Free Material Optimization with Control of the Fundamental Eigenfrequency

M. Stingl,∗ M. Kočvara† and G. Leugering‡

October 4, 2007

Keywords. structural optimization, material optimization, semidefinite programming, nonlinear programming

AMS subject classifications. 74B05, 74P05, 90C90, 90C30, 90C22

Abstract

The goal of this paper is to formulate and solve free material optimization problems with constraints on the minimal eigenfrequency of a structure. A natural formulation of this problem as linear semidefinite program turns out to be numerically intractable. As alternative, we propose a new approach, which is based on a nonlinear semidefinite low-rank approximation of the semidefinite dual. Throughout this article, an algorithm is introduced and convergence properties are investigated. The article is concluded by numerical experiments proving the effectiveness of the new approach.

1 Introduction

Free material optimization (FMO) is a branch of structural optimization that gains more and more interest in the recent years. The underlying FMO model was introduced in [3] and has been studied in several further articles as, for example, [2, 23]. The design variable in FMO is the full elastic stiffness tensor that can vary from point to point. The method is supported by powerful optimization and numerical techniques, which allow for scenarios with complex bodies, fine finite-element meshes and several load cases. On the other hand, the basic FMO model has certain drawbacks. For example, structures may fail due to high stresses, or due to lack of stability of the optimal structure (compare [13, 14] for further discussion). In order to prevent this undesirable behavior, additional requirements have to be taken into account. Typically, such modifications lead to additional constraints on the set of admissible materials and/or the set of admissible displacements, destroying the favorable mathematical structure of the original problem (see [13, 14]). The particular cause of structural failure we want to investigate in the scope of this article is vibration resonance. Throughout this article,

∗Institute of Applied Mathematics, University of Erlangen, Martensstr. 3, 91058 Erlangen, Germany (stingl@am.uni-erlangen.de)
†School of Mathematics, University of Birmingham, Edgbaston Birmingham B15 2TT, UK (kocvara@maths.bham.ac.uk)
‡Institute of Applied Mathematics, University of Erlangen, Martensstr. 3, 91058 Erlangen, Germany (leugering@am.uni-erlangen.de)
we give an appropriate formulation of an FMO problem, which takes care of this phenomenon, derive various discretized formulations and propose an efficient algorithm for the solution of the same.

In contrast to the original FMO model, which is based on a static PDE system, vibration of a body is a dynamic process. In the first part of the paper we demonstrate how we can bypass this additional challenge by use of a reformulation as (time-independent) generalized eigenvalue problem. As result we obtain an extended FMO problem with an eigenfrequency condition. For the latter problem we are able to prove the existence of an optimal solution. In the second section we explain, how an existing discretization scheme (proposed in [21]) can be extended to cover the additional eigenfrequency condition. In the third section we give a first formulation of the discretized FMO problem with vibration control as a linear semidefinite program. Moreover we explain, why this formulation is not suited to serve as a basis for efficient numerical calculations. In the framework of the fourth section we develop an algorithm, which is based on a low-rank approximation of the semidefinite dual. The low-rank approach is motivated by ideas recently introduced by Monteiro and Burer (see [5]) for the solution of linear semidefinite programs. The article is concluded by numerical studies.

Throughout this article we use the following notation: We denote by $S^N$ the space of symmetric $N \times N$-matrices equipped with the standard inner product $\langle \cdot, \cdot \rangle$ defined by $\langle A, B \rangle := \text{Tr}(AB)$ for any pair of matrices $A, B \in S^N$. We denote by $S^N_+$ the cone of all positive semidefinite matrices in $S^N$ and use the abbreviation $A \succeq 0$ for matrices $A \in S^N_+$. Moreover, for $A, B \in S^d$, we say that $A \succeq \succ_S B$ if and only if $A - B \succeq 0$, and similarly for $A \prec \prec_S B$.

### 2 A mathematical model

Material optimization deals with optimal design of elastic structures, where the design variables are material properties. The material can even vanish in certain areas, thus one often speaks of topology optimization.

Let $\Omega \subset \mathbb{R}^2$ be a two-dimensional bounded domain$^1$ with a Lipschitz boundary. By $u(x) = (u_1(x), u_2(x))$ we denote the displacement vector at a point $x$ of the body under load $f$, and by

$$ e_{ij}(u(x)) = \frac{1}{2} \left( \frac{\partial u_i(x)}{\partial x_j} + \frac{\partial u_j(x)}{\partial x_i} \right) $$

the (small-)strain tensor. We assume that our system is governed by linear Hooke’s law, i.e., the stress is a linear function of the strain

$$ \sigma_{ij}(x) = E_{ijkl}(x) e_{kl}(u(x)) \quad \text{(in tensor notation),} $$

where $E$ is the elastic stiffness tensor. The symmetries of $E$ allow us to write the 2nd order tensors $e$ and $\sigma$ as vectors

$$ e = (e_{11}, e_{22}, \sqrt{2} e_{12})^T \in \mathbb{R}^3, \quad \sigma = (\sigma_{11}, \sigma_{22}, \sqrt{2} \sigma_{12})^T \in \mathbb{R}^3. $$

Correspondingly, the 4th order tensor $E$ can be written as a symmetric $3 \times 3$ matrix

$$ E = \begin{pmatrix} E_{1111} & E_{1122} & \sqrt{2} E_{1112} \\ E_{1222} & \sqrt{2} E_{2212} & \sqrt{2} E_{2212} \\ \text{sym.} & 2 E_{1212} & \end{pmatrix}. \tag{1} $$

$^1$The entire presentation is given for two-dimensional bodies, to keep the notation simple. Analogously, all this can be done for three-dimensional solids.
In this notation, Hooke's law reads as $\sigma(x) = E(x)e(u(x))$.

Using the elastic stiffness tensor $E$ and a given external load function $f \in [L_2(\Gamma)]^2$, (where $\Gamma$ is a part of $\partial \Omega$ that is not fixed by Dirichlet boundary conditions) we obtain the following basic boundary value problem of linear elasticity:

Find $u \in [H^1(\Omega)]^2$ such that

\[
\begin{align*}
\text{div}(\sigma) &= 0 \quad \text{in } \Omega \\
\sigma \cdot n &= f \quad \text{on } \Gamma \\
u &= 0 \quad \text{on } \Gamma_0 \\
\sigma &= E \cdot e(u) \quad \text{in } \Omega
\end{align*}
\]

Here $\Gamma$ and $\Gamma_0$ are open disjunctive subsets of $\partial \Omega$. The corresponding weak form, the so called weak equilibrium equation, reads as:

Find $u \in V$, such that

\[
\int_{\Omega} \langle E(x)e(u(x)), e(v(x)) \rangle \, dx = \int_{\Gamma} f(x) \cdot v(x) \, dx, \quad \forall v \in V,
\]

where $V = \{ u \in [H^1(\Omega)]^2 | u = 0 \text{ on } \Gamma_0 \} \supset [H_0^1(\Omega)]^2$ reflects the Dirichlet boundary conditions.

