Hannukainen, Antti; Malinen, Jarmo; Ojalammi, Antti

Efficient solution of symmetric eigenvalue problems from families of coupled systems

Published in:
SIAM Journal on Numerical Analysis

DOI:
10.1137/18M1202323

Published: 01/01/2019

Please cite the original version:
EFFICIENT SOLUTION OF SYMMETRIC EIGENVALUE PROBLEMS FROM FAMILIES OF COUPLED SYSTEMS

ANTTI HANNUKAINEN†, JARMO MALINEN†, AND ANTI OJALAMMI†

Abstract. Efficient solution of the lowest eigenmodes is studied for a family of related eigenvalue problems with common 2 × 2 block structure. It is assumed that the upper diagonal block varies between different versions while the lower diagonal block and the range of the coupling blocks remain unchanged. Such block structure naturally arises when studying the effect of a subsystem to the eigenmodes of the full system. The proposed method is based on interpolation of the resolvent function after some of its singularities have been removed by a spectral projection. Singular value decomposition can be used to further reduce the dimension of the computational problem. Error analysis of the method indicates exponential convergence with respect to the number of interpolation points. Theoretical results are illustrated by two numerical examples related to finite element discretization of the Laplace operator.

Key words. eigenvalue problem, subspace method, dimension reduction, acoustics

AMS subject classification. 65F15

DOI. 10.1137/18M1202323

1. Introduction. There is often a need to study the effect of a subsystem to the vibration modes of the whole system. For example, consider the modal computations of a vocal tract constrained into a magnetic resonance imaging (MRI) scanner [17, 26]. In this case, the system consists of the vocal tract air volume (i.e., the interior system) that changes during speech and the air volume of the MRI head coil (i.e., the exterior system) that stays unchanged; see Figure 1. For a high resolution description of speech production, it is desirable to compute the resonances for a very large number of vocal tract shapes. In order to speed up these computations, there is a strong incentive to precompute the effect of the unchanging exterior system and use it efficiently.

Modal analysis of systems consisting of interior and exterior parts leads to an algebraic eigenvalue problem

\[ A\mathbf{x} = \lambda M\mathbf{x} \quad \text{with the normalization} \quad \mathbf{x}^T M\mathbf{x} = 1 \]

that can accordingly be decomposed as

\[ \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \lambda \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}. \]

Here the matrix blocks \( A_{11} \) and \( A_{22} \) refer to interior and exterior systems, respectively, and the matrix blocks \( A_{12} \) and \( A_{21} = A_{12}^T \) are related to the coupling between the two systems. The same descriptions hold for the matrix \( M \). In the following, we assume that the matrices \( A \) and \( M \) are large, sparse, symmetric, and positive definite,
implying the same properties for $A_{ii}$ and $M_{ii}$ for $i = 1, 2$. This assumption is satisfied, e.g., when problem (1.1) is related to the finite element discretization of an elliptic PDE.

In this article, a novel method is proposed for efficiently solving a large number of different versions of problem (1.2) for the smallest eigenvalues $\lambda \in (0, \Lambda)$, $\Lambda > 0$, together with the corresponding eigenvectors. In applications, the number of eigenvalues in $(0, \Lambda)$ is typically much smaller than the dimension of the full problem. It is assumed that the matrices $A_{22}, M_{22}$ and subspaces $\text{range}(A_{21}), \text{range}(M_{21})$ remain unchanged while the matrix $A_{11}$ varies between different versions of the problem (1.2). In the proposed method, the matrix blocks related to the exterior system are replaced by ones with considerably smaller dimension. As shown in section 7, the time required to solve eigenvalues of interest for the acoustic system shown in Figure 1 is reduced from 25 to 5 seconds, not accounting for precomputation time. In a family of 1,000 different vocal tract samples this constitutes a saving of over 5 hours. After 40 eigensolves, the proposed method is faster even when the precomputational time is taken into account.

There exists a considerable amount of literature on the solution of large, sparse, symmetric, and positive definite eigenvalue problems; see [29]. The state-of-the-art solution method for this class of problems is the Lanczos iteration, which is a Rayleigh–Ritz method based on solving the eigenvalue problem in a Krylov subspace. When the interest lies in the smallest eigenvalues, the convergence of the iteration is sped up by using the shift-and-invert strategy, i.e., considering the eigenvalue problem related to the matrix $(A + \sigma M)^{-1}$ for some $\sigma \in \mathbb{R}$ instead.

As such, the Lanczos iteration is not well suited for including precomputations involving the exterior system. Computing the lowest eigenmodes using the shift-and-invert strategy requires the action of $(A + \sigma M)^{-1}$ in each iteration step. As several
linear systems need to be solved, the matrix \((A + \sigma M)\) is typically factorized, e.g., using the \(LDL^T\) factorization. Unfortunately, all factorizations have to be recomputed for different versions of \((1.2)\). In doing so, the block structure of the problem should be taken into account; see section 5 for an example in recycling information in computing block Cholesky factorizations. However, such a strategy does not easily allow for dimension reduction in (the exterior part of) the eigenvalue problem.

We propose a condensed pole interpolation (CPI) method that is based on the Rayleigh–Ritz procedure. In CPI, a subspace related to the exterior part of the problem is precomputed by a combination of a spectral projection, Chebyshev interpolation of the resolvent after removal of poles, and dimension reduction using singular value decomposition (SVD). This subspace is constructed only once, and it can be reused for different versions of \(A_{11}\). For each version of \((1.2)\), one solves a much smaller symmetric, positive definite eigenvalue problem using, e.g., the Lanczos iteration with the shift-and-invert strategy. Dimension reduction using SVD in the context of eigenvalue problems has been studied, e.g., in [19, 11, 16].

Our approach has some similarities with the component mode synthesis (CMS) introduced in the 1960s as a substructuring method for engineering simulations [18, 3]. An error estimate for the original CMS is given in [13, 15], an error indicator has been studied in [22, 7], and more efficient variants have been introduced, e.g., in [28, 23, 24, 4]. The CMS method has been further developed into an automated multilevel substructuring method having a much smaller precomputational cost without loss of accuracy; see [5, 21]. The rational filtering domain decomposition eigenvalue solver (RF-DDES) has recently been proposed in [20] for computing eigenvalues in a spectral interval of interest by using a Neumann series approximation of the resolvent function. The underlying philosophy of CPI is similar to RF-DDES.

The outline of the work is as follows. The required background is reviewed in section 2, and CPI is introduced together with its error analysis in section 3. Eigenvalue error estimates are given in section 4, and the optimal selection of the two parameter values, required by CPI, is discussed in section 5. Further dimension reduction is the subject of section 6. Finally, the theoretical treatment is illustrated in numerical examples in section 7.

2. Background. Let \(A, M \in \mathbb{R}^{n \times n}\) be symmetric, positive definite matrices. Let \((\lambda, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^n \setminus \{0\}, \lambda > 0\), be a solution of the full symmetric eigenvalue problem \(A\mathbf{x} = \lambda M\mathbf{x}\) such that \(\mathbf{x}^T M \mathbf{x} = 1\). For such \(\lambda\)’s, we write

\[
\lambda \in \sigma(A, M) := \{z \in \mathbb{C} : A - zM \text{ is not invertible}\}.
\]

From now on, denote \(\mathbf{x} = [\mathbf{x}_1 \mathbf{x}_2]^T\), where \(\mathbf{x}_1 \in \mathbb{R}^{n_1}\), \(\mathbf{x}_2 \in \mathbb{R}^{n_2}\) with \(n = n_1 + n_2\).

We call this the standard splitting of \(\mathbb{R}^n\), where \(\mathbb{R}^{n_1}\) and \(\mathbb{R}^{n_2}\) are called interior and exterior spaces, respectively. Using the standard splitting, the full eigenvalue problem has the structure

\[
\begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix}
\begin{bmatrix}
\mathbf{x}_1 \\
\mathbf{x}_2
\end{bmatrix}
= \lambda
\begin{bmatrix}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{bmatrix}
\begin{bmatrix}
\mathbf{x}_1 \\
\mathbf{x}_2
\end{bmatrix},
\]

where the blocks \(A_{11}, M_{11} \in \mathbb{R}^{n_1 \times n_1}\), \(A_{22}, M_{22} \in \mathbb{R}^{n_2 \times n_2}\), \(A_{12}, M_{12} = \mathbb{R}^{n_1 \times n_2}\), and \(A_{21}, M_{21} = \mathbb{R}^{n_2 \times n_1}\).

The topic of this work is the solution of different versions of the eigenvalue problem \((2.2)\), where the matrices \(A_{11}, M_{11}\) are free to vary but the matrices \(A_{22}, M_{22}\) and subspaces \(\text{range}(A_{21}), \text{range}(M_{21})\) stay the same. In this case, one can afford even
expensive precomputations for the unchanging components as a part of the eigenvalue solution method.

As an example, consider the acoustic system shown in Figure 1. In this case, the eigenvalue problem (2.2) arises from finite element discretization of the variational eigenvalue problem: Find \((\lambda', u) \in \mathbb{R} \times \mathcal{V}\) such that

\[(2.3) \quad (\nabla u, \nabla v) = \lambda'(u, v) \quad \text{for all} \quad v \in \mathcal{V},\]

where \((\cdot, \cdot)\) is the inner product of \(L^2(\Omega)\), and the subspace \(\mathcal{V} \subset H^1(\Omega)\) enforces the homogeneous Dirichlet boundary condition at least on a part of the boundary \(\partial \Omega\). Then the resulting \(A, M \in \mathbb{R}^{n \times n}\) are symmetric, positive definite stiffness and mass matrices, respectively.

The standard splitting in (2.3) arises from decomposition of the domain \(\Omega\) into nonoverlapping subdomains \(\Omega_1\) and \(\Omega_2\), corresponding to varying and unchanging parts of the system, respectively. The interior interface between the two subdomains is denoted by \(\Gamma = (\partial \Omega_1 \cup \partial \Omega_2) \setminus \partial \Omega\). The vectors \(x_1 \in \mathbb{R}^{n_1}\) and \(x_2 \in \mathbb{R}^{n_2}\) correspond to the degrees of freedom of the finite element space on \(\Omega\) corresponding to \(\Omega_1\) and \(\Omega_2\), respectively. In addition to \(n_1\) and \(n_2\), we define a third characterizing integer

\[(2.4) \quad n_\Gamma := \dim \left( \text{range}(\begin{bmatrix} M_21 & A_21 \end{bmatrix}) \right)\]

which gives the number of degrees of freedom over which the interior and the exterior systems interact on the interface \(\Gamma\). The FEM discretization of the full domain \(\Omega\) can be carried out in many ways, and the interface \(\Gamma\) need not be consistent with the FEM mesh. However, the three numbers \(n_1, n_2, n_\Gamma\) can always be extracted from the standard splitting.

