A Numerical Algorithm for Block-Diagonal Decomposition of Matrix $*$-Algebras, Part I: Proposed Approach and Application to Semidefinite Programming*

Kazuo Murota[†] Yoshihiro Kanno[‡]
Masakazu Kojima[§] Sadayoshi Kojima[¶]

September 2007 / June 2008 / May 2009

Abstract

Motivated by recent interest in group-symmetry in semidefinite programming, we propose a numerical method for finding a finest simultaneous block-diagonalization of a finite number of matrices, or equivalently the irreducible decomposition of the generated matrix $*$-algebra. The method is composed of numerical-linear algebraic computations such as eigenvalue computation, and automatically makes full use of the underlying algebraic structure, which is often an outcome of physical or geometrical symmetry, sparsity, and structural or numerical degeneracy in the given matrices. The main issues of the proposed approach are presented in this paper under some assumptions, while the companion paper (Part II) gives an algorithm with full generality. Numerical examples of truss and frame designs are also presented.

Keywords: matrix $*$-algebra, block-diagonalization, group symmetry, sparsity, semidefinite programming

---

*This is a revised manuscript of the technical report "A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras," issued in September 2007 as METR 2007-52, Department of Mathematical Informatics, University of Tokyo, and also as Research Report B-445, Department of Mathematical and Computing Sciences, Tokyo Institute of Technology.

[†]Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan. murota@mist.i.u-tokyo.ac.jp

[‡]Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo, Tokyo 113-8656, Japan. kanno@mist.i.u-tokyo.ac.jp

[§]Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo 152-8552, Japan. kojima@is.titech.ac.jp

[¶]Department of Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo 152-8552, Japan. sadayosi@is.titech.ac.jp

1 Introduction

This paper is motivated by recent studies on group symmetries in semidefinite programs (SDPs) and sum of squares (SOS) and SDP relaxations [1, 5, 7, 12, 14]. A common and essential problem in these studies can be stated as follows: Given a finite set of $n \times n$ real symmetric matrices $A_1, A_2, \ldots, A_N$, find an $n \times n$ orthogonal matrix P that provides them with a simultaneous block-diagonal decomposition, i.e., such that $P^\top A_1 P, P^\top A_2 P, \ldots, P^\top A_N P$ become block-diagonal matrices with a common block-diagonal structure. Here $A_1, A_2, \ldots, A_N$ correspond to data matrices associated with an SDP. We say that the set of given matrices $A_1, A_2, \ldots, A_N$ is decomposed into a set of block-diagonal matrices or that the SDP is decomposed into an SDP with the block-diagonal data matrices. Such a block-diagonal decomposition is not unique in general; for example, any symmetric matrix may trivially be regarded as a one-block matrix. As diagonal-blocks of the decomposed matrices get smaller, the transformed SDP could be solved more efficiently by existing software packages developed for SDPs [3, 28, 29, 34]. Naturally we are interested in a finest decomposition. A more specific account of the decomposition of SDPs will be given in Section 2.1.

There are two different but closely related theoretical frameworks with which we can address our problem of finding a block-diagonal decomposition for a finite set of given $n \times n$ real matrices. The one is group representation theory [23, 27] and the other matrix $*$-algebra [32]. They are not only necessary to answer the fundamental theoretical question of the existence of such a finest block-diagonal decomposition but also useful in its computation. Both frameworks have been utilized in the literature [1, 5, 7, 12, 14] cited above.

Kanno et al. [14] introduced a class of group symmetric SDPs, which arise from topology optimization problems of trusses, and derived symmetry of central paths which play a fundamental role in the primal-dual interior-point method [33] for solving them. Gatermann and Parrilo [7] investigated the problem of minimizing a group symmetric polynomial. They proposed to reduce the size of SOS–SDP relaxations for the problem by exploiting the group symmetry and decomposing the SDP. On the other hand, de Klerk et al. [4] applied the theory of matrix $*$-algebra to reduce the size of a class of group symmetric SDPs. Instead of decomposing a given SDP into a block-diagonal form by using its group symmetry, their method transforms the problem to an equivalent SDP through a $*$-algebra isomorphism. We also refer to Kojima et al. [16] as a paper where matrix $*$-algebra was studied in connection with SDPs. Jansson et al. [12] brought group symmetries into equality-inequality constrained polynomial optimization problems and their SDP relaxation. More recently, de Klerk and Sotirov [5] dealt with quadratic assignment problems, and showed how to exploit their group symmetries to reduce the size of their SDP relaxations (see Remark 4.3 for more details).

All existing studies [1, 5, 7, 12] on group symmetric SDPs mentioned above assume that the algebraic structure such as group symmetry and matrix $*$-algebra behind a given SDP is known in advance before computing a decomposition of the SDP. Such an algebraic structure arises naturally from the physical or geometrical structure underlying the SDP, and so the assumption is certainly practical and reasonable. When we assume symmetry of an SDP (or the data matrices $A_1, A_2, \ldots, A_N$) with reference to a group G, to be specific, we are in fact considering the class of SDPs that enjoy the same group symmetry. As a consequence, the resulting transformation matrix P is universal in the sense that it is valid for the decomposition of all SDPs belonging to the class. This universality is often useful, but at the same time we should note that the given SDP is just a specific instance in the class. A further decomposition may possibly be obtained by exploiting an additional algebraic structure, if any, which is not captured by the assumed group symmetry but possessed by the given problem. Such an additional algebraic structure is often induced from sparsity of the data matrices of the SDP, as we see in the topology optimization problem of trusses in Section 5. The possibility of a further decomposition due to sparsity will be illustrated in Sections 2.2 and 5.1.

In the present papers, consisting of Parts I and II, we propose a numerical method for finding a finest simultaneous block-diagonal decomposition of a finite number of $n \times n$ real matrices $A_1, A_2, \ldots, A_N$. The method does not require any algebraic structure to be known in advance, and is based on numerical linear algebraic computations such as eigenvalue computation. It is free from group representation theory or matrix $*$-algebra during its execution, although its validity relies on matrix $*$-algebra theory. This main feature of our method makes it possible to compute a finest block-diagonal decomposition by taking into account the underlying physical or geometrical symmetry, the sparsity of the given matrices, and some other implicit or overlooked symmetry.

Our method is based on the following ideas. We consider the matrix $*$-algebra $\mathcal{T}$ generated by $A_1, A_2, \ldots, A_N$ with the identity matrix, and make use of a well-known fundamental fact (see Theorem 3.1) about the decomposition of $\mathcal{T}$ into simple components and irreducible components. The key observation is that the decomposition into simple components can be computed from the eigenvalue (or spectral) decomposition of a randomly chosen symmetric matrix in $\mathcal{T}$. Once the simple components are identified, the decomposition into irreducible components can be obtained by "local" coordinate changes within each eigenspace, to be explained in Section 3. In Part I we present the main issues of the proposed approach by considering a special case where (i) the given matrices $A_1, A_2, \ldots, A_N$ are symmetric and (ii) each irreducible component of $\mathcal{T}$ is isomorphic to a full matrix algebra of some order (i.e., of type $\mathbb{R}$ to be defined in Section 3.1). The general case, technically more involved, will be covered by Part II of this paper.

Part I of this paper is organized as follows. Section 2 illustrates our motivation of simultaneous block-diagonalization and the notion of the finest block-diagonal decomposition. Section 3 describes the theoretical background of our algorithm based on matrix $*$-algebra. In Section 4, we present an algorithm for computing the finest simultaneous block-diagonalization, as well as a suggested practical variant thereof. Numerical results are shown in Section 5; Section 5.1 gives illustrative small examples, Section 5.2 shows SDP problems arising from topology optimization of symmetric trusses, and Section 5.3 deals with a quadratic SDP problem arising from topology optimization of symmetric frames.

2 Motivation

2.1 Decomposition of semidefinite programs

In this section it is explained how simultaneous block diagonalization can be utilized in semidefinite programming.

Let $A_p \in \mathcal{S}_n$ $(p = 0, 1, \ldots, m)$ and $b = (b_p)_{p=1}^m \in \mathbb{R}^m$ be given matrices and a given vector, where $\mathcal{S}_n$ denotes the set of $n \times n$ symmetric real matrices. The standard form of a primal-dual pair of *semidefinite programming* (SDP) problems can be formulated as

$$\left.\begin{array}{ll} \min & A_0 \bullet X \\ \text{s.t.} & A_p \bullet X = b_p, \quad p = 1, \ldots, m, \\ & \mathcal{S}_n \ni X \succeq O; \end{array}\right\} \qquad (2.1)$$

$$\left.\begin{array}{ll} \max & b^\top y \\ \text{s.t.} & Z + \displaystyle\sum_{p=1}^{m} A_p y_p = A_0, \\ & \mathcal{S}_n \ni Z \succeq O. \end{array}\right\} \qquad (2.2)$$

Here X is the decision (or optimization) variable in (2.1), Z and y_p $(p = 1, \ldots, m)$ are the decision variables in (2.2), $A \bullet X = \mathrm{tr}(AX)$ for symmetric matrices A and X, $X \succeq O$ means that X is positive semidefinite, and $^\top$ denotes the transpose of a vector or a matrix.

Suppose that $A_0, A_1, \ldots, A_m$ are transformed into block-diagonal matrices by an $n \times n$ orthogonal matrix P as

$$P^\top A_p P = \begin{pmatrix} A_p^{(1)} & O \\ O & A_p^{(2)} \end{pmatrix}, \quad p = 0, 1, \ldots, m,$$

where $A_p^{(1)} \in \mathcal{S}_{n_1}$, $A_p^{(2)} \in \mathcal{S}_{n_2}$, and $n_1 + n_2 = n$. The problems (2.1) and

(2.2) can be reduced to

$$\left.\begin{array}{ll} \min & A_0^{(1)} \bullet X_1 + A_0^{(2)} \bullet X_2 \\ \text{s.t.} & A_p^{(1)} \bullet X_1 + A_p^{(2)} \bullet X_2 = b_p, \quad p = 1, \ldots, m, \\ & \mathcal{S}_{n_1} \ni X_1 \succeq O, \quad \mathcal{S}_{n_2} \ni X_2 \succeq O; \end{array}\right\} \tag{2.3}$$

$$\left.\begin{array}{ll} \max & b^\top y \\ \text{s.t.} & Z_1 + \sum_{p=1}^{m} A_p^{(1)} y_p = A_0^{(1)}, \\ & Z_2 + \sum_{p=1}^{m} A_p^{(2)} y_p = A_0^{(2)}, \\ & \mathcal{S}_{n_1} \ni Z_1 \succeq O, \quad \mathcal{S}_{n_2} \ni Z_2 \succeq O. \end{array}\right\} \tag{2.4}$$

Note that the number of variables of (2.3) is smaller than that of (2.1). The constraint on the $n \times n$ symmetric matrix in (2.2) is reduced to the constraints on the two matrices in (2.4) with smaller sizes.

It is expected that the computational time required by the primal-dual interior-point method is reduced significantly if the problems (2.1) and (2.2) can be reformulated as (2.3) and (2.4). This motivates us to investigate a numerical technique for computing a simultaneous block diagonalization in the form of

$$P^\top A_p P = \mathrm{diag}\,(A_p^{(1)}, A_p^{(2)}, \ldots, A_p^{(t)}) = \bigoplus_{j=1}^{t} A_p^{(j)}, \quad A_p^{(j)} \in \mathcal{S}_{n_j}, \tag{2.5}$$

where $A_p \in \mathcal{S}_n$ $(p = 0, 1, \ldots, m)$ are given symmetric matrices. Here $\bigoplus$ designates a direct sum of the summand matrices, which contains the summands as diagonal blocks.

2.2 Group symmetry and additional structure due to sparsity

With reference to a concrete example, we illustrate the use of group symmetry and also the possibility of a finer decomposition based on an additional algebraic structure due to sparsity.