In free material optimization (FMO), the design variable is the elastic stiffness tensor $E$ which is a function of the space variable $x$ (see [3]). The only constraints on $E$ are that it is physically reasonable, i.e., that $E$ is symmetric and positive semidefinite. This gives rise to the following definition

\[
E_0 := \{ E \in L^\infty(\Omega)^{3 \times 3} | E = E^T, E \succeq 0 \text{ a.e. in } \Omega \}.
\]

The choice of $L^\infty$ is due to the fact that we want to allow for material/no-material situations. A frequently used measure for the stiffness of the material tensor is its trace. In order to avoid arbitrarily stiff material, we add pointwise stiffness restrictions of the form $0 \leq \rho \leq \text{Tr}(E) \leq \overline{\rho}$, where $\overline{\rho}$ is a finite real number. Moreover we prescribe the total stiffness/volume by the constraint $\nu(E) = \overline{\nu}$. Here the volume $\nu(E)$ is defined as $\int_{\Omega} \text{Tr}(E) \, dx$ and $\overline{\nu} \in \mathbb{R}$ is an upper bound on overall resources. Accordingly, we define the set of admissible materials as

\[
E := \{ E \in L^\infty(\Omega)^{3 \times 3} | E = E^T, E \succeq \rho I, \text{Tr}(E) \leq \overline{\rho} \text{ a.e. in } \Omega, \nu(E) = \overline{\nu} \}.
\]

We are now able to present the minimum compliance single-load FMO problem

\[
\inf_{E \in E} \int_{\Gamma} f(x) \cdot u_E(x) \, dx
\]

subject to $u_E$ solves (3).

The objective, the so called compliance functional, measures how good the structure can carry the load $f$. In [21] it is shown that problem (4) can be given equivalently as

\[
\inf_{E \in E} c(E)
\]

where $c(E)$ is a closed formula for the compliance given by

\[
c(E) = \sup_{u \in V} -a_E(u, u) + 2 \int_{\Gamma} f \cdot u \, dx.
\]
Problem (4) has been extensively studied in [15, 21]. The most successful method for the solution of problem (4), based on dualization of the original problem [2, 21] gave rise to a software package MOPED, which was recently applied to real-world applications and lead to significant improvements. On the other hand the underlying FMO model has certain limitations. Probably the most serious drawback of problem (4) is that it does not count with possible instability of the structure (compare [13]). One possible source of such instability is vibration resonance. In the sequel we develop a generalized FMO model, which is more robust with respect to this phenomenon. As we will see, the modification in the model results in an additional constraint on the set of admissible materials $\mathcal{E}$.

Vibration of a body - as a dynamic process - can be modeled by the following time-dependent PDE:

$$\text{div}(\sigma(x,t)) = \rho E(x)\ddot{u}(x,t), \quad (x,t) \in \Omega \times [0, T],$$

with boundary conditions

$$\frac{\partial}{\partial n}\sigma(x,t) = 0 \text{ on } \Gamma \times [0, T],$$
$$u(x,t) = 0 \text{ on } \Gamma_0 \times [0, T].$$

Here the material density term $\rho$ is defined by $\rho E(x) = \text{tr}(E(x))$ and $E(x)$ is the material tensor introduced earlier in this section. As in this case there is no external force applied to the system, we call the solutions of (5) free motions. Using the assumption of Hooke’s law and introducing the differential operator $S_E(\cdot) := \text{div}(Ee(\cdot))$ we obtain from (5):

$$S_E(u(x,t)) = \rho E(x)\ddot{u}(x,t), \quad (x,t) \in \Omega \times [0, T],$$

$$\frac{\partial}{\partial n}S_E(u(x,t)) = 0 \text{ on } \Gamma \times [0, T],$$
$$u(x,t) = 0 \text{ on } \Gamma_0 \times [0, T].$$

Using Fourier transform we derive the following characterization of solutions of system (6):

**Proposition 2.1.** The solutions of system (6) are of the form

$$u(x,t) = \sum_{j=1}^{\infty} \left[ a_j \cos(\sqrt{\lambda_j}t) + b_j \sin(\sqrt{\lambda_j}t) \right] w_j(x),$$

where $a_j, b_j, \lambda_j, w_j$ are free real parameters and $\lambda_j, w_j$ are the solutions (eigenvalues and eigenvectors, respectively) of the generalized eigenvalue problem

$$-S_E(w_j(x)) = \lambda_j \rho E(x)w_j(x), \quad x \in \Omega,$$
$$\frac{\partial}{\partial n}\sigma(x) = 0 \text{ on } \Gamma,$$
$$u(x) = 0 \text{ on } \Gamma_0.$$

Introducing the bilinear forms

$$a_E(w,v) = \int_{\Omega} \langle E(x)e(w(x)), e(v(x)) \rangle dx,$$
$$b_E(w,v) = \int_{\Omega} \rho E(x)w(x)v(x) dx,$$
the weak form associated with (8) is:

\[ \text{Find } \lambda \in \mathbb{R}, u \in V, u \neq 0, \text{ such that: } a_E(u, v) = \lambda b_E(u, v) \quad \forall v \in V. \] (10)

We use the following definition of the smallest well defined eigenvalue.

**Definition 2.2.** For each \( E \in \mathcal{E}_0 \), let \( \lambda_{\text{min}}(E) \) denote the smallest well defined eigenvalue of system (10), i.e.

\[ \lambda_{\text{min}} = \min \{ \lambda \mid \exists u \in V: \text{Equation (10) holds for } (\lambda, u) \text{ and } u \notin \ker(b_E) \}, \]

where \( \ker(b_E) = \{ z \mid b_E(z, v) = 0 \quad \forall v \in V \} \).

It is well known from engineering literature that the stability of a structure with respect to vibration phenomena can be improved by raising its fundamental eigenfrequency. This is our motivation to consider the following problem: We search for a material distribution \( E \) such that the minimal well defined eigenvalue of system (10) is larger than a prescribed positive lower bound. Denoting this value by \( \hat{\lambda} \), we obtain the constraint

\[ \lambda_{\text{min}}(E) \geq \hat{\lambda}. \] (11)

In Appendix A, Corollary A.3 it is shown that equation (11) can be equivalently reformulated as

\[ \inf_{u \in V, \|u\| = 1} a_E(u, u) - \hat{\lambda} b_E(u, u) \geq 0. \] (12)

Introducing the function

\[ \lambda^\hat{\lambda}_{\text{min}} : L^\infty(\Omega)^{3 \times 3} \rightarrow \mathbb{R} \cup \{ \infty \}, \]

\[ E \mapsto \inf_{u \in V, \|u\| = 1} a_E(u, u) - \hat{\lambda} b_E(u, u), \]

we are able to state the minimum compliance single-load FMO problem with vibration control

\[ \inf_{E \in \mathcal{E}} \int_{\Gamma} f(x) \cdot u_E(x) dx \] (13)

subject to

\[ u_E \text{ solves (3)}, \]

\[ \lambda^\hat{\lambda}_{\text{min}}(E) \geq 0. \]

Next we want to investigate the well-posedness of problem (13). We start with the following lemma:

**Lemma 2.3.** The function \( \lambda^\hat{\lambda}_{\text{min}} \) is upper semicontinuous and concave.

**Proof.** We first note that for fixed \( u \in V \) the mapping \( E \mapsto a_E(u, u) - \hat{\lambda} b_E(u, u) \) is affine and consequently continuous w. r. t. \( E \). Consequently, \( \lambda^\hat{\lambda}_{\text{min}} \) is the infimum of affine functionals and thus concave. Moreover, \( \lambda^\hat{\lambda}_{\text{min}} \) is upper semicontinuous, as it is the infimum of (infinitely many) continuous functionals [8, Proposition III.1.2]. \( \square \)

Using Lemma 2.3 we are able to give details about the structure of the feasible set of problem (13):
Lemma 2.4. The set $E^{\hat{\lambda}} = \{ E \in \mathcal{E} \mid \lambda_{\min}^{\hat{\lambda}} \geq 0 \}$ is weakly-∗ compact.

