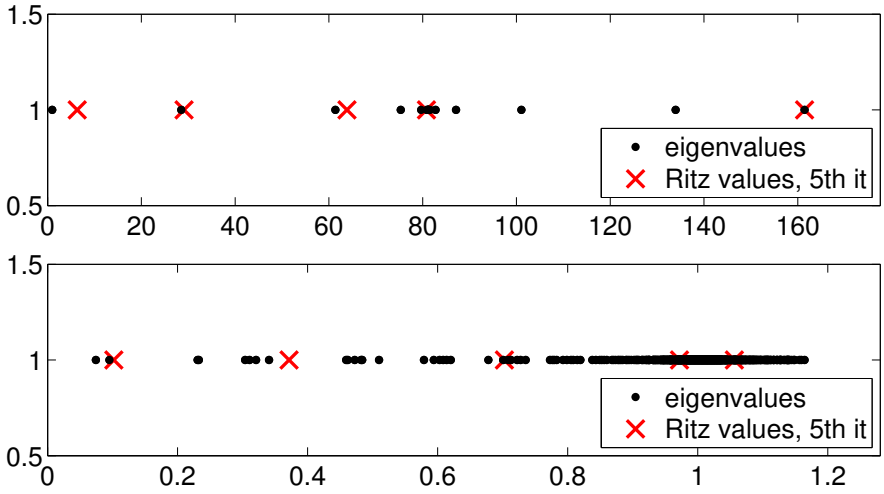

FIG. 14. *Illustration of the Ritz values computed at the fifth PCG iteration. Top: Problem with the Laplace operator preconditioning. We observe four Ritz values approximating the eigenvalues at the lower end of the spectrum and one Ritz value very closely approximating the largest eigenvalue. Bottom: Problem with the ICHOL preconditioning. We do not yet observe a good approximation of any of the eigenvalues, but we can see that the extremal Ritz values approach the ends of the spectral interval.*

TABLE 2

Detail of the points of increase (Ritz values) and the weights (see (2.13)*) of the distribution function* $\omega^{\mathbf{C}}(\lambda)$ *associated with the problem with ICHOL preconditioning. The effective condition number is for the given example determined by* $\lambda_6^{\mathbf{C}}$ *and* $\lambda_{3969}^{\mathbf{C}}$*; see the bottom part of Figure* 15.

Index	1	2	3	4
Eigenvalues	0.074	0.095	0.231	0.233
Total weight	8×10^{-5}	6.4×10^{-3}	8×10^{-7}	10^{-8}

Index	5	6	7–3969
Eigenvalues	0.304	$\lambda_6^{\mathbf{C}} = 0.311$	$0.321 - \lambda_{3969}^{\mathbf{C}} = 1.1643$
Total weight	6×10^{-5}	1.5×10^{-3}	0.992

must start at that point. Similarly, we can predict that the case with the incomplete Cholesky preconditioner will behave as it does.

The presented spectral information determines the behavior of the PCG iterations. Knowing the spectral information a priori, one can *quantitatively predict* the convergence for the Laplace and incomplete Cholesky preconditioners shown in Figure 13.

5. Concluding remarks. We have analyzed the operator $\mathcal{L}^{-1}\mathcal{A}$ generated by preconditioning second order elliptic PDEs with the inverse of the Laplacian. Previously, it has been proven that the range of the coefficient function k of the elliptic PDE is contained in the spectrum of $\mathcal{L}^{-1}\mathcal{A}$ but only for operators defined on infinitely dimensional spaces. In this paper, we show that a substantially stronger result holds in the discrete case of conforming finite elements. More precisely, we show that the eigenvalues of the matrix $\mathbf{L}^{-1}\mathbf{A}$, where $\mathbf{L}$ and $\mathbf{A}$ are the associated stiffness matrices, lie in resolution dependent intervals around the nodal values of the coefficient function that tend to the nodal values as the resolution increases. Moreover, there is a pairing (possibly nonunique) of the eigenvalues and the nodal values of the coefficient function due to Hall's theory of bipartite graphs. Finally, we demonstrate that the conjugate