Consider an $n \times n$ matrix of the form

$$A = \begin{bmatrix} B & E & E & C \\ E & B & E & C \\ E & E & B & C \\ C^\top & C^\top & C^\top & D \end{bmatrix} \tag{2.6}$$

with an $n_{\mathrm{B}} \times n_{\mathrm{B}}$ symmetric matrix $B \in \mathcal{S}_{n_{\mathrm{B}}}$ and an $n_{\mathrm{D}} \times n_{\mathrm{D}}$ symmetric matrix $D \in \mathcal{S}_{n_{\rm D}}$. Obviously we have $A = A_1 + A_2 + A_3 + A_4$ with

$$A_1 = \begin{bmatrix} B & O & O & O \\ O & B & O & O \\ O & O & B & O \\ O & O & O & O \end{bmatrix}, \quad A_2 = \begin{bmatrix} O & O & O & C \\ O & O & O & C \\ O & O & O & C \\ C^\top & C^\top & C^\top & O \end{bmatrix}, \tag{2.7}$$

$$A_3 = \begin{bmatrix} O & O & O & O \\ O & O & O & O \\ O & O & O & O \\ O & O & O & D \end{bmatrix}, \quad A_4 = \begin{bmatrix} O & E & E & O \\ E & O & E & O \\ E & E & O & O \\ O & O & O & O \end{bmatrix}. \tag{2.8}$$

Let P be an $n \times n$ orthogonal matrix defined by

$$P = \left[\begin{array}{cc|cc} I_{n_{\rm B}}/\sqrt{3} & O & I_{n_{\rm B}}/\sqrt{2} & I_{n_{\rm B}}/\sqrt{6} \\ I_{n_{\rm B}}/\sqrt{3} & O & -I_{n_{\rm B}}/\sqrt{2} & I_{n_{\rm B}}/\sqrt{6} \\ I_{n_{\rm B}}/\sqrt{3} & O & O & -2I_{n_{\rm B}}/\sqrt{6} \\ O & I_{n_{\rm D}} & O & O \end{array}\right], \tag{2.9}$$

where $I_{n_{\rm B}}$ and $I_{n_{\rm D}}$ denote identity matrices of orders $n_{\rm B}$ and $n_{\rm D}$, respectively. With this P the matrices A_p are transformed to block-diagonal matrices as

$$P^\top A_1 P = \left[\begin{array}{cc|c|c} B & O & O & O \\ O & O & O & O \\ \hline O & O & B & O \\ \hline O & O & O & B \end{array}\right] = \begin{bmatrix} B & O \\ O & O \end{bmatrix} \oplus B \oplus B, \tag{2.10}$$

$$P^\top A_2 P = \left[\begin{array}{cc|c|c} O & \sqrt{3}C & O & O \\ \sqrt{3}C^\top & O & O & O \\ \hline O & O & O & O \\ \hline O & O & O & O \end{array}\right] = \begin{bmatrix} O & \sqrt{3}C \\ \sqrt{3}C^\top & O \end{bmatrix} \oplus O \oplus O, \tag{2.11}$$

$$P^\top A_3 P = \left[\begin{array}{cc|c|c} O & O & O & O \\ O & D & O & O \\ \hline O & O & O & O \\ \hline O & O & O & O \end{array}\right] = \begin{bmatrix} O & O \\ O & D \end{bmatrix} \oplus O \oplus O, \tag{2.12}$$

$$P^\top A_4 P = \left[\begin{array}{cc|c|c} 2E & O & O & O \\ O & O & O & O \\ \hline O & O & -E & O \\ \hline O & O & O & -E \end{array}\right] = \begin{bmatrix} 2E & O \\ O & O \end{bmatrix} \oplus (-E) \oplus (-E). \tag{2.13}$$

Note that the partition of P is not symmetric for rows and columns; we have $(n_{\rm B}, n_{\rm B}, n_{\rm B}, n_{\rm D})$ for row-block sizes and $(n_{\rm B}, n_{\rm D}, n_{\rm B}, n_{\rm B})$ for column-block sizes. As is shown in (2.10)–(2.13), A_1, A_2, A_3 and A_4 are decomposed simultaneously in the form of (2.5) with $t = 3$, $n_1 = n_{\rm B} + n_{\rm D}$, and $n_2 = n_3 = n_{\rm B}$. Moreover, the second and third blocks coincide, i.e., $A_p^{(2)} = A_p^{(3)}$, for each p.

The decomposition described above coincides with the standard decomposition [23, 27] for systems with group symmetry. The matrices A_p above are symmetric with respect to S_3, the symmetric group of order $3! = 6$, in that

$$T(g)^\top A_p T(g) = A_p, \quad \forall g \in G, \quad \forall p \tag{2.14}$$

holds for $G = S_3$. Here the family of matrices $T(g)$, indexed by elements of G, is an orthogonal matrix representation of G in general. In the present example, the S_3-symmetry formulated in (2.14) is equivalent to

$$T_i^\top A_p T_i = A_p, \quad i = 1, 2, \quad p = 1, 2, 3, 4$$

with

$$T_1 = \begin{bmatrix} O & I_{n_B} & O & O \\ I_{n_B} & O & O & O \\ O & O & I_{n_B} & O \\ O & O & O & I_{n_D} \end{bmatrix}, \quad T_2 = \begin{bmatrix} O & I_{n_B} & O & O \\ O & O & I_{n_B} & O \\ I_{n_B} & O & O & O \\ O & O & O & I_{n_D} \end{bmatrix}.$$

According to group representation theory, a simultaneous block-diagonal decomposition of A_p is obtained through the decomposition of the representation T into irreducible representations. In the present example, we have

$$P^\top T_1 P = \left[\begin{array}{cc|cc} I_{n_B} & O & O & O \\ O & I_{n_D} & O & O \\ \hline O & O & -I_{n_B} & O \\ O & O & O & I_{n_B} \end{array}\right], \tag{2.15}$$

$$P^\top T_2 P = \left[\begin{array}{cc|cc} I_{n_B} & O & O & O \\ O & I_{n_D} & O & O \\ \hline O & O & -I_{n_B}/2 & \sqrt{3} I_{n_B}/2 \\ O & O & -\sqrt{3} I_{n_B}/2 & -I_{n_B}/2 \end{array}\right], \tag{2.16}$$

where the first two blocks correspond to the unit (or trivial) representation (with multiplicity $n_B + n_D$) and the last two blocks to the two-dimensional irreducible representation (with multiplicity n_B).

The transformation matrix P in (2.9) is universal in the sense that it brings any matrix A satisfying $T_i^\top A T_i = A$ for $i = 1, 2$ into the same block-diagonal form. Put otherwise, the decomposition given in (2.10)–(2.13) is the finest possible decomposition that is valid for the class of matrices having the S_3-symmetry. It is noted in this connection that the underlying group G, as well as its representation $T(g)$, is often evident in practice, reflecting the geometrical or physical symmetry of the problem in question.

The universality of the decomposition explained above is certainly a nice feature of the group-theoretic method, but what we need is the decomposition of a single specific instance of a set of matrices. For example suppose that $E = O$ in (2.6). Then the decomposition in (2.10)–(2.13) is not the finest possible, but the last two identical blocks, i.e., $A_p^{(2)}$ and $A_p^{(3)}$, can be decomposed further into diagonal matrices by the eigenvalue (or spectral) decomposition of B. Although this example is too simple to be convincing, it is sufficient to suggest the possibility that a finer decomposition may possibly be obtained from an additional algebraic structure that is not ascribed to the assumed group symmetry. Such an additional algebraic structure often stems from sparsity, as is the case with the topology optimization problem of trusses treated in Section 5.2.

Mathematically, such an additional algebraic structure could also be described as a group symmetry by introducing a larger group. This larger group, however, would be difficult to identify in practice, since it is determined as a result of the interaction between the underlying geometrical or physical symmetry and other factors, such as sparsity and parameter dependence. The method of block-diagonalization proposed here will automatically exploit such algebraic structure in the course of numerical computation. Numerical examples in Section 5.1 will demonstrate that the proposed method can cope with different kinds of additional algebraic structures for the matrix (2.6).

3 Mathematical Basis

We introduce some mathematical facts that will serve as a basis for our algorithm.

3.1 Matrix $*$-algebras

Let $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ be the real number field, the complex field, and the quaternion field, respectively. The quaternion field $\mathbb{H}$ is a vector space $\{a + \imath b + \jmath c + kd : a, b, c, d \in \mathbb{R}\}$ over $\mathbb{R}$ with basis $1, \imath, \jmath$ and k, equipped with the multiplication defined as follows:

$$\imath = \jmath k = -k\jmath, \quad \jmath = k\imath = -\imath k, \quad k = \imath\jmath = -\jmath\imath, \quad \imath^2 = \jmath^2 = k^2 = -1$$

and for all $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ and $x, y, u, v \in \mathbb{H}$,

$$(\alpha x + \beta y)(\gamma u + \delta v) = \alpha\gamma xu + \alpha\delta xv + \beta\gamma yu + \beta\delta yv.$$

For a quaternion $h = a + \imath b + \jmath c + kd$, its conjugate is defined as $\bar{h} = a - \imath b - \jmath c - kd$, and the norm of h is defined as $|h| = \sqrt{h\bar{h}} = \sqrt{\bar{h}h} = \sqrt{a^2 + b^2 + c^2 + d^2}$. We can consider $\mathbb{C}$ as a subset of $\mathbb{H}$ by identifying the generator $\imath$ of the quaternion field $\mathbb{H}$ with the imaginary unit of the complex field $\mathbb{C}$.

Let $\mathcal{M}_n$ denote the set of $n \times n$ real matrices over $\mathbb{R}$. A subset $\mathcal{T}$ of $\mathcal{M}_n$ is said to be a *$*$-subalgebra* (or a *matrix $*$-algebra*) over $\mathbb{R}$ if $I_n \in \mathcal{T}$ and

$$A, B \in \mathcal{T}; \alpha, \beta \in \mathbb{R} \implies \alpha A + \beta B, AB, A^\top \in \mathcal{T}. \tag{3.1}$$

Obviously, $\mathcal{M}_n$ itself is a matrix $*$-algebra. There are two other basic matrix $*$-algebras: the real representation of complex matrices $\mathcal{C}_n \subset \mathcal{M}_{2n}$ defined by

$$\mathcal{C}_n = \left\{ \begin{bmatrix} C(z_{11}) & \cdots & C(z_{1n}) \\ \vdots & \ddots & \vdots \\ C(z_{n1}) & \cdots & C(z_{nn}) \end{bmatrix} : z_{11}, z_{12}, \ldots, z_{nn} \in \mathbb{C} \right\}$$

with

$$C(a + \imath b) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix},$$

and the real representation of quaternion matrices $\mathcal{H}_n \subset \mathcal{M}_{4n}$ defined by

$$\mathcal{H}_n = \left\{ \begin{bmatrix} H(h_{11}) & \cdots & H(h_{1n}) \\ \vdots & \ddots & \vdots \\ H(h_{n1}) & \cdots & H(h_{nn}) \end{bmatrix} : h_{11}, h_{12}, \ldots, h_{nn} \in \mathbb{H} \right\}$$

with

$$H(a + \imath b + \jmath c + kd) = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}.$$

For two matrices A and B, their direct sum, denoted as $A \oplus B$, is defined as

$$A \oplus B = \begin{bmatrix} A & O \\ O & B \end{bmatrix},$$

and their tensor product, denoted as $A \otimes B$, is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nn}B \end{bmatrix},$$

where A is assumed to be $n \times n$. Note that $A \otimes B = \Pi^\top (B \otimes A) \Pi$ for some permutation matrix Π.

We say that a matrix $*$-algebra $\mathcal{T}$ is *simple* if $\mathcal{T}$ has no ideal other than $\{O\}$ and $\mathcal{T}$ itself, where an *ideal* of $\mathcal{T}$ means a submodule $\mathcal{I}$ of $\mathcal{T}$ such that

$$A \in \mathcal{T}, B \in \mathcal{I} \implies AB, BA \in \mathcal{I}.$$

A linear subspace W of $\mathbb{R}^n$ is said to be *invariant* with respect to $\mathcal{T}$, or $\mathcal{T}$*-invariant*, if $AW \subseteq W$ for every $A \in \mathcal{T}$. We say that $\mathcal{T}$ is *irreducible* if no $\mathcal{T}$-invariant subspace other than $\{\mathbf{0}\}$ and $\mathbb{R}^n$ exists. It is mentioned that $\mathcal{M}_n$, $\mathcal{C}_n$ and $\mathcal{H}_n$ are typical examples of irreducible matrix $*$-algebras. If $\mathcal{T}$ is irreducible, it is simple (cf. Lemma A.4).

We say that matrix $*$-algebras $\mathcal{T}_1$ and $\mathcal{T}_2$ are *isomorphic* if there exists a bijection ϕ from $\mathcal{T}_1$ to $\mathcal{T}_2$ with the following properties:

$$\phi(\alpha A + \beta B) = \alpha\phi(A) + \beta\phi(B), \quad \phi(AB) = \phi(A)\phi(B), \quad \phi(A^\top) = \phi(A)^\top.$$

If $\mathcal{T}_1$ and $\mathcal{T}_2$ are isomorphic, we write $\mathcal{T}_1 \simeq \mathcal{T}_2$. For a matrix $*$-algebra $\mathcal{T}$ and an orthogonal matrix P, the set

$$P^\top \mathcal{T} P = \{P^\top A P : A \in \mathcal{T}\}$$

forms another matrix $*$-algebra isomorphic to $\mathcal{T}$. For a matrix $*$-algebra $\mathcal{T}'$, the set

$$\mathcal{T} = \{\mathrm{diag}\,(B, B, \ldots, B) : B \in \mathcal{T}'\}$$

forms another matrix $*$-algebra isomorphic to $\mathcal{T}'$.