Proof. The weakly-∗ compactness of $E$ has been shown in the proof of Theorem 2.1 in [2]. Thus the only thing we have to show is that $E^{\hat{\lambda}}$ is closed. But this follows immediately from the closedness of $E$ and Lemma 2.3. ∎

Now the following Theorem can be proved exactly in the same way as Theorem 2.1 in [2]:

Theorem 2.5. If the set $E^{\hat{\lambda}}$ is non-empty, then problem (13) has at least one solution.

We conclude this section by two remarks.

Remark 2.6. From the equivalence of (11) and (12) and Lemma 2.4 we immediately conclude that the function $\lambda_{\min} : L^\infty(\Omega)^{3 \times 3} \to \mathbb{R} \cup \{\infty\}$ is upper semicontinuous and quasiconcave. Using this and the fact that the compliance functional, given by the formula

$$c(E) = \sup_{u \in V} -\frac{1}{2} a_\mathcal{E}(u, u) + \int_{\Gamma} f \cdot u \, dx$$

is convex and lower semicontinuous w. r. t. $E$ (see [21]), we may repeat the arguments above in order to verify existence of at least one optimal solution for the following problems:

$$\inf_{E \in \mathcal{E}} \, v(E), \quad \text{subject to} \quad c(E) \leq \delta, \quad \lambda_{\min} \geq \hat{\lambda},$$

and

$$\inf_{E \in \mathcal{E}} -\lambda_{\min}(E), \quad \text{subject to} \quad c(E) \leq \delta, \quad v(E) = \bar{v}.$$ 

Here $\delta \in \mathbb{R}$ is an upper bound on the compliance of the structure.

Remark 2.7. All results presented above remain true in case of more general Dirichlet boundary conditions of the form

$$u_i = 0 \text{ on } \Gamma_0 \text{ for } i = 1 \text{ and/or } 2.$$

3 Discretization

In order to solve the infinite-dimensional problem (13) numerically, we have to find an appropriate discretization scheme. We extend the finite element discretization schemes proposed in [2, 21]. Rather than presenting the full convergence analysis of our discretization scheme here, we just note that the finite element approach and convergence analysis presented in [21] applies to our problem without changes.

To keep notation simple, we use the same symbols for the discrete objects (vectors) as for the “continuum” ones (functions). Suppose that $\Omega$ can be partitioned into $M$
squares called $\Omega$, which are all of the same size $h \in \mathbb{R}$ (otherwise we use the standard isoparametric concept, see [7]). Let us denote by $n$ the number of nodes (vertices of the elements). We approximate the matrix function $E(x)$ by a function that is constant on each element, i.e., characterized by a vector of matrices $E = (E_1, \ldots, E_M)$ of its element values. Hence the discrete counterpart of the set of admissible materials in algebraic form is

$$\tilde{\mathcal{E}} = \left\{ E \in (\mathbb{R}^3)^M \left| E_i \equiv 0, \quad \rho \leq \text{Tr}(E_i) \leq \bar{\rho}, \quad i = 1, \ldots, M, \quad \sum_{i=1}^{M} \text{Tr}(E_i) = \tilde{\nu}\right. \right\}. \quad (16)$$

Here $\tilde{\nu}$ is the upper bound on resources introduced in (4).

Further assume that the displacement vector $u(x)$ is approximated by a continuous function that is bilinear in each coordinate on every element. Such a function can be written as $u(x) = \sum_{i=1}^{n} u_i \vartheta_i(x)$ where $u_i$ is the value of $u$ at $i$th node and $\vartheta_i$ is the basis function associated with $i$th node (for details, see [7]). Recall that the displacement function is vector valued with 2 components. Consequently any function in the discrete set of admissible displacements can be identified with a vector in $\mathbb{R}^N$, where $N = 2n - \#(\text{components of } u \text{ fixed by Dirichlet b. c.})$ and we obtain

$$\tilde{V} = \mathbb{R}^N. \quad (17)$$

With the (reduced) family of basis functions $\vartheta_k, k = 1, 2, \ldots, N$, we define a $3 \times 2$ matrix

$$\tilde{B}_k^\top = \begin{pmatrix} \frac{\partial \vartheta_k}{\partial x_1} & 0 & \frac{1}{2} \frac{\partial \vartheta_k}{\partial x_2} \\ \frac{\partial \vartheta_k}{\partial x_2} & 0 & \frac{1}{2} \frac{\partial \vartheta_k}{\partial x_1} \end{pmatrix}. \quad (18)$$

Now, for an element $\Omega_i$, let $\mathcal{D}_i$ be an index set of nodes belonging to this element. Next we want to derive the discrete counterpart part of $a_E(\cdot, \cdot)$. We use a Gauss formula for the evaluation of the integral over each element $\Omega_i$, assume that there are $G$ Gauss integration points on each element and denote by $x_i^{G,j}$ the $j$th integration point on the $i$th element. Next we construct block matrices $B_{i,k} \in \mathbb{R}^{3 \times N}$ composed of $(3 \times 2)$ blocks $\tilde{B}_j(x_i^{G,j}), j \in \mathcal{D}_i$ at the $j$-th position and zeros otherwise. Then the discrete counterpart of $a_E(\cdot, \cdot)$, usually known as the *stiffness matrix* is

$$A(E) = \sum_{i=1}^{M} A_i(E), \quad A_i(E) = \sum_{k=1}^{G} B_{i,k}^\top E_i B_{i,k}. \quad (18)$$

Note that the matrices $A_i$ are often called *element stiffness matrices*. Now, assuming the load function $f$ to be linear on each element and identifying such a function with a vector $f \in \mathbb{R}^N$, the discrete objective functional and equilibrium condition read as

$$f^\top u, \quad A(E)u = f, \quad (19)$$

respectively. Similarly, we use the representation of the discrete displacement functions in the basis of $\vartheta_k, k = 1, 2, \ldots, N$ to derive the discrete version of the bilinear form $b_E(\cdot, \cdot)$. Defining vectors $V_{i,k} \in \mathbb{R}^N, i = 1, 2, \ldots, M, k = 1, 2, \ldots, G$ with $\vartheta_j(x_i^k), j \in \mathcal{D}_i$ at the $j$-th position and zeros otherwise, the *mass matrix* is given by

$$M(E) = \sum_{i=1}^{M} M_i(E), \quad M_i(E) = \text{Tr}(E_i) M_i, \quad M_i = \sum_{k=1}^{G} V_{i,k}^\top V_{i,k}. \quad (20)$$
As above, $M(E)$ is composed by a sum of matrices $M_i$, the element mass matrices. Finally, the discrete counterpart of the condition of the lower eigenfrequency of the structure (12) reads as

$$\inf_{u \in \mathbb{R}^n, \|u\| = 1} u^\top \left( A(E) - \hat{\lambda}M(E) \right) u \geq 0. \quad (21)$$

After discretization, problem (13) becomes

$$\min_{u \in \mathbb{R}^N, E \in \mathcal{E}} f^\top u \quad (22)$$
subject to

$$A(E)u = f,$$
$$\inf_{u \in \mathbb{R}^n, \|u\| = 1} u^\top \left( A(E) - \hat{\lambda}M(E) \right) u \geq 0.$$  

Problem (22) is a mathematical programming problem with linear matrix inequality constraints and standard nonlinear constraints; in the following section we will show how this problem can be turned into a standard linear semidefinite program.

4 A linear SDP approach

In the recent years excellent software packages, most of them based on the interior point idea, have been developed for the solution of linear SDP problems. For an overview, compare, for example, [22] or [16]. In the sequel we give an alternative formulation of the discrete FMO problem (22) as linear semidefinite program.