### 2.1. Subspace methods.

Most solution methods for eigenvalue problems are of Rayleigh–Ritz type in which the eigenvalue problem is projected to a given subspace of \(\mathbb{R}^n\) [29]. For this purpose, let \(Q \in \mathbb{R}^{n \times k}\), \(k \leq n\), be a method matrix with linearly independent column vectors that is used for defining the method subspace \(\mathcal{V} := \text{range}(Q)\). Poor conditioning in numerical realizations is avoided by choosing the column vectors of \(Q\) orthonormal in an appropriate inner product; see section 6.

In the Rayleigh–Ritz procedure, the eigenvalue problem in \(\mathcal{V}\) is posed as follows: find \((\tilde{\lambda}, \tilde{x}) \in \mathbb{R} \times \mathbb{R}^k\) such that

\[(2.5) \quad Q^T A Q \tilde{x} = \tilde{\lambda} Q^T M Q \tilde{x}.\]

The set of approximate eigenvalues \(\tilde{\lambda}\) is denoted by \(\sigma_Q(A, M) := \sigma(Q^T A Q, Q^T M Q)\) as in (2.1). In fact, the set \(\sigma_Q(A, M)\) depends only on the method subspace \(V\).

**Lemma 2.1.** Let \(A, M \in \mathbb{R}^{n \times n}\) be symmetric and positive definite. In addition, let \(Q_1, Q_2 \in \mathbb{R}^{n \times k}\), \(k < n\) be such that \(\text{range}(Q_1) = \text{range}(Q_2)\). Then

\[(2.6) \quad \sigma_{Q_1}(A, M) = \sigma_{Q_2}(A, M).\]

Hence, we can write \(\sigma_Q(A, M) = \sigma_V(A, M)\), where \(V = \text{range}(Q)\). The aim is to find a low dimensional subspace \(V\) such that \(\sigma_V(A, M)\) is a reasonable approximation for a relevant part of \(\sigma(A, M)\). Those eigenvalues can be computed using, e.g., the shift-and-invert Lanczos iteration [14].
2.2. Estimate for the relative eigenvalue error. The relative error between corresponding eigenvalues in \( \sigma(A, M) \) and \( \sigma_V(A, M) \) is estimated by studying approximation of eigenvectors in the method subspace \( V \):

**Proposition 2.2.** Let \( A, M \in \mathbb{R}^{n \times n} \) be symmetric and positive definite matrices. Let \( (\lambda, x) \in \sigma(A, M) \times \mathbb{R}^n \setminus \{0\} \) be an eigenpair of (1.1) corresponding to a simple eigenvalue \( \lambda \) such that \( x^T M x = 1 \). In addition, assume that the spectral gap condition in [25, Thm. 2.7] is satisfied. Then for any \( V \subset \mathbb{R}^n \) there exists \( \tilde{\lambda} \in \sigma_V(A, M) \) and \( C(\lambda) := C(\lambda; A, M, V) \) such that

\[
\frac{|\lambda - \tilde{\lambda}|}{\lambda} \leq C(\lambda) \min_{v \in V} \|x - v\|_A^2, \quad \text{where} \quad \|x\|_A := \|A^{1/2}x\|_2.
\]

This proposition is a special case of a Hilbert space result in [25]; see also [2, 12, 6]. We normalize the eigenvectors \( x \) as \( x^T M x = 1 \) instead of using \( x^T A x = 1 \) from [25]. Because

\[
\min_{v \in V} \|x - v\|_A^2 = \lambda^{-1} \min_{v \in V} \left\| \frac{x}{\|M^{1/2}x\|_2} - v \right\|_A^2
\]

the different normalizations can be absorbed in \( C(\lambda) \). Otherwise the multiplier \( \lambda^{-1} \) will appear in (4.9). The eigenvector error can be similarly related to the angle between the exact eigenvector and the method subspace; see [25]. Except for the term \( \lambda^{-1} \) due to normalization, the dependency of \( C(\lambda; A, M, V) \) on its parameters is explained in [25, Thm. 2.7].

Henceforth, the method subspace is required to satisfy

\[
(2.8) \quad V = \left\{ \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \left| \begin{array}{c} v_1 \in \mathbb{R}^{n_1} \text{ and } v_2 \in V_2 \end{array} \right. \right\}, \quad \text{where} \quad V_2 \subset \mathbb{R}^{n_2}.
\]

Let \( x \) in Proposition 2.2 be decomposed as in (2.2), i.e., \( x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \). Choosing \( v \) in (2.7) so that \( v_1 = x_1 \) and using Proposition 2.2 leads to

\[
(2.9) \quad \frac{|\lambda - \tilde{\lambda}|}{\lambda} \leq C(\lambda) \min_{v_2 \in V_2} \|x_2 - v_2\|_{A_{22}}^2.
\]

We conclude that a subspace \( V_2 \) should accurately represent the \( x_2 \)-component of eigenvectors \( x \) for \( \lambda \in (0, \Lambda) \). In the proposed method, this approximation is guaranteed by constructing \( V_2 \) using a combination of spectral projection and Chebyshev interpolation of the resolvent.

**Remark 2.3.** In the case of multiple eigenvalues, the relative error in eigenvalue \( \lambda \) is related to the maximum over the corresponding eigenspace \( E_\lambda \):

\[
\frac{|\lambda - \tilde{\lambda}|}{\lambda} \leq C(\lambda) \max_{x \in E_\lambda} \min_{v \in V, x^T M x = 1} \|x - v\|_A^2.
\]

All upcoming results generalize to multiple eigenvalues by replacing \( x \) with \( \hat{x} \) such that

\[
\max_{x \in E_\lambda} \min_{v \in V, x^T M x = 1} \|x - v\|_A = \min_{v \in V} \|\hat{x} - v\|_A.
\]

For notational convenience, we assume in the following that all eigenvalues are simple.
2.3. Method matrix in component mode synthesis. In CMS, the domain Ω is decomposed into several subdomains, and the matrix A is partitioned according to the degrees of freedom corresponding to the subdomains and the subdomain interfaces. After partitioning, the matrix A is block diagonalized using an appropriate elimination matrix. In the last step, the block corresponding to the subdomain degrees of freedom is truncated by using a select number of eigenvectors of each local eigenvalue problem.

It is straightforward to adapt CMS to deal with the standard splitting in (2.2) and to perform the dimension reduction only on the exterior domain. The method matrix Q is constructed as a product of an elimination matrix

\[ G := \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} \]

that block diagonalizes the matrix A and a matrix containing eigenvectors related to the \( K \ll n_2 \) smallest eigenvalues of the subproblem \( A_{22}x_2 = \lambda M_{22}x_2 \). The resulting method matrix is

\[
Q := \begin{bmatrix} I & 0 \\ -A_{22}^{-1}A_{21} & I \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & [v_1, \ldots, v_K] \end{bmatrix}.
\]

When using Q in (2.10), numerical experiments in section 7 indicate similar performance as reported in [13, 15] and larger computational effort compared to CPI.

3. Condensed pole interpolation method. Assume that A and M are now represented through the standard splitting as in (2.2). The topic of this section is the construction of the subspace \( V_2 \) in (2.8). Consider the eigenvalue \( \lambda > 0 \) as fixed, and define an additional bound \( \tilde{\Lambda} \) satisfying

\[
0 < \lambda < \Lambda < \tilde{\Lambda} = \gamma \Lambda \quad \text{for} \quad \gamma > 1,
\]

where \((0, \Lambda)\) is the spectral interval of interest.

3.1. Eigenvector basis for the exterior subspace. Let \((\mu_k, v_k) \in \mathbb{R} \times \mathbb{R}^{n_2} \setminus \{0\}\) be solutions of the symmetric exterior eigenvalue problem such that

\[
A_{22}v_k = \mu_k M_{22}v_k \quad \text{and} \quad v_j^T M_{22}v_k = \delta_{j,k} \quad \text{such that} \quad 1 \leq j, k \leq n_2.
\]

(Note that \(M_{22}\) is positive definite since \(M\) is.) For \(\tilde{\Lambda} > 0\), let \(P_{\tilde{\Lambda}} \in \mathbb{R}^{n_2 \times n_2}\) be the \(M_{22}\)-orthogonal projection matrix

\[
P_{\tilde{\Lambda}} := \sum_{\{k : \mu_k \in (0, \tilde{\Lambda})\}} v_k v_k^T M_{22} \quad \text{satisfying} \quad M_{22}P_{\tilde{\Lambda}} = P_{\tilde{\Lambda}}^T M_{22}.
\]

We further restrict \(V_2\) in (2.8) to subspaces of the type

\[
V_2 = \text{range}(P_{\tilde{\Lambda}}) \oplus W_2,
\]

where the complementing subspace \(W_2 \subset \mathbb{R}^{n_2}, W_2 \perp \text{range}(P_{\tilde{\Lambda}})\) in the \(A_{22}\)-inner product will be chosen so that the eigenvalue error given by (2.9) can be conveniently bounded from above.

3.2. Error estimate based on projection and interpolation. For \(\lambda \in \sigma(A, M) \setminus \sigma(A_{22}, M_{22})\) and the corresponding eigenvector \(x = [x_1, x_2]^T\), (2.2) gives

\[
x_2 = (A_{22} - \lambda M_{22})^{-1} Z(\lambda) x_1, \quad \text{where} \quad Z(\lambda) := \lambda M_{21} - A_{21}.
\]
EIGENPROBLEMS FROM FAMILIES OF COUPLED SYSTEMS

Clearly,
\[
\min_{v_2 \in V_2} \| x_2 - v_2 \|_{A_22} = \min_{v_2 \in V_2} \| P_{\tilde{\Lambda}} x_2 + (I - P_{\tilde{\Lambda}}) x_2 - v_2 \|_{A_22} = \min_{v_2 \in V_2} \| (I - P_{\tilde{\Lambda}})(A_{22} - \lambda M_{22})^{-1} Z(\lambda) x_1 - v_2 \|_{A_22},
\]
(3.5)
since \( P_{\tilde{\Lambda}}(A_{22} - \lambda M_{22})^{-1} Z(\lambda) x_1 \in \text{range}(P_{\tilde{\Lambda}}) \subset V_2 \) by (3.4). Because also \( W_2 \subset V_2 \), it follows that
\[
\min_{v_2 \in V_2} \| x_2 - v_2 \|_{A_22} \leq \min_{w_2 \in W_2} \| (I - P_{\tilde{\Lambda}})(A_{22} - \lambda M_{22})^{-1} Z(\lambda) x_1 - w_2 \|_{A_22}.
\]
(3.6)

To introduce the CPI method, we proceed to construct the complementing subspace \( W_2 \) for (3.6) depending on \( \tilde{\Lambda} \), some distinct interpolation points \( \{ \xi_i \}_{i=1}^N \subset (0, \Lambda) \), and subspaces \( \text{range}(M_{21}) \) and \( \text{range}(A_{21}) \) related to the standard splitting of the original matrices \( A \text{ and } M \), i.e.,
\[
W_2 = W_2 \left( \tilde{\Lambda}, \{ \xi_i \}_{i=1}^N, \text{range} \left( \begin{bmatrix} M_{21} & A_{21} \end{bmatrix} \right) \right).
\]
Let
\[
f_{\tilde{\Lambda}}(\xi) := (I - P_{\tilde{\Lambda}})(A_{22} - \xi M_{22})^{-1} \in \mathbb{R}^{n_2 \times n_2} \quad \text{for } \xi \in (0, \Lambda) \setminus \sigma(A_{22}, M_{22}).
\]
Then \( A_{22} v_k - \xi M_{22} v_k = (\mu_k - \xi) M_{22} v_k \), where \( (\mu_k, v_k) \), \( k = 1, \ldots, n_2 \), are given by (3.2). This implies
\[
(A_{22} - \xi M_{22})^{-1} M_{22} v_k = (\mu_k - \xi)^{-1} v_k.
\]
(3.7)
Hence,
\[
f_{\tilde{\Lambda}}(\xi) M_{22} v_k = (I - P_{\tilde{\Lambda}})(A_{22} - \xi M_{22})^{-1} M_{22} v_k = (\mu_k - \xi)^{-1}(I - P_{\tilde{\Lambda}}) v_k = (\mu_k - \xi)^{-1} v_k - \sum_{\{j : \mu_j \in (0, \tilde{\Lambda})\}} v_j v_j^T M_{22} v_k \]
(3.8)
= \begin{cases} (\mu_k - \xi)^{-1} v_k & \text{for } k \text{ satisfying } \mu_k > \tilde{\Lambda}, \\ 0 & \text{otherwise.} \end{cases}
\]
Because \( \mathbb{R}^{n_2} = \text{span}_{k=1,\ldots,n_2} \{ M_{22} v_k \} \), we conclude that \( f_{\tilde{\Lambda}} \) is, in fact, an analytic function on the whole interval \( (0, \tilde{\Lambda}) \) that contains the original domain \( (0, \Lambda) \setminus \sigma(A_{22}, M_{22}) \). Hence, the assumption \( \lambda \notin \sigma(A_{22}, M_{22}) \) can be removed. As \( f_{\tilde{\Lambda}} \) is analytic, it can be approximated in various ways such as series expansions or interpolation.

In the CPI method, the complementing subspace \( W_2 \) is chosen so that the right-hand side of (3.6) can be bounded using interpolation error estimates. Let \( \{ \xi_i \}_{i=1}^N \subset (0, \Lambda) \) be a set of distinct interpolation points and
\[
W_2 := \text{span}_{i=1,\ldots,N} \{ f_{\tilde{\Lambda}}(\xi_i) [M_{21} \ A_{21}] \}.
\]
(3.9)
In addition, let
\[
K(l) := \# \{ \mu_k \mid \mu_k < l \}.
\]
(3.10)
The dimensions of the spaces \( W_2 \) and \( V_2 \) depend on the number of interpolation points \( N \), the number \( K(\tilde{\Lambda}) \) of exterior eigenvalues \( \mu_k \) smaller than \( \tilde{\Lambda} \), and \( n_\Gamma \) defined in (2.4). Then \( \dim(W_2) \leq N n_\Gamma \) and \( \dim(V_2) \leq N n_\Gamma + K(\tilde{\Lambda}) \).
3.3. Outline of the method. The CPI method introduced above is based on solving the original eigenvalue problem by restricting the exterior system to the space $V_2 = \text{range}(P_\Lambda) \oplus W_2$, where

$$W_2 = \text{span}_{i=1,\ldots,N} \{ f_{\Lambda}(\xi_i) [M_{21} \ A_{21}] \}, \quad f_{\Lambda}(\xi) = (I - P_{\Lambda})(A_{22} - \xi M_{22})^{-1},$$

and $\xi_i \in (0, \Lambda)$, $i = 1, 2, \ldots, N$, are the interpolation points chosen as in (4.5). Practical realization of CPI requires a method matrix $Q$ which, by (2.8), has the structure

$$Q = \begin{bmatrix} I & 0 \\ 0 & Q_{22} \end{bmatrix},$$

where $I \in \mathbb{R}^{n_1 \times n_1}$ is the identity matrix and $V_2 = \text{range}(Q_{22})$. The column vectors of $Q_{22}$ form a basis of the space $V_2$ and are constructed with the aid of the sample vectors $q_{ij} \in \mathbb{R}^{n_2}$. Let $\{p_1, \ldots, p_r\} \subset \mathbb{R}^{n_2}$ be a set of (possibly linearly dependent) vectors such that

$$\text{span}\{p_1, \ldots, p_r\} = \text{range}([M_{21} \ A_{21}]).$$

The sample vectors are computed by solving the linear systems

$$(A_{22} - \xi_i M_{22})q_{ij} = p_j,$$

and the complementing subspace $W_2$ is given by

$$W_2 = \text{span}\{(I - P_{\Lambda})q_{ij} \mid i = 1, \ldots, N \text{ and } j = 1, \ldots, r\}.$$

Practical realization of the CPI method consists of the following steps:

1. Compute the smallest eigenpairs $(\mu_k, v_k)$ of the exterior system $A_{22}v_k = \mu_k M_{22}v_k$ satisfying $\mu_k \leq \Lambda$.
2. Compute sample vectors $q_{11}, \ldots, q_{Nr}$ as solutions of $(A_{22} - \xi_i M_{22})q_{ij} = p_j$ for $i = 1, \ldots, N$ and $j = 1, \ldots, r$.
3. Collect the eigenvectors from step 1 and the sample vectors from step 2 into matrix $B = [v_1, \ldots, v_K, (I - P_{\Lambda})q_{11}, \ldots, (I - P_{\Lambda})q_{Nr}]$. Use SVD to compute an orthonormal basis for $V_2$ from $B$. Use the basis vectors as columns of $\tilde{Q}$.
4. Solve the eigenvalue problem $\tilde{Q}^T A \tilde{Q} \tilde{x} = \tilde{\lambda} \tilde{Q}^T M \tilde{Q} \tilde{x}$ using, e.g., the Lanczos iteration.

Step 3 will be modified to include an additional dimension reduction of $V_2$ in section 6, which leads to a considerably smaller eigenvalue problem while maintaining the desired accuracy.

4. Bound for the relative eigenvalue error. We proceed to give an upper bound for the relative error.

**Lemma 4.1.** Let $V_2$ be as defined in (2.8), (3.3), and (3.9). Denote the Lagrange interpolating polynomials by

$$\ell_i(\lambda) = \prod_{\substack{1 \leq j \leq N \\backslash j \neq i}} \frac{\lambda - \xi_j}{\xi_i - \xi_j} \quad \text{for } i = 1, \ldots, N,$$
where \( \{ \xi_i \}_{i=1}^N \subset (0, \Lambda) \) are the interpolation points used to define \( W_2 \). Then for any \( \lambda \in \sigma(A, M) \cap (0, \Lambda) \) there exists \( \tilde{\lambda} \in \sigma_V(A, M) \) such that

\[
\frac{|\lambda - \tilde{\lambda}|}{\lambda} \leq C(\lambda) \sum_{k, \mu_k > \tilde{\lambda}} \mu_k c_k(\lambda)^2 \beta_k(\lambda)^2 \quad \text{for all} \quad \tilde{\lambda} > \Lambda,
\]

where the coefficients \( \beta_k(\lambda) \) and \( c_k(\xi) \) are defined by

\[
Z(\lambda)x_1 = \sum_{k=1}^{n_2} \beta_k(\lambda) M_{22} v_k \quad \text{and} \quad c_k(\xi) := \left( \frac{1}{\mu_k - \xi} - \sum_{j=1}^N \frac{\ell_j(\xi)}{\mu_k - \xi_j} \right).
\]

Observe that the coefficients \( c_k(\xi) \) are the error functions in Lagrange interpolation at points \( \{ \xi_i \}_{i=1}^N \) of the rational function \((\mu_k - \lambda)^{-1}\), and they are analytic functions for all \( \xi \notin \sigma(A_{22}, M_{22}) \). Note that if \( \{ k : \mu_k > \tilde{\lambda} \} = \emptyset \), then (4.3) gives \( \lambda = \tilde{\lambda} \).

Proof. To obtain an upper bound for relative eigenvalue error in (2.9), we choose

\[
w_2 = \sum_{i=1}^N \ell_i(\lambda) f_{\Lambda}(\xi_i) Z(\lambda)x_1 \in W_2
\]

in (3.6), giving

\[
\frac{|\lambda - \tilde{\lambda}|}{\lambda} \leq C(\lambda) \left\| \left( f_{\Lambda}(\lambda) - \sum_{i=1}^N \ell_i(\lambda) f_{\Lambda}(\xi_i) \right) Z(\lambda)x_1 \right\|^2_{A_{22}}.
\]

The term \( Z(\lambda)x_1 \in \mathbb{R}^{n_2} \) has the expansion

\[
Z(\lambda)x_1 = \sum_{k=1}^{n_2} \beta_k(\lambda) M_{22} v_k,
\]

where \( v_k \) are given by (3.2). Using (3.8) gives

\[
\left( f_{\Lambda}(\lambda) - \sum_{i=1}^N \ell_i(\lambda) f_{\Lambda}(\xi_i) \right) Z(\lambda)x_1 = \sum_{k, \mu_k > \tilde{\lambda}} c_k(\lambda) \beta_k(\lambda) v_k,
\]

where \( \tilde{\lambda} = \gamma \Lambda \) for \( \gamma > 1 \). By using the exterior eigenvector basis equation (3.2), we get

\[
\left\| \sum_{k, \mu_k > \tilde{\lambda}} c_k(\lambda) \beta_k(\lambda) v_k \right\|^2_{A_{22}} = \sum_{k, \mu_k > \tilde{\lambda}} \mu_k c_k(\lambda)^2 \beta_k(\lambda)^2,
\]

which completes the proof.

To estimate the relative error from (4.3), it only remains to bound \( c_k(\lambda) \) and \( \beta_k(\lambda) \) from above. In order to obtain a good upper bound for the functions \( c_k(\lambda) \), it
is beneficial to choose the interpolation points as zeroes of the Chebyshev polynomials on the interval \((0, \Lambda)\),

\[
\xi_i = \frac{\Lambda}{2} \left[ 1 + \cos \left( \frac{2i - 1}{2N} \pi \right) \right], \quad i = 1, \ldots, N.
\]

Then the functions \(c_k(\lambda)\) defined in (4.1) can be uniformly bounded on \((0, \Lambda)\) by the standard Lagrange error estimates

\[
\sup_{\lambda \in (0, \Lambda)} |c_k(\lambda)| \leq \frac{\Lambda^N}{2^{2N-1} N!} \sup_{\xi \in (0, \Lambda)} \left| \frac{d^N}{d\xi^N} (\mu_k - \xi)^{-1} \right| \leq \frac{\Lambda^N}{2^{2N-1} N!} \sup_{\xi \in (0, \Lambda)} (\mu_k - \xi)^{-N-1} = \frac{\Lambda^N}{2^{2N-1} (\mu_k - \Lambda)^{N+1}}.
\]

To bound the coefficients \(\beta_k(\lambda)\), we need a technical lemma.

**Lemma 4.2.** Let \(M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}\) be a symmetric, positive definite matrix. Then

\[
C_M \|x\|_{M_{22}}^2 \geq \|x\|_{M_{22}}^2 \quad \text{for all} \quad x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},
\]

where

\[
C_M^{-1} := \min \left\{ \eta > 0 : \eta \in \sigma(I - M_{22}^{-1} M_{12}^T M_{11}^{-1} M_{12} M_{22}^{-1/2}) \right\}.
\]

**Proof.** Since \(M\) is positive definite, so are \(M_{11}\) and \(M_{22}\). Defining \(M_k := \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}\) for \(k \geq 0\) we observe that \(x^T M x \geq k \|x\|_{M_{22}}^2\) for all \(x\) if and only if \(M_k \geq 0\). Since \(M_0 = M\), the set \(\{k > 0 : M_k \geq 0\}\) is nonempty by the continuity of the eigenvalues of the matrix elements and the fact that the set of invertible matrices is open. Hence, we can define \(C_M := \max \{k > 0 : M_k \geq 0\}\). Similarly, we may reason that the matrix \(M_k\) for \(k = C_M\) is not invertible but it satisfies \(M_k \geq 0\).

For any \(\eta \in \mathbb{R}\) the matrix \(M_\eta\) is not invertible if and only if \(M_\eta x = 0\) for some \(x \neq 0\) if and only if

\[
\eta \in 1 - \sigma(M_{21} M_{11}^{-1} M_{12} M_{22}^{-1}) \subset 1 - \sigma(M_{21} M_{11}^{-1} M_{12} M_{22}^{-1}) \cup \{0\} = \sigma(I - M_{22}^{-1/2} M_{12}^T M_{11}^{-1} M_{12} M_{22}^{-1/2}) \cup \{1\} \subset (-\infty, 1]
\]

since \(M_{22}^{-1/2} M_{12}^T M_{11}^{-1} M_{12} M_{22}^{-1/2} \geq 0\). We used here the fact that \(\sigma(AB) \cup \{0\} = \sigma(BA) \cup \{0\}\) for all square matrices \(A\) and \(B\). Defining now \(C_M\) by (4.7), we observe that \(C_M^{-1} \leq C_M\), and the proof is thus complete. \(\square\)

**Remark 4.3.** Note that if \(M_{12} = 0\), then \(C_M = 1\). We leave it to the reader to verify that the estimate in Lemma 4.2 is, in fact, sharp. This can be seen by checking that an equality \(C_M^{-1} = \hat{C}_M\) holds in the proof.

We have now completely specified the CPI method together with its error estimate, and we are in position to state our first main result.

**Theorem 4.4.** Let \(A\) and \(M\) be as defined in (1.1), \(\{\xi_i\}_{i=1}^N\) be the Chebyshev interpolation points of the interval \((0, \Lambda)\) given in (4.5), and \(V_2 = \text{range}(P_\Lambda) \oplus W_2\), where \(W_2\) and \(P_\Lambda\) are as defined in (3.3) and (3.9), respectively. Finally, let \(V\) be as defined in (2.8).
Finally, we use (4.1) to estimate the coefficients $c_k$ using Lemma 4.1. The estimate for $\mu$ in (4.8) and (4.9), we have

$$C_M(x^T M x) \geq \sum_{\mu_k > \tilde{\lambda}} \left( \frac{\beta_k(\lambda)}{\mu_k - \lambda} \right)^2.$$  

Combining these with (4.6) gives the estimate

$$\sum_{\mu_k > \tilde{\lambda}} \mu_k c_k(\lambda)^2 \beta_k(\lambda)^2 \leq \sum_{\mu_k > \tilde{\lambda}} \mu_k \left( \frac{\beta_k(\lambda)}{\mu_k - \lambda} \right)^2 \frac{\Lambda^{2N}(\mu_k - \lambda)^2}{4^{2N-1}(\mu_k - \lambda)^2N+2} \leq 4C_M \left( \frac{\Lambda}{4} \right)^{2N} \frac{\tilde{\lambda}(\tilde{\lambda} - \lambda)^2}{(\lambda - \lambda)^2N+2} \leq 4C_M \Lambda^{2N}(\lambda - \lambda)^2 \cdot \frac{L}{4(\lambda - \lambda)^2} \cdot \frac{(\lambda - \lambda)^2}{(\gamma - 1)^2} \cdot \frac{1}{4(\gamma - 1)} \cdot \frac{1}{4(\gamma - 1)} \cdot \frac{1}{4(\gamma - 1)} \cdot \frac{1}{4(\gamma - 1)} \cdot \frac{1}{4(\gamma - 1)} \cdot \frac{1}{4(\gamma - 1)}.$$ 

We used here the fact that the function

$$x \mapsto \frac{x(x - \lambda)^2}{(x - \lambda)^2N+2}$$

is decreasing for $x > \lambda$, and hence its maximum over $[\tilde{\lambda}, \infty)$ is attained at $x = \tilde{\lambda}$. Finally, we use $\lambda > 0$ to obtain the final estimate,

$$\frac{|\lambda - \tilde{\lambda}|}{\lambda} \leq 4C_M(\lambda)^2 \gamma^{3} \left( \frac{1}{4(\gamma - 1)^2} \right)^{2N} = C_M(\lambda)^2 \gamma^{3} \left( \frac{1}{4(\gamma - 1)^2} \right)^{2N+2}. $$

This completes the proof.
If we fix $\gamma > 5/4$ in Theorem 4.4, the right-hand side of (4.8) converges to zero as $N \to \infty$. Observe that the set $\{k : \mu_k > \gamma \Lambda\} = \emptyset$ for $\gamma$ large enough for any $\Lambda > 0$. Then the sum in inequality (4.9) vanishes, and again $\lambda = \lambda$ follows.

5. **Computational cost.** The error estimate given in Theorem 4.4 allows one to choose the values for $N$ and $\gamma$ in an optimal way, depending on the target error level and a model for the computational cost required to solve the eigenvalue problem. Solving the smallest elements $\lambda \in \sigma_V(A, M)$ using the method matrix $Q$ (as given in (3.11)) amounts to solving $\hat{A}\hat{x} = \lambda\hat{M}\hat{x}$ in which

$$
\hat{A} = \begin{bmatrix}
A_{11} & A_{12}
Q_{12}^T A_{12} & Q_{12}^T A_{22}
Q_{12}^T & Q_{12}^T A_{22}
\end{bmatrix}
$$

and

$$
\hat{M} = \begin{bmatrix}
M_{11} & M_{12}
M_{22} & M_{22}
Q_{12}^T M_{12} & Q_{12}^T M_{22}
\end{bmatrix}.
$$

By section 3.3, $Q_{22} \in \mathbb{R}^{n_2 \times \dim(V_2)}$, and hence

$$
Q_{22}^T A_{22} Q_{22} \in \mathbb{R}^{\dim(V_2) \times \dim(V_2)}, \quad \text{where } \dim(V_2) \leq K(\gamma \Lambda) + n_1 N.
$$

Denote the Cholesky factorizations of the matrices $\hat{A}$, $\hat{M}$ by $\hat{A} = R^T R$ and $\hat{M} = L^T L$, respectively. Eigenvalues $\hat{\lambda} \in \sigma_V(A, M) \cap (0, \Lambda)$ can be solved, e.g., by applying the Lanczos iteration to

$$
L \hat{A}^{-1} L^T y = \hat{\lambda}^{-1} y, \quad \text{where } \hat{M} = L^T L \quad \text{and} \quad y = L x.
$$

This requires repeated multiplications by $L \hat{A}^{-1} L^T$ which can be efficiently carried out using Cholesky factorizations. The factorizations should be computed by taking advantage of the block structure, for example, by writing $\hat{A} = R^T R$ so that

$$(5.1) \quad R = \begin{bmatrix}
R_{11} & R_{11}^T \hat{A}_{12} \\
0 & R_{22}
\end{bmatrix}, \hat{A}_{11} = R_{11}^T R_{11}, \text{ and } \hat{A}_{22} - \hat{A}_{12}^T \hat{A}_{11}^{-1} \hat{A}_{12} = R_{22}^T R_{22},$$

where the Schur complement of $\hat{A}$ with respect to $A_{11}$ has also been Cholesky factorized. A similar formula can be used for $\hat{M}$ to produce $L$ in block form. The matrix $R_{22}$ can be computed as a low-rank update to the factorization of $A_{22}$. This leads to

$$
L R^{-1} R^{-T} L^T y = \hat{\lambda}^{-1} y.
$$

The cost of the matrix-vector multiplication by $L R^{-1} R^{-T} L^T$ is of lower order compared to computing the factorizations which we discuss next. The block structure can be used in the spirit of (5.1) to recycle the factorization of $A_{22}$ in the shift-and-invert Lanczos iteration.

For each version of problem (1.1), one has to recompute the Cholesky factorizations in (5.1). The cost of factorizing $A_{11}$ does not depend on the choice of $V_2$. Hence, we only model the cost of computing the Cholesky factorization for the Schur complement $\hat{A}_{22} - \hat{A}_{12}^T \hat{A}_{11}^{-1} \hat{A}_{12}$. Depending on the underlying problem, the Schur complement can be sparse or dense. Thus, the cost of computing the factorization is modeled as proportional to the $r$th power of $\dim(V_2)$ as

$$(5.2) \quad \text{cost}(\gamma, N) := (K(\gamma \Lambda) + n_1 N)^r,$$

where $K(\ell)$ is as defined in (3.10) and the parameter $r \in [0, 3]$ depends on the sparsity of the Schur complement.
5.1. Optimization of $N$ and $\gamma$. A typical application for CPI is the solution of eigenvalues for the Dirichlet Laplacian in $\Omega \subset \mathbb{R}^d$ using the finite element method. In this case, an asymptotically accurate description for $K(l)$ is given by the Weyl law [30, Chap. 8] as $K(l) \approx C(d) \text{vol}(\Omega_2)l^{d/2}$, where $C(d) = (2\pi)^{-d} \text{vol} B_d$ and $B_d$ is a $d$-dimensional unit ball. Motivated by Theorem 4.4, we define a normalized error level

$$\eta = \frac{|\lambda - \hat{\lambda}|}{\bar{\Lambda}} \leq 64C_M \Lambda C(\lambda) \eta$$

by Theorem 4.4. An optimal value combination for parameters $N$ and $\gamma$ for a normalized error level $\eta$ is obtained by minimizing the cost function (5.2) under the constraint $ntol(\gamma, N) = \eta$.

**Theorem 5.1.** Let $\eta > 0$, and let $ntol(\gamma, N)$ be as given in (5.3). Define the computational cost model function as

$$g(\gamma, N) := \left( C(d) \text{vol}(\Omega_2)(\gamma \Lambda)^{d/2} + n_{\Gamma} N \right)^{r},$$

where $C(d) = (2\pi)^{-d} \text{vol} B_d$, and $B_d$ is a $d$-dimensional unit ball. Then the computational cost model function is minimized under the constraint $ntol(N, \gamma) = \eta$ by choosing $N(\gamma)$ and $\gamma$ such that

$$N(\gamma) = \frac{dC(d)\Lambda^{d/2} \text{vol}(\Omega_2)}{2n_{\Gamma}} \gamma^{d/2 - 1}(\gamma - 1) \ln 4(\gamma - 1) - 2$$

and

$$\gamma^3 \left( \frac{1}{4(\gamma - 1)} \right)^{2N(\gamma)+2} = \eta.$$

This theorem follows by minimizing $C(d) \text{vol}(\Omega_2)(\gamma \Lambda)^{d/2} + n_{\Gamma} N$ under the constraint $ntol(\gamma, N) = \eta$ by using the method of Lagrange multipliers. Observe that (5.4) and (5.5) do not depend on the exponent $r$ in the approximate cost function $g(\gamma, N)$.

**Remark 5.2.** From (5.4) and (5.5) one can numerically solve $N$ and $\gamma$ as a function of $\eta$. The resulting $N$ is typically not an integer but it can be rounded up while preserving the desired normalized tolerance $\eta$. For $d = 2$, the graphical approximation given in Figure 2 can be used. Denote

$$\hat{\eta} := \eta^{\frac{2\pi n_{\Gamma}}{\bar{\Lambda}}} \text{ and } \hat{N}(\gamma) := (\gamma - 1) \ln 4(\gamma - 1).$$

Since $27/64 \leq \gamma^3 (4(\gamma - 1))^{-2} \leq 1/2$ for $\gamma \in [2, 5]$ we have

$$\hat{\eta}(\gamma) \approx 2 \left( \frac{1}{4(\gamma - 1)} \right)^{\hat{N}(\gamma)}.$$

Using this approximation to determine $\gamma$ eliminates $n_{\Gamma}$ and $\Lambda$ from the graphical procedure. The value for $N(\gamma)$ is recovered from $N(\gamma) = \frac{\Lambda}{2\pi n_{\Gamma}} \hat{N}(\gamma)$ and rounded up.
6. Dimension reduction. It usually turns out that the space \( V_2 \) is excessively large for the chosen error level, and it can be replaced by \( \tilde{V}_2 \subset V_2 \) of considerably smaller dimension while maintaining desired accuracy. We proceed to discuss how such \( \tilde{V}_2 \) can be constructed.

For ease of presentation, assume now that the eigenvalues \( \{\mu_k\} \) in (3.2) are given in nondecreasing order, and write \( K = K(\tilde{\Lambda}) \) as in (3.10). Recall the definition of \( B \) from section 3.3

\[
B = [v_1, \ldots, v_K, (I - P_{\tilde{\Lambda}})q_1, \ldots, (I - P_{\tilde{\Lambda}})q_N],
\]

which satisfies \( V_2 = \text{range}(B) \). Note that the matrix \( B \) may have a nontrivial null space. Hence, the \( Q_{22} \)-block in the method matrix \( Q \) defined by (3.11) is obtained by computing a basis for \( \text{range}(B) \) using, e.g., SVD of \( B \).

Denoting a low-rank approximation of \( B \) by \( \tilde{B} \), the corresponding method matrix for the dimension reduced version of CPI is given by

\[
\tilde{Q} := \begin{bmatrix}
I & 0 \\
0 & Q_{22}
\end{bmatrix}, \quad \text{where ker}(\tilde{Q}_{22}) = \{0\} \text{ and range}(\tilde{Q}_{22}) = \text{range}(B).
\]

Further, let \( \tilde{V} := \text{range}(\tilde{Q}) \). We proceed to give an error estimate for the dimension reduced version of CPI.

**Lemma 6.1.** Let \( (\lambda, x) \in (0, \Lambda) \times \mathbb{R}^n \setminus \{0\} \) be an eigenpair of (1.1) with \( x = [x_1 \ x_2]^T \) according to the standard splitting. Let \( \{p_1, \ldots, p_r\} \subset \mathbb{R}^{n_2} \) be vectors satisfying (3.12). Let \( B \in \mathbb{R}^{n_2 \times (K+N_r)} \) be as defined in (6.1) and

\[
u_2 := P_{\tilde{\Lambda}}x_2 + \sum_{i=1}^{N} \ell_i(\lambda)f_{\tilde{\Lambda}}(\xi_i)Z(\lambda)x_1, \quad \text{where} \quad Z(\lambda) = \lambda M_{21} - A_{21}.
\]

Define the dimension reduced method matrix and the corresponding subspace by (6.2).

Then there exists \( \tilde{\lambda} \in \sigma_{\tilde{V}}(A,M) \) such that

\[
|\lambda - \tilde{\lambda}| \leq 2C_M C(\Lambda) \Lambda(4\gamma)^3 \left( \frac{1}{4(\gamma-1)} \right)^{2N+2} + 2 \left\| R \left( B - \tilde{B} \right) \right\|^2 \min_{\tilde{\alpha} = u_2} \|\tilde{\alpha}\|^2,
\]
where $A_{22} = R^T R$ is the Cholesky factorization of $A_{22}$, and the constants $N, \gamma, C_M$, and $C(\lambda)$ are as in Theorem 4.4.

In computations, one would choose an optimal combination of $N$ and $\gamma$ as described in section 5 for the untruncated version of CPI.

**Proof.** The original error estimate (4.8) in Theorem 4.4 was derived by implicitly constructing $u_2$ in (3.5) and (4.2) in order to bound the right-hand side of (2.9), i.e.,

$$
\min_{v_2 \in V_2} \| x_2 - v_2 \|_{A_{22}} \leq \| x_2 - u_2 \|_{A_{22}}.
$$

In that theorem, the latter term is further estimated by

$$
\| x_2 - u_2 \|_{A_{22}}^2 \leq C_M \Lambda(4\gamma)^3 \left(\frac{1}{4(\gamma - 1)}\right)^{2N+2}.
$$

The proof of the current claim follows from this by a perturbation argument. Let $\alpha$ be such that $u_2 = B\alpha$. Since $V_2 = \text{range}(B)$, we have

$$
\min_{v_2 \in V_2} \| x_2 - v_2 \|_{A_{22}} \leq \| x_2 - \tilde{B}\alpha \|_{A_{22}} = \| x_2 - B\alpha + (B - \tilde{B})\alpha \|_{A_{22}}
\leq \| x_2 - u_2 \|_{A_{22}} + \| (B - \tilde{B})\alpha \|_{A_{22}}
\leq \| x_2 - u_2 \|_{A_{22}} + \| R(B - \tilde{B})\|\|\alpha\|.
$$

The claim follows by squaring this estimate and applying Proposition 2.2 with $\tilde{V}$ in place of $V$.

To make practical use of Lemma 6.1 to achieve a given target level for the relative eigenvalue error, we start by bounding the first term in (6.4) by choosing the parameter value combination $N, \gamma$ using (5.4) and (5.5). It remains to bound the second term in (6.4) so that

$$
\left\| R \left( B - \tilde{B} \right) \right\| \min_{\beta\alpha = u_2} \| \hat{\alpha} \| \leq \| R(B - \tilde{B})\|\|\alpha\| < \sqrt{tol}
$$

for a given truncation error level $tol > 0$. A vector $\alpha$ satisfying $Q_{22}\alpha = u_2$ and upper bound for $\|\alpha\|$ are given below. Given such $\alpha$, we then use the SVD $RB = \sum_{i=1}^{K_c} \sigma_i u_i w_i^T$ (where $\sigma_i$ are ordered in nonincreasing order) to construct the lowest rank $\tilde{B}$ satisfying $\| R(B - \tilde{B})\|\|\alpha\| < \sqrt{tol}$ as

$$
\tilde{B} = R^{-1} \sum_{i=1}^{K_c} \sigma_i u_i w_i^T, \quad \text{where} \quad K_c := \max\{i \mid \sigma_i > \sqrt{tol}/\|\alpha\|\}.
$$

The method matrix block $\tilde{Q}_{22}$ for the dimension reduced CPI, defined in (6.2), is obtained as

$$
\tilde{Q}_{22} := R^{-1} [u_1 \ u_2 \ \ldots \ u_{K_c}],
$$

where the column vectors of $\tilde{Q}_{22}$ are orthonormal in the $A_{22}$-inner product. Because $\alpha \neq 0$, the number $K_c = K_c(\|\alpha\|, tol)$ is always defined, and the truncation error level $tol > 0$ can be chosen arbitrarily small.

We make use of the Lebesgue constant $\Lambda_N$ for Chebyshev interpolation points of $(0, \Lambda)$ (see, e.g., [10]), given by

$$
\Lambda_N := \max_{t \in (0, \Lambda)} \sum_{i=1}^{N} \mid \ell_i(t) \mid \quad \text{that satisfy} \quad \Lambda_N = \frac{2}{\pi} \log N + O(1).
$$
Lemma 6.2. Make the same assumptions and use the same notation as in Lemma 6.1. Then there exists $\alpha$ satisfying $u_2 = Q_{22}\alpha$ such that
\begin{equation}
\|\alpha\|^2 \leq 1 + \Lambda_N^2 \|\theta\|^2,
\end{equation}
where the coefficient vector $\theta := [\theta_1 \quad \theta_2 \quad \ldots \quad \theta_r]^T$ satisfies
\begin{equation}
Z(\lambda)x_1 = \sum_{j=1}^r \theta_j p_j.
\end{equation}
Proof. Define the coefficients $\tau_k$ by the expansion
\begin{equation}
P_\Lambda x_2 = \sum_{k=1}^K \tau_k v_k \quad \text{so that} \quad \|P_\Lambda x_2\|^2_{M_{22}} = \sum_{k=1}^K \tau_k^2 \leq 1.
\end{equation}
Indeed, this holds by the $M_{22}$-orthogonality of $\{v_k\}$ (see (3.3)) and the normalization $x^TMx = 1$ implying $\|x_2\|_{M_{22}} \leq 1$ and hence $\|P_\Lambda x_2\|_{M_{22}} \leq 1$.
Define $\alpha_{i,j} := \ell_i(\lambda)\theta_j$, where $i = 1, \ldots, N$ and $j = 1, \ldots, r$. Then $u_2$ in (6.3) can be written in the form
\begin{equation}
u_2 = Q_{22}\alpha = \sum_{k=1}^K \tau_k v_k + \sum_{i=1}^N \sum_{j=1}^r \alpha_{i,j} (I - P_\Lambda) q_{i,j}
\end{equation}
by (3.13)–(6.1). So, we can choose
\[ \alpha := [\tau_1 \quad \ldots \quad \tau_K \quad \alpha_{1,1} \quad \ldots \quad \alpha_{1,r} \quad \ldots \quad \alpha_{N,1} \quad \ldots \quad \alpha_{N,r}]^T. \]
By (6.9) and the definition of $\alpha_{i,j}$ we have
\[ \|\alpha\|^2 = \sum_{k=1}^K \tau_k^2 + \sum_{i=1}^N \sum_{j=1}^r \alpha_{i,j}^2 \leq 1 + \sum_{i=1}^N \ell_i(\lambda)^2 \sum_{j=1}^r \theta_j^2. \]
Observing that $\sum_{i=1}^N \ell_i(\lambda)^2 \leq \Lambda_N^2$ completes the proof.

Lemma 6.3. Make the same assumptions and use the same notation as in Lemma 6.1. Assume that the matrices $A_{21}$, $M_{21}$ are ordered so that only their first $r_1$, $r_2$ column vectors are nonvanishing, respectively. Let
\begin{equation}
\{p_1, \ldots, p_r\} = \{A_{21}e_1, \ldots, A_{21}e_{r_1}, M_{21}e_1, \ldots, M_{21}e_{r_2}\}.
\end{equation}
Then there exists a coefficient vector $\theta := [\theta_1 \quad \theta_2 \quad \ldots \quad \theta_r]^T$ such that $Z(\lambda)x_1 = \sum_{j=1}^r \theta_j p_j$ and
\[ \|\theta\|^2 \leq (1 + \lambda^2)\|M^{-1}\|. \]
Proof. Since $Z(\lambda) = \lambda M_{21} - A_{21}$ holds, it follows from (6.10) that
\[ Z(\lambda)x_1 = \sum_{j=1}^{r_1} \lambda x_{1,j} M_{21} e_j - \sum_{j=1}^{r_2} x_{1,j} A_{21} e_j \quad \text{and} \quad \sum_{j=1}^r \theta_j p_j = \sum_{j=1}^{r_1} \theta_j M_{21} e_j + \sum_{j=r_1+1}^{r_2} \theta_j A_{21} e_j. \]
Hence, one solution of (6.8) is \( \theta = \begin{bmatrix} \lambda x_{1,1} & \ldots & \lambda x_{1,r_1} & -x_{1,1} & \ldots & -x_{1,r_2} \end{bmatrix} \), and it satisfies the estimate

\[
\|\theta\|^2 = \lambda^2 \sum_{j=1}^{r_1} x_{1,j}^2 + \sum_{j=1}^{r_2} x_{2,j}^2 \leq (1 + \lambda^2) \|x_1\|^2
\]

(6.11)

\[
\leq (1 + \lambda^2) \|x\|^2 = (1 + \lambda^2) \frac{x^T x}{x^T M x}
\]

\[
\leq (1 + \lambda^2) \max_{v \in \mathbb{R}^n, v \neq 0} \frac{v^T v}{v^T M v} = (1 + \lambda^2) \|M^{-1}\|
\]

because \( x^T M x = 1 \). The proof is now complete.

The combination of Lemmas 6.1, 6.2, and 6.3 yields the following result.

**Theorem 6.4.** Let \( (\lambda, x) \in (0, \Lambda) \times \mathbb{R}^n \setminus \{0\} \) be an eigenpair of (1.1). Let the vectors \( \{p_1, \ldots, p_n\} \) be defined by (6.10) and the matrix \( Q_{22} \in \mathbb{R}^{n_2 \times (K+N)} \) as in (6.1). Define \( R \) by the Cholesky factorization \( R^T R = A_{22} \) and let \( \sum_{i=1}^{n_2} \sigma_i u_i w_i^T \) be the SVD of \( R Q_{22} \). For any truncation error level \( \text{tol} > 0 \), define

\[
K_c(\text{tol}) := \max \{ i \mid \sigma_i^2 (1 + \Lambda_N^2) (1 + \lambda^2) \|M^{-1}\| > \text{tol} \},
\]

where \( \Lambda_N \) is given by (6.6). Define the method matrix \( \tilde{Q} \) and the subspace \( \tilde{V} \) by (6.2) and (6.5).

Then there exists \( \tilde{\lambda} \in \sigma_{\tilde{V}}(A, M) \) such that

\[
\frac{\lambda - \tilde{\lambda}}{\lambda} \leq 2C_M C(\lambda) \Lambda(4\gamma)^3 \left( \frac{1}{4(\gamma - 1)} \right)^{2N+2} + 2\text{tol}
\]

for any parameter value combination \( N, \gamma \) where the constants \( C_M \) and \( C(\lambda) \) are as in Theorem 4.4.

A typical application of CPI is the solution of the lowest eigenmodes of the Laplace operator in a bounded domain \( \Omega \subset \mathbb{R}^d \) using the finite element method. When piecewise linear basis functions are used on quasi-uniform simplicial meshes, it is known that \( \|M^{-1}\| \leq C h^{-d} \), where the constant \( C \) is independent on the mesh size \( h \); see [9, section 6]. A reasonable value for the cut-off index in Theorem 6.4 can be computed with the help of this estimate.

**7. Model problems.** We proceed to illustrate theoretical results by two numerical examples. Both examples involve the eigenpairs \( (\lambda', u) \in (0, \Lambda) \times \mathcal{V} \) of the variational eigenvalue problem

\[
(\nabla u, \nabla v) = \lambda' (u, v) \quad \text{for all} \quad v \in \mathcal{V}
\]

(7.1)

of the Laplace operator where \( \mathcal{V} \subset H^1(\Omega) \) is a subspace where the homogeneous Dirichlet boundary condition holds at least on a part of the boundary \( \partial \Omega \). Problem (7.1) is discretized using the finite element method with piecewise linear basis functions leading to the algebraic eigenvalue problem (1.1) that is the subject matter of this article.

**7.1. Computational considerations.** Let us begin by describing an implementation of CPI. The problem data consists of the spectral interval of interest \( (0, \Lambda) \), the specified upper bound for the relative eigenvalue error, and the symmetric positive
definite stiffness and mass matrices $A$ and $M$. Without loss of generality, the basis functions can be assumed to be ordered so that $A$ and $M$ obey the standard splitting given in (2.2) corresponding to the interior and the exterior systems. The purpose is to compute spectral approximations

$$
\sigma_V(A, M) \cap (0, \Lambda) \approx \sigma(A, M) \cap (0, \Lambda)
$$

for several versions of (1.1) sharing the same exterior system. Note that the dimension of the eigenvalue problem may vary between different versions as long as $\text{range}(A_{21})$ and $\text{range}(M_{21})$ remain fixed. Thus, the finite element mesh of the exterior part stays constant while the mesh of the interior part may vary.

As discussed in section 5, an effective choice of $\gamma$ and $N$ requires an a priori information on the eigenvalue distribution of problem (7.1) that is encoded in the function $K(l)$ in (3.10). We model $K(l)$ by the Weyl law as in section 5. Values for $\gamma$ and $N$ are then chosen using Theorem 5.1. For practical reasons, we set $C(\lambda) = C_M = 1$ in (4.8). For a given finite element mesh size $h > 0$, the term $\|a\|$ is approximated by setting $\|M^{-1}\| = h^{-d}$ and applying Lemmas 6.2 and 6.3.

CPI consists of the following steps:

1. Using the target relative error level, determine $N$ and $\gamma$ using, e.g., (5.4) and (5.5). Compute the Chebyshev points $\{\xi_i\}_{i=1}^N$ using (4.5).
2. Compute the $K = K(\gamma \Lambda)$ smallest eigenpairs $(\mu_k, \nu_k)$ of the exterior system $A_{22} \nu_k = \mu_k M_{22} \nu_k$.
3. Let $p_j$, $j = 1, \ldots, r_1$, and $q_j$, $j = r_1 + 1, \ldots, r$, with $r = r_1 + r_2$, be the nonzero columns of $M_{21}$ and $A_{21}$, respectively. Compute the sample vectors $q_{1j}, \ldots, q_{N_r}$ as solutions of $(A_{22} - \xi_i M_{22}) q_{ij} = p_j$.
4. Collect eigenvectors from step 1 and sample vectors from step 2 into matrix $B = [v_1, \ldots, v_K, (I - P_\Lambda)q_{11}, \ldots, (I - P_\Lambda)q_{N_r}]$. Compute the SVD $RB = \sum_{i} \sigma_i u_i w_i^T$, where $\sigma_i$ are ordered in nonincreasing order and $R^T R = A_{22}$.
5. Choose the cut-off index as $K_c = \max\{i \mid \sigma_i \geq \|a\|^{-1} tol\}$. Construct the method matrix using vectors $u_1, \ldots, u_{K_c}$ from step 3 as

$$
\tilde{Q} = \begin{bmatrix} I & 0 \\ 0 & \tilde{Q}_{22} \end{bmatrix}, \quad \tilde{Q}_{22} = R^{-1} \begin{bmatrix} u_1 & \ldots & u_{K_c} \end{bmatrix}.
$$

6. Solve the eigenvalue problem $\tilde{Q}^T A \tilde{Q} \tilde{x} = \tilde{\lambda} \tilde{Q}^T M \tilde{Q} \tilde{x}$, e.g., using the Lanczos iteration.

In step 2, one has to determine a tolerance for computing the exterior eigenvectors $v_1, \ldots, v_K$. We proceed to analyze the effect of exterior eigenvector error to the accuracy of the eigenvalues computed using the CPI method. Our analysis relies on a perturbation argument identical to one used in Lemma 6.1. When the exterior eigenvectors are incorrectly computed, the space $V_2$ is replaced by $\tilde{V}_2$ defined as

$$
(7.2) \quad \tilde{V}_2 = \text{range}(\tilde{P}_\Lambda) \oplus \tilde{W}_2, \quad \text{where} \quad \tilde{W}_2 = \text{span}_{i=1,\ldots,N} \{\tilde{f}_\Lambda(\xi_i) [M_{21} \ A_{21}]\}.
$$

The matrix $\tilde{P}_\Lambda$ is the $A_{22}$-orthogonal projection onto $\text{span}_{k=1,\ldots,K} \{\tilde{v}_k\}$ and $\tilde{f}_\Lambda(\xi) = (I - \tilde{P}_\Lambda)(A_{22} - \xi M_{22})^{-1}$. Note that $P_\Lambda$ defined in (3.3) is also an $A_{22}$-orthogonal projection. Following [6], the eigenvector error is measured using the gap, i.e., the maximum angle between the exact and the approximate eigenspace in the $A_{22}$-norm.
as

\begin{equation}
(7.3) \quad d_H := \max \left( \max_{\mathbf{u} \in \text{range}(P_{\lambda})} \| (I - \hat{P}_{\lambda}) \mathbf{u} \|_{A_{22}}, \max_{\hat{\mathbf{u}} \in \text{range}(P_{\lambda})} \| (I - \hat{P}_{\lambda}) \hat{\mathbf{u}} \|_{A_{22}} \right).
\end{equation}

**Corollary 7.1.** Let \((\lambda, \mathbf{x}) \in (0, \Lambda) \times \mathbb{R}^n \setminus \{0\}\) be an eigenpair of (1.1) with \(\mathbf{x} = [x_1, x_2]^T\) according to the standard splitting. By \(\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_K\) denote the set of approximate exterior eigenvectors and by \(\hat{P}_{\lambda}\) the \(A_{22}\)-orthogonal projection onto \(\text{span}_{k=1, \ldots, K} \{\hat{\mathbf{v}}_k\}\). Let the method subspace \(\tilde{V}\) and \(\tilde{V}_2\) in (7.2) be related as \(V\) and \(V_2\) in (2.8). Then there exists \(\hat{\lambda} \in \sigma_{\tilde{V}}(A, M)\) such that

\[
\frac{\| \lambda - \hat{\lambda} \|}{\lambda} \leq 2C_M C(\lambda) d_H \left( 4 + \frac{1}{\alpha} \right)^{N+2} + C' \max_{k=1, \ldots, N} \left( \frac{\mu_k - \lambda}{\mu_k - \xi} \right)^2 d_H^2
\]

for any parameter value combination \(N, \gamma\) with \(C' := 2(1 + \Lambda N)^2\). The constants \(C_M\) and \(C(\lambda)\) are as in Theorem 4.4, \(d_H\) is the maximal angle defined by (7.3), and \(\Lambda N\) is defined in (6.6).

The contribution of inaccurate exterior eigenvectors to relative error in eigenvalues depends inversely on \(\text{dist}\{\xi_i, \mu_k\}\). As Chebyshev interpolation points are not nested and \(\mu_k\)’s have been determined in step 2, one may be able to adjust the number of interpolation points \(N\) so that \(\text{dist}\{\xi_i, \mu_k\}\) increases.

**Proof.** Let \(\hat{\mathbf{u}}_2 = \hat{P}_{\lambda} \mathbf{x}_2 + \sum_{i=1}^N \ell_i(\lambda) \hat{f}_A(\xi_i) Z(\lambda) \mathbf{x}_1\) and \(\mathbf{u}_2\) be as in Lemma 6.1. Then

\[
\mathbf{u}_2 - \hat{\mathbf{u}}_2 = (\hat{P}_{\lambda} - P_{\lambda}) \mathbf{x}_2 + \sum_{i=1}^N \ell_i(\lambda)) (P_{\lambda} - \hat{P}_{\lambda}) (A_{22} - \xi_i M_{22})^{-1} Z(\lambda) \mathbf{x}_1.
\]

Since \(\| \hat{P}_{\lambda} - P_{\lambda} \|_{A_{22}} = d_H\) by [8], we have

\[
\| \mathbf{u}_2 - \hat{\mathbf{u}}_2 \|_{A_{22}} \leq d_H \| \mathbf{x}_2 \|_{A_{22}} + d_H \sum_{i=1}^N \| \ell_i(\lambda) \|_{A_{22}} (A_{22} - \xi_i M_{22})^{-1} Z(\lambda) \mathbf{x}_1 \|_{A_{22}}
\]

and further,

\[
\| \mathbf{u}_2 - \hat{\mathbf{u}}_2 \|_{A_{22}} \leq d_H \| \mathbf{x}_2 \|_{A_{22}} + d_H A_N \max_{i=1, \ldots, N} \| (A_{22} - \xi_i M_{22})^{-1} Z(\lambda) \mathbf{x}_1 \|_{A_{22}}.
\]

It now follows from a similar argument that was used to derive (4.9) that

\[
\| (A_{22} - \xi_i M_{22})^{-1} Z(\lambda) \mathbf{x}_1 \|_{A_{22}} ^2 = \sum_{k=1}^{N} \mu_k \left( \frac{\beta_k^2(\lambda)}{\mu_k - \xi} \right) ^2 = \max_{k=1, \ldots, n} \left( \frac{\mu_k - \lambda}{\mu_k - \xi} \right) ^2 \| \mathbf{x}_2 \|_{A_{22}} ^2.
\]

The claim follows by a perturbation argument as in Lemma 6.1.

**Remark 7.2.** An alternative approach for step 3 is to directly solve for \(\hat{q}_{ij} := (I - P_{\lambda})^l q_{ij}\) using the saddle point formulation

\[
\begin{bmatrix}
(A_{22} - \xi_i M_{22}) & M_{22} B_1 \\
B_1^T M_{22} & 0
\end{bmatrix}
\begin{bmatrix}
\hat{q}_{ij} \\
\nu
\end{bmatrix}
= \begin{bmatrix}
\hat{p}_j \\
0
\end{bmatrix},
\]
where $B_1 = [v_1 \ldots v_K]$ and $v_k$’s are computed in step 2. This formulation preserves most of the sparse structure of the linear systems (3.13) and is numerically stable when an interpolation point $\xi_i$ is close to $\mu_k$. All numerical experiments were performed without paying attention to this issue.

Due to memory constraints, it is not always feasible to store vectors $q_{ij}$ in step 3 or to explicitly construct $B \in \mathbb{R}^n_2 \times (K + Nr)$ in step 4 when $n_2$ or $K + Nr$ is prohibitively large. In the construction of the method matrix, step 5, only vectors $u_1, \ldots, u_K$ corresponding to the largest singular values of $RB$ are needed. Steps 3 and 4 can be combined into an iterative solution of the largest singular values of $RB$ and the corresponding vectors $u_i$ using action of $(RB)(RB)^T$ without storing $q_{ij}$.

### 7.2. 2D rectangle.

We consider numerical solution of (7.1) in the rectangular domain shown in Figure 1. The homogeneous Dirichlet boundary condition is used. The domain is uniformly discretized with 67,872 triangular elements and 34,241 nodes. With the boundary conditions, this resulted in $n = 33,633$ degrees of freedom with $n_1 = 163$, $n_1 = 14,638$, and $n_2 = 18,995$ in (2.2). The spectral interval of interest $(0, \Lambda)$ with $\Lambda = 135$ allows us to compute 15 of the lowest eigenvalues $\lambda_1, \lambda_2, \ldots$ in nondecreasing order. The numerically obtained largest relative eigenvalue errors (without using the SVD-based dimension reduction process of section 6) and its upper bound from (4.8) are shown in Figure 3.

The largest relative eigenvalue errors with several values of $\gamma$ and $N$ are compared to the theoretical estimate (equation (4.8)) in Figure 4. For validation of the cost model in (5.2), the computational time to solve (2.5) with several $N$ and $\gamma$ is illustrated in Figure 5.

The effect of the SVD-based dimension reduction of section 6 is demonstrated in Figure 6. The relative eigenvalue error and $\text{dim}(\tilde{V}_2)$ are given as a function of the truncation error level $\text{tol} > 0$.

Finally, the CPI method was benchmarked against the CMS implementation detailed in section 2.3. Values for $N$ and $\gamma$ were chosen from (5.4) and (5.5) given a list of target error levels. The value of $\text{tol}$ was adjusted for each target error level. The comparison is illustrated in Figure 7.
**Fig. 4.** Largest relative eigenvalue errors for 15 lowest eigenvalues of the 2D rectangle using several values of $N$ and $\gamma$ from the theoretical estimate equation (4.8) (left) and from numerical experiments (right). The value of $K(\gamma \Lambda)$ varies from $K = 18$ when $\gamma = 2$ to $K = 48$ for $\gamma = 5$.

**Fig. 5.** The computation time (in seconds) required by MATLAB R2017a `eigs` to solve the 15 smallest eigenvalues of (2.5) for several values of $N$ and $\gamma$. The time is averaged from 50 computations. An Intel Xeon E5-1630 CPU with 32GB of RAM was used.

**Fig. 6.** The effect of the cut-off threshold $\sigma$ in SVD to the relative error and the size of the method matrix $Q$. The parameter values $\gamma = 4$ and $N = 6$ were used.
7.3. 3D acoustic example. The computational domain shown in Figure 1 consists of a human vocal tract geometry $\Omega_1$ and a mock-up model of the MRI head coil $\Omega_2$. The vocal tract geometries were automatically extracted from MRI data as explained in [1, 27], and the interface $\Gamma$ was attached. The vocal tract was embedded into a head model purchased from Turbosquid [31].

The homogeneous Dirichlet boundary condition was posed on the areas marked in Figure 8, and the Neumann condition was used on other parts of the boundary. The interface $\Gamma$ is a spherical surface separating $\Omega_1$ and $\Omega_2$. We use three versions of the vocal tract geometry corresponding to Finnish vowels [a], [i], and [u] as visualized in Figure 9. The domain $\Omega_2$ contains 522,517 tetrahedral elements and 101,222 nodes, and the interface has $n_\Gamma = 950$ degrees of freedom. Having set the boundary conditions, we have $n_2 = 97,375$. 

---

**Fig. 7.** The CPI method compared with CMS in the 2D rectangular domain. The $x$-axis corresponds to the amount of degrees of freedom related to $\Omega_2$.

**Fig. 8.** A visualization of the special degrees of freedom for the acoustic example. Dirichlet boundary condition is posed on the red area, and the interface $\Gamma$ is marked by the yellow faces.

**Fig. 9.** The three vocal tract geometries ([a], [i], [u]) together with the fixed exterior domain $\Omega_2$. 

---

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
The computation time (in seconds) required by MATLAB eigs to solve the original eigenvalue problem equation (1.1) versus solving (2.5), where \( \tilde{Q} \) is constructed by dimension reduced CPI. The forty lowest eigenvalues were solved, and their maximum relative eigenvalue error is given. The same software and hardware were used as specified in Figure 5.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>( n_1 )</th>
<th>CPI</th>
<th>Original</th>
<th>Maximum relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i]</td>
<td>16293</td>
<td>4.44</td>
<td>26.05</td>
<td>7.05 \cdot 10^{-8}</td>
</tr>
<tr>
<td>[i]</td>
<td>13939</td>
<td>4.57</td>
<td>28.49</td>
<td>7.28 \cdot 10^{-8}</td>
</tr>
<tr>
<td>[u]</td>
<td>15139</td>
<td>4.66</td>
<td>25.66</td>
<td>7.36 \cdot 10^{-8}</td>
</tr>
</tbody>
</table>

The eigenvalues \( \lambda \) in (7.1) and resonant frequencies satisfy \( 2\pi f = c\lambda^{1/2} \), where \( c = 344.2 \) is the speed of sound at temperature 294K. Setting the target relative eigenvalue error to \( 10^{-6} \) corresponds to relative error of \( 10^{-3} \) for resonant frequencies. The spectral interval of interest \( \lambda \in (0,3,000) \) matches frequencies up to 3kHz. Equations (5.4) and (5.5) give \( N = 3 \), \( \gamma = 8 \) for a relative eigenvalue error level \( 10^{-6} \).

Writing

\[
\begin{bmatrix}
A_{11} & A_{12} \tilde{Q}_{22} \\
A_{21} \tilde{Q}_{22} & \tilde{Q}_{22}^T A_{22} \tilde{Q}_{22}
\end{bmatrix}
\begin{bmatrix}
\tilde{x}
\end{bmatrix}
= \lambda
\begin{bmatrix}
M_{11} & M_{12} \tilde{Q}_{22} \\
M_{21} \tilde{Q}_{22} & \tilde{Q}_{22}^T M_{22} \tilde{Q}_{22}
\end{bmatrix}
\begin{bmatrix}
\tilde{x}
\end{bmatrix}
\]

for a given a method matrix \( \tilde{Q} \), we can use the same blocks \( \tilde{Q}_{22}^T A_{22} \tilde{Q}_{22} \) and \( \tilde{Q}_{22}^T M_{22} \tilde{Q}_{22} \) for different versions, and only the off-diagonal blocks have to be updated.

A conservative choice for truncation index \( K_c \) was made according to target relative error \( 10^{-6} \) resulting into subspace \( \tilde{V}_2 \) with \( \text{dim}(\tilde{V}_2) = 1,601 \). The forty lowest resonant frequencies could be computed to the given tolerance. The computation times and maximum relative eigenvalue errors are listed in Table 1. The sample vectors in step 3 were solved in parallel for each sample point using different cores of the CPU. The precomputations needed to form the CPI method matrix took 13 minutes and 15 seconds. As shown in Table 1, the time difference between solving the original and the reduced eigenvalue problem is around 20 seconds. Thus, 40 different versions of the acoustic eigenvalue problems need to be solved in order to break even in terms of computational time.

A comparison between CPI and CMS was also performed in the 3D example. The values \( \gamma = 8 \), \( N = 3 \) were used in CPI with a varying cut-off threshold \( \sigma \) in order to produce comparable dimensions for the subspace \( \tilde{V}_2 \). The largest relative eigenvalue error for both methods was measured using different subspace dimensions. Additionally, the relative eigenvalue error for each of the 40 smallest eigenvalues was compared when the subspace dimension was 1,600. For comparison, MATLAB eigs uses by default a subspace of dimension at most \( 2k \), where \( k \) is the number of eigenvalues to be computed. The results are shown in Figure 10. Computing this many eigenvalues of the exterior system for the CMS method required 27 minutes and 41 seconds, which is approximately twice the time required by the CPI method. With CPI and the parameters used, 550 eigenvalues needed to be used for the exterior system.

**8. Conclusions.** We have presented a CPI method for efficient solution of symmetric structured eigenvalue problems by constructing a particular method subspace \( V \). Error analysis for the CPI method shows convergence of relative eigenvalue error at rate \( C\rho^N \), where \( N \) is the number of interpolation points in the spectral interval of
interest $(0, \Lambda)$, and $\rho < 1$ depends on the oversampling parameter $\gamma > 5/4$. Optimal parameter values for $\gamma$ and $N$ are chosen based on a cost model. A dimension reduced version of the CPI method with convergence analysis is given. Numerical experiments on the finite element discretized Laplace operator $(d = 2, 3)$ show that the method has practical value and indicate faster convergence in comparison to the convergence estimate equation (4.8). The authors have observed that the CPI method becomes increasingly competitive against MATLAB `eigs` for large values of $\Lambda$, excluding precomputations. In addition, the performance is improved for families of matrices where $n_1$ and $n_1/n_2$ are small.

The CPI method requires precomputation of a basis for the method subspace $V$. This involves solution of exterior eigenvectors $(\mu_k, v_k)$ satisfying $\mu_k \leq \gamma \Lambda$ and the solution of $n_1 N$ linear systems. Additional precomputation cost is produced by SVD in dimension reduction. The cost model used for optimizing $\gamma$ and $N$ does not account for any precomputation costs. In the acoustic eigenvalue problem used for benchmarking, the CPI method becomes competitive in comparison to `eigs` after 40 eigensolves if precomputation time is taken into account. Even costly precomputations are justified in processes where eigensolution must be repeated several times by the designer.

The computational speed of the CPI method comes at a price; the algorithm is memory intensive compared to `eigs`. In the acoustic eigenvalue problem involving the 40 lowest eigenvalues, the method subspace is of dimension 1,601, whereas the subspace used by `eigs` is of dimension 80. The feasibility of precomputations may be limited by storage more severely than by computation time, in particular, if dimension reduction is not used.

Application of the CPI method to domain decomposition in the finite element context presents a topic for future work. In fact, the presented analysis already extends to the setting where $A_{22}$ and $M_{22}$ are block diagonal matrices as is the case when multiple subdomains are treated.

Acknowledgments. The geometry for the exterior model in section 7.3 is loosely based on the MRI head coil design provided by Siemens Healthineers. The authors are grateful for the comments of the reviewers.
## EIGENPROBLEMS FROM FAMILIES OF COUPLED SYSTEMS

### REFERENCES

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title</th>
<th>Journal/Book</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M. C. C. Bampton and R. R. Craig</td>
<td>Coupling of substructures for dynamic analyses</td>
<td>AIAA J., 6 (1968)</td>
<td>pp. 1313–1319</td>
</tr>
<tr>
<td>A. Böttcher and I. Spitkovsky</td>
<td>A gentle guide to the basics of two projections theory</td>
<td>Linear Algebra Appl., 432 (2010), pp. 1412–1459</td>
<td></td>
</tr>
</tbody>
</table>

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.