From a standard result of the theory of matrix $*$-algebra (e.g., [32, Chapter X]) we can see the following structure theorem for a matrix $*$-subalgebra over $\mathbb{R}$. This theorem is stated in [16, Theorem 5.4] with a proof, but, in view of its fundamental role in this paper, we give an alternative proof in Appendix.

Theorem 3.1. *Let $\mathcal{T}$ be a $*$-subalgebra of $\mathcal{M}_n$ over $\mathbb{R}$.*

(A) *There exist an orthogonal matrix $\hat{Q} \in \mathcal{M}_n$ and simple $*$-subalgebras $\mathcal{T}_j$ of $\mathcal{M}_{\hat{n}_j}$ for some $\hat{n}_j$ $(j = 1, 2, \ldots, \ell)$ such that*

$$\hat{Q}^\top \mathcal{T} \hat{Q} = \{\mathrm{diag}\,(S_1, S_2, \ldots, S_\ell) : S_j \in \mathcal{T}_j \ (j = 1, 2, \ldots, \ell)\}.$$

(B) *If $\mathcal{T}$ is simple, there exist an orthogonal matrix $P \in \mathcal{M}_n$ and an irreducible $*$-subalgebra $\mathcal{T}'$ of $\mathcal{M}_{\bar{n}}$ for some $\bar{n}$ such that*

$$P^\top \mathcal{T} P = \{\mathrm{diag}\,(B, B, \ldots, B) : B \in \mathcal{T}'\}.$$

(C) *If $\mathcal{T}$ is irreducible, there exists an orthogonal matrix $P \in \mathcal{M}_n$ such that $P^\top \mathcal{T} P = \mathcal{M}_n, \mathcal{C}_{n/2}$ or $\mathcal{H}_{n/4}$.* ■

The three cases in Theorem 3.1(C) above will be referred to as case $\mathbb{R}$, case $\mathbb{C}$ or case $\mathbb{H}$, according to whether $\mathcal{T} \simeq \mathcal{M}_n$, $\mathcal{T} \simeq \mathcal{C}_{n/2}$ or $\mathcal{T} \simeq \mathcal{H}_{n/4}$. We also speak of type $\mathbb{R}$, type $\mathbb{C}$ or type $\mathbb{H}$ for an irreducible matrix $*$-algebra.

It follows from the above theorem that, with a single orthogonal matrix P, all the matrices in $\mathcal{T}$ can be transformed simultaneously to a block-diagonal form as

$$P^\top A P = \bigoplus_{j=1}^{\ell} \bigoplus_{i=1}^{\bar{m}_j} B_j = \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes B_j) \tag{3.2}$$

with $B_j \in \mathcal{T}'_j$, where $\mathcal{T}'_j$ denotes the irreducible $*$-subalgebra corresponding to the simple subalgebra $\mathcal{T}_j$; we have $\mathcal{T}'_j = \mathcal{M}_{\bar{n}_j}, \mathcal{C}_{\bar{n}_j/2}$ or $\mathcal{H}_{\bar{n}_j/4}$ for some

$\bar{n}_j$, where the structural indices ℓ, $\bar{n}_j$, $\bar{m}_j$ and the algebraic structure of $\mathcal{T}'_j$ for $j = 1, \ldots, \ell$ are uniquely determined by $\mathcal{T}$. It may be noted that $\hat{n}_j$ in Theorem 3.1 (A) is equal to $\bar{m}_j \bar{n}_j$ in the present notation. Conversely, for any choice of $B_j \in \mathcal{T}'_j$ for $j = 1, \ldots, \ell$, the matrix of (3.2) belongs to $P^\top \mathcal{T} P$.

We denote by

$$\mathbb{R}^n = \bigoplus_{j=1}^{\ell} U_j \tag{3.3}$$

the decomposition of $\mathbb{R}^n$ that corresponds to the simple components. In other words, $U_j = \mathrm{Im}(\hat{Q}_j)$ for the $n \times \hat{n}_j$ submatrix $\hat{Q}_j$ of $\hat{Q}$ that corresponds to $\mathcal{T}_j$ in Theorem 3.1 (A). Although the matrix $\hat{Q}$ is not unique, the subspace U_j is determined uniquely and $\dim U_j = \hat{n}_j = \bar{m}_j \bar{n}_j$ for $j = 1, \ldots, \ell$.

In Part I of this paper we consider the case where

$$\text{(i) } A_1, A_2, \ldots, A_N \text{ are symmetric matrices, and} \tag{3.4}$$

$$\text{(ii) Each irreducible component of } \mathcal{T} \text{ is of type } \mathbb{R}. \tag{3.5}$$

By so doing we can we present the main ideas of the proposed approach more clearly without involving complicating technical issues. An algorithm that works in the general case will be given in Part II of this paper [22].

Remark 3.1. Case $\mathbb{R}$ seems to be the primary case in engineering applications. For instance the $\mathrm{T_d}$-symmetric truss treated in Section 5.2 falls into this category. When the $*$-algebra $\mathcal{T}$ is given as the family of matrices invariant to a group G as $\mathcal{T} = \{A \mid T(g)^\top A T(g) = A, \ \forall g \in G\}$ for some orthogonal representation T of G, case $\mathbb{R}$ is guaranteed if every real-irreducible representation of G is absolutely irreducible. Dihedral groups and symmetric groups, appearing often in applications, have this property. The achiral tetrahedral group $\mathrm{T_d}$ is also such a group. ■

Remark 3.2. Throughout this paper we assume that the underlying field is the field $\mathbb{R}$ of real numbers. In particular, we consider SDP problems (2.1) and (2.2) defined by real symmetric matrices A_p $(p = 0, 1, \ldots, m)$, and accordingly the $*$-algebra $\mathcal{T}$ generated by these matrices over $\mathbb{R}$. An alternative approach is to formulate everything over the field $\mathbb{C}$ of complex numbers, as, e.g., in [30]. This possibility is discussed in Section 6. ■

3.2 Simple components from eigenspaces

Let $A_1, \ldots, A_N \in \mathcal{S}_n$ be $n \times n$ symmetric real matrices, and $\mathcal{T}$ be the $*$-subalgebra over $\mathbb{R}$ generated by $\{I_n, A_1, \ldots, A_N\}$. Note that (3.2) holds for every $A \in \mathcal{T}$ if and only if (3.2) holds for $A = A_p$ for $p = 1, \ldots, N$.

A key observation for our algorithm is that the decomposition (3.3) into simple components can be computed from the eigenvalue (or spectral) decomposition of a single matrix A in $\mathcal{T} \cap \mathcal{S}_n$ if A is sufficiently generic with respect to eigenvalues.

Let A be a symmetric matrix in $\mathcal{T}$, $\alpha_1, \ldots, \alpha_k$ be the distinct eigenvalues of A with multiplicities denoted as $m_1, \ldots, m_k$, and $Q = [Q_1, \ldots, Q_k]$ be an orthogonal matrix consisting of the eigenvectors, where Q_i is an $n \times m_i$ matrix for $i = 1, \ldots, k$. Then we have

$$Q^\top A Q = \mathrm{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k}). \tag{3.6}$$

Put $K = \{1, \ldots, k\}$ and for $i \in K$ define $V_i = \mathrm{Im}(Q_i)$, which is the eigenspace corresponding to α_i.

Let us say that $A \in \mathcal{T} \cap \mathcal{S}_n$ is *generic in eigenvalue structure* (or simply *generic*) if all the matrices $B_1, \ldots, B_\ell$ appearing in the decomposition (3.2) of A are free from multiple eigenvalues and no two of them share a common eigenvalue. For a generic matrix A the number k of distinct eigenvalues is equal to $\sum_{j=1}^{\ell} \bar{n}_j$ and the list (multiset) of their multiplicities $\{m_1, \ldots, m_k\}$ is the union of $\bar{n}_j$ copies of $\bar{m}_j$ over $j = 1, \ldots, \ell$. It is emphasized that the genericity is defined with respect to $\mathcal{T}$ (and not to $\mathcal{M}_n$).

The eigenvalue decomposition of a generic A is consistent with the decomposition (3.3) into simple components of $\mathcal{T}$, as follows.

Proposition 3.2. *Let $A \in \mathcal{T} \cap \mathcal{S}_n$ be generic in eigenvalue structure. For any $i \in \{1, \ldots, k\}$ there exists $j \in \{1, \ldots, \ell\}$ such that $V_i \subseteq U_j$. Hence there exists a partition of $K = \{1, \ldots, k\}$ into ℓ disjoint subsets:*

$$K = K_1 \cup \cdots \cup K_\ell \tag{3.7}$$

such that

$$U_j = \bigoplus_{i \in K_j} V_i, \qquad j = 1, \ldots, \ell. \tag{3.8}$$

Note that $m_i = \bar{m}_j$ for $i \in K_j$ and $|K_j| = \bar{n}_j$ for $j = 1, \ldots, \ell$.

The partition (3.7) of K can be determined as follows. Let $\sim$ be the equivalence relation on K defined as the symmetric and transitive closure of the binary relation:

$$i \sim i' \quad \Longleftrightarrow \quad \exists p \ (1 \le p \le N) : Q_i^\top A_p Q_{i'} \ne O, \tag{3.9}$$

where $i \sim i$ for all $i \in K$ by convention.

Proposition 3.3. *The partition* (3.7) *coincides with the partition of K into equivalence classes induced by $\sim$.*

Proof. This is not difficult to see from the general theory of matrix $*$-algebra, but a proof is given here for completeness. Denote by $\{L_1, \ldots, L_{\ell'}\}$ the equivalence classes with respect to $\sim$.

If $i \sim i'$, then $Q_i^\top A_p Q_{i'} \ne O$ for some p. This means that for any $I \subseteq K$ with $i \in I$ and $i' \in K \setminus I$, the subspace $\bigoplus_{i'' \in I} V_{i''}$ is not invariant under A_p.

Hence $V_{i'}$ must be contained in the same simple component as V_i. Therefore each L_j must be contained in some $K_{j'}$.

To show the converse, define a matrix $\tilde{Q}_j = (Q_i \mid i \in L_j)$, which is of size $n \times \sum_{i \in L_j} m_i$, and an $n \times n$ matrix $E_j = \tilde{Q}_j \tilde{Q}_j^\top$ for $j = 1, \ldots, \ell'$. Each matrix E_j belongs to $\mathcal{T}$, as shown below, and it is idempotent (i.e., ${E_j}^2 = E_j$) and $E_1 + \cdots + E_{\ell'} = I_n$. On the other hand, for distinct j and j' we have $\tilde{Q}_j^\top A_p \tilde{Q}_{j'} = O$ for all p, and hence $\tilde{Q}_j^\top M \tilde{Q}_{j'} = O$ for all $M \in \mathcal{T}$. This implies that $E_j M = M E_j$ for all $M \in \mathcal{T}$. Therefore $\mathrm{Im}(E_j)$ is a union of simple components, and hence L_j is a union of some $K_{j'}$'s.

It remains to show that $E_j \in \mathcal{T}$. Since α_i's are distinct, for any real numbers $u_1, \ldots, u_k$ there exists a polynomial f such that $f(\alpha_i) = u_i$ for $i = 1, \ldots, k$. Let f_j be such f for $(u_1, \ldots, u_k)$ defined as $u_i = 1$ for $i \in L_j$ and $u_i = 0$ for $i \in K \setminus L_j$. Then $E_j = \tilde{Q}_j \tilde{Q}_j^\top = Q \cdot f_j(\mathrm{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k})) \cdot Q^\top = Q \cdot f_j(Q^\top A Q) \cdot Q^\top = f_j(A)$. This shows $E_j \in \mathcal{T}$. □

A generic matrix A can be obtained as a random linear combination of generators, as follows. For a real vector $r = (r_1, \ldots, r_N)$ put

$$A(r) = r_1 A_1 + \cdots + r_N A_N.$$

We denote by $\mathrm{span}\{\cdots\}$ the set of linear combinations of the matrices in the braces.

Proposition 3.4. *If* $\mathrm{span}\{I_n, A_1, \ldots, A_N\} = \mathcal{T} \cap \mathcal{S}_n$, *there exists an open dense subset* R *of* $\mathbb{R}^N$ *such that* $A(r)$ *is generic in eigenvalue structure for every* $r \in R$.