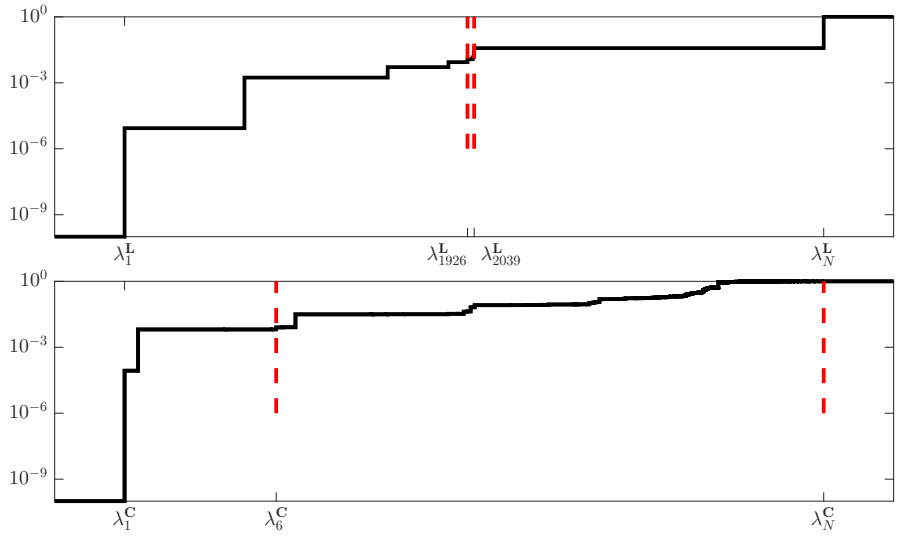

FIG. 15. *Distribution functions: Top: Laplace operator preconditioning. Bottom: ICHOL preconditioning. Red dashed lines represent the position of eigenvalues associated with the effective condition numbers after five iterations.*

gradient method utilizes the structure of the spectrum (more precisely, of the associated distribution function) to accelerate the iterations. In fact, even though the condition number involved, for instance, with incomplete Cholesky preconditioning is significantly smaller than for the Laplacian preconditioner, the performance measured by the number of iterations[10] when using Cholesky is much worse. In this case, the accelerated performance of the Laplacian preconditioner can be fully explained by an analysis of the distribution functions.

As mentioned above, in the numerical experiments of this paper we use Lagrange elements. However, the results (Theorem 3.1 and Corollary 3.2) (see, in particular, the derivation in (3.8) and (3.9)) are valid for any conforming approximation.

Acknowledgments. The authors are grateful to Marie Kubínová for pointing out to us Theorem 3.5, which greatly simplifies the proof of Theorem 3.1, and to Jan Papež for his early experiments and discussion concerning the motivating example. The authors are also grateful to the anonymous referees for their careful reading of the manuscript and their helpful comments.

REFERENCES

[1] O. AXELSSON, *Iterative Solution Methods*, Cambridge University Press, Cambridge, UK, 1994, https://doi.org/10.1017/cbo9780511624100.

[2] O. AXELSSON AND G. LINDSKOG, *On the eigenvalue distribution of a class of preconditioning methods*, Numer. Math., 48 (1986), pp. 479–498, https://doi.org/10.1007/bf01389447.

[3] O. AXELSSON AND G. LINDSKOG, *On the rate of convergence of the preconditioned conjugate gradient method*, Numer. Math., 48 (1986), pp. 499–523, https://doi.org/10.1007/bf01389448.

[4] J. A. BONDY AND U. S. R. MURTY, *Graph Theory with Applications*, Elsevier, New York, 1976.

[10] As mentioned above, the evaluation of the cost of practical computations must also take into account the cost of individual iterations, which can be different for different preconditioners, depending on the problem, its discretization, and the particular software implementation.

[5] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, SIAM, Philadelphia, 2002, https://doi.org/10.1137/1.9780898719208, reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 25001)].

[6] J. W. Daniel, *The conjugate gradient method for linear and nonlinear operator equations*, SIAM J. Numer. Anal., 4 (1967), pp. 10–26, https://doi.org/10.1137/0704002.

[7] A. L. Dulmage and N. S. Mendelsohn, *Coverings of bipartite graphs*, Canad. J. Math., 10 (1958), pp. 517–534, https://doi.org/10.4153/CJM-1958-052-0.

[8] M. Engeli, T. Ginsburg, H. Rutishauser, and E. Stiefel, *Refined Iterative Methods for Computation of the Solution and the Eigenvalues of Self-Adjoint Boundary Value Problems*, Mitt. Inst. Angew. Math. Zürich 8, Birkhäuser, Basel, 1959, https://doi.org/10.1007/978-3-0348-7224-9.

[9] V. Faber, T. A. Manteuffel, and S. V. Parter, *On the theory of equivalent operators and application to the numerical solution of uniformly elliptic partial differential equations*, Adv. in Appl. Math., 11 (1990), pp. 109–163, https://doi.org/10.1016/0196-8858(90)90007-L.

[10] T. Gergelits and Z. Strakoš, *Composite convergence bounds based on Chebyshev polynomials and finite precision conjugate gradient computations*, Numer. Algorithms, 65 (2014), pp. 759–782, https://doi.org/10.1007/s11075-013-9713-z.

[11] A. Greenbaum, *Behavior of slightly perturbed Lanczos and conjugate-gradient recurrences*, Linear Algebra Appl., 113 (1989), pp. 7–63, https://doi.org/10.1016/0024-3795(89)90285-1.

[12] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469, https://doi.org/10.1137/s0895479894275030.

[13] A. Greenbaum and Z. Strakoš, *Predicting the behavior of finite precision Lanczos and conjugate gradient computations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 121–137, https://doi.org/10.1137/0613011.

[14] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent Advances in Iterative Methods, IMA Vol. Math. Appl. 60, Springer, New York, 1994, pp. 95–118, https://doi.org/10.1007/978-1-4613-9353-5_7.

[15] A. Günnel, R. Herzog, and E. Sachs, *A note on preconditioners and scalar products in Krylov subspace methods for self-adjoint problems in Hilbert space*, Electron. Trans. Numer. Anal., 41 (2014), pp. 13–20.

[16] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer-Verlag, New York, 1994, https://doi.org/10.1007/978-3-319-28483-5.

[17] P. Hall, *On representatives of subsets*, J. Lond. Math. Soc., s1-10 (1935), pp. 26–30, https://doi.org/10.1112/jlms/s1-10.37.26.

[18] R. Herzog and E. Sachs, *Superlinear convergence of Krylov subspace methods for self-adjoint problems in Hilbert space*, SIAM J. Numer. Anal., 53 (2015), pp. 1304–1324, https://doi.org/10.1137/140973050.

[19] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–436, https://doi.org/10.6028/jres.049.044.

[20] R. Hiptmair, *Operator preconditioning*, Comput. Math. Appl., 52 (2006), pp. 699–706, https://doi.org/10.1016/j.camwa.2006.10.008.

[21] A. Jennings, *Influence of the eigenvalue spectrum on the convergence rate of the conjugate gradient method*, J. Inst. Math. Appl., 20 (1977), pp. 61–72, https://doi.org/10.1093/imamat/20.1.61.

[22] A. Jennings and G. M. Malik, *The solution of sparse linear equations by the conjugate gradient method*, Internat. J. Numer. Methods Engrg., 12 (1978), pp. 141–158, https://doi.org/10.1002/nme.1620120114.

[23] C. Lanczos, *Solution of systems of linear equations by minimized iterations*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 33–53, https://doi.org/10.6028/jres.049.006.

[24] J. Liesen and Z. Strakoš, *Krylov Subspace Methods: Principles and Analysis*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, UK, 2012, https://doi.org/10.1093/acprof:oso/9780199655410.001.0001.

[25] A. Logg, K.-A. Mardal, and G. N. Wells, eds., *Automated Solution of Differential Equations by the Finite Element Method*, Springer, Berlin, Heidelberg, 2012, https://doi.org/10.1007/978-3-642-23099-8.

[26] J. Málek and Z. Strakoš, *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs*, SIAM Spotlights 1, SIAM, Philadelphia, 2015, https://doi.org/10.1137/1.9781611973846.

[27] K. A. Mardal and R. Winther, *Preconditioning discretizations of systems of partial differ-*

ential equations, Numer. Linear Algebra Appl., 18 (2011), pp. 1–40, https://doi.org/10.1002/nla.716.

[28] G. Meurant, *The Lanczos and Conjugate Gradient Algorithms. From Theory to Finite Precision Computations*, Software Environ. Tools 19, SIAM, Philadelphia, 2006, https://doi.org/10.1137/1.9780898718140.

[29] G. Meurant and Z. Strakoš, *The Lanczos and conjugate gradient algorithms in finite precision arithmetic*, Acta Numer., 15 (2006), pp. 471–542, https://doi.org/10.1017/S096249290626001X.

[30] P. Morin, R. H. Nochetto, and K. G. Siebert, *Convergence of adaptive finite element methods*, SIAM Rev., 44 (2002), pp. 631–658, https://doi.org/10.1137/S0036144502409093.

[31] B. F. Nielsen and K.-A. Mardal, *Analysis of the minimal residual method applied to ill posed optimality systems*, SIAM J. Sci. Comput., 35 (2013), pp. A785–A814, https://doi.org/10.1137/120871547.

[32] B. F. Nielsen, A. Tveito, and W. Hackbusch, *Preconditioning by inverting the Laplacian: An analysis of the eigenvalues*, IMA J. Numer. Anal., 29 (2009), pp. 24–42, https://doi.org/10.1093/imanum/drm018.

[33] C. C. Paige, *The Computation of Eigenvalues and Eigenvectors of Very Large and Sparse Matrices*, Ph.D. thesis, London University, London, England, 1971.

[34] C. C. Paige, *Accuracy and effectiveness of the Lanczos algorithm for the symmetric eigenproblem*, Linear Algebra Appl., 34 (1980), pp. 235–258, https://doi.org/10.1016/0024-3795(80)90167-6.

[35] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Comput. Sci. Sci. Comput., Academic Press, Boston, 1990.

[36] Z. Strakoš, *On the real convergence rate of the conjugate gradient method*, Linear Algebra Appl., 154/156 (1991), pp. 535–549, https://doi.org/10.1016/0024-3795(91)90393-B.

[37] A. van der Sluis and H. A. van der Vorst, *The rate of convergence of conjugate gradients*, Numer. Math., 48 (1986), pp. 543–560, https://doi.org/10.1007/BF01389450.

[38] J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Landmarks Math., Princeton University Press, Princeton, NJ, 1996, https://doi.org/10.2307/j.ctt1wq8zhp, translated from the German print published in 1932 and with a preface by Robert T. Beyer, Twelfth printing, Princeton Paperbacks.

[39] Yu. V. Vorobyev, *Methods of Moments in Applied Mathematics*, translated from the Russian by Bernard Seckler, Gordon and Breach Science Publishers, New York, 1965, https://doi.org/10.2307/2004791.