

ARTICLE TYPE

Calculating the minimal/maximal eigenvalue of symmetric parametrized matrices using projection

Koen Ruymbeek* | Karl Meerbergen | Wim Michiels

¹Department of Computer Science, KU Leuven, Belgium

Correspondence

*Koen Ruymbeek, Email:
koen.ruymbeek@cs.kuleuven.be

Present Address

Celestijnenlaan 200A, 3001 Leuven.

Summary

In applications of linear algebra including nuclear physics and structural dynamics, there is a need to deal with uncertainty in the matrices. We focus on matrices that depend on a set of parameters ω and we are interested in the smallest eigenvalue of a matrix pencil (A, B) with A, B symmetric and B positive definite. If ω can be interpreted as the realisation of random variables, one may be interested in statistical moments of the smallest eigenvalue. In order to obtain statistical moments, we need a fast evaluation of the eigenvalue as a function of ω . Since this is costly for large matrices, we are looking for a small parametrized eigenvalue problem whose smallest eigenvalue makes a small error with the minimal eigenvalue of the large eigenvalue problem. The advantage, in comparison with a global polynomial approximation (on which, e.g., the polynomial chaos approximation relies), is that we do not suffer from the possible non-smoothness of the minimal eigenvalue. The small scale eigenvalue problem is obtained by projection of the large scale problem. Our main contribution is that for constructing the subspace we use multiple eigenvectors as well as derivatives of eigenvectors. We provide theoretical results and document numerical experiments regarding the beneficial effect of adding multiple eigenvectors and derivatives.

KEYWORDS:

Generalised eigenvalue problems, symmetric matrices, global approximation, minimal eigenvalue

1 | INTRODUCTION

Let $A(\omega), B(\omega)$ be $n \times n$ symmetric matrices which smoothly depend on a parameter ω in a compact subset $\Omega \subset \mathbb{R}^d$. We impose that $B(\omega)$ is positive definite, for $\omega \in \Omega$, so all eigenvalues $\lambda(\omega)$ of the generalized eigenvalueproblem $(A(\omega), B(\omega))$,

$$A(\omega)x(\omega) = \lambda(\omega)B(\omega)x(\omega),$$

are real, see, e.g.,¹ Ch. 9. The objective is to calculate in an efficient way, an accurate global approximation of the minimal eigenvalue $\lambda_1(\omega)$ over the whole parameterspace. Standard eigensolvers like the Lanczos' method², Jacobi-Davidson³ or LOBPCG⁴ can be applied for large-scale problems for some points in the parameterspace but such an approach becomes expensive when considering a large number of sample points.

This paper focuses on large scale eigenvalue problems. One class of examples is the estimation of the coercivity constant of parametrized elliptic partial differential equations⁹. Another is the computation of the inf-sup constant of PDEs from the smallest singular value of $A(\omega)$, which is computed from the minimal eigenvalue of $A(\omega)^T A(\omega)$. The minimal eigenvalue problem

also plays a role in the characterization of pseudospectra^{10, 11} (**reviewer 2: remark 5**). Related problems are the determination of statistical moments of the minimal eigenvalue when ω is the realization of random variables and the minimization or maximization of the j th largest eigenvalue of a parameter dependent Hermitian matrix; see¹² and the references therein.

There are several methods for this kind of problem. The first category concerns polynomial-based methods. In citeAndreev2012 the authors approximate $\lambda_1(\omega)$ by sparse tensor products of Legendre polynomials and in⁶ polynomial chaos is advocated, a technique which is frequently used when the parameters are stochastic. These techniques have difficulties with the possible lack of smoothness of the minimal eigenvalue. Another approach is the so-called Successive Constraint Method (SCM), see⁷ and the references therein. In paper⁷, lower and upper bounds are calculated for each sample in a finite subset of the parameterspace. The main disadvantage of this method is that it does not provide a way to calculate the smallest eigenvalue of a parameter sample other than the samples in the initial subset.

The idea of this paper is to approximate the minimal eigenvalue of a large scale matrix pencil by the minimal eigenvalue of a smaller matrix, which we call the reduced problem. The reduced problem is obtained by projection of the large scale matrix on a well-chosen subspace. The reduced problem adopts the same smoothness properties or lack thereof as the original large scale problem, which allows us to use less samples than methods based on smooth approximations such as polynomials. The novelty of this paper is in the construction of the subspace to build the reduced eigenvalue problem. The subspace is constructed from samples of the associated eigenvector of the large scale matrix in the parameterspace. The selection of sample points is based on a greedy method⁸, with the aim to minimize the residual norm of the large scale eigenvalue problem with a minimum amount of samples. For each sampling point, a large scale eigenvalue problem is solved. Eigenvalue solvers such as Lanczos' method or Jacobi-Davidson usually compute more information than just one eigenvector approximation. The goal is to explore whether other (freely available) information can be used with the aim to further reduce the number of sample points and, consequently, the computational cost. In particular, we will study higher order Hermite interpolation sampling by inserting partial derivatives of the eigenvector towards the parameters in the subspace as well, and the addition of more than one eigenvector approximation for each sample point. Calculating the maximal eigenvalue or an eigenvalue that meets another condition (second maximal/minimal eigenvalue) can be done in a similar way as the method proposed here.

The plan of the paper is as follows. In section 2 we introduce the necessary notation and we give some motivating examples. We explain theoretically why we add the first eigenvectors as well as the partial derivatives of the first eigenvector in section 3. Section 4 gives insight how eigenvectors change over the parameters. Subsequently we explain how we build up our subspace and we present the algorithm. In section 6 we illustrate the theory and algorithms numerically. We finish this paper with some concluding remarks and research possibilities.

2 | NOTATION AND MOTIVATING EXAMPLES

We denote vectors by small bold letters and matrices by large bold letters. A subspace of \mathbb{R}^n is denoted by calligraphic letters. We already mentioned that we restrict ourselves to the case where both \mathbf{A} and \mathbf{B} are symmetric, \mathbf{B} is positive definite; \mathbf{A} and \mathbf{B} depend analytically on the parameter for all $\omega \in \Omega$. From this last requirement, it follows that the inner product defined by

$$(\mathbf{x}, \mathbf{y})_{\mathbf{B}} := \mathbf{x}^T \mathbf{B} \mathbf{y}, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

and the associated \mathbf{B} -norm

$$\|\mathbf{x}\|_{\mathbf{B}} := (\mathbf{x}, \mathbf{x})_{\mathbf{B}}, \mathbf{x} \in \mathbb{R}^n$$

are well-defined. We call $(\lambda_i, \mathbf{x}_i)$, $i = 1, \dots, n$ an *eigenpair* of the generalised eigenvalue problem (\mathbf{A}, \mathbf{B}) if

$$\mathbf{A} \mathbf{x}_i = \lambda_i \mathbf{B} \mathbf{x}_i, \quad \mathbf{x}_i \neq \mathbf{0}$$

where λ_i and \mathbf{x}_i are respectively called an *eigenvalue* and an associated *eigenvector*. In this case it is proven in¹³ that the eigenvalues are real and we assume further that

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n. \quad (1)$$

Eigenvectors are always assumed to be \mathbf{B} -orthonormal, meaning that $\mathbf{x}_i^T \mathbf{B} \mathbf{x}_j = \delta_{i,j}$, $i, j = 1, \dots, n$. (**reviewer 2: remark 1**).

We further define

$$\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \quad \mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \quad (2)$$

with $\mathbf{X}^T \mathbf{B} \mathbf{X} = \mathbf{I}_n$ then we can prove that

$$\mathbf{X}^T \mathbf{A} \mathbf{X} = \mathbf{\Lambda}. \quad (3)$$

It immediately follows that if \mathbf{A} is also positive definite, the eigenvalues of the couple (\mathbf{A}, \mathbf{B}) are strictly positive. Furthermore, if the eigenvalue is simple for $\boldsymbol{\omega} \in \Omega \subset \mathbb{R}^d$, there is an environment around $\boldsymbol{\omega}$ where this eigenvalue is differentiable as we enumerate the eigenvalues for all parameter values in increasing order. (reviewer 2: remark 2) If the eigenvalue has a multiplicity $m > 1$ for a given $\boldsymbol{\omega}$ and is simple in an open set around $\boldsymbol{\omega}$, then it can be decomposed in this open set in each direction into m differentiable curves and something similar can be done for \mathbf{B} -normalized eigenvectors, see e.g.¹⁴ ch. 9. This does not mean that it can also be decomposed in d -dimensional surfaces, see Example 1.