Proof. Let B_{pj} denote the matrix B_j in the decomposition (3.2) of $A = A_p$ for $p = 1, \ldots, N$. For $j = 1, \ldots, \ell$ define $f_j(\lambda) = f_j(\lambda; r) = \det(\lambda I - (r_1 B_{1j} + \cdots + r_N B_{Nj}))$, which is a polynomial in $\lambda, r_1, \ldots, r_N$. By the assumption on the linear span of generators, $f_j(\lambda)$ is free from multiple roots for at least one $r \in \mathbb{R}^N$, and it has a multiple root only if r lies on the algebraic set, say, Σ_j defined by the resultant of $f_j(\lambda)$ and $f_j'(\lambda)$. For distinct j and j', $f_j(\lambda)$ and $f_{j'}(\lambda)$ do not share a common root for at least one $r \in \mathbb{R}^N$, and they have a common root only if r lies on the algebraic set, say, $\Sigma_{jj'}$ defined by the resultant of $f_j(\lambda)$ and $f_{j'}(\lambda)$. Then we can take $R = \mathbb{R}^N \setminus \left[(\cup_j \Sigma_j) \cup \left(\cup_{j,j'} \Sigma_{jj'} \right) \right]$. □

We may assume that the coefficient vector r is normalized, for example, to $\|r\|_2 = 1$, where $\|r\|_2 = \sqrt{\sum_{p=1}^N {r_p}^2}$. Then the above proposition implies that $A(r)$ is generic for almost all values of r, or with probability one if r is chosen at random. It should be clear that we can adopt any normalization scheme (other than $\|r\|_2 = 1$) for this statement.

3.3 Transformation for irreducible components

Once the transformation matrix Q for the eigenvalue decomposition of a generic matrix A is known, the transformation P for $\mathcal{T}$ can be obtained through "local" transformations within eigenspaces corresponding to distinct eigenvalues, followed by a "global" permutation of rows and columns.

Proposition 3.5. *Let $A \in \mathcal{T} \cap \mathcal{S}_n$ be generic in eigenvalue structure, and $Q^\top AQ = \mathrm{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k})$ be the eigenvalue decomposition as in* (3.6)*. Then the transformation matrix P in* (3.2) *can be chosen in the form of*

$$P = Q \cdot \mathrm{diag}\,(P_1, \ldots, P_k) \cdot \Pi \qquad (3.10)$$

with orthogonal matrices $P_i \in \mathcal{M}_{m_i}$ for $i = 1, \ldots, k$, and a permutation matrix $\Pi \in \mathcal{M}_n$.

Proof. For simplicity of presentation we focus on a simple component, which is equivalent to assuming that for each $A' \in \mathcal{T}$ we have $P^\top A'P = I_{\bar{m}} \otimes B'$ for some $B' \in \mathcal{M}_k$, where $\bar{m} = m_1 = \cdots = m_k$. Since P may be replaced by $P(I_{\bar{m}} \otimes S)$ for any orthogonal S, it may be assumed further that $P^\top AP = I_{\bar{m}} \otimes D$, where $D = \mathrm{diag}\,(\alpha_1, \ldots, \alpha_k)$, for the particular generic matrix A. Hence $\Pi P^\top AP\Pi^\top = D \otimes I_{\bar{m}}$ for a permutation matrix Π. Comparing this with $Q^\top AQ = D \otimes I_{\bar{m}}$ and noting that α_i's are distinct, we see that

$$P\Pi^\top = Q \cdot \mathrm{diag}\,(P_1, \ldots, P_k)$$

for some $\bar{m} \times \bar{m}$ orthogonal matrices $P_1, \ldots, P_k$. This gives (3.10). □

4 Algorithm for Simultaneous Block-Diagonalization

On the basis of the theoretical considerations in Section 3, we propose in this section an algorithm for simultaneous block-diagonalization of given symmetric matrices $A_1, \ldots, A_N \in \mathcal{S}_n$ by an orthogonal matrix P:

$$P^\top A_p P = \bigoplus_{j=1}^{\ell} \bigoplus_{i=1}^{\bar{m}_j} B_{pj} = \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes B_{pj}), \quad p = 1, \ldots, N, \qquad (4.1)$$

where $B_{pj} \in \mathcal{M}_{\bar{n}_j}$ for $j = 1, \ldots, \ell$ and $p = 1, \ldots, N$. Our algorithm consists of two parts corresponding to (A) and (B) of Theorem 3.1 for the $*$-subalgebra $\mathcal{T}$ generated by $\{I_n, A_1, \ldots, A_N\}$. The former (Section 4.1) corresponds to the decomposition of $\mathcal{T}$ into simple components and the latter (Section 4.2) to the decomposition into irreducible components. A practical variant of the algorithm is described in Section 4.3. Recall that we assume (3.5).

4.1 Decomposition into simple components

We present here an algorithm for the decomposition into simple components. Algorithm 4.1 below does not presume $\mathrm{span}\{I_n, A_1, \ldots, A_N\} = \mathcal{T} \cap \mathcal{S}_n$, although its correctness relies on this condition.

Algorithm 4.1.

Step 1: Generate random numbers $r_1, \ldots, r_N$ (with $\|r\|_2 = 1$), and set $A = \sum_{p=1}^{N} r_p A_p$.

Step 2: Compute the eigenvalues and eigenvectors of A. Let $\alpha_1, \ldots, \alpha_k$ be the distinct eigenvalues of A with their multiplicities denoted by $m_1, \ldots, m_k$. Let $Q_i \in \mathbb{R}^{n \times m_i}$ be the matrix consisting of orthonormal eigenvectors corresponding to α_i, and define the matrix $Q \in \mathbb{R}^{n \times n}$ by $Q = (Q_i \mid i = 1, \ldots, k)$. This means that

$$Q^\top A Q = \mathrm{diag}\,(\alpha_1 I_{m_1}, \ldots, \alpha_k I_{m_k}).$$

Step 3: Put $K = \{1, \ldots, k\}$, and let $\sim$ be the equivalence relation on K induced from the binary relation:

$$i \sim i' \quad \Longleftrightarrow \quad \exists p \ (1 \leq p \leq N) : Q_i^\top A_p Q_{i'} \neq O. \tag{4.2}$$

Let

$$K = K_1 \cup \cdots \cup K_\ell \tag{4.3}$$

be the partition of K consisting of the equivalence classes with respect to $\sim$. Define matrices $Q[K_j]$ by

$$Q[K_j] = (Q_i \mid i \in K_j), \quad j = 1, \ldots, \ell,$$

and set

$$\hat{Q} = (Q[K_1], \ldots, Q[K_\ell])\,.$$

Compute $\hat{Q}^\top A_p \hat{Q}$ $(p = 1, \ldots, N)$, which results in a simultaneous block-diagonalization with respect to the partition (3.7).

Example 4.1. Suppose that the number of distinct eigenvalues of A is five, i.e., $K = \{1, 2, 3, 4, 5\}$, and that the partition of K is obtained as $K_1 = \{1, 2, 3\}$, $K_2 = \{4\}$, and $K_3 = \{5\}$, where $\ell = 3$. Then $A_1, \ldots, A_N$ are decomposed simultaneously as

$$\hat{Q}^\top A_p \hat{Q} = \begin{array}{c} \begin{array}{ccccc} m_1 & m_2 & m_3 & m_4 & m_5 \end{array} \\ \left[\begin{array}{ccc|c|c} * & * & * & O & O \\ * & * & * & O & O \\ * & * & * & O & O \\ \hline O & O & O & * & O \\ \hline O & O & O & O & * \end{array}\right] \end{array} \tag{4.4}$$

for $p = 1, \ldots, N$. ■

For the correctness of the above algorithm we have the following.

Proposition 4.2. *If the matrix A generated in Step 1 is generic in eigenvalue structure, the orthogonal matrix $\hat{Q}$ constructed by Algorithm 4.1 gives the transformation matrix $\hat{Q}$ in Theorem 3.1 (A) for the decomposition of $\mathcal{T}$ into simple components.*

Proof. This follows from Propositions 3.2 and 3.3. □

Proposition 3.4 implies that the matrix A in Step 1 is generic with probability one if $\mathrm{span}\{I_n, A_1, \ldots, A_N\} = \mathcal{T} \cap \mathcal{S}_n$. This condition, however, is not always satisfied by the given matrices $A_1, \ldots, A_N$. In such a case we can generate a basis of $\mathcal{T} \cap \mathcal{S}_n$ as follows. First choose a linearly independent subset, say, $\mathcal{B}_1$ of $\{I_n, A_1, \ldots, A_N\}$. For $k = 1, 2, \ldots$ let $\mathcal{B}_{k+1}$ ($\supseteq \mathcal{B}_k$) be a linearly independent subset of $\{(AB + BA)/2 \mid A \in \mathcal{B}_1, B \in \mathcal{B}_k\}$. If $\mathcal{B}_{k+1} = \mathcal{B}_k$ for some k, we can conclude that $\mathcal{B}_k$ is a basis of $\mathcal{T} \cap \mathcal{S}_n$. Note that the dimension of $\mathcal{T} \cap \mathcal{S}_n$ is equal to $\sum_{j=1}^{\ell} \bar{n}_j(\bar{n}_j + 1)/2$, which is bounded by $n(n+1)/2$. It is mentioned here that $\mathcal{S}_n$ is a linear space equipped with an inner product $A \bullet B = \mathrm{tr}(AB)$ and the Gram–Schmidt orthogonalization procedure works.

Proposition 4.3. *If a basis of $\mathcal{T} \cap \mathcal{S}_n$ is computed in advance, Algorithm 4.1 gives, with probability one, the decomposition of $\mathcal{T}$ into simple components.*

4.2 Decomposition into irreducible components

According to Theorem 3.1 (B), the block-diagonal matrices $\hat{Q}^\top A_p \hat{Q}$ obtained by Algorithm 4.1 can further be decomposed. By construction we have $\hat{Q} = Q\hat{\Pi}$ for some permutation matrix $\hat{\Pi}$. In the following we assume $\hat{Q} = Q$ to simplify the presentation.

By Proposition 3.5 this finer decomposition can be obtained through a transformation of the form (3.10), which consists of "local coordinate changes" by a family of orthogonal matrices $\{P_1, \ldots, P_k\}$, followed by a permutation by Π.

The orthogonal matrices $\{P_1, \ldots, P_k\}$ should be chosen in such a way that if $i, i' \in K_j$, then

$$P_i^\top Q_i^\top A_p Q_{i'} P_{i'} = b_{ii'}^{(pj)} I_{\bar{m}_j} \tag{4.5}$$

for some $b_{ii'}^{(pj)} \in \mathbb{R}$ for $p = 1, \ldots, N$. Note that the solvability of this system of equations in P_i ($i = 1, \ldots, k$) and $b_{ii'}^{(pj)}$ ($i, i' = 1, \ldots, k$; $j = 1, \ldots, \ell$; $p = 1, \ldots, N$) is guaranteed by (4.1) and Proposition 3.5. Then with $\tilde{P} = Q \cdot \mathrm{diag}\,(P_1, \ldots, P_k)$ and $B_{pj} = (b_{ii'}^{(pj)} \mid i, i' \in K_j)$ we have

$$\tilde{P}^\top A_p \tilde{P} = \bigoplus_{j=1}^{\ell} (B_{pj} \otimes I_{\bar{m}_j}) \tag{4.6}$$

for $p = 1, \ldots, N$. Finally we apply a permutation of rows and columns to obtain (4.1).

Example 4.2. Recall Example 4.1. We consider the block-diagonalization of the first block $\hat{A}_p = Q[K_1]^\top A_p Q[K_1]$ of (4.4), where $m_1 = m_2 = m_3 = 2$ and $K_1 = \{1, 2, 3\}$. We first compute orthogonal matrices P_1, P_2 and P_3 satisfying

$$\operatorname{diag}(P_1, P_2, P_3)^\top \cdot \hat{A}_p \cdot \operatorname{diag}(P_1, P_2, P_3) = \begin{bmatrix} b_{11}^{(p1)} I_2 & b_{12}^{(p1)} I_2 & b_{13}^{(p1)} I_2 \\ b_{21}^{(p1)} I_2 & b_{22}^{(p1)} I_2 & b_{23}^{(p1)} I_2 \\ b_{31}^{(p1)} I_2 & b_{32}^{(p1)} I_2 & b_{33}^{(p1)} I_2 \end{bmatrix}.$$

Then a permutation of rows and columns yields a block-diagonal form $\operatorname{diag}(B_{p1}, B_{p1})$ with $B_{p1} = \begin{bmatrix} b_{11}^{(p1)} & b_{12}^{(p1)} & b_{13}^{(p1)} \\ b_{21}^{(p1)} & b_{22}^{(p1)} & b_{23}^{(p1)} \\ b_{31}^{(p1)} & b_{32}^{(p1)} & b_{33}^{(p1)} \end{bmatrix}$. ■

The family of orthogonal matrices $\{P_1, \ldots, P_k\}$ satisfying (4.5) can be computed as follows. Recall from (4.2) that for $i, i' \in K$ we have $i \sim i'$ if and only if $Q_i^\top A_p Q_{i'} \neq O$ for some p. It follows from (4.5) that $Q_i^\top A_p Q_{i'} \neq O$ means that it is nonsingular.

Fix j with $1 \leq j \leq \ell$. We consider a graph $G_j = (K_j, E_j)$ with vertex set K_j and edge set $E_j = \{(i, i') \mid i \sim i'\}$. This graph is connected by the definition of K_j. Let T_j be a spanning tree, which means that T_j is a subset of E_j such that $|T_j| = |K_j| - 1$ and any two vertices of K_j are connected by a path in T_j. With each $(i, i') \in T_j$ we can associate some $p = p(i, i')$ such that $Q_i^\top A_p Q_{i'} \neq O$.

To compute $\{P_i \mid i \in K_j\}$, take any $i_1 \in K_j$ and put $P_{i_1} = I_{\bar{m}_j}$. If $(i, i') \in T_j$ and P_i has been determined, then let $\hat{P}_{i'} = (Q_i^\top A_p Q_{i'})^{-1} P_i$ with $p = p(i, i')$, and normalize it to $P_{i'} = \hat{P}_{i'} / \|q\|$, where q is the first-row vector of $\hat{P}_{i'}$. Then $P_{i'}$ is an orthogonal matrix that satisfies (4.5). By repeating this we can obtain $\{P_i \mid i \in K_j\}$.

Remark 4.1. A variant of the above algorithm for computing $\{P_1, \ldots, P_k\}$ is suggested here. Take a second random vector $r' = (r'_1, \ldots, r'_N)$, independently of r, to form $A(r') = r'_1 A_1 + \cdots + r'_N A_N$. For $i, i' \in K_j$ we have, with probability one, that $(i, i') \in E_j$ if and only if $Q_i^\top A(r') Q_{i'} \neq O$. If P_i has been determined, we can determine $P_{i'}$ by normalizing $\hat{P}_{i'} = (Q_i^\top A(r') Q_{i'})^{-1} P_i$ to $P_{i'} = \hat{P}_{i'} / \|q\|$, where q is the first-row vector of $\hat{P}_{i'}$. ■

Remark 4.2. The proposed method relies on numerical computations to determine the multiplicities of eigenvalues, which in turn determine the block-diagonal structures. As such the method is inevitably faced with numerical noises due to rounding errors. A scaling technique to remedy this difficulty is suggested in Remark 5.1 for truss optimization problems. ■

Remark 4.3. The idea of using a random linear combination in constructing simultaneous block-diagonalization can also be found in a recent paper of de Klerk and Sotirov [5]. Their method, called "block diagonalization heuristic" in Section 5.2 of [5], is different from ours in two major points.

First, the method of [5] assumes explicit knowledge about the underlying group G, and works with the representation matrices, denoted $T(g)$ in (2.14). Through the eigenvalue (spectral) decomposition of a random linear combination of $T(g)$ over $g \in G$, the method finds an orthogonal matrix P such that $P^\top T(g) P$ for $g \in G$ are simultaneously block-diagonalized, just as in (2.15) and (2.16). Then G-symmetric matrices A_p, satisfying (2.14), will also be block-diagonalized.

Second, the method of [5] is not designed to produce the finest possible decomposition of the matrices A_p, as is recognized by the authors themselves. The constructed block-diagonalization of $T(g)$ is not necessarily the irreducible decomposition, and this is why the resulting decomposition of A_p is not guaranteed to be finest possible. We could, however, apply the algorithm of Section 4.2 of the present paper to obtain the irreducible decomposition of the representation $T(g)$. Then, under the assumption (3.5), the resulting decomposition of A_p will be the finest decomposition that can be obtained by exploiting the G-symmetry. ■

Remark 4.4. Eberly and Giesbrecht [6] proposed an algorithm for the simple-component decomposition of a separable matrix algebra (not a $*$-algebra) over an arbitrary infinite field. Their algorithm is closely related to our algorithm in Section 3.2. In particular, their "self-centralizing element" corresponds to our "generic element". Their algorithm, however, is significantly different from ours in two ways: (i) treating a general algebra (not a $*$-algebra) it employs a transformation of the form $S^{-1}AS$ with a nonsingular matrix S instead of an orthogonal transformation, and (ii) it uses companion forms and factorization of minimum polynomials instead of eigenvalue decomposition. The decomposition into irreducible components, which inevitably depends on the underlying field, is not treated in [6]. ■

4.3 A practical variant of the algorithm

In Propositions 3.4 we have considered two technical conditions:

1. $\mathrm{span}\{I_n, A_1, \ldots, A_N\} = \mathcal{T} \cap \mathcal{S}_n$,

2. $r \in R$, where R is an open dense set,

to ensure genericity of $A = \sum_{p=1}^{N} r_p A_p$ in eigenvalue structure. The genericity of A guarantees, in turn, that our algorithm yields the finest simultaneous block-diagonalization (see Proposition 4.2). The condition $r \in R$ above can be met with probability one through a random choice of r.

To meet the first condition we could generate a basis of $\mathcal{T} \cap \mathcal{S}_n$ in advance, as is mentioned in Proposition 4.3. However, an explicit computation of a basis seems too heavy to be efficient. It should be understood that the above two conditions are introduced as sufficient conditions to avoid degeneracy in eigenvalues. By no means are they necessary for the success of the algorithm. With this observation we propose the following procedure as a practical variant of our algorithm.

We apply Algorithm 4.1 to the given family $\{A_1, \ldots, A_N\}$ to find an orthogonal matrix Q and a partition $K = K_1 \cup \cdots \cup K_\ell$. In general there is no guarantee that this corresponds to the decomposition into simple components, but in any case Algorithm 4.1 terminates without getting stuck. The algorithm does not hang up either, when a particular choice of r does not meet the condition $r \in R$. Thus we can always go on to the second stage of the algorithm for the irreducible decomposition.

Next, we are to determine a family of orthogonal matrices $\{P_1, \ldots, P_k\}$ that satisfies (4.5). This system of equations is guaranteed to be solvable if A is generic (see Proposition 3.5). In general we may possibly encounter a difficulty of the following kinds:

1. For some $(i, i') \in T_j$ the matrix $Q_i^\top A_p Q_{i'}$ with $p = p(i, i')$ is singular and hence $P_{i'}$ cannot be computed. This includes the case of a rectangular matrix, which is demonstrated in Example 4.3 below.

2. For some p and $(i, i') \in E_j$ the matrix $P_i^\top Q_i^\top A_p Q_{i'} P_{i'}$ is not a scalar multiple of the identity matrix.

Such inconsistency is an indication that the decomposition into simple components has not been computed correctly. Accordingly, if either of the above inconsistency is detected, we restart our algorithm by adding some linearly independent matrices of $\mathcal{T} \cap \mathcal{S}_n$ to the current set $\{A_1, \ldots, A_N\}$. It is mentioned that the possibility exists, though with probability zero, that r is chosen badly to yield a nongeneric A even when $\mathrm{span}\{I_n, A_1, \ldots, A_N\} = \mathcal{T} \cap \mathcal{S}_n$ is true.

It is expected that we can eventually arrive at the correct decomposition after a finite number of iterations. With probability one, the number of restarts is bounded by the dimension of $\mathcal{T} \cap \mathcal{S}_n$, which is $O(n^2)$. When it terminates, the modified algorithm always gives a legitimate simultaneous block-diagonal decomposition of the form (4.1).

There is some subtlety concerning the optimality of the obtained decomposition. If a basis of $\mathcal{T} \cap \mathcal{S}_n$ is generated, the decomposition coincides, with probability one, with the canonical finest decomposition of the $*$-algebra $\mathcal{T}$. However, when the algorithm terminates before it generates a basis of $\mathcal{T} \cap \mathcal{S}_n$, there is no theoretical guarantee that the obtained decomposition is the finest possible. Nevertheless, it is very likely in practice that the obtained decomposition coincides with the finest decomposition.

Example 4.3. Here is an example that requires an additional generator to be added. Suppose that we are given

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and let $\mathcal{T}$ be the matrix $*$-algebra generated by $\{I_4, A_1, A_2\}$. It turns out that the structural indices in (4.1) are: $\ell = 2$, $\bar{m}_1 = \bar{m}_2 = 1$, $\bar{n}_1 = 1$ and $\bar{n}_2 = 3$. This means that the list of eigenvalue multiplicities of $\mathcal{T}$ is $\{1, 1, 1, 1\}$. Note also that $\dim(\mathcal{T} \cap \mathcal{S}_4) = \bar{n}_1(\bar{n}_1 + 1)/2 + \bar{n}_2(\bar{n}_2 + 1)/2 = 7$.

For $A(r) = r_1 A_1 + r_2 A_2$ we have

$$A(r) \begin{bmatrix} 1 & 0 \\ 0 & (r_1 - r_2)/c \\ 0 & r_1/c \\ 0 & 0 \end{bmatrix} = (r_1 + r_2) \begin{bmatrix} 1 & 0 \\ 0 & (r_1 - r_2)/c \\ 0 & r_1/c \\ 0 & 0 \end{bmatrix}, \tag{4.7}$$

where $c = \sqrt{(r_1 - r_2)^2 + {r_1}^2}$. This shows that $A(r)$ has a multiple eigenvalue $r_1 + r_2$ of multiplicity two, as well as two other simple eigenvalues. Thus for any r the list of eigenvalue multiplicities of $A(r)$ is equal to $\{2, 1, 1\}$, which differs from $\{1, 1, 1, 1\}$ for $\mathcal{T}$.

The discrepancy in the eigenvalue multiplicities cannot be detected during the first stage of our algorithm. In Step 2 we have $k = 3$, $m_1 = 2$, $m_2 = m_3 = 1$. The orthogonal matrix Q is partitioned into three submatrices Q_1, Q_2 and Q_3, where Q_1 (nonunique) may possibly be the 4×2 matrix shown in (4.7), and Q_2 and Q_3 consist of a single column. Since $Q^\top A_p Q$ is of the form

$$Q^\top A_p Q = \left[\begin{array}{cc|c|c} 1 & 0 & 0 & 0 \\ 0 & * & * & * \\ \hline 0 & * & * & * \\ \hline 0 & * & * & * \end{array}\right]$$

for $p = 1, 2$, we have $\ell = 1$ and $K_1 = \{1, 2, 3\}$ in Step 3. At this moment an inconsistency is detected, since $m_1 \neq m_2$ inspite of the fact that $i = 1$ and $i' = 2$ belong to the same block K_1.

We restart the algorithm, say, with an additional generator

$$A_3 = \frac{1}{2}(A_1 A_2 + A_2 A_1) = \frac{1}{2} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -2 \end{bmatrix}$$

to consider $\tilde{A}(r) = r_1 A_1 + r_2 A_2 + r_3 A_3$ instead of $A(r) = r_1 A_1 + r_2 A_2$. Then $\tilde{A}(r)$ has four simple eigenvalues for generic values of $r = (r_1, r_2, r_3)$, and accordingly we have $\{1,1,1,1\}$ as the list of eigenvalue multiplicities of $\tilde{A}(r)$, which agrees with that of $\mathcal{T}$.

In Step 2 of Algorithm 4.1 we now have $k = 4$, $m_1 = m_2 = m_3 = m_4 = 1$. The orthogonal matrix Q is partitioned into four 4×1 submatrices, and $Q^\top A_p Q$ is of the form

$$Q^\top A_p Q = \left[\begin{array}{c|c|c|c} 1 & 0 & 0 & 0 \\ \hline 0 & * & * & * \\ \hline 0 & * & * & * \\ \hline 0 & * & * & * \end{array}\right]$$

for $p = 1, 2, 3$, from which we obtain $K_1 = \{1\}$, $K_2 = \{2, 3, 4\}$ with $\ell = 2$ in Step 3. Thus we have arrived at the correct decomposition consisting of a 1×1 block and a 3×3 block. Note that the correct decomposition is obtained in spite of the fact that $\{I_4, A_1, A_2, A_3\}$ does not span $\mathcal{T} \cap \mathcal{S}_4$. ■

5 Numerical examples

5.1 Effects of additional algebraic structures

It is demonstrated here that our method automatically reveals inherent algebraic structures due to parameter dependence as well as to group symmetry. The S_3-symmetric matrices $A_1, \ldots, A_4$ in (2.7) and (2.8) are considered in three representative cases.

Case 1:

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \end{bmatrix}, \quad E = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix},$$

Case 2:

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \end{bmatrix}, \quad E = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix},$$

Case 3:

$$B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 \end{bmatrix}, \quad E = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

We have $n_{\mathrm{B}} = 2$ and $n_{\mathrm{D}} = 1$ in the notation of Section 2.2.

Case 1 is a generic case under S_3-symmetry. The simultaneous block-diagonalization is of the form

$$P^\top A_p P = B_{p1} \oplus (I_2 \otimes B_{p2}), \quad p = 1, \ldots, 4, \tag{5.1}$$

with $B_{p1} \in \mathcal{M}_3$, $B_{p2} \in \mathcal{M}_2$; i.e., $\ell = 2$, $\bar{m}_1 = 1$, $\bar{m}_2 = 2$, $\bar{n}_1 = 3$, $\bar{n}_2 = 2$ in (4.1). By (2.10)–(2.13), a possible choice of these matrices is

$$B_{11} = \begin{bmatrix} B & O \\ O & O \end{bmatrix}, \; B_{21} = \begin{bmatrix} O & \sqrt{3}C \\ \sqrt{3}C^\top & O \end{bmatrix}, \; B_{31} = \begin{bmatrix} O & O \\ O & D \end{bmatrix}, \; B_{41} = \begin{bmatrix} 2E & O \\ O & O \end{bmatrix},$$

and $B_{12} = B$, $B_{22} = B_{32} = O$, $B_{42} = -E$. Our implementation of the proposed method yields the same decomposition but with different matrices. For instance, we have obtained

$$B_{12} = \begin{bmatrix} -0.99954 & -0.04297 \\ -0.04297 & 2.99954 \end{bmatrix}, \quad B_{42} = \begin{bmatrix} -1.51097 & 0.52137 \\ 0.52137 & -3.48903 \end{bmatrix}.$$

Here it is noted that the obtained B_{12} and B_{42} are related to B and E as

$$\begin{bmatrix} B_{12} & O \\ O & B_{12} \end{bmatrix} = \tilde{P}^{\top} \begin{bmatrix} B & O \\ O & B \end{bmatrix} \tilde{P}, \quad \begin{bmatrix} B_{42} & O \\ O & B_{42} \end{bmatrix} = \tilde{P}^{\top} \begin{bmatrix} -E & O \\ O & -E \end{bmatrix} \tilde{P}$$

for an orthogonal matrix $\tilde{P}$ expressed as $\tilde{P} = \begin{bmatrix} \tilde{P}_{11} & \tilde{P}_{12} \\ \tilde{P}_{21} & \tilde{P}_{22} \end{bmatrix}$ with

$$\tilde{P}_{11} = -\tilde{P}_{22} = \begin{bmatrix} 0.12554 & -0.12288 \\ -0.12288 & -0.12554 \end{bmatrix}, \tilde{P}_{12} = \tilde{P}_{21} = \begin{bmatrix} 0.70355 & -0.68859 \\ -0.68859 & -0.70355 \end{bmatrix}.$$

In Case 2 we have a commutativity relation $BE = EB$. This means that B and E can be simultaneously diagonalized, and a further decomposition of the second factor in (5.1) should result. Instead of (5.1) we have

$$P^{\top} A_p P = B_{p1} \oplus (I_2 \otimes B_{p2}) \oplus (I_2 \otimes B_{p3}), \quad p = 1, \ldots, 4,$$

with $B_{p1} \in \mathcal{M}_3$, $B_{p2} \in \mathcal{M}_1$ and $B_{p3} \in \mathcal{M}_1$; i.e., $\ell = 3$, $\bar{m}_1 = 1$, $\bar{m}_2 = \bar{m}_3 = 2$, $\bar{n}_1 = 3$, $\bar{n}_2 = \bar{n}_3 = 1$ in (4.1). The proposed method yields $B_{12} = [\ 3.00000\ ]$, $B_{42} = [\ -4.00000\ ]$, $B_{13} = [\ -1.00000\ ]$ and $B_{43} = [\ -2.00000\ ]$, successfully detecting the additional algebraic structure caused by $BE = EB$.

Case 3 contains a further degeneracy that the column vector of C is an eigenvector of B and E. This splits the 3×3 block into two, and we have

$$P^{\top} A_p P = B_{p1} \oplus B_{p4} \oplus (I_2 \otimes B_{p2}) \oplus (I_2 \otimes B_{p3}), \quad p = 1, \ldots, 4,$$

with $B_{p1} \in \mathcal{M}_2$, $B_{pj} \in \mathcal{M}_1$ for $j = 2, 3, 4$; i.e., $\ell = 4$, $\bar{m}_1 = \bar{m}_4 = 1$, $\bar{m}_2 = \bar{m}_3 = 2$, $\bar{n}_1 = 2$, $\bar{n}_2 = \bar{n}_3 = \bar{n}_4 = 1$ in (4.1). For instance, we have indeed obtained

$$B_{11} \oplus B_{14} = \left[\begin{array}{cc|c} 0.48288 & 1.10248 & 0 \\ 1.10248 & 2.51712 & 0 \\ \hline 0 & & -1.00000 \end{array}\right].$$

Also in this case the proposed method works, identifying the additional algebraic structure through numerical computation.

The three cases are compared in Table 1.

Table 1: Block-diagonalization of S_3-symmetric matrices in (2.7) and (2.8).

	Case 1		Case 2		Case 3	
	$\bar{n}_j$	$\bar{m}_j$	$\bar{n}_j$	$\bar{m}_j$	$\bar{n}_j$	$\bar{m}_j$
$j = 1$	3	1	3	1	2	1
$j = 4$	—	—	—	—	1	1
$j = 2$	2	2	1	2	1	2
$j = 3$	—	—	1	2	1	2

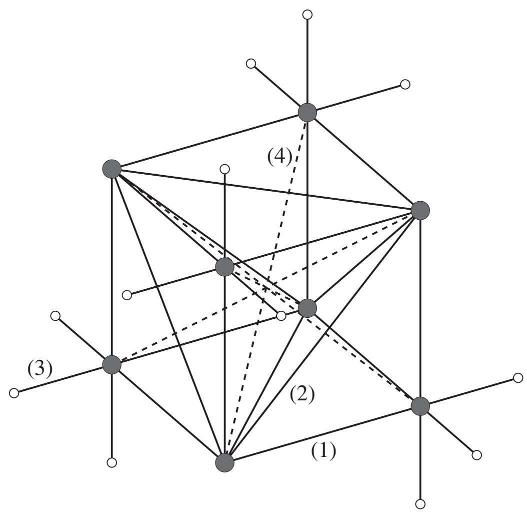

Figure 1: A cubic (or T_d-symmetric) space truss.

5.2 Optimization of symmetric trusses

Use and significance of our method are illustrated here in the context of semidefinite programming for truss design treated in [25]. Group-symmetry and sparsity arise naturally in truss optimization problems [1, 14]. It will be confirmed that the proposed method yields the same decomposition as the group representation theory anticipates (Case 1 below), and moreover, it gives a finer decomposition if the truss structure is endowed with an additional algebraic structure due to sparsity (Case 2 below).

The optimization problem we consider here is as follow. An initial truss configuration is given with fixed locations of nodes and members. Optimal cross-sectional areas, minimizing total volume of the structure, are to be found subject to the constraint that the eigenvalues of vibration are not smaller than a specified value.

To be more specific, let $n^{\rm d}$ and $n^{\rm m}$ denote the number of degrees of freedom of displacements and the number of members of a truss, respectively.

Let $K \in \mathcal{S}_{n^{\mathrm{d}}}$ denote the stiffness matrix, and $M_{\mathrm{S}} \in \mathcal{S}_{n^{\mathrm{d}}}$ and $M_0 \in \mathcal{S}_{n^{\mathrm{d}}}$ the mass matrices for the structural and nonstructural masses, respectively; see, e.g., [35] for the definitions of these matrices. The ith eigenvalue Ω_i of vibration and the corresponding eigenvector $\phi_i \in \mathbb{R}^{n^{\mathrm{d}}}$ are defined by

$$K\phi_i = \Omega_i(M_{\mathrm{S}} + M_0)\phi_i, \quad i = 1, 2, \dots, n^{\mathrm{d}}. \tag{5.2}$$

Note that, for a truss, K and M_{S} can be written as

$$K = \sum_{j=1}^{n^{\mathrm{m}}} K_j \eta_j, \qquad M_{\mathrm{S}} = \sum_{j=1}^{n^{\mathrm{m}}} M_j \eta_j \tag{5.3}$$

with sparse constant symmetric matrices K_j and M_j, where η_j denotes the cross-sectional area of the jth member. With the notation $l = (l_j) \in \mathbb{R}^{n^{\mathrm{m}}}$ for the vector of member lengths and $\bar{\Omega}$ for the specified lower bound of the fundamental eigenvalue, our optimization problem is formulated as

$$\left.\begin{array}{ll} \min & \displaystyle\sum_{j=1}^{n^{\mathrm{m}}} l_j \eta_j \\ \text{s.t.} & \Omega_i \geq \bar{\Omega}, \quad i = 1, \dots, n^{\mathrm{d}}, \\ & \eta_j \geq 0, \quad j = 1, \dots, n^{\mathrm{m}}. \end{array}\right\} \tag{5.4}$$

It is pointed out in [25] that this problem (5.4) can be reduced to the following dual SDP (cf. (2.2)):

$$\left.\begin{array}{ll} \max & -\displaystyle\sum_{j=1}^{n^{\mathrm{m}}} l_j \eta_j \\ \text{s.t.} & \displaystyle\sum_{j=1}^{n^{\mathrm{m}}} (K_j - \bar{\Omega} M_j)\eta_j - \bar{\Omega} M_0 \succeq O, \\ & \eta_j \geq 0, \quad j = 1, \dots, n^{\mathrm{m}}. \end{array}\right\} \tag{5.5}$$

We now consider this SDP for the cubic truss shown in Fig. 1. The cubic truss contains 8 free nodes, and hence $n^{\mathrm{d}} = 24$. As for the members we consider two cases:

Case 1: $n^{\mathrm{m}} = 34$ members including the dotted ones;

Case 2: $n^{\mathrm{m}} = 30$ members excluding the dotted ones.

A regular tetrahedron is constructed inside the cube. The lengths of members forming the edges of the cube are 2 m. The lengths of the members outside the cube are 1 m. A nonstructural mass of 2.1×10^5 kg is located at each node indicated by a filled circle in Fig. 1. The lower bound of the eigenvalues is specified as $\bar{\Omega} = 10.0$. All the remaining nodes are pin-supported (i.e., the locations of those nodes are fixed in the three-dimensional space, while the rotations of members connected to those nodes are not prescribed).

Thus, the geometry, the stiffness distribution, and the mass distribution of this truss are all symmetric with respect to the geometric transformations by elements of (full or achiral) tetrahedral group $\mathrm{T_d}$, which is isomorphic to the symmetric group S_4. The $\mathrm{T_d}$-symmetry can be exploited as follows.

First, we divide the index set of members $\{1, \dots, n^{\mathrm{m}}\}$ into a family of orbits, say J_p with $p = 1, \dots, m$, where m denotes the number of orbits. We have $m = 4$ in Case 1 and $m = 3$ in Case 2, where representative members belonging to four different orbits are shown as (1)–(4) in Fig. 1. It is mentioned in passing that the classification of members into orbits is an easy task for engineers. Indeed, this is nothing but the so-called *variable-linking technique*, which has often been employed in the literature of structural optimization in obtaining symmetric structural designs [20].

Next, with reference to the orbits we aggregate the data matrices as well as the coefficients of the objective function in (5.5) to A_p $(p = 0, 1, \dots, m)$ and b_p $(p = 1, \dots, m)$, respectively, as

$$A_0 = -\bar{\Omega} M_0; \quad A_p = \sum_{j \in J_p} (-K_j + \bar{\Omega} M_j), \quad b_p = \sum_{j \in J_p} l_j, \quad p = 1, \dots, m.$$

Then (5.5) can be reduced to

$$\left. \begin{array}{ll} \max & -\displaystyle\sum_{p=1}^{m} b_p y_p \\ \mathrm{s.t.} & A_0 - \displaystyle\sum_{p=1}^{m} A_p y_p \succeq O, \\ & y_p \geq 0, \quad p = 1, \dots, m \end{array} \right\} \tag{5.6}$$

as long as we are interested in a symmetric optimal solution, where $y_p = \eta_j$ for $j \in J_p$. Note that the matrices A_p $(p = 0, 1, \dots, m)$ are symmetric in the sense of (2.14) for $G = \mathrm{T_d}$. The two cases share the same matrices A_1, A_2, A_3, and A_0 is proportional to the identity matrix.

The proposed method is applied to A_p $(p = 0, 1, \dots, m)$ for their simultaneous block-diagonalization. The practical variant described in Section 4.3 is employed. In either case it has turned out that additional generators are not necessary, but random linear combinations of the given matrices A_p $(p = 0, 1, \dots, m)$ are sufficient to find the block-diagonalization. The assumption (3.5) has turned out to be satisfied.

In Case 1 we obtain the decomposition into $1+2+3+3 = 9$ blocks, one block of size 2, two identical blocks of size 2, three identical blocks of size 3, and three identical blocks of size 4, as summarized on the left of Table 2. This result conforms with the group-theoretic analysis. The tetrahedral group $\mathrm{T_d}$, being isomorphic to S_4, has two one-dimensional irreducible representations,

Table 2: Block-diagonalization of cubic truss optimization problem.

	Case 1: $m = 4$		Case 2: $m = 3$	
	block size $\bar{n}_j$	multiplicity $\bar{m}_j$	block size $\bar{n}_j$	multiplicity $\bar{m}_j$
$j = 1$	2	1	2	1
$j = 2$	2	2	2	2
$j = 3$	2	3	2	3
$j = 4$	4	3	2	3
$j = 5$	—	—	2	3

one two-dimensional irreducible representation, and two three-dimensional irreducible representations [23, 27]. The block indexed by $j = 1$ corresponds to the unit representation, one of the one-dimensional irreducible representations, while the block for the other one-dimensional irreducible representation is missing. The block with $j = 2$ corresponds to the two-dimensional irreducible representation, hence $\bar{m}_2 = 2$. Similarly, the blocks with $j = 3, 4$ correspond to the three-dimensional irreducible representation, hence $\bar{m}_3 = \bar{m}_4 = 3$.