Proposition 4.1. Problem (22) is equivalent to the following linear semidefinite program

$$\min_{\alpha \in \mathbb{R}, u \in \mathbb{R}^N, E \in \mathcal{E}} \alpha \quad (23)$$
subject to

$$\begin{pmatrix} \alpha & -f \\ -f & A(E) \end{pmatrix} \succeq 0,$$
$$A(E) - \hat{\lambda}M(E) \preceq 0.$$  

Proof. After introducing an auxiliary variable $\alpha$, the assertion follows immediately from Proposition 3.1 in [1].

In the remainder of this section we point out, why the formulation as linear SDP is impractical for the efficient solution of FMO problems with vibration control, due to the large number of variables and the size of the matrix constraints. Our observations are based on practical experience and complexity estimates. We have solved Example 1 from section 7 by state-of-the-art linear SDP solvers. The fastest solver needed about 50 hours on a high end computer with a processor speed of approximately 3 GHz. Using this number as a reference and taking into account that the computational complexity of all currently available linear SDP solvers depends at least quadratically (sometimes even cubically) on the size of the matrix constraints and typically cubically...
on the number of variables, it becomes obvious that formulation (23) is not suited to serve as a basis for the efficient solution of FMO problems of practical size.

Remark 4.2. Note that a formulation similar to (23) has been successfully applied to problems of Truss Topology design as well as variable thickness sheet problems in the past. The interested reader is referred to [17, 1, 13]. Note however that in both cases, the number of optimization variables is considerably lower.

5 The dual problem and a low-rank approximation

The goal of this section is to find an alternative formulation to problem (23), which is numerically tractable. Our strategy is as follows: First, we derive the Lagrange dual to problem (23) and then we present a low-rank approximation to the same, which is (under certain assumptions) equivalent to the original problem.

The following theorem allows us to identify problem (22) as the Lagrange dual of a convex semidefinite program:

Theorem 5.1. Problem (23) is equivalent to the Lagrange dual of the problem

\[
\max_{u \in \mathbb{R}^N, \alpha \in \mathbb{R}, W \succeq 0, \beta_l \geq 0, \beta_u \geq 0} \quad 2f^T u - \alpha V + \rho \sum_{i=1}^M \beta_l^i - \rho \sum_{i=1}^M \beta_u^i \quad \text{subject to (24)}
\]

where \( g_i(u, \alpha, W, \beta_l^i, \beta_u^i) = 0, \ i = 1, 2, \ldots, M, \)

Moreover there is no duality gap and the optimal material matrices

\( E_i^*, i = 1, 2, \ldots, M \)

take the role of Lagrangian multipliers associated with the non-linear inequality constraints

\( g_i(u, \alpha, W, \beta_l^i, \beta_u^i) \leq 0, \ i = 1, 2, \ldots, M. \)

Proof. Rewriting the eigenfrequency constraints as in problem (23) and taking into account that \( A(E) \) is positive definite for all \( E \in \tilde{E} \), we observe that problem (22) can be written equivalently as

\[
\min_{E \in \tilde{E}} \max_{u \in \mathbb{R}^N} 2f^T u - u^T A(E) u \quad \text{subject to (25)}
\]

\( A(E) - \hat{\lambda} M(E) \succeq 0. \)
The Lagrangian associated with problem (25) can be written in the form

$$\mathcal{L}(E, u, \alpha, W, \beta^l, \beta^u) := 2f^\top u - u^\top A(E)u$$

$$+ \sum_{i=1}^{M} \beta_i^l (p - \text{Tr}(E_i)) + \sum_{i=1}^{M} \beta_i^u (\text{Tr}(E_i) - p)$$

$$+ \alpha \left( \sum_{i=1}^{M} \text{Tr}(E_i) - V \right) + (W, \hat{\lambda}M(E) - A(E)),$$

where

$$(E, u, \alpha, W, \beta^l, \beta^u) \in (\mathbb{S}^N_+)^M \times \mathbb{R}^N \times \mathbb{S}^N_+ \times \mathbb{R}_+^{2M}.$$  

Now problem (25) can be formulated as

$$\min_{E \geq 0} \max_{u \in \mathbb{R}^N, \alpha \in \mathbb{R}_+, W \in \mathbb{S}_+^N, \beta^l, \beta^u \geq 0} \mathcal{L}(E, u, \alpha, W, \beta^l, \beta^u).$$  

(27)

The Lagrange dual to (27) is

$$\max_{u \in \mathbb{R}^N, \alpha \in \mathbb{R}_+, W \in \mathbb{S}_+^N, \beta^l, \beta^u \geq 0} \min_{E \geq 0} \mathcal{L}(E, u, \alpha, W, \beta^l, \beta^u).$$  

(28)

Taking into account that

$$u^\top A(E)u = u^\top \left( \sum_{i=1}^{M} \sum_{j=1}^{G} B_{ij} E_i B_{ij}^\top \right) u = \sum_{i=1}^{M} \left( E_i, \sum_{j=1}^{G} B_{ij}^\top uu^\top B_{ij} \right),$$

$$\langle W, A(E) \rangle = \sum_{i=1}^{M} \left( E_i, \sum_{j=1}^{G} B_{ij}^\top W B_{ij} \right),$$

$$\langle W, M(E) \rangle = \sum_{i=1}^{M} \left( E_i, I \right) \langle W, M_i \rangle = \sum_{i=1}^{M} \left( E_i, \langle W, M_i \rangle I \right)$$

and $\text{Tr}(E_i) = \langle E_i, I \rangle$ for all $i = 1, 2, \ldots, M$, we obtain

$$\mathcal{L}(E, u, \alpha, W, \beta^l, \beta^u) = 2f^\top u - \alpha V + \rho \sum_{i=1}^{M} \beta_i^l - \bar{p} \sum_{i=1}^{M} \beta_i^u$$

$$- \sum_{i=1}^{M} \left( E_i, \sum_{j=1}^{G} B_{ij} uu^\top B_{ij} + \sum_{j=1}^{G} B_{ij}^\top W B_{ij} - \hat{\lambda} \langle W, M_i \rangle I - (\alpha + \beta_i^l - \beta_i^u) I \right).$$

Using this form and interpreting $E_i$ ($i = 1, 2, \ldots, M$) as Lagrangian multipliers, problem (28) takes the form

$$\max_{u \in \mathbb{R}^N, \alpha \in \mathbb{R}_+, W \in \mathbb{S}_+^N, \beta^l, \beta^u \geq 0} 2f^\top u - \alpha V + \rho \sum_{i=1}^{M} \beta_i^l - \bar{p} \sum_{i=1}^{M} \beta_i^u$$

subject to

$$\sum_{j=1}^{G} B_{ij}^\top (uu^\top + W) B_{ij} - \left( \hat{\lambda} \langle W, M_i \rangle + \alpha + \beta_i^l - \beta_i^u \right) I \ll 0, \quad i = 1, 2, \ldots, M.$$
but this is problem (24). Finally, taking into account that problem (24) is convex and that the Slater condition holds (in order to construct a strictly feasible point, use arbitrary \( W \succ 0, \alpha \in \mathbb{R}^N, \beta^i \geq 0, \beta^u \geq 0 \) and choose \( \alpha \) large enough, such that all inequalities in problem (24) are strictly feasible), the fact that the duality gap is zero follows from [4, Theorem 5.81].

Later we will make use of the following proposition. The prove is straightforward, but rather technical and therefore postponed to Appendix I of this article.

**Proposition 5.2.** A tuple \((u^*, \alpha^*, W^*, \beta^*, \beta^u; E^*) \in \mathbb{R}^{N+1} \times \mathbb{S}^N_+ \times \mathbb{R}^{2M} \times (\mathbb{S}^N_+)^M\) is a KKT point of (24) if and only if the conditions

\[
g_i(u^*, \alpha^*, W^*, \beta^*, \beta^u) \leq 0 \quad (i = 1, 2, \ldots, M),
\]

\[
\rho \leq \text{Tr}(E^*_i) \leq \overline{\rho} \quad (i = 1, 2, \ldots, M),
\]

\[
\beta^u_i \big( \rho - \text{Tr}(E^*_i) \big) = 0, \quad \beta^u_i \big( \text{Tr}(E^*_i) - \overline{\rho} \big) = 0, \quad (i = 1, 2, \ldots, M),
\]

\[
\sum_{i=1}^{M} \text{Tr}(E^*_i) = V, \quad A(E^*) - \lambda^* M(E^*) \succ 0, \quad \langle A(E^*) - \lambda^* M(E^*), W^* \rangle = 0,
\]

\[
A(E^*)u^* = f, \quad f^\top u^* = \alpha^* V - \rho \sum_{i=1}^{M} \beta^i + \overline{\rho} \sum_{i=1}^{M} \beta^i
\]

are satisfied.

Theorem 5.1 makes sure that we can retrieve the solution of (22) by calculating a primal-dual solution of (24). Consequently, we could apply any convex semidefinite programming solver, which is able to generate primal-dual solutions. Note however that the computational complexity of problem (24) is not much better than that of the linear SDP problem (23). This is due to the large size of the positive semidefiniteness constraint \( W \succ 0 \). For this reason, we follow the idea of Monteiro and Burer (see [5]), in order to construct a low-rank approximation of (24). Suppose for a moment that we know a primal solution \( E^* \) of problem (23). Then we define

\[
R_0 := \dim \left( \ker \left( A(E^*) - \hat{\lambda} M(E^*) \right) \right),
\]

which is equal to the dimension of the multiplicity of the smallest eigenvalue \( \hat{\lambda} \) of the generalized eigenvalue problem \( A(E^*)v = \hat{\lambda} M(E^*)v \). Now, assuming that Slater’s condition holds for (23), we observe that the matrix \( W^* \) in problem (24) takes the role of the Lagrangian multiplier associated with the inequality constraint

\[
A(E) - \hat{\lambda} M(E) \succ 0.
\]

Moreover it follows from the complementarity slackness condition that there exists an optimal multiplier \( W^* \), with the property

\[
\text{rank}(W^*) \leq R_0.
\]

This is our motivation to substitute \( W \) by \( \sum_{\ell=1}^{L} w_\ell w_\ell^\top \) in problem (24). Doing this, we obtain

\[
\max_{\substack{w_1, w_2, \ldots, w_L, u \in \mathbb{R}^N \atop u \geq 0, \beta^2 \geq 0, \beta^u \geq 0}} \left\{ \sum_{\ell=1}^{L} 2f^\top w_\ell - \alpha V + \rho \sum_{i=1}^{M} \beta^i + \overline{\rho} \sum_{i=1}^{M} \beta^u \right\} \quad \text{subject to} \quad \tilde{g}_i(u, \alpha, w_1, w_2, \ldots, w_L, \beta^i, \beta^u) \leq 0, \quad i = 1, \ldots, M,
\]

11
where \( \tilde{g}_i(u, \alpha, w_1, w_2, \ldots, w_L, \beta^i, \beta^u) \) is defined as

\[
\sum_{j=1}^G B_{ij}^T \left( uu^T + \sum_{l=1}^L w_l w_l^T \right) B_{ij} - \left( \tilde{\lambda} \left( \sum_{l=1}^L w_l w_l^T, M_i \right) + \alpha + \beta^i - \beta^u \right) I
\]

for all \( i = 1, 2, \ldots, M \). The following theorem provides a relation between problems (24) and (30).

**Theorem 5.3.** Let \( E^*_0 \) be a solution of (23) and \( R_0 \) defined by (29). Then there exists \( L \leq R_0 \) such that for all (global) solutions \( (u^*, \alpha^*, w_1^*, \ldots, w_L^*, \beta^i, \beta^u) \) of (30) the tuple \( (u^*, \alpha^*, W^*, \beta^i, \beta^u) \) with \( W^* := \sum_{l=1}^L w_l^T(w_l^*)^T \) is a solution of (24). Moreover each vector of Lagrangian multipliers \( E^* \) associated with the inequality constraints

\[
\tilde{g}_i(u, \alpha, w_1, w_2, \ldots, w_L, \beta^i, \beta^u) \leq 0, \quad i = 1, 2, \ldots, M,
\]

forms an optimal solution of (23).

In order to prove the assertions of Theorem 5.3, we make use of the following Lemmas:

**Lemma 5.4.** Any local minimum of problem (30) is a local minimum of problem (24) with an additional rank constraint of the form

\[
\text{rank}(W) \leq L
\]

and vice versa.

**Proof.** The assertion of Lemma 5.4 can be proven exactly in the same way as Proposition 2.3 in [6]. \( \square \)

**Lemma 5.5.** Robinson’s constraint qualification is satisfied by problem (30) in any feasible point.

**Proof.** Using [4, formula (5.195)] we can write Robinson’s constraint qualification for an arbitrary feasible point \( (\tilde{u}, \tilde{\alpha}, \tilde{w}_1, \ldots, \tilde{w}_L, \tilde{\beta}^i, \tilde{\beta}^u) \in \mathbb{R}^{(L+1)N+2M+1} \) as follows: There exists a direction \( h \in \mathbb{R}^{(L+1)N+2M+1} \) such that the inequality

\[
\tilde{g}_i(\tilde{u}, \tilde{\alpha}, \tilde{w}_1, \ldots, \tilde{w}_L, \tilde{\beta}^i, \tilde{\beta}^u) + \nabla \tilde{g}_i(\tilde{u}, \tilde{\alpha}, \tilde{w}_1, \ldots, \tilde{w}_L, \tilde{\beta}^i, \tilde{\beta}^u) \cdot h < 0
\]

holds for all \( i = 1, 2, \ldots, M \). Obviously, the direction \((0, 1, 0, \ldots, 0)\) with 1 in the position of the variable \( \alpha \) satisfies (31). \( \square \)

A simple consequence of Lemma 5.5 is that each local minimum of problem (30) can be extended to a KKT-point. Now we are able to prove Theorem 5.3:

**Proof.** Let \( L := R_0 \) and \( \hat{x}^* := (u^*, \alpha^*, w_1^*, \ldots, w_L^*, \beta^i, \beta^u) \) be a (global) solution of (30). Then we conclude from Lemma 5.4 that \( x^* := (u^*, \alpha^*, \sum w_l^T w_l^T, \beta^i, \beta^u) \) is a global solution of problem (24) with an additional rank constraint of the form \( \text{rank}(W) \leq L \). But then we conclude from the definition of \( R_0 \) that \( x^* \) is a global solution of (24). Moreover Lemma 5.5 guarantees the existence of optimal Lagrangian multipliers \( \tilde{E}^* \in \mathbb{S}_+^M \) associated with the inequality constraints

\[
\tilde{g}_i(u, \alpha, w_1, w_2, \ldots, w_L, \beta^i, \beta^u) \leq 0, \quad i = 1, 2, \ldots, M.
\]
Now we define a Lagrangian for problem (24) as follows:

\[ \tilde{L}(x, E) = \begin{cases} 
    f_0(x) + \sum_{i=1}^{M} \langle E_i, g_i(x) \rangle & \text{if } x \in C, E_i \neq 0 \ (I = 1, 2, \ldots, M) \\
    -\infty & \text{if } x \in C, E_i = 0 \text{ for some } i \\
    -\infty & \text{if } x \notin C 
\end{cases} \]

where \( f_0 \) is the objective of (24) and \( C \) is the convex set

\[ C := \mathbb{R}^{N+1} \times \mathbb{S}_+^M \times \mathbb{R}_+^{2M}. \]

As \( x^* \) is a global solution of (24) and

\[ \tilde{L}(x^*, E^*) = f_0(x^*), \]

we conclude that \((x^*, E^*)\) is a saddle point of \( \tilde{L} \). Now we obtain for example from [19, Theorem 28.3] that \( E^* \) is a solution of the dual problem to (24) and this is equivalent to (23).

Theorem 5.3 allows to replace problem (24) by a low-rank problem with appropriate rank. The advantage of the low-rank problem is that (whenever the multiplicity of the smallest generalized eigenvalue of the stencil \( A(E^*) \mid M(E^*) \)) is not too large) the dimension of the optimization variable is significantly lower than in the original problem. Moreover there is no large semidefinite constraint in the low-rank problem. On the other hand, problem (30) is a non-convex semidefinite program, which can still be considered large-scale. For such a problem it is generally difficult (if not impossible) to calculate a global solution. Even in the case a global solution has been found, it is not a trivial problem to detect globality. For the first problem, from theoretical point of view, we can not do much more than using an optimization algorithm with strong local convergence properties and provide a good start point. We will see in section 7 that this is not a big problem in practice, as the local algorithm of our choice typically identifies the global optimum, provided the rank is large enough. This observation coincides with the experience reported by Burer and Monteiro in [5] for linear semidefinite programs approximated by low-rank problems. For the second problem, we develop a practical globality test in the sequel. We start with the following proposition, which provides a characterization of KKT-points for problem (30). The proof uses almost exactly the same arguments as the proof of Proposition 5.2 and is therefore omitted.

**Proposition 5.6.** A tuple \((u^*, \alpha^*, w^*, \beta^l, \beta^u ; E^*) \in \mathbb{R}^{N+1} \times \mathbb{R}^{N \times L} \times \mathbb{R}_+^{2M} \times (\mathbb{S}_+^M)^M\) is a KKT point of (30) if and only if the conditions

\[ g_i(u^*, \alpha^*, \sum_{\ell=1}^L w^\ell_i w^\ell_i^T, \beta^l, \beta^u) \leq 0 \ (i = 1, 2, \ldots, M), \]

\[ \rho \leq \text{Tr}(E^*_i) \leq \overline{\rho} \ (i = 1, 2, \ldots, M), \sum_{i=1}^M \text{Tr}(E^*_i) = V, \]

\[ \beta^l (\rho - \text{Tr}(E^*_i)) = 0, \quad \beta^u (\text{Tr}(E^*_i) - \overline{\rho}) = 0, \ (i = 1, 2, \ldots, M), \]

\[ (A(E^*) - \lambda^* M(E^*)) \sum_{\ell=1}^L w^\ell_i w^\ell_i^T = 0, \]

\[ A(E^*) u^* = f, \quad f^T u^* = \alpha^* V - \rho \sum_{i=1}^M \beta^l_i + \overline{\rho} \sum_{i=1}^M \beta^u_i \]

are satisfied.

The following corollary is a direct consequence of Proposition 5.2 and Proposition 5.6 and provides a globality test for an arbitrary KKT point of problem (30):
Corollary 5.7. Suppose that the vector \( x^* = (u^*, \alpha^*, w_1^*, \ldots, w_L^*, \beta^{	imes}, \beta^{\mu*}; E^*) \in \mathbb{R}^{(L+1)N+2M+1} \times \mathbb{S}^3 \) is a KKT point of problem (30) and
\[
A(E^*) - \lambda M(E^*) \succeq 0.
\]
Then \((u^*, \alpha^*, \sum_{\ell=1}^L w_\ell^T w_\ell, \beta^{\times}, \beta^{\mu*}; E^*)\) is a KKT point of problem (24), and thus a primal-dual solution pair of problems (23) and (24).

Moreover the following corollary can be derived directly from the KKT conditions in Proposition 5.6 and provides an interpretation of the solution vector.

Corollary 5.8. Let \((u^*, \alpha^*, \sum_{\ell=1}^L w_\ell^T w_\ell, \beta^{\times}, \beta^{\mu*}; E^*)\) be a KKT point of problem (24). Then \(u^*\) is the optimal displacement field associated with the material \(E^*\) and the vectors \(w_1^*, w_2^*, \ldots, w_L^*\) are eigenmodes associated with the generalized eigenvalue problem \(A(E^*)v = \hat{\lambda}M(E^*)v\).

6 A low-rank algorithm

Based on the results of Theorem 5.3 and Corollary 5.7 we present a low-rank algorithm for the free material optimization problem with control of the lowest eigenfrequency:

Algorithm 6.1.

Input: Problem (24), \(L = 1\).

1. Solve (30) with rank \(L\).

Output: \((\tilde{u}, \tilde{\alpha}, \tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_L, \tilde{\beta}^{\times}, \tilde{\beta}^{\mu}; \tilde{E})\).

2. Check optimality of \((\tilde{u}, \tilde{\alpha}, \tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_L, \tilde{\beta}^{\times}, \tilde{\beta}^{\mu}; \tilde{E})\):
   - If \(A(\tilde{E}) - \lambda M(\tilde{E}) \succeq 0\) STOP;
   - \((\tilde{u}, \tilde{\alpha}, \sum_{\ell=1}^L \tilde{w}_\ell^T \tilde{w}_\ell, \tilde{\beta}^{\times}, \tilde{\beta}^{\mu}; \tilde{E})\) is optimal.

3. Increment \(L\) and GOTO step 1.

Output: \((\tilde{u}, \tilde{\alpha}, \sum_{\ell=1}^L \tilde{w}_\ell^T \tilde{w}_\ell, \tilde{\beta}^{\times}, \tilde{\beta}^{\mu}; \tilde{E})\).

Theorem 6.2. Let \(E_0^*\) be a solution of problem (23). Then Algorithm 6.1 converges in at most
\[
R_0 = \dim \left( \ker \left( A(E_0^*) - \lambda M(E_0^*) \right) \right)
\]
steps.

Proof. The assertion of Theorem 6.2 is a direct consequence of Theorem 5.3.

The algorithm of our choice used to solve low-rank problems of the form (30) is based on a generalized augmented Lagrangian method for the solution of nonlinear semidefinite programs of the form

\[
\min_{x \in \mathbb{R}^n} f(x) \tag{32}
\]
subject to
\[
G_j(x) \preceq 0, \quad j \in \mathcal{J} = \{1, 2, \ldots, J\};
\]
where \( f : \mathbb{R}^n \to \mathbb{R} \), and \( G_j(x) : \mathbb{R}^n \to \mathbb{S}_m^j \) \((j \in \mathcal{J})\) are twice continuously differentiable mappings. Local as well as global convergence properties under standard assumptions are discussed in detail in [20]. The algorithm, implemented in the code PENNON [12], has been recently applied to non-linear SDP problems arising from various applications; compare, for example, [13, 14] and [10, 11].

7 Numerical experiments

Goals of the numerical experiments The goals of the numerical experiments presented throughout this section are

- studying the practical behaviour of Algorithm 6.1, when applied to FMO problems of practical size,
- studying the practical ability of the local algorithm applied in Step 1 of Algorithm 6.1 to detect global optima,
- comparison of the performance of the low-rank algorithm and the direct solution of the primal SDP problem (23).

All experiments have been performed on a Sun Opteron machine with 8 Gbyte of memory and a processor speed of approximately 3 GHz.

Example 1 In our first example a rectangular 2D-plate is clamped on its left boundary and subjected to a load from the right (see Figure 1). The design space is discretized by 5,000 finite elements. Without eigenfrequency condition the lowest eigenfrequency in the optimal design is of order \(10^{-9}\). With eigenfrequency condition, we manage to raise the lowest eigenfrequency to \(5 \cdot 10^{-4}\). The corresponding multiplicity of the lowest eigenfrequency is two in this example. The optimal compliance value is 33 percent worse than in the pure compliance problem. The low-rank algorithm needs about 10 minutes to calculate the optimal result. The primal algorithm on the other hand needs already about 50 hours on the same example. Optimal densities are visualized in Figure 2. In Figure 3, the displacement field is plotted along with the two eigenmodes.

Example 2 Our second example models a 2D-bridge, which is clamped on the lower left and lower right corner. The bridge is subjected to vertical forces at the bottom (see Figure 4). The design space is again discretized by 5,000 finite elements. Also in this example the minimal eigenfrequency in the pure compliance problem is of order \(10^{-9}\). With eigenfrequency condition, the fundamental frequency is raised to \(1 \cdot 10^{-3}\). The corresponding multiplicity of the lowest eigenfrequency is again two. The optimal compliance value is 25 percent worse than in the pure compliance problem. The low-rank algorithm needs about 8 minutes to calculate the optimal result. The primal algorithm on the other hand needs already about 50 hours on the same example. Optimal densities are visualized in Figure 5. In Figure 6 we show the optimal displacement field along with the two eigenmodes corresponding to the lowest eigenfrequency.

Example 3 In our third example we consider a rectangular 3D-plate, which is clamped on its left boundary and subjected to a load from the right (see Figure 7). The design space is discretized by 5,000 finite elements. The optimal design calculated for the pure compliance problem results in a fundamental eigenfrequency of order \(10^{-9}\). With
Figure 1: Basic test problem – mesh, boundary conditions

Figure 2: Optimal density plots w/o / w vibration control

Figure 3: Displacement field (t) and eigenmodes (b)
Figure 4: Basic test problem – mesh, boundary conditions

Figure 5: Optimal density plots w/o / w vibration control

Figure 6: Displacement field (t) and eigenmodes (b)
Figure 7: Basic test problem – mesh, boundary conditions

Figure 8: Optimal density plots w/o / w vibration control

eigenfrequency condition, the lowest eigenfrequency is raised to $1 \cdot 10^{-4}$. The corresponding multiplicity of the lowest eigenfrequency is three. The optimal compliance value is only 5 percent worse compared to the pure compliance problem. The low-rank algorithm needs about 3 hours to calculate the optimal result. Again we visualize optimal densities (see Figure 8) and the corresponding displacement field along with the optimal eigenmodes (see Figure 9).

A Appendix I

We consider the following situation: Let $V$ be a Hilbert space, equipped with an inner product $(\cdot, \cdot)_V$ and a corresponding norm $\| \cdot \|_V$. Let further $a: V \times V \to \mathbb{R}$ be a bounded, symmetric and $V$-elliptic bilinear form. An abstract eigenvalue problem is defined as follows: Find $\lambda \in \mathbb{R}$ and $u \in V, u \neq 0$, such that

$$a(u, v) = \lambda (u, v)_V \quad \forall v \in V.$$  \hspace{1cm} (33)

The following theorem deals with existence of solutions of the latter problem (see, for example, [9] or [18]):

**Theorem A.1.** There exists an increasing sequence of positive eigenvalues of problem (33) tending to $\infty$:

$$0 < \lambda_1 \leq \lambda_2 \leq \ldots, \quad \lambda_k \to \infty \text{ for } k \to \infty$$
and an orthonormal basis \( \{ w_n \} \) of \( V \) consisting of the normalized eigenfunctions associated with \( \lambda_n \):

\[
a(w_n, v) = \lambda_n(w_n, v) \quad \forall v \in V, \|w_n\|_V = 1.
\]

Furthermore the following formula holds true for \( \lambda_1 \):

\[
\lambda_1 = \min_{v \in V, v \neq 0} \frac{a(v, v)}{\|v\|_V^2}.
\]

Now we want to apply Theorem A.1 to the generalized eigenvalue problem (10). Obviously the following inclusion holds true for the bilinear forms defined in (9)

\[
\ker(b_E) \subset \ker(a_E).
\]

Based on this observation, we define a Banach space \( \overline{V} \) as the factor space \( V \setminus \ker(b_E) \). On \( \overline{V} \times \overline{V} \) we further define the inner product

\[
(\overline{u}, \overline{v})_{\overline{V}} := b_E(u, v),
\]

where \( u, v \) are arbitrary representatives of the equivalence classes \( \overline{u} \) and \( \overline{v} \), respectively. The inner product \( (\cdot, \cdot)_{\overline{V}} \) induces the norm \( \|\overline{u}\|_{\overline{V}} := \sqrt{b_E(\overline{u}, \overline{u})} \) on \( \overline{V} \). Consequently, \( \overline{V} \) is a Hilbert space. Next we define the bilinear form

\[
\overline{a}_E(\overline{u}, \overline{v}) := a_E(u, v),
\]
where again \( u, v \) are arbitrary representatives of the equivalence classes \( \overline{u} \) and \( \overline{v} \). Defining the eigenvalue problem: Find \( \lambda \in \mathbb{R} \) and \( u \in \mathcal{V} \), \( u \neq 0 \), such that

\[
\pi_E(\overline{u}, \overline{v}) = \lambda(\overline{u}, \overline{v}) \quad \forall \overline{v} \in \mathcal{V},
\]  

we are able to state the following corollary:

**Corollary A.2.** There exists an increasing sequence of well defined eigenvalues of problem (10) tending to \( \infty \):

\[
0 \leq \lambda_1 \leq \lambda_2 \leq \ldots, \quad \lambda_k \to \infty \quad \text{for} \quad k \to \infty
\]

and associated eigenfunctions \( w_n, n = 1, 2, \ldots \), orthonormal w. r. t. the inner product \((\cdot, \cdot)_\mathcal{V}\), such that:

\[
a_E(w_n, v) = \lambda_n b_E(w_n, v) \quad \forall v \in \mathcal{V}.
\]

Furthermore the following formula holds true for \( \lambda_1 \):

\[
\lambda_1 = \min_{v \in \mathcal{V}, v \notin \ker(b_E)} \frac{a_E(v, v)}{b_E(v, v)} = \min_{v \in \mathcal{V}, \|v\|_2 = 1} a_E(v, v).
\]

**Proof.** Suppose for a moment that \( \pi_E \) is \( \overline{\mathcal{V}} \)-elliptic. Then we are able to apply Theorem A.1 to problem (37) and all assertions of Corollary A.2 follow immediately from (34), (35) and (36). On the other hand, if \( \pi_E \) fails to be \( \overline{\mathcal{V}} \)-elliptic, we define a bilinear form \( \pi'_E \) by

\[
\pi'_E(\overline{u}, \overline{v}) := \pi_E(\overline{u}, \overline{v}) + \mu(u, v)_\mathcal{V}.
\]

Obviously \( \pi'_E \) is \( \overline{\mathcal{V}} \)-elliptic for any (arbitrary small) positive \( \mu \in \mathbb{R} \). Applying Theorem A.1 to \( \overline{\mathcal{V}} \) and \( \pi'_E \) we obtain the following estimates for the eigenvalues of problem (37)

\[
-\mu < \lambda_1 \leq \lambda_2 \leq \ldots, \quad \lambda_k \to \infty \quad \text{for} \quad k \to \infty.
\]

As \( \mu \) can be chosen arbitrarily small we conclude \( 0 \leq \lambda_1 \) and the proof is complete.

Let us finally define the bilinear form \( c_{E,\hat{\lambda}}(u, v) := a_E(u, v) - \hat{\lambda}b_E(u, v) \) and consider the eigenvalue problem:

Find \( \lambda \in \mathbb{R}, u \in \mathcal{V}, u \neq 0 \) such that :

\[
c_{E,\hat{\lambda}}(u, v) = \lambda(u, v) \quad \forall v \in \mathcal{V}.
\]

(38)

Then the following result proves the equivalence of (11) and (12).

**Corollary A.3.** The following assertions are equivalent:

a) The minimal well defined eigenvalue of the generalized eigenvalue problem (10) is nonnegative.

b) The minimal eigenvalue of the eigenvalue problem (38) is nonnegative.

**Proof.** Using Corollary A.2 and equation (34) we have

\[
a) \quad \Rightarrow \quad \frac{a_E(v, v)}{b_E(v, v)} \geq \hat{\lambda} \quad \forall v \in \mathcal{V}, v \notin \ker(b_E) \\
\Rightarrow \quad a_E(v, v) - \hat{\lambda}b_E(v, v) \geq 0 \quad \forall v \in \mathcal{V}, v \notin \ker(b_E) \\
(34) \quad \Rightarrow \quad a_E(v, v) - \hat{\lambda}b_E(v, v) \geq 0 \quad \forall v \in \mathcal{V}, v \neq 0 \\
\Rightarrow \quad \frac{a_E(v, v) - \hat{\lambda}b_E(v, v)}{\|v\|^2} \geq 0 \quad \forall v \in \mathcal{V}, v \neq 0 \\
\Rightarrow \quad a_E(v, v) - \hat{\lambda}b_E(v, v) \geq 0 \quad \forall v \in \mathcal{V}, \|v\| = 1 \quad \Leftrightarrow \quad b).
\]

\]
Appendix II

In the sequel we give a proof of Proposition 5.2:

Proof. We start with the KKT conditions of problem (24) in standard form: Let

\[
F := \{ (u, \alpha, W, \beta^l, \beta^u) \in \mathbb{R}^{N+1} \times \mathbb{S}^N \times \mathbb{R}^{2M} \mid W \succ 0, \beta^l \geq 0, \beta^u \geq 0 \}
\]

and \(x^* := (u^*, \alpha^*, W^*, \beta^l^*, \beta^u^*) \in F\). Then a pair \((x^*, E^*)\) is a KKT point of (24) if and only if the conditions

\[
g_i(x^*) \preceq 0, \quad E_i^* \succeq 0, \quad i = 1, 2, \ldots, M,
\]

\[
\nabla_{u, \alpha} f(x^*) - \sum_{i=1}^{M} (E_i^*, \nabla_{u, \alpha} g_i(x^*)) = 0,
\]

\[
(W^*, \beta^l^*, \beta^u^*)^\top - \text{Proj}_F \left( (W^*, \beta^l^*, \beta^u^*)^\top - \nabla_{W, \beta^l, \beta^u} f(x^*) \right)
\]

\[
- \sum_{i=1}^{M} (E_i^*, \nabla_{W, \beta^l, \beta^u} g_i(x^*)) = 0,
\]

\[
\left\langle E_i^*, g_i(x^*) \right\rangle = 0, \quad i = 1, 2, \ldots, M,
\]

are satisfied. Now we have

- \(\frac{\partial}{\partial u} f(x^*) = 2f\),
- \(\frac{\partial}{\partial \alpha} f(x^*) = -V\),
- \(\frac{\partial}{\partial W} f(x^*) = 0\),
- \(\frac{\partial}{\partial \beta^l_i} f(x^*) = \rho_i, \quad (i=1,2,\ldots,M)\),
- \(\frac{\partial}{\partial \beta^u_i} f(x^*) = -\bar{\rho}, \quad (i=1,2,\ldots,M)\),
- \(\frac{\partial}{\partial \alpha} \left( \sum_{i=1}^{M} \left( E_i^*, g_i(x^*) \right) \right) = \sum_{i=1}^{M} \sum_{j=1}^{G} \frac{\partial}{\partial \alpha} \left( E_i^*, B_{ij}^T uu^T B_{ij} \right)|_{u=u^*} = \sum_{i=1}^{M} \sum_{j=1}^{G} 2B_{ij}E_i^*B_{ij}^T u^* = 2A(E^*)u^*\),
- \(\frac{\partial}{\partial \alpha} \left( \sum_{i=1}^{M} \left( E_i^*, \nabla g_i(x^*) \right) \right) = \sum_{i=1}^{M} (E_i^*, -I) = -\sum_{i=1}^{M} \text{Tr}(E_i^*)\),
• \[ \frac{\partial}{\partial W} \left( \sum_{i=1}^M (E_i^*, \nabla g_i(x^*)) \right) = \right. \]
\[ = \sum_{i=1}^M (E_i^*, \frac{\partial}{\partial W} (\sum_{j=1}^G B_{ij}^T W B_{ij}) - \lambda^* (W, M_i) I) \big|_{W=W^*} = \right. \]
\[ = \sum_{i=1}^M \sum_{j=1}^G B_{ij} E_i^* B_{ij}^T - \lambda^* M_i E_i^*, I) = A(E^*) - \lambda^* M(E^*), \]

• \[ \frac{\partial}{\partial W} \left( \sum_{i=1}^M (E_i^*, \nabla g_i(x^*)) \right) = \langle E_i^*, -I \rangle = -\text{Tr}(E_i^*), \]

• \[ \frac{\partial}{\partial \beta} \left( \sum_{i=1}^M (E_i^*, \nabla g_i(x^*)) \right) = \langle E_i^*, I \rangle = \text{Tr}(E_i^*). \]

Hence (40) is equivalent to
\[ A(E^*) u^* = f, \quad \sum_{i=1}^M \text{Tr}(E_i^*) = V, \]

and (41) is equivalent to
\[ \rho \leq \text{Tr}(E_i^*) \leq \overline{\rho} \quad (i = 1, 2, \ldots, M), \]

\[ \beta^T \left( \rho - \text{Tr}(E_i^*) \right) = 0, \quad \beta^T u^* (\text{Tr}(E_i^*) - \overline{\rho}) = 0, \quad (i = 1, 2, \ldots, M), \]

\[ A(E^*) - \lambda^* M(E^*) \succ 0, \quad \langle A(E^*) - \lambda^* M(E^*), W^* \rangle = 0. \]

Next, we see from (39) that (42) is equivalent to
\[ \sum_{i=1}^M \langle E_i^*, g_i(x^*) \rangle = 0. \]

We further calculate
\[ \sum_{i=1}^M \langle E_i^*, g_i(x^*) \rangle = \]
\[ = \sum_{i=1}^M \left( \langle E_i^*, \sum_{j=1}^G B_{ij}^T (u^* u^T + W^*) B_{ij} - \lambda^* M_i \rangle + \alpha^* + \beta^T - \beta^u \rangle I \right) \]
\[ = -\alpha^* \sum_{i=1}^M \text{Tr}(E_i^*) + u^T A(E^*) u^* + (W^*, A(E^*) - \tilde{\lambda} M(E^*)) \]
\[ + \sum_{i=1}^M \beta^T \text{Tr}(E_i^*) - \sum_{i=1}^M \beta^T \text{Tr}(E_i^*) \]
\[ = -\alpha^* V + f^T u^* + \rho \sum_{i=1}^M \beta^T - \overline{\rho} \sum_{i=1}^M \beta^u. \]

and the proof is complete. \(\Box\)
Acknowledgements

This work has been partially supported by the EU Commission in the Sixth Framework Program, Project No. 30717 PLATO-N.

References