Remark that in view of (1), \mathbf{X} is not continuous in the parameter since in an open set around a point where not all eigenvalues are simple, the order of the eigenvectors may change.

Example 1. Let $\Omega = [-0.5, 0.5] \times [-0.5, 0.5]$ be the parameterspace and let the matrices be

$$\mathbf{A}(\boldsymbol{\omega}) = \mathbf{W}(\boldsymbol{\omega}) \begin{bmatrix} \omega_1 + 1 & \omega_2 & & & \\ \omega_2 & -\omega_1 + 1 & & & \\ & & 3 & & \\ & & & \ddots & \\ & & & & n \end{bmatrix} \mathbf{W}(\boldsymbol{\omega})^T, \mathbf{B}(\boldsymbol{\omega}) = \mathbf{I}_n$$

with $\mathbf{W}(\boldsymbol{\omega}) = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n]$ an orthonormal $(n \times n)$ -matrix for all $\boldsymbol{\omega} \in \Omega$. The eigenvalues are $\lambda_1(\boldsymbol{\omega}) = -\sqrt{\omega_1^2 + \omega_2^2} + 1$, $\lambda_2(\boldsymbol{\omega}) = \sqrt{\omega_1^2 + \omega_2^2} + 1$ and the other eigenvalues are equal to $3, \dots, n$. The two smallest eigenvalues are displayed in Figure 1. We observe that the eigenvalues cannot be decomposed in 2 smooth surfaces.

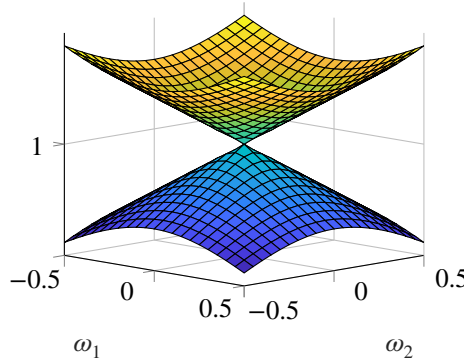


FIGURE 1 Surfaces that represent $\lambda_1(\boldsymbol{\omega})$ and $\lambda_2(\boldsymbol{\omega})$ in Example 1

Example 2. This example is taken from^{7, Example 4.4}. The parameterspace is $\Omega = [0.02, 0.5] \times [2, 8]$ and $\mathbf{A}(\boldsymbol{\omega}) = \sum_{i=1}^3 \theta_i(\boldsymbol{\omega}) \mathbf{A}_i$ with $\theta_i(\boldsymbol{\omega}), i = 1, 2, 3$ analytic functions and \mathbf{B} is constant. The dimension of the problem is 1311. The matrices originate from a finite element discretisation of a PDE. In Figure 2, we show the minimal eigenvalue as a function of the parameters. More details can be found in^{7, Example 4.4}.

It is immediately clear from the previous examples that an approximation method (including polynomial approximations) that relies on the smoothness of the surface, is not appropriate. We consider projection methods and we describe in the next section why they are well suited in this case.

We end this section by introducing the concept of reduced eigenvalue problems. Let $\mathcal{V} \subset \mathbb{R}^n$ be a subspace spanned by the columns of an orthonormal matrix $\mathbf{V} := [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m] \in \mathbb{R}^{n \times m}$ with m the dimension of the subspace. We call $(\mathbf{V}^T \mathbf{A}(\boldsymbol{\omega}) \mathbf{V}, \mathbf{V}^T \mathbf{B}(\boldsymbol{\omega}) \mathbf{V})$ the *reduced eigenvalue problem* on \mathcal{V} . We denote by $(\lambda_i^{\mathcal{V}}(\boldsymbol{\omega}), \mathbf{x}_i^{\mathcal{V}}(\boldsymbol{\omega})), i = 1, \dots, m$ an eigenpair of this reduced eigenvalue problem.

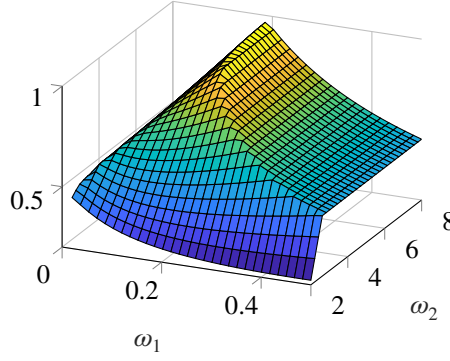


FIGURE 2 Surface of $\lambda_1(\omega)$ in Example 2

3 | HERMITE INTERPOLATION BY PROJECTION

In this section, we derive Hermite interpolation properties of the reduced eigenvalue problem when the subspace is built with the first eigenvector and its partial derivatives in some sample points. Concretely, we show that the eigenvalue itself is interpolated as well as its first and second partial derivatives. To prove this, we need a characterisation of the first and second derivative of an eigenvalue and of the first derivative of an eigenvector. **Since eigenvectors depend on the normalisation, we first outline the adopted setting, in which they are uniquely defined.** Let λ_i^* be a simple eigenvalue with associated eigenvector \mathbf{x}_i^* of $\omega^* \in \Omega$ then there is an environment Ω^* around ω^* and smooth functions

$$\begin{aligned} \lambda_i : \Omega^* &\rightarrow \mathbb{R} & \text{and} & & \mathbf{x}_i : \Omega^* &\rightarrow \mathbb{R}^n \\ \omega &\mapsto \lambda_i(\omega) & & & \omega &\mapsto \mathbf{x}_i(\omega) \end{aligned} \quad (4)$$

such that

$$\left\{ \begin{array}{ll} \mathbf{A}(\omega)\mathbf{x}_i(\omega) - \lambda_i(\omega)\mathbf{B}(\omega)\mathbf{x}_i(\omega) = 0, & \forall \omega \in \Omega^* \quad ((\lambda_i(\omega), \mathbf{x}_i(\omega)) \text{ is eigenpair}) \\ \mathbf{x}_i(\omega)^T \mathbf{B}(\omega) \mathbf{x}_i(\omega) = 1, & \forall \omega \in \Omega^* \quad (\text{normalisation condition}) \\ \lambda_i(\omega^*) = \lambda_i^* & \\ \mathbf{x}_i(\omega^*) = \mathbf{x}_i^* & \end{array} \right. \quad (5)$$

(reviewer 2: remark 2)

In the remainder of this section and the next section, we characterise the derivatives of the functions in (4). The derivative of a simple eigenvalue λ_i with associated eigenvector \mathbf{x}_i is

$$\frac{\partial \lambda_i}{\partial \omega_j} = \mathbf{x}_i^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_i, \quad j = 1, \dots, d. \quad (6)$$

This is a generalisation of the more known result for standard eigenvalue problems, see e.g.^{14, Ch. 9}. For the second derivative, it is sufficient to differentiate equation (6). In this way, we get the following expression for the second derivative

$$\frac{\partial^2 \lambda_i}{\partial \omega_j \partial \omega_k} = 2\mathbf{x}_i^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \frac{\partial \mathbf{x}_i}{\partial \omega_k} + \mathbf{x}_i^T \left(\frac{\partial^2 \mathbf{A}}{\partial \omega_j \partial \omega_k} - \frac{\partial \lambda_i}{\partial \omega_k} \frac{\partial \mathbf{B}}{\partial \omega_j} - \lambda_i \frac{\partial^2 \mathbf{B}}{\partial \omega_j \partial \omega_k} \right) \mathbf{x}_i, \quad j, k = 1, \dots, d.$$

We characterize the derivative of the eigenvector associated with a simple eigenvalue as the solution of a system of linear equations. In this proof we need following lemma.

Lemma 1. It holds that

$$\mathbf{x}_i^T \mathbf{B} \frac{\partial \mathbf{x}_i}{\partial \omega_j} = -\frac{1}{2} \mathbf{x}_i^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{x}_i, \quad j = 1, \dots, d. \quad (7)$$

Proof. This follows from taking the derivative of the normalisation condition in (5). □

Theorem 1. The derivative of the eigenvector \mathbf{x}_i associated with the simple eigenvalue λ_i is characterized as the solution of the following system:

$$\begin{bmatrix} \lambda_i \mathbf{B} - \mathbf{A} & \mathbf{B} \mathbf{x}_i \\ \mathbf{x}_i^T \mathbf{B} & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{x}_i}{\partial \omega_j} \\ \frac{\partial \lambda_i}{\partial \omega_j} \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_i \\ \mathbf{x}_i^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{x}_i \\ -\frac{1}{2} \end{bmatrix}, \quad j = 1, \dots, d. \quad (8)$$

Proof. We start by differentiating $(\lambda_i \mathbf{B} - \mathbf{A}) \mathbf{x}_i = 0$ to get

$$\left(\frac{\partial \lambda_i}{\partial \omega_j} \mathbf{B} + \lambda_i \frac{\partial \mathbf{B}}{\partial \omega_j} - \frac{\partial \mathbf{A}}{\partial \omega_j} \right) \mathbf{x}_i + (\lambda_i \mathbf{B} - \mathbf{A}) \frac{\partial \mathbf{x}_i}{\partial \omega_j} = 0, \quad j = 1, \dots, d.$$

As $(\lambda_i \mathbf{B} - \mathbf{A})$ is a singular matrix, the equation

$$(\lambda_i \mathbf{B} - \mathbf{A}) \frac{\partial \mathbf{x}_i}{\partial \omega_j} = - \left(\frac{\partial \lambda_i}{\partial \omega_j} \mathbf{B} + \lambda_i \frac{\partial \mathbf{B}}{\partial \omega_j} - \frac{\partial \mathbf{A}}{\partial \omega_j} \right) \mathbf{x}_i, \quad j = 1, \dots, d \quad (9)$$

is not sufficient to characterise the partial derivative of an eigenvector. The missing information is the information in the direction of \mathbf{x}_i which is given in (7). The combination of (9) and (7) leads to the proof of the theorem. \square

We prove in the next property that if both the eigenvector and its partial derivatives are present in the subspace \mathcal{V} for $\omega^* \in \Omega$, the minimal eigenvalue of the reduced eigenvalue problem on \mathcal{V} is a Hermite interpolant of degree two of the minimal eigenvalue of the pencil $(\mathbf{A}(\omega^*), \mathbf{B}(\omega^*))$. Furthermore we state also the well-known property that the minimal eigenvalue of a reduced eigenvalue problem is bounded from below by the original eigenvalue problem.

Property 1. 1. (Hermite interpolation) If $(\lambda_i, \mathbf{x}_i)$ is a simple eigenpair and if $\mathbf{x}_i \in \mathcal{V}$, then

- $(\lambda_i^\mathcal{V}, \mathbf{x}_i^\mathcal{V}) = (\lambda_i, \mathbf{V}^T \mathbf{x}_i)$ is an eigenpair of $(\mathbf{V}^T \mathbf{A} \mathbf{V}, \mathbf{V}^T \mathbf{B} \mathbf{V})$.
- $\frac{\partial \lambda_i^\mathcal{V}}{\partial \omega_j} = \frac{\partial \lambda_i}{\partial \omega_j}, \quad j = 1, \dots, d.$

If $\mathbf{x}_i, \frac{\partial \mathbf{x}_i}{\partial \omega_j} \in \mathcal{V}$, then

$$\frac{\partial^2 \lambda_i^\mathcal{V}}{\partial \omega_j \partial \omega_k} = \frac{\partial^2 \lambda_i}{\partial \omega_j \partial \omega_k}, \quad k = 1, \dots, d$$

2. If $\mathcal{V}_1 \subset \mathcal{V}_2$ then it holds that

$$\lambda_1 \leq \lambda_1^{\mathcal{V}_2} \leq \lambda_1^{\mathcal{V}_1}, \quad \forall \omega \in \Omega, \quad j = 1, \dots, d.$$

This means that if we extend the subspace, the estimation will be at least as good.

Proof. Assertion 1:

The first two statements are well-known results in the case of a standard eigenvalue problem, see e.g.¹² and the references therein. The proofs for the generalized eigenvalue problem are stated for the sake of completeness. For the first statement, it is sufficient to see that $\mathbf{x}_i = \mathbf{V} \mathbf{V}^T \mathbf{x}_i$ to obtain $\mathbf{V}^T \mathbf{A} \mathbf{V} \mathbf{V}^T \mathbf{x}_i = \lambda_i \mathbf{V}^T \mathbf{B} \mathbf{V} \mathbf{V}^T \mathbf{x}_i$.

For the second, the result follows from

$$\begin{aligned} \frac{\partial \lambda_i}{\partial \omega_j} &= \mathbf{x}_i^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_i \\ &= (\mathbf{V} \mathbf{x}_i^\mathcal{V})^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{V} \mathbf{x}_i^\mathcal{V} \\ &= (\mathbf{x}_i^\mathcal{V})^T \left(\mathbf{V}^T \frac{\partial \mathbf{A}}{\partial \omega_j} \mathbf{V} - \lambda_i \mathbf{V}^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{V} \right) \mathbf{x}_i^\mathcal{V} \\ &= \frac{\partial \lambda_i^\mathcal{V}}{\partial \omega_j}. \end{aligned}$$

If also $\frac{\partial \mathbf{x}_i}{\partial \omega_j} \in \mathcal{V}$, then there exists a vector \mathbf{z} such that $\frac{\partial \mathbf{x}_i}{\partial \omega_j} = \mathbf{V}\mathbf{z}$. We prove that $\mathbf{z} = \frac{\partial \mathbf{x}_i^{\mathcal{V}}}{\partial \omega_j}$, i.e. the projection of $\frac{\partial \mathbf{x}_i}{\partial \omega_j}$ on \mathcal{V} equals the derivative of the first eigenvector of the reduced eigenvalue problem. This derivative is uniquely determined by the system in (8), from which it follows that

$$\begin{cases} (\lambda_i \mathbf{B} - \mathbf{A}) \mathbf{V}\mathbf{z} + \frac{\partial \lambda_i^{\mathcal{V}}}{\partial \omega_j} \mathbf{B}\mathbf{V}\mathbf{x}_i^{\mathcal{V}} = \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i^{\mathcal{V}} \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{V}\mathbf{x}_i^{\mathcal{V}} \\ (\mathbf{x}_i^{\mathcal{V}})^T \mathbf{V}^T \mathbf{B}\mathbf{z} = -\frac{\left(\mathbf{x}_i^{\mathcal{V}}, \mathbf{V}^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{V}\mathbf{x}_i^{\mathcal{V}} \right)}{2} \end{cases}$$

which means that $[\mathbf{z}, \frac{\partial \lambda_i^{\mathcal{V}}}{\partial \omega_j}]^T$ is the vector such that

$$\begin{cases} \mathbf{V}^T (\lambda_i \mathbf{B} - \mathbf{A}) \mathbf{V}\mathbf{z} + \frac{\partial \lambda_i^{\mathcal{V}}}{\partial \omega_j} \mathbf{V}^T \mathbf{B}\mathbf{V}\mathbf{x}_i^{\mathcal{V}} = \mathbf{V}^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i^{\mathcal{V}} \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{V}\mathbf{x}_i^{\mathcal{V}} \\ (\mathbf{x}_i^{\mathcal{V}})^T \mathbf{V}^T \mathbf{B}\mathbf{V}\mathbf{z} = -\frac{\left(\mathbf{x}_i^{\mathcal{V}}, \mathbf{V}^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{V}\mathbf{x}_i^{\mathcal{V}} \right)}{2} \end{cases}.$$

This is the system that uniquely determines the derivative of the first eigenvector of the reduced eigenvalue problem, so $\mathbf{z} = \frac{\partial \mathbf{x}_i^{\mathcal{V}}}{\partial \omega_j}$. From equation (7), we obtain

$$\begin{aligned} \frac{\partial^2 \lambda_i}{\partial \omega_j \partial \omega_k} &= 2 (\mathbf{x}_i^{\mathcal{V}})^T \mathbf{V}^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i^{\mathcal{V}} \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{V} \frac{\partial \mathbf{x}_i^{\mathcal{V}}}{\partial \omega_k} + \dots \\ &\quad (\mathbf{x}_i^{\mathcal{V}})^T \mathbf{V}^T \left(\frac{\partial^2 \mathbf{A}}{\partial \omega_j \partial \omega_k} - \frac{\partial \lambda_i^{\mathcal{V}}}{\partial \omega_k} \frac{\partial \mathbf{B}}{\partial \omega_j} - \lambda_i^{\mathcal{V}} \frac{\partial^2 \mathbf{B}}{\partial \omega_j \partial \omega_k} \right) \mathbf{V}\mathbf{x}_i^{\mathcal{V}}. \\ &= 2 (\mathbf{x}_i^{\mathcal{V}})^T \left(\mathbf{V}^T \frac{\partial \mathbf{A}}{\partial \omega_j} \mathbf{V} - \lambda_i^{\mathcal{V}} \mathbf{V}^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{V} \right) \frac{\partial \mathbf{x}_i^{\mathcal{V}}}{\partial \omega_k} + \dots \\ &\quad (\mathbf{x}_i^{\mathcal{V}})^T \left(\mathbf{V}^T \frac{\partial^2 \mathbf{A}}{\partial \omega_j \partial \omega_k} \mathbf{V} - \frac{\partial \lambda_i^{\mathcal{V}}}{\partial \omega_k} \mathbf{V}^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{V} - \lambda_i^{\mathcal{V}} \mathbf{V}^T \frac{\partial^2 \mathbf{B}}{\partial \omega_j \partial \omega_k} \mathbf{V} \right) \mathbf{x}_i^{\mathcal{V}} \\ &= \frac{\partial^2 \lambda_i^{\mathcal{V}}}{\partial \omega_j \partial \omega_k}. \end{aligned}$$

Assertion 2: This results is well known, we refer to¹³ for a proof. □

4 | CHARACTERISATION OF THE DERIVATIVE OF AN EIGENVECTOR

The previous section showed that adding the derivative of the eigenvector to the subspace leads to higher-order Hermite interpolation in the eigenvalue. **The aim of this section is to get more insight in the behaviour of eigenvalues as a function of the parameters by deriving an analytic formula for the derivative of the eigenvector. The next result can also be found in¹⁵, but we give here our proof, which is based on the diagonalisation of $\lambda_1 \mathbf{B} - \mathbf{A}$ in (8), to show the connection with Theorem 1, (reviewer 1: remark 2).**

Theorem 2. If \mathbf{x}_i is an eigenvector associated with a simple eigenvalue λ_i (reviewer 2: remark 3), then we have

$$\frac{\partial \mathbf{x}_i}{\partial \omega_j} = -\frac{\mathbf{x}_i^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{x}_i}{2} \mathbf{x}_i + \sum_{k=1, k \neq i}^n \left(\frac{\left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_i \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_i}{\lambda_i - \lambda_k} \right) \mathbf{x}_k, \quad j = 1, \dots, d. \quad (10)$$

Proof. Without loss of generality, we prove the statement for $i = 1$. From (3) it follows immediately that system (8) can be written as

$$\left[\begin{array}{c|c} \lambda_1 (\mathbf{X}^{-1})^T \mathbf{X}^{-1} - (\mathbf{X}^{-1})^T \mathbf{A} \mathbf{X}^{-1} & (\mathbf{X}^{-1})^T \mathbf{e}_1 \\ \hline \mathbf{e}_1^T \mathbf{X}^{-1} & 0 \end{array} \right] \left[\begin{array}{c} \frac{\partial \mathbf{x}_1}{\partial \omega_j} \\ \frac{\partial \lambda_1}{\partial \omega_j} \end{array} \right] = \left[\begin{array}{c} \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_1 \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_1 \\ -\frac{\mathbf{x}_1^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{x}_1}{2} \end{array} \right].$$

Using $\mathbf{X}\mathbf{X}^T\mathbf{B} = \mathbf{I}$, we obtain

$$\left[\begin{array}{c|c} \lambda_1 (\mathbf{X}^{-1})^T - (\mathbf{X}^{-1})^T \mathbf{A} & (\mathbf{X}^{-1})^T \mathbf{e}_1 \\ \hline \mathbf{e}_1^T & 0 \end{array} \right] \left[\begin{array}{c} \mathbf{X}^T \mathbf{B} \frac{\partial \mathbf{x}_1}{\partial \omega_j} \\ \frac{\partial \lambda_1}{\partial \omega_j} \end{array} \right] = \left[\begin{array}{c} \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_1 \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_1 \\ -\frac{\mathbf{x}_1^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{x}_1}{2} \end{array} \right].$$

By multiplying from the left with the non-singular matrix $\begin{bmatrix} \mathbf{X}^T & \mathbf{0} \\ 0 & 1 \end{bmatrix}$ we get

$$\left[\begin{array}{c|c} \lambda_1 \mathbf{I} - \mathbf{A} & \mathbf{e}_1 \\ \hline \mathbf{e}_1^T & 0 \end{array} \right] \left[\begin{array}{c} \mathbf{X}^T \mathbf{B} \frac{\partial \mathbf{x}_1}{\partial \omega_j} \\ \frac{\partial \lambda_1}{\partial \omega_j} \end{array} \right] = \left[\begin{array}{c} \mathbf{X}^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_1 \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_1 \\ -\frac{\mathbf{x}_1^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{x}_1}{2} \end{array} \right].$$

From this expression it follows that

$$\frac{\partial \lambda_1}{\partial \omega_j} = \mathbf{x}_1^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_1 \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_1$$

$$\mathbf{x}_k^T \mathbf{B} \frac{\partial \mathbf{x}_1}{\partial \omega_j} = \begin{cases} -\frac{\mathbf{x}_k^T \frac{\partial \mathbf{B}}{\partial \omega_j} \mathbf{x}_1}{2}, & k = 1 \\ \frac{\mathbf{x}_k^T \left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_1 \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_1}{\lambda_1 - \lambda_k}, & k \neq 1. \end{cases}$$

Finally, multiplying with \mathbf{X} from the left proves the statement. \square

We observe that the weight of $\mathbf{x}_k, k = 2, \dots, n$ in the expression for $\frac{\partial \mathbf{x}_1}{\partial \omega_j}$ depends on how close λ_k is to λ_1 and depends on the projection of $\left(\frac{\partial \mathbf{A}}{\partial \omega_j} - \lambda_1 \frac{\partial \mathbf{B}}{\partial \omega_j} \right) \mathbf{x}_1, j = 1, \dots, d$ on \mathbf{x}_k . In general, the closer λ_k is to λ_1 the higher is the impact of \mathbf{x}_k . We also deduce that if \mathbf{B} does not depend on the variables then the partial derivative is \mathbf{B} -orthonormal to \mathbf{x}_1 . Example 3 shows that the norm of the partial derivative of the eigenvector cannot be bounded uniformly on the set Ω , despite \mathbf{A} and \mathbf{B} being analytic and \mathbf{B} positive definite for $\boldsymbol{\omega} \in \Omega$.

Example 3. We reconsider example 1 with $n = 2$ and $\mathbf{W} = \mathbf{I}_2$. In Figure 3, we plot the eigenvalues $\lambda_1(\boldsymbol{\omega})$ and $\lambda_2(\boldsymbol{\omega})$ in (a) and (b) and the norm of $\frac{\partial \mathbf{x}_1}{\partial \omega_2}$ in (c) and (d) as a function of ω_2 , for fixed values of ω_1 . One can observe that for $\omega_2 = 0$, the norm of the partial derivative to ω_2 becomes larger if ω_1 goes to 0. It can be shown by Theorem 2 that this norm is unbounded. For $\omega_2 = 0$, the eigenvectors are equal to

$$\mathbf{x}_1(\omega_1, 0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{ and } \mathbf{x}_2(\omega_1, 0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

from which we can deduce that

$$\begin{aligned} \frac{\partial \mathbf{x}_1(\boldsymbol{\omega})}{\partial \omega_2} &= \mathbf{x}_2(\omega_1, 0) \frac{\left(\frac{\partial \mathbf{A}(\omega_1, 0)}{\partial \omega_2} \mathbf{x}_1(\omega_1, 0), \mathbf{x}_2(\omega_1, 0) \right)}{\lambda_1(\omega_1, 0) - \lambda_2(\omega_1, 0)} \\ &= \mathbf{x}_2(\omega_1, 0) \frac{\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)}{\lambda_1(\omega_1, 0) - \lambda_2(\omega_1, 0)} \\ &= \mathbf{x}_2(\omega_1, 0) \frac{1}{-2\omega_1}. \end{aligned}$$

We conclude that if ω_1 goes to 0, the norm of the partial derivative goes to infinity.

Although the analytic formula for the partial derivative gives much insight how eigenvectors depend on the other eigenvectors and eigenvalues, it is less suitable to compute derivatives. The reason is that we need all eigenvectors and eigenvalues and that in practice it is far too costly to calculate these. In practice we calculate the partial derivatives by solving system (8). **Note that**

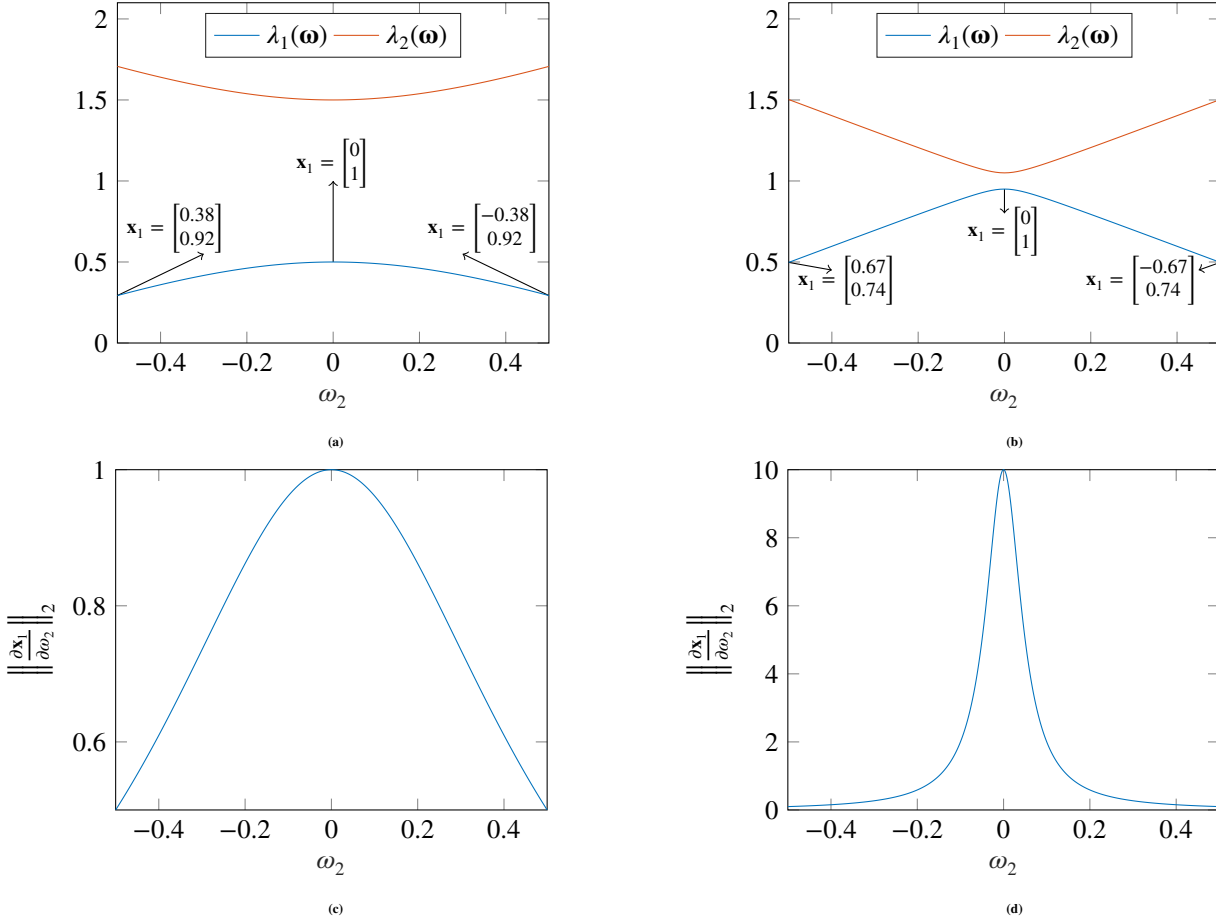


FIGURE 3 (a) and (b): $\lambda_1(\omega)$ and $\lambda_2(\omega)$ with $\omega_1 = 0.5$ (left) and $\omega_1 = 0.05$ (right), (c) and (d): norm of $\frac{\partial \mathbf{x}_1(\omega)}{\partial \omega_2}$ with $\omega_1 = 0.5$ (left) and $\omega_1 = 0.05$ (right)

the cost of solving (8) for all parameters is not negligible, since it depends on the dimension and the sparsity pattern of the matrices. Therefore we will propose an alternative construction of the subspace that adds approximations to the derivatives in the next section. (**reviewer 1: remark 3**)

5 | COMPUTING A GLOBAL APPROXIMATION

5.1 | Subspace expansion

Property 1 presents the key theory for the construction of the reduced problem: adding eigenvectors and partial derivatives leads to Hermite interpolation of degree two for the eigenvalue. The eigenvalue and associated eigenvector are computed by a subspace projection method. The computation of the derivatives of the eigenvector from (8) is usually computationally expensive. Therefore we look for cheaper alternatives. Eq. (10) suggests that the derivatives mainly depend on the eigenvectors associated with eigenvalues closest to λ_1 . In particular, this is the case when the spectrum is well-separated with only a few eigenvalues near λ_1 . This implies that adding a few eigenvectors to the subspace in an interpolation point may serve as a computationally cheap surrogate for adding one eigenvector and its derivatives. (**reviewer 1: remark 3**) Approximations to the associated eigenvectors of the eigenvalues nearest λ_1 are usually available from subspace methods. Indeed, when we use the Lanczos' method, the subspace produces reasonably good approximations to these eigenvectors, at no additional cost. Moreover, it should be noted that there is another reason why adding the second and third eigenvector may be beneficial: the eigenvector of the 2nd minimal eigenvalue may be a good approximation to the eigenvector associated with the minimal eigenvalue at another parameter value $\omega_1 \in \Omega$. This means we probably do not need to solve a costly eigenvalue problem in ω_1 . Example 4 gives a good illustration

of such a situation. In addition, we see in the next section that adding the second eigenvector helps making a sharp upper bound for the error.

Example 4. We reconsider example 1 where we take n arbitrarily large. If $\omega_2 = 0$, then one can verify that

$$\mathbf{x}_1(\omega_1, 0) = \begin{cases} \mathbf{w}_1(\omega_1, 0) & , \omega_1 \leq 0 \\ \mathbf{w}_2(\omega_1, 0) & , \omega_1 > 0 \end{cases}, \quad \mathbf{x}_2(\omega_1, 0) = \begin{cases} \mathbf{w}_2(\omega_1, 0) & , \omega_1 \leq 0 \\ \mathbf{w}_1(\omega_1, 0) & , \omega_1 > 0 \end{cases}.$$

Take $\omega^1 < 0$ and $\omega^2 > 0$ two points in the first dimension of Ω . If both \mathbf{w}_1 and \mathbf{w}_2 do not strongly depend on ω , this means that $\mathbf{x}_1(\omega^1, 0) \approx \mathbf{x}_2(\omega^2, 0)$. Therefore we conclude that $\mathcal{V}_1 := \text{span}\{\mathbf{x}_1(\omega^i, 0), i = 1, 2\} \approx \text{span}\{\mathbf{x}_i(\omega^1, 0), i = 1, 2\} =: \mathcal{V}_2$, but \mathcal{V}_2 is less costly to calculate than \mathcal{V}_1 . This is illustrated in Figure 4.

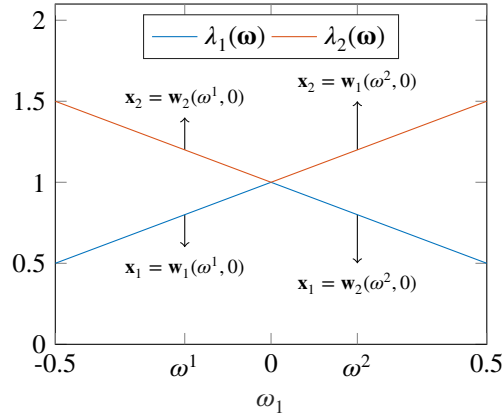


FIGURE 4 Illustration corresponding to Example 4. If both \mathbf{w}_1 and \mathbf{w}_2 do not depend much on ω_1 , then $\mathcal{V}_1 := \text{span}\{\mathbf{x}_1(\omega^i), i = 1, 2\} \approx \text{span}\{\mathbf{x}_i(\omega^1), i = 1, 2\} =: \mathcal{V}_2$ but \mathcal{V}_2 is less costly to calculate than \mathcal{V}_1 .

5.2 | Upper bound of the error

Before we can state the algorithm, we need a measure to qualify if the subspace is rich enough or not. The subspace is rich enough if for every point in the parameterspace the difference between the minimal eigenvalue of the reduced problem and the minimal eigenvalue of the original large problem is smaller than a given tolerance. This means we need a way to measure the error for any given $\omega \in \Omega$. An upper bound can be obtained only using the first eigenpair via an extension of the Bauer-Fike theorem (see¹⁶ and more recently¹) to generalized eigenvalue problems. For this, we define the residual of an approximate eigenpair $(\tilde{\lambda}, \tilde{\mathbf{x}})$ as

$$\mathbf{r} = \mathbf{A}\tilde{\mathbf{x}} - \tilde{\lambda}\mathbf{B}\tilde{\mathbf{x}}.$$

The proof of this extension is omitted as it is completely analogous to the original proof for the standard case.

Theorem 3. (Bauer-Fike for generalized eigenvalue problems) Let $(\tilde{\lambda}, \tilde{\mathbf{x}})$ be an approximate eigenpair of (\mathbf{A}, \mathbf{B}) where $\tilde{\mathbf{x}}$ is of \mathbf{B} -norm unity and $\tilde{\lambda} = (\mathbf{A}\tilde{\mathbf{x}}, \tilde{\mathbf{x}})$. Then, there exists an eigenvalue λ of (\mathbf{A}, \mathbf{B}) such that

$$|\tilde{\lambda} - \lambda| \leq \frac{\|\mathbf{r}\|_2}{\sqrt{\lambda_1(\mathbf{B})}}. \quad (11)$$

The residual norm is very cheap to calculate for a set of parameter samples if the dependency of \mathbf{A} and \mathbf{B} is affine. More precisely, if $\mathbf{A}(\omega)$ and $\mathbf{B}(\omega)$ can be written as $\mathbf{A}(\omega) = \sum_{i=1}^{m_0} \theta_{A,i}(\omega) \mathbf{A}_i$ resp. $\mathbf{B}(\omega) = \sum_{i=1}^{m_1} \theta_{B,i}(\omega) \mathbf{B}_i$ with $\theta_{A,i}, \theta_{B,j}, i = 1, \dots, m_0, j = 1, \dots, m_1$ scalar functions then the residual is in this case

$$\begin{aligned} \mathbf{r}(\omega) &= \mathbf{A}(\omega)\tilde{\mathbf{x}}(\omega) - \tilde{\lambda}(\omega)\mathbf{B}(\omega)\tilde{\mathbf{x}}(\omega) \\ &= \sum_{i=1}^{m_0} \theta_{A,i}(\omega) (\mathbf{A}_i \tilde{\mathbf{v}}) \mathbf{x}^\nu(\omega) - \tilde{\lambda}(\omega) \sum_{i=1}^{m_1} \theta_{B,i}(\omega) (\mathbf{B}_i \tilde{\mathbf{v}}) \mathbf{x}^\nu(\omega). \end{aligned}$$

This means that the precomputation of $\mathbf{A}_i, i = 1, \dots, m_0$ and $\mathbf{B}_i, i = 1, \dots, m_1$ reduces the cost of calculating the residual norm. The drawback of (11) is that this is usually a very crude bound. If we have also information about the second smallest eigenvalue, we can use the Kato-Temple theorem^{17, 18} for a sharper bound. We rephrase the theorem for the generalised eigenvalue problem.

Theorem 4. (Kato-Temple for generalized eigenvalue problems) Let $(\tilde{\lambda}, \tilde{\mathbf{x}})$ be an approximate eigenpair of (\mathbf{A}, \mathbf{B}) where $\tilde{\mathbf{x}}$ is of \mathbf{B} -norm unity and $\tilde{\lambda} = (\mathbf{A}\tilde{\mathbf{x}}, \tilde{\mathbf{x}})$. Assume that we know an interval $]\alpha, \beta[$ that contains $\tilde{\lambda}$ and one eigenvalue λ of (\mathbf{A}, \mathbf{B}) . Then it holds

$$-\frac{\|\mathbf{r}\|_2^2}{\lambda_1(\mathbf{B})(\tilde{\lambda} - \alpha)} \leq \tilde{\lambda} - \lambda \leq \frac{\|\mathbf{r}\|_2^2}{\lambda_1(\mathbf{B})(\beta - \tilde{\lambda})}.$$

If we now define δ as the distance to the second smallest eigenvalue, then we get

$$|\tilde{\lambda} - \lambda_1| \leq \frac{\|\mathbf{r}\|_2^2}{\lambda_1(\mathbf{B})\delta}, \quad (12)$$

which is in general a sharper upper bound than the one from Theorem 3.

In the Bauer-Fike as well as in the Kato-Temple theorem the bound depends on the first eigenvalue of \mathbf{B} . As it is expensive to calculate this eigenvalue for every ω , we only calculate it for one parameter value and use it for the whole domain. As already mentioned, we approximate the second, respectively third eigenvalue by the second, respectively third Ritz value from the Krylov subspace. **Whether we use the Bauer-Fike theorem or the Temple-Kato theorem depends on how accurately we can determine the gap δ between the first two eigenvalues. If we add the first m eigenvectors ($m > 1$) to the subspace, we interpolate the first m eigenvalues over the parameterspace, hence the gap δ can be estimated well so we make in this case use of the Kato-Temple theorem. If only one eigenvector is added, an accurate approximation of the gap cannot be guaranteed, in which case the Bauer-Fike theorem is more suitable. (reviewer 2: remark 4)**

5.3 | Algorithm

We now have all ingredients for an algorithm to compute the reduced problem. A remaining crucial element is the choice of sample points, which we now discuss. The proposed method is inspired by the reduced basis method for solving parametric PDEs, see⁸.

We assume that the parameter space Ω is a Cartesian product of intervals in \mathbb{R} , so

$$\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_d = [a^1, b^1] \times [a^2, b^2] \times \dots \times [a^d, b^d].$$

We first make an initial subspace \mathcal{V} constructed with the first eigenvector(s) and the partial derivatives for all sample points in an initial set Ω_{init} (see line 1-4 in Algorithm 1) and we check then if this constructed subspace is large enough by testing the upper bound on some training set Ω_{train} . Whether we use the Kato-Temple theorem 4 or the Bauer-Fike theorem 3 depends on the number of eigenvectors we add per interpolation point, see previous section. If the upper bound at all points in the training set Ω_{train} is not below a certain tolerance, we add the first eigenvectors and partial derivatives of the point in the training set where the upper bound is maximal (see line 13-22 in Algorithm 1). Once the upper bound is below the tolerance, these sample points are removed from the training set (see line 7-10 in Algorithm 1). We are allowed to do this because this implies that the error on the eigenvalue is below the tolerance and furthermore the error cannot increase by the second assertion in Theorem 1. The initial set of sample points is decomposed as

$$\Omega_{\text{init}} := \Omega_1^{\text{init}} \times \Omega_2^{\text{init}} \times \dots \times \Omega_d^{\text{init}} \quad (13)$$

where

$$\Omega_i^{\text{init}} = \{a_i + (j-1)h_i | j = 1, 2, \dots, n_i\}, h_i = \frac{b^i - a^i}{n_i - 1}, i = 1, 2, \dots, d. \quad (14)$$

This is an initial set of $n_{\text{init}} = n_1 n_2 \dots n_d$ points. The training set Ω_{train} is composed in a similar way but using a much finer grid. The algorithm is stated in pseudocode in Algorithm 1. As we do experiments with adding partial derivatives to the subspace and with not adding them to the subspace, we state 'optionally' after the lines where we calculate the partial derivatives.

5.4 | Saturation assumption

Although the training set becomes smaller in every iteration, calculating the upper bound for all sample points in the training set can become expensive if many iterations are needed. Instead of updating all upper bounds, we select only those that potentially

Algorithm 1 Aim: Calculating a subspace \mathcal{V} such that $|\lambda_1^{\mathcal{V}}(\omega) - \lambda_1(\omega)| < \text{tol}, \forall \omega \in \Omega$

Input:

1. Matrix $\mathbf{A}(\omega)$ and $\mathbf{B}(\omega)$
2. All partial derivatives $\frac{\partial \mathbf{A}(\omega)}{\partial \omega_i}$ and $\frac{\partial \mathbf{B}(\omega)}{\partial \omega_i}, i = 1, 2, \dots, d$
3. Choose an initial set Ω_{init} and a training set Ω_{train} of sample points.
4. n_{max} maximal number of iterations

Output: $\lambda_1^{\mathcal{V}}(\omega)$ such that $\max_{\omega \in \Omega_{\text{train}}} |\lambda_1^{\mathcal{V}}(\omega) - \lambda_1(\omega)| < \text{tol}$

- 1: Calculate $(\lambda_1(\omega), \mathbf{x}_1(\omega)), \omega \in \Omega_{\text{init}}$
 - 2: Estimate $(\tilde{\lambda}_j(\omega), \tilde{\mathbf{x}}_j(\omega)), j = 2, 3, \dots, m, \omega \in \Omega_{\text{init}}$
 - 3: Calculate partial derivatives $\frac{\partial \mathbf{x}_1(\omega)}{\partial \omega_j}, j = 1, 2, \dots, d, \omega \in \Omega_{\text{init}}$ from system (8) if $|\lambda_1 - \tilde{\lambda}_2| > 10^{-8}$ (If λ_1 is simple) (Optionally)
 - 4: $\mathcal{V} := \text{span}\{\mathbf{x}_1(\omega), \tilde{\mathbf{x}}_2(\omega) \dots, \tilde{\mathbf{x}}_m(\omega), \frac{\partial \mathbf{x}_1(\omega)}{\partial \omega_1}, \frac{\partial \mathbf{x}_1(\omega)}{\partial \omega_2}, \dots, \frac{\partial \mathbf{x}_1(\omega)}{\partial \omega_d} | \omega \in \Omega_{\text{init}}\}$
 - 5: $u_{\text{old}}(\omega) = 1, \omega \in \Omega_{\text{train}}$
 - 6: **for** $i = 1, 2, \dots, n_{\text{max}}$ **do**
 - 7: Update the training set
 - 8: Calculate new upper bound $u_{\text{new}}(\omega)$ for all $\omega \in \Omega_{\text{train}}$ using Theorem 3 if $m = 1$ or Theorem 4 if $m > 1$
 - 9: Upper bound $u(\omega) = \min(u_{\text{old}}(\omega), u_{\text{new}}(\omega))$
 - 10: $\Omega_{\text{valid}} = \{\omega | \omega \in \Omega_{\text{train}} \text{ and } u(\omega) < \text{tol}\}$
 - 11: $\Omega_{\text{train}} = \Omega_{\text{train}} \setminus \Omega_{\text{valid}}$
 - 12: **if** Ω_{train} empty **then**
 - 13: break
 - 14: **else**
 - 15: Adding vectors to the subspace
 - 16: $\omega^i = \arg \max_{\omega \in \Omega_{\text{train}}} u(\omega)$
 - 17: Calculate $(\lambda_1(\omega^i), \mathbf{x}_1(\omega^i))$
 - 18: Estimate $(\tilde{\lambda}_j(\omega^i), \tilde{\mathbf{x}}_j(\omega^i)), j = 2, 3, \dots, m$
 - 19: **if** $|\lambda_1 - \tilde{\lambda}_2| > 10^{-8}$ **then** (Check if λ_1 is simple)
 - 20: Calculate partial derivatives $\frac{\partial \mathbf{x}_1(\omega^i)}{\partial \omega_j}, j = 1, 2, \dots, d$ from system (8) (optionally)
 - 21:
 - 22: $\mathcal{V} = \mathcal{V} \cup \{\mathbf{x}_1(\omega^i), \tilde{\mathbf{x}}_2(\omega^i) \dots, \tilde{\mathbf{x}}_m(\omega^i), \} \left(\bigcup \left\{ \frac{\partial \mathbf{x}_1(\omega^i)}{\partial \omega_1}, \frac{\partial \mathbf{x}_1(\omega^i)}{\partial \omega_2}, \dots, \frac{\partial \mathbf{x}_1(\omega^i)}{\partial \omega_d} \right\} \text{(optionally)} \right)$
 - 23: **else**
 - 24: $\mathcal{V} = \mathcal{V} \cup \{\mathbf{x}_1(\omega^i), \tilde{\mathbf{x}}_2(\omega^i) \dots, \tilde{\mathbf{x}}_m(\omega^i)\}$
 - 25: **end if**
 - 26: Make basis for \mathcal{V}
 - 27: $\Omega_{\text{train}} = \Omega_{\text{train}} \setminus \omega^i$
 - 28: **end if**
 - 29: **end for**
-

generate the highest upper bound for the error. We therefore make use of the so-called saturation assumption, see¹⁹. Let u^k be the upper bound at iteration k . The saturation assumption says that there exists a $C > 0$ such that

$$u^l(\omega) < Cu^k(\omega), \forall l > k, \forall \omega \in \Omega. \quad (15)$$

In our method the saturation assumption is fulfilled for $C = 1$ as we only update the upper bound u^k if the upper bound has decreased. We use this in the following way. We first sort the parameter values in Ω_{train} by the upper bound of the previous iteration in descending order. We recompute the upper bound and update after each computation the maximal upper bound u_{max} . If for a certain ω , it holds that $u^k(\omega) < u_{\text{max}}$ we are allowed to skip all the next sample points in the sequence by (15). The next ω for which we add the eigenvectors and all partial derivatives, is the ω which has the upper bound u_{max} .

6 | IMPLEMENTATION DETAILS AND NUMERICAL RESULTS

All algorithms written in the previous sections are implemented in Matlab version R2017a. All experiments are performed on an Intel i5-6300U with 2.5 GHz and 8 GB RAM. In all examples, the linear systems in building the Krylov space and system of the form (8) to compute the derivative of an eigenvector could still be solved efficiently using a direct method. Practically, because the used matrices are sparse, it is efficient to use the backslash-operator in Matlab for solving the systems of the form (8). We use the `eigs`-command in Matlab to calculate the minimal eigenvalue which uses the shift-and-invert Arnoldi's method with shift 0. This method makes a Krylov subspace and we adapt the code such that it also returns this space. We use the other vectors in the Krylov subspace to approximate the second eigenvector. In all examples a tolerance for the error of 10^{-5} is used. We applied our algorithm to examples where the smallest eigenvalue corresponds to calculating the coercivity constant when solving PDEs⁷ and to one example from structural mechanics. We compare the results from four cases: adding only the first eigenvector, adding the first two eigenvectors, adding the first eigenvector and all partial derivatives and finally the first two eigenvectors and all partial derivatives. In the next tables, we put the dimension of the found subspace, the number of sample points, the total computing time and the average time we needed to compute the partial derivatives and eigenvectors.

For Example 2, the initial set is built by discretising both dimensions of the parameterspace into 4 to get an initial subset of 16 points. For the training set, we discretise the first respectively the second dimension of the parameterspace into 25 respectively 40 points, to have a total of 1000 points in our training set. The results can be found in Table 1. We see that adding other information than only the first eigenvector is beneficial. The number of points where we need to calculate the second eigenvector and the partial derivative is much lower which results in a smaller computational time. The main benefit from adding the second eigenvector is that we can use the more strict upper bound of Kato-Temple. We also noticed that the time needed to compute the derivatives is much lower than calculating the eigenvalues. We see we get the best results when we add both the eigenvectors and the partial derivatives into the subspace.

TABLE 1 The results for Example 2

	1 eigv	2 eigv	1 eigv + partial deriv.	2 eigv + partial deriv.
dimension \mathcal{V}	96	108	138	132
nbr points	96	54	46	33
total time	59.040s	38.923s	49.997s	29.293s
time derivative (per vector)	/	/	0.002s	0.002s
time eigenv (per vector)	0.080s	0.042s	0.079s	0.044s

The next example is also taken from^{7, Example 4.3}. The parameterspace is $\Omega = [-0.1, 0.1] \times [0.2, 0.3]$ and the matrix \mathbf{A} is of the form $\mathbf{A}(\omega) = \sum_{i=1}^{16} \theta_i(\omega) \mathbf{A}_i$ with $\theta_i(\omega)$ analytic functions and matrix \mathbf{B} is constant. The dimension of the problem is 2183. As training set we choose a grid where we discretise the interval for ω_1 into 40 and ω_2 into 25 points. As initial set we discretise both intervals into 3. The results are stated in Table 2. We see that in this case the needed subspace is low-dimensional which results in a small computational time. Similar conclusions can be drawn for this example.

TABLE 2 Results

	1 eigv	2 eigv	1 eigv + partial deriv.	2 eigv + partial deriv.
dimension \mathcal{V}	37	30	48	40
nbr points	37	15	16	10
total time	32.038s	14.617s	24.372s	8.499s
time derivative (per vector)	/	/	0.005s	0.004s
time eigenv (per vector)	0.034s	0.015s	0.026s	0.012s

In the previous examples matrix \mathbf{B} was constant, since the article from which we took the examples considered only eigenvalue problems with constant \mathbf{B} . The next example originates from an application in structural mechanics and here matrix \mathbf{B} depends on ω .

The matrices in this example are the stiffness and mass matrices arising from a finite element approximation of a vibrating beam that can deflect in both directions perpendicular to its own axis. The first parameter concerns the size of the beam along one of the axis (in interval $[0.1, 1]$) and the second parameter deals with the stiffness and density of the material (in interval $[100, 1000]$). For the initial set, we discretise both dimensions into 3. The matrices have size 1404 and they both depend on at least one of the parameters. In this problem, the stiffness and mass matrix both dependent on the two parameters. One is usually interested in the minimal eigenfrequencies which boils down to calculating the minimal eigenvalue. In Figure 5 we depict the minimal eigenvalue over the parameterspace and Table 3 summarizes the results. We observe that only a small subspace was needed and that we benefit from including partial derivatives in the subspace since the computation of the eigenvectors takes much more time than solving the system to compute the partial derivatives.

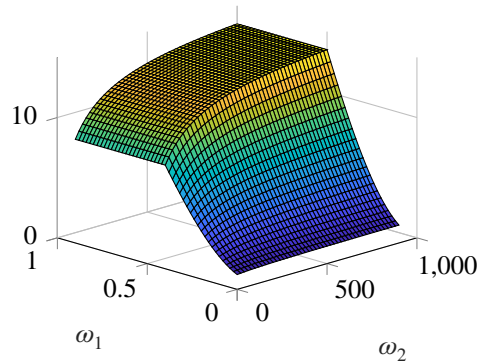


FIGURE 5 The minimal eigenvalue $\lambda_1(\omega)$ over the parameter for the example from structural mechanics. We see the minimal eigenvalue does not smoothly depend on the variable.

TABLE 3 Results for example from structural mechanics

	1 eigv	2 eigv	1 eigv + partial deriv.	2 eigv + partial deriv.
dimension \mathcal{V}	17	22	32	31
nbr points	17	11	12	9
total time	193.727s	105.806s	123.125s	83.531s
time derivative (per vector)	/	/	0.005s	0.005s
time eigenv (per vector)	1.743s	0.871s	1.787s	0.890s

7 | CONCLUSION

The main contribution of this paper is that we showed that for calculating extreme eigenvalues of eigenpairs (\mathbf{A}, \mathbf{B}) where both matrices are symmetric and \mathbf{B} is positive definite, it is beneficial to add, besides the first eigenvector, also its partial derivatives, or alternatively, the Ritz vector associated with the second and third Ritz values. We have proved that in the case of the addition of partial derivatives, the eigenvalues satisfy a Hermite interpolation property of order 2. Numerical examples confirm this statement and showed that this property can also be used in practice.

The question arises what we do if the system of linear equations for calculating the partial derivatives cannot be efficiently solved by direct solvers. A possibility is to first estimate the partial derivative from the Krylov subspace \mathcal{V}_K obtained when calculating the first eigenvector as this space contains approximations of the vectors from which the Ritz-values are the closest to the minimal eigenvalue. Let \mathbf{V}_K be a basis of this subspace then we first project the system in (8) on \mathcal{V}_K to obtain

$$\begin{bmatrix} \mathbf{V}_K^T (\lambda_1 \mathbf{B} - \mathbf{A}) \mathbf{V}_K & \mathbf{V}_K^T \mathbf{B} \mathbf{x}_1 \\ \mathbf{x}_1^T \mathbf{B} \mathbf{V}_K & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial \mathbf{x}_1}{\partial \omega_i} \mathbf{v}_K \\ \frac{\partial \lambda_1}{\partial \omega_i} \mathbf{v}_K \end{bmatrix} = \begin{bmatrix} \mathbf{V}_K^T \left(\frac{\partial \mathbf{A}}{\partial \omega_i} - \lambda_1 \frac{\partial \mathbf{B}}{\partial \omega_i} \right) \mathbf{x}_1 \\ \mathbf{x}_1^T \frac{\partial \mathbf{B}}{\partial \omega_i} \mathbf{x}_1 - \frac{\lambda_1}{2} \end{bmatrix} \quad (16)$$

which has the same dimension as \mathcal{V}_K . We can use this approximation as a starting vector in an iterative method like GMRES to compute the derivative.

ACKNOWLEDGEMENTS

The authors thank the referees for their helpful remarks. This work was supported by the project KU Leuven Research Council grant OT/14/074 and C14/17/072 and by the project G0A5317N of the Research Foundation-Flanders (FWO - Vlaanderen).

References

1. Saad Y. Numerical Methods for Large Eigenvalue Problems. Numerical Methods for Large Eigenvalue Problems. 2011;.
2. Lanczos C. An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators. Journal of Research of the National of Standards. 1950;**45**(4):255–285.
3. G Sleijpen GL, and Van der Vorst Ha. A Jacobi-Davidson Iteration Method for Linear Eigenvalue Problems. SIAM Journal on Matrix Analysis and Applications. 1996;**17**(2):401–425.
4. Knyazev AV. Toward the Optimal Preconditioned Eigensolver: Locally Optimal Block Preconditioned Conjugate Gradient Method. SIAM Journal on Scientific Computing. 2001;**23**(2):517–541. Available from: <http://epubs.siam.org/doi/10.1137/S1064827500366124>.
5. Andreev R, and Schwab C. In: Graham IG, Hou TY, Lakkis O, and Scheichl R, editors. Sparse Tensor Approximation of Parametric Eigenvalue Problems. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012. p. 203–241. Available from: https://doi.org/10.1007/978-3-642-22061-6_7.
6. Ghanem R, and Ghosh D. Efficient characterization of the random eigenvalue problem in a polynomial chaos decomposition. International Journal for Numerical Methods in Engineering. 2007 oct;**72**(4):486–504.
7. Sirković P, and Kressner D. Subspace Acceleration for Large-Scale Parameter-Dependent Hermitian Eigenproblems. SIAM Journal on Matrix Analysis and Applications. 2016 jan;**37**(2):695–718. Available from: <http://epubs.siam.org/doi/10.1137/15M1017181>.
8. Quarteroni A, Manzoni A, and Negri F. Reduced Basis Methods for Partial Differential Equations: An introduction. vol. 92 of UNITEXT. Cham: Springer; 2016. Available from: <http://link.springer.com/10.1007/978-3-319-15431-2>.
9. Rozza G, Huynh DBP, and Patera AT. Reduced Basis Approximation and a Posteriori Error Estimation for Affinely Parametrized Elliptic Coercive Partial Differential Equations. Archives of Computational Methods in Engineering. 2008 sep;**15**(3):229–275. Available from: <http://link.springer.com/10.1007/s11831-008-9019-9>.
10. Karow M, Kokiopoulou E, and Kressner D. On the computation of structured singular values and pseudospectra. Systems & Control Letters. 2010 feb;**59**(2):122–129. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0167691109001509>.
11. Sirković P. A reduced basis approach to large-scale pseudospectra computation. Numerical Linear Algebra with Applications. 2019;**26**(2):e2222. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/nla.2222>.

12. Kangal F, Meerbergen K, Mengi E, and Michiels W. A subspace method for large-scale eigenvalue optimization. *SIAM Journal on Matrix Analysis and Applications*. 2018;**39**(1):48–82.
13. Parlett BN. *The Symmetric Eigenvalue Problem*. siam; 1998.
14. Lax PD. *Linear algebra and its applications*. Wiley-Interscience; 2007.
15. Seyranian AP, and Mailybaev AA. *Multiparameter stability theory with mechanical applications*. vol. 13. World Scientific; 2003.
16. Bauer FL, and Fike CT. Norms and exclusion theorems. *Numerische Mathematik*. 1960 dec;**2**(1):137–141.
17. Kato T. On the Upper and Lower Bounds of Eigenvalues. *Journal of the Physical Society of Japan*. 1949 jul;**4**(4-6):334–339. Available from: <http://journals.jps.jp/doi/10.1143/JPSJ.4.334>.
18. Temple G. The Accuracy of Rayleigh’s Method of Calculating the Natural Frequencies of Vibrating Systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 1952 feb;**211**(1105):204–224. Available from: <http://rspa.royalsocietypublishing.org/cgi/doi/10.1098/rspa.1952.0034>.
19. Hesthaven JS, Stamm B, and Zhang S. Efficient greedy algorithms for high-dimensional parameter spaces with applications to empirical interpolation and reduced basis methods. *ESAIM: Mathematical Modelling and Numerical Analysis*. 2014;**48**(1):259–283. Available from: <http://www.esaim-m2an.org/10.1051/m2an/2013100>.