In Case 2 sparsity plays a role to split the last block into two, as shown on the right of Table 2. We now have 12 blocks in contrast to 9 blocks in Case 1. Recall that the sparsity is due to the lack of the dotted members. It is emphasized that the proposed method successfully captures the additional algebraic structure introduced by sparsity.

Remark 5.1. Typically, actual trusses are constructed by using steel members, where the elastic modulus and the mass density of members are $E = 200.0$ GPa and $\rho = 7.86 \times 10^3$ kg/m^3, respectively. Note that the matrices K_j and M_j defining the SDP problem (5.6) are proportional to E and ρ, respectively. In order to avoid numerical instability in our block-diagonalization algorithm, E and ρ are scaled as $E = 1.0 \times 10^{-2}$ GPa and $\rho = 1.0 \times 10^8$ kg/m^3, so that the largest eigenvalue in (5.2) becomes sufficiently small. For example, if we choose the member cross-sectional areas as $\eta_j = 10^{-2}$ m^2 for $j = 1, \ldots, n^{\mathrm{m}}$, the maximum eigenvalue is 1.59×10^4 rad^2/s^2 for steel members, which is natural from the mechanical point of view. In contrast, by using the fictitious parameters mentioned above, the maximum eigenvalue is reduced to 6.24×10^{-2} rad^2/s^2, and then our block-diagonalization algorithm can be applied without any numerical instability. Note that the transformation matrix obtained by our algorithm for block-diagonalization of $A_0, A_1, \ldots, A_m$ is independent of the values of E and ρ. Hence, it is recommended for numerical stability to compute transformation matrices for the scaled matrices $\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_m$ by choosing appropriate fictitious values of E and ρ. It is easy to find a candidate of such fictitious values, because we know that the maximum eigenvalue can be reduced by decreasing E and/or increasing ρ. Then the obtained transformation matrices can be used to decompose the original matrices $A_0, A_1, \dots, A_m$ defined with the actual material parameters. ■

5.3 Quadratic semidefinite programs for symmetric frames

Effectiveness of our method is demonstrated here for the SOS–SDP relaxation of a quadratic SDP arising from a frame optimization problem. Quadratic (or polynomial) SDPs are known to be difficult problems, although they are, in principle, tractable by means of SDP relaxations. The difficulty may be ascribed to two major factors: (i) SDP relaxations tend to be large in size, and (ii) SDP relaxations often suffer from numerical instability. The block-diagonalization method makes the size of the SDP relaxation smaller, and hence mitigates the difficulty arising from the first factor.

The frame optimization problem with a specified fundamental eigenvalue $\bar{\Omega}$ can be treated basically in the same way as the truss optimization problem in Section 5.2, except that some nonlinear terms appear in the SDP problem.

First, we formulate the frame optimization problem in the form of (5.4), where "$\eta_j \geq 0$" is replaced by "$0 \leq \eta_j \leq \bar{\eta}_j$" with a given upper bound for η_j. Recall that η_j represents the cross-sectional area of the jth element and n^{m} denotes the number of members. We choose η_j ($j = 1, \dots, n^{\mathrm{m}}$) as the design variables.

As for the stiffness matrix K, we use the Euler–Bernoulli beam element [35] to define

$$K = \sum_{j=1}^{n^{\mathrm{m}}} K_j^{\mathrm{a}} \eta_j + \sum_{j=1}^{n^{\mathrm{m}}} K_j^{\mathrm{b}} \xi_j, \tag{5.7}$$

where K_j^{a} and K_j^{b} are sparse constant symmetric matrices, and ξ_j is the moment of inertia of the jth member. The mass matrix M_{S} due to the structural mass remains the same as in (5.3), being a linear function of η. Each member of the frame is assumed to have a circular solid cross-section with radius r_j. Then we have $\eta_j = \pi r_j^2$ and $\xi_j = \frac{1}{4}\pi r_j^4$.

Just as (5.4) can be reduced to (5.5), our frame optimization problem can be reduced to the following problem:

$$\left.\begin{array}{ll} \max & -\displaystyle\sum_{j=1}^{n^{\mathrm{m}}} l_j \eta_j \\ \text{s.t.} & \displaystyle\frac{1}{4\pi} \sum_{j=1}^{n^{\mathrm{m}}} K_j^{\mathrm{b}} \eta_j^2 + \sum_{j=1}^{n^{\mathrm{m}}} (K_j^{\mathrm{a}} - \bar{\Omega} M_j) \eta_j - \bar{\Omega} M_0 \succeq O, \\ & 0 \leq \eta_j \leq \bar{\eta}_j, \quad j = 1, \dots, n^{\mathrm{m}}, \end{array}\right\} \tag{5.8}$$

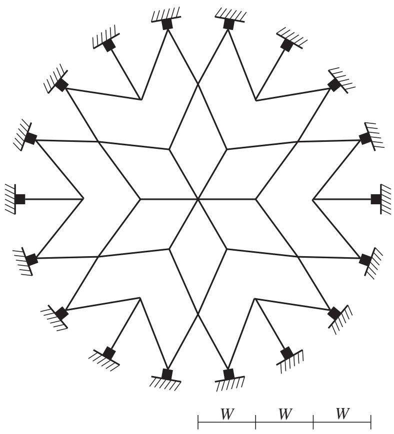

Figure 2: A D_6-symmetric plane frame.

which is a quadratic SDP. See [13] for details.

Suppose that the frame structure is endowed with geometric symmetry; Fig. 2 shows an example with D_6-symmetry. According to the symmetry the index set of the members $\{1, \ldots, n^{\mathrm{m}}\}$ is partitioned into orbits $\{J_p \mid p = 1, \ldots, m\}$. For symmetry of the problem, $\bar{\eta}_j$ should be constant on each orbit J_p and we put $d_p = \bar{\eta}_j$ for $j \in J_p$. By the variable-linking technique, (5.8) is reduced to the following quadratic SDP:

$$\left.\begin{array}{ll} \max & \displaystyle\sum_{p=1}^{m} b_p y_p \\ \text{s.t.} & F_0 - \displaystyle\sum_{p=1}^{m} F_p y_p - \sum_{p=1}^{m} G_p y_p^2 \succeq O, \\ & 0 \leq y_p \leq d_p, \quad p = 1, \ldots, m, \end{array}\right\} \tag{5.9}$$

where $F_0 = -\bar{\Omega} M_0$ and

$$F_p = \sum_{j \in J_p} (-K_j^{\mathrm{a}} + \bar{\Omega} M_j), \quad G_p = -\frac{1}{4\pi} \sum_{j \in J_p} K_j^{\mathrm{b}}, \quad b_p = -\sum_{j \in J_p} l_j, \quad p = 1. \ldots, m.$$

Suppose further that an orthogonal matrix P is found that simultaneously block-diagonalizes the coefficient matrices as

$$\begin{aligned} P^\top F_p P &= \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes \widetilde{F}_{pj}), \quad p = 0, 1, \ldots, m, \\ P^\top G_p P &= \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes \widetilde{G}_{pj}), \quad p = 1, \ldots, m. \end{aligned}$$

Then the inequality $F_0 - \sum_{p=1}^{m} F_p y_p - \sum_{p=1}^{m} G_p y_p^2 \succeq O$ in (5.9) is decomposed into a set of smaller-sized quadratic matrix inequalities

$$\widetilde{F}_{0j} - \sum_{p=1}^{m} \widetilde{F}_{pj} y_p - \sum_{p=1}^{m} \widetilde{G}_{pj} y_p^2 \succeq O, \quad j = 1, \ldots, \ell.$$

Then the problem (5.9) is rewritten equivalently to

$$\left.\begin{array}{ll} \max & \sum_{p=1}^{m} b_p y_p \\ \text{s.t.} & \widetilde{F}_{0j} - \sum_{p=1}^{m} \widetilde{F}_{pj} y_p - \sum_{p=1}^{m} \widetilde{G}_{pj} y_p^2 \succeq O, \quad j = 1, \ldots, \ell, \\ & 0 \leq y_p \leq d_p, \quad p = 1, \ldots, m. \end{array}\right\} \tag{5.10}$$

The original problem (5.9) can be regarded as a special case of (5.10) with $\ell = 1$.

We now briefly explain the SOS–SDP relaxation method [15], which we shall apply to the quadratic SDP of (5.10). It is an extension of the SOS–SDP relaxation method of Lasserre [21] for a polynomial optimization problem to a polynomial SDP. See also [10, 11, 17].

We use the notation $y^\alpha = y_1^{\alpha_1} y_2^{\alpha_2} \cdots y_m^{\alpha_m}$ for $\alpha \in \mathbb{Z}_+^m$ and $y = (y_1.y_2, \ldots, y_m)^\top \in \mathbb{R}^m$, where $\mathbb{Z}_+^m$ denotes the set of m-dimensional nonnegative integer vectors. An $n \times n$ polynomial matrix means a polynomial in y with coefficients of $n \times n$ matrices, i.e., an expression like $H(y) = \sum_{\alpha \in \mathcal{H}} H_\alpha y^\alpha$ with a nonempty finite subset $\mathcal{H}$ of $\mathbb{Z}_+^m$ and a family of matrices $H_\alpha \in \mathcal{M}_n$ indexed by $\alpha \in \mathcal{H}$. We refer to $\deg(H(y)) = \max\left\{\sum_{p=1}^{m} \alpha_p \mid \alpha \in \mathcal{H}\right\}$ as the degree of $H(y)$. The set of $n \times n$ polynomial matrices in y is denoted by $\mathcal{M}_n[y]$, whereas $\mathcal{S}_n[y]$ denotes the set of $n \times n$ symmetric polynomial matrices, i.e., the set of $H(y)$'s with $H_\alpha \in \mathcal{S}_n$ ($\alpha \in \mathcal{H}$). For $n = 1$, we have $\mathcal{S}_1[y] = \mathcal{M}_1[y]$, which coincides with the set $\mathbb{R}[y]$ of polynomials in y with real coefficients.

A polynomial SDP is an optimization problem defined in terms of a polynomial $a(y) \in \mathbb{R}[y]$ and a number of symmetric polynomial matrices $B_j(y) \in \mathcal{S}_{n_j}[y]$ ($j = 1, \ldots, L$) as

$$\text{PSDP: } \min \quad a(y) \quad \text{s.t. } B_j(y) \succeq O, \quad j = 1, \ldots, L. \tag{5.11}$$

We assume that PSDP has an optimal solution with a finite optimal value ζ^*. The quadratic SDP (5.10) under consideration is a special case of PSDP with $L = \ell + 2m$ and

$$\begin{aligned} B_j(y) &= \widetilde{F}_{0j} - \sum_{p=1}^{m} \widetilde{F}_{pj} y_p - \sum_{p=1}^{m} \widetilde{G}_{pj} y_p^2, \quad j = 1, \ldots, \ell, \\ B_{\ell+p}(y) &= y_p, \quad B_{\ell+m+p}(y) = d_p - y_p, \quad p = 1, \ldots, m. \end{aligned}$$

PSDP is a nonconvex problem, and we shall resort to an SOS–SDP relaxation method.

We introduce SOS polynomials and SOS polynomial matrices. For each nonnegative integer ω define

$$\begin{aligned} \mathbb{R}[y]^2_\omega &= \left\{ \sum_{i=1}^k g_i(y)^2 \mid g_i(y) \in \mathbb{R}[y],\ \deg(g_i(y)) \le \omega\ (i = 1, \dots, k) \text{ for some } k \right\}, \\ \mathcal{M}_n[y]^2_\omega &= \left\{ \sum_{i=1}^k G_i(y)^\top G_i(y) \mid G_i(y) \in \mathcal{M}_n[y],\ \deg(G_i(y)) \le \omega\ (i = 1, \dots, k) \text{ for some } k \right\}. \end{aligned}$$

With reference to PSDP in (5.11) let $\omega_0 = \lceil \deg(a(y))/2 \rceil$, $\omega_j = \lceil \deg(B_j(y))/2 \rceil$, and $\omega_{\max} = \max\{\omega_j \mid j = 0, 1, \dots, L\}$, where $\lceil \cdot \rceil$ means rounding-up to the nearest integer. For $\omega \ge \omega_{\max}$, we consider an SOS optimization problem

$$\left. \begin{array}{lll} \mathrm{SOS}(\omega): & \max & \zeta \\ & \text{s.t.} & a(y) - \displaystyle\sum_{j=1}^L W_j(y) \bullet B_j(y) - \zeta \in \mathbb{R}[y]^2_\omega, \\ & & W_j(y) \in \mathcal{M}_{n_j}[y]^2_{(\omega-\omega_j)}, \quad j = 1, \dots, L. \end{array} \right\} \quad (5.12)$$

We call ω the relaxation order. Let ζ_ω denote the optimal value of $\mathrm{SOS}(\omega)$.

The sequence of $\mathrm{SOS}(\omega)$ (with $\omega = \omega_{\max}, \omega_{\max}+1, \dots$) serves as tractable convex relaxation problems of PSDP. The following facts are known:

(i) $\zeta_\omega \le \zeta_{\omega+1} \le \zeta^*$ for $\omega \ge \omega_{\max}$, and ζ_ω converges to ζ^* as $\omega \to \infty$ under a moderate assumption on PSDP.

(ii) $\mathrm{SOS}(\omega)$ can be solved numerically as an SDP, which we will write in SeDuMi format as

$$\mathrm{SDP}(\omega): \ \min \quad c(\omega)^\top x \ \text{ s.t. } \quad A(\omega)x = b(\omega),\ x \succeq 0.$$

Here $c(\omega)$, $b(\omega)$ and denote vectors, and $A(\omega)$ a matrix. We note that their construction depend on not only on the data polynomial matrices $B_j(y)$ $(j = 1, \dots, L)$, but also the relaxation order ω.

(iii) The sequence of solutions of $\mathrm{SDP}(\omega)$ provides approximate optimal solutions of PSDP with increasing accuracy under the moderate assumption.

(iv) The size of $A(\omega)$ increases as we take larger ω.

(v) The size of $A(\omega)$ increases as the size n_j of $B_j(y)$ gets larger ($j = 1, \dots, L$).

Table 3: Computational results of quadratic SDP (5.10) for frame optimization.

quadratic SDP (5.10) with $m = 5$	(a) no symmetry used	(b) symmetry used	(c) symmetry + sparsity used
number of SDP blocks ℓ	1	6	8
SDP block sizes	57	3, 4, 5, 7 9×2, 10×2	1×6, 2, 2×6, 3, 3, 5, 6×2, 7×2
SDP(ω) with $\omega = 3$			
size of $A(\omega)$	461×1,441,267	461×131,938	461×68,875
number of SDP blocks	12	17	19
maximum SDP block size	1197×1197	210×210	147×147
average SDP block size	121.9×121.9	62.6×62.6	46.1×46.1
relative error $\epsilon_{\rm obj}$	6.2×10^{-9}	4.7×10^{-10}	2.4×10^{-9}
cpu time (s) for SDP(ω)	2417.6	147.4	59.5

See [15, 17] for more details about the SOS–SDP relaxation method for polynomial SDP.

Now we are ready to present our numerical results for the frame optimization problem. We consider the plane frame in Fig. 2 with 48 beam elements ($n^{\rm m} = 48$), which is symmetric with respect to the dihedral group D_6. A uniform nonstructural concentrated mass is located at each free node. The index set of members $\{1, \ldots, n^{\rm m}\}$ is divided into five orbits $J_1, \ldots, J_5$. In the quadratic SDP formulation (5.9) we have $m = 5$ and the size of the matrices F_p and G_p is $3 \times 19 = 57$. We compare three cases:

(a) Neither symmetry nor sparsity is exploited.

(b) D_6-symmetry is exploited, but sparsity is not.

(c) Both D_6-symmetry and sparsity are exploited by the proposed method.

In our computation we used a modified version of SparsePOP [31] to generate an SOS–SDP relaxation problem from the quadratic SDP (5.10), and then solved the relaxation problem by SeDuMi 1.1 [26, 28] on a 2.66 GHz Dual-Core Intel Xeon cpu with 4GB memory.

Table 3 shows the numerical data in three cases (a), (b) and (c). In case (a) we have a single ($\ell = 1$) quadratic inequality of size 57 in the quadratic SDP (5.10). In case (b) we have $\ell = 6$ distinct blocks of sizes 3, 4, 5, 7, 9 and 10 in (5.10), where 9×2 and 10×2 in the table mean that the blocks of sizes 9 and 10 appear with multiplicity 2. This is consistent with the group-theoretic fact that D_6 has four one-dimensional and two two-dimensional irreducible representations. In case (c) we have $\ell = 8$ quadratic inequalities of sizes 1, 2, 2, 3, 3, 5, 6 and 7 in (5.10).

In all cases, SDP(ω) with the relaxation order $\omega = 3$ attains an approximate optimal solution of the quadratic SDP (5.10) with high accuracy. The accuracy is monitored by ϵ_{obj}, which is a computable upper bound on the relative error $|\zeta^* - \zeta_\omega|/|\zeta_\omega|$ in the objective value. The computed solutions to the relaxation problem SDP(ω) turned out to be feasible solutions to (5.10).

We observe that our block-diagonalization works effectively. It considerably reduces the size of the relaxation problem SDP(ω), which is characterized in terms of factors such as the size of $A(\omega)$, the maximum SDP block size and the average SDP block size in Table 3. Smaller values in these factors in cases (b) and (c) than in case (a) contribute to discreasing the cpu time for solving SDP(ω) by SeDuMi. The cpu times in cases (b) and (c) are, respectively, $147.4/2417.6 \approx 1/16$ and $59.5/2417.6 \approx 1/40$ of that in case (a). Thus our block-diagonalization method significantly enhances the computational efficiency.

6 Discussion

Throughout this paper we have assumed that the underlying field is the field $\mathbb{R}$ of real numbers. Here we discuss an alternative approach to formulate everything over the field $\mathbb{C}$ of complex numbers. We denote by $\mathcal{M}_n(\mathbb{C})$ the set of $n \times n$ complex matrices and consider a $*$-algebra $\mathcal{T}$ over $\mathbb{C}$. It should be clear that $\mathcal{T}$ is a $*$-algebra over $\mathbb{C}$ if it is a subset of $\mathcal{M}_n(\mathbb{C})$ such that $I_n \in \mathcal{T}$ and it satisfies (3.1) with "$\alpha, \beta \in \mathbb{R}$" replaced by "$\alpha, \beta \in \mathbb{C}$" and "$A^\top$" by "$A^*$" (the conjugate transpose of A). Simple and irreducible $*$-algebras over $\mathbb{C}$ are defined in an obvious way.

The structure theorem for a $*$-algebra over $\mathbb{C}$ takes a simpler form than Theorem 3.1 as follows [32] (see also [2, 8]).

Theorem 6.1. *Let $\mathcal{T}$ be a $*$-subalgebra of $\mathcal{M}_n(\mathbb{C})$.*

(A) *There exist a unitary matrix $\hat{Q} \in \mathcal{M}_n(\mathbb{C})$ and simple $*$-subalgebras $\mathcal{T}_j$ of $\mathcal{M}_{\hat{n}_j}(\mathbb{C})$ for some $\hat{n}_j$ $(j = 1, 2, \ldots, \ell)$ such that*

$$\hat{Q}^* \mathcal{T} \hat{Q} = \{\mathrm{diag}\,(S_1, S_2, \ldots, S_\ell) : S_j \in \mathcal{T}_j \ (j = 1, 2, \ldots, \ell)\}.$$

(B) *If $\mathcal{T}$ is simple, there exist a unitary matrix $P \in \mathcal{M}_n(\mathbb{C})$ and an irreducible $*$-subalgebra $\mathcal{T}'$ of $\mathcal{M}_{\bar{n}}(\mathbb{C})$ for some $\bar{n}$ such that*

$$P^* \mathcal{T} P = \{\mathrm{diag}\,(B, B, \ldots, B) : B \in \mathcal{T}'\}.$$

(C) *If $\mathcal{T}$ is irreducible, then $\mathcal{T} = \mathcal{M}_n(\mathbb{C})$.*

The proposed algorithm can be adapted to the complex case to yield the decomposition stated in this theorem. Note that the assumption like (3.5) is not needed in the complex case because of the simpler statement in (C) above.

When given real symmetric matrices A_p we could regard them as Hermitian matrices and apply the decomposition over $\mathbb{C}$. The resulting decomposition is at least as fine as the one over $\mathbb{R}$, since unitary transformations contain orthogonal transformations as special cases. The diagonal blocks in the decomposition over $\mathbb{C}$, however, are complex matrices in general, and this can be a serious drawback in some applications where real eigenvalues play critical roles. This is indeed the case with structural analysis, as described in Section 5.2, of truss structures having cyclic symmetry and also with bifurcation analysis [9] of symmetric systems.

As for SDPs, the formulation over $\mathbb{C}$ is a feasible alternative. When given an SDP problem over $\mathbb{R}$ we could regard it as an SDP problem over $\mathbb{C}$ and apply the decomposition over $\mathbb{C}$. A dual pair of SDP problems over $\mathbb{C}$ can be defined by (2.1) and (2.2) with Hermitian matrices A_p $(p = 0, 1, \ldots, m)$ and a real vector $b = (b_p)_{p=1}^m \in \mathbb{R}^m$. The decision variables X and Z are Hermitian matrices, and y_p $(p = 1, \ldots, m)$ are real numbers. The interior-point method was extended to this case [24, 28]. Such embedding into $\mathbb{C}$, however, entails significant loss in computational efficiency.

Acknowledgments

The authors thank Etienne de Klerk for communicating a reference [6], Dima Pasechnik for discussion, and Takanori Maehara for pointing out a flaw in the original proof of Proposition 3.5 and for suggesting the proof of the structure theorem outlined in Appendix. This work is supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

A Proof of the Structure Theorem

A proof of the structure theorem over $\mathbb{R}$, Theorem 3.1, is outlined here. We follow the terminology of Lam [19], and quote three fundamental theorems. For a division ring D we denote by $\mathcal{M}_n(D)$ the set of $n \times n$ matrices with entries from D.

Theorem A.1 (Wedderburn–Artin [19, Theorem 3.5 & pp. 38–39])**.** (1) *Let R be any semisimple ring. Then*

$$R \simeq \mathcal{M}_{n_1}(D_1) \times \cdots \times \mathcal{M}_{n_r}(D_r) \tag{A.1}$$

for suitable division rings $D_1, \ldots, D_r$ and positive integers $n_1, \ldots, n_r$. The number r is uniquely determined, as are the pairs $(D_1, n_1), \ldots, (D_r, n_r)$ (up to a permutation).

(2) *If k is a field and R is a finite-dimensional semisimple k-algebra, each D_i above is a finite-dimensional k-division algebra.*

Theorem A.2 (Frobenius [19, Theorem 13.12])**.** *Let D be a division algebra over $\mathbb{R}$. Then, as an $\mathbb{R}$-algebra, D is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$ (division algebra of real quaternions).*

Theorem A.3 (Special case of [19, Theorem 3.3 (2)])**.** *Let D be a division algebra over $\mathbb{R}$. Then, $\mathcal{M}_n(D)$ has a unique irreducible representation in $\mathcal{M}_{kn}(\mathbb{R})$ up to equivalence, where $k = 1, 2, 4$ according to whether D is isomorphic to $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$.*

Theorem 3.3 (2) in [19] says in fact much more that any irreducible representation of a matrix algebra over some division ring is equivalent to a left regular representation. This general claim is used in [19] to prove the uniqueness of the decomposition in the Wedderburn-Artin theorem. Thus logically speaking, the claim of Theorem A.3 could be understood as a part of the statement of the Wedderburn-Artin theorem. However this theorem is usually stated as a theorem for the intrinsic structure of the algebra R, and the uniqueness of an irreducible representation of simple algebra is hidden behind. Thus we have stated Theorem A.3 to make sure what we have known extrinsically for the argument we present here.

Let $\mathcal{T}$ be a $*$-subalgebra of $\mathcal{M}_n$ over $\mathbb{R}$. We prepare some lemmas.

Lemma A.4. *If $\mathcal{T}$ is irreducible, then it is simple.*

Proof. Let $\mathcal{I}$ be an ideal of $\mathcal{T}$. Since $W = \mathrm{span}\{Ax \mid A \in \mathcal{I},\ x \in \mathbb{R}^n\}$ is a $\mathcal{T}$-invariant subspace and $\mathcal{T}$ is irreducible, we have $W = \{\mathbf{0}\}$ or $W = \mathbb{R}^n$. In the former case we have $\mathcal{I} = \{O\}$. In the latter case, for an orthonormal basis $e_1, \ldots, e_n$ there exist some $A_{ij} \in \mathcal{I}$ and $x_{ij} \in \mathbb{R}^n$ such that $e_i = \sum_j A_{ij} x_{ij}$ for $i = 1, \ldots, n$. Then $I_n = \sum_{i=1}^n e_i e_i^\top = \sum_{i=1}^n \sum_j A_{ij}(x_{ij} e_i^\top) \in \mathcal{I}$. This shows $\mathcal{I} = \mathcal{T}$. $\square$

Lemma A.5. *There exists an orthogonal matrix* Q *such that*

$$Q^\top AQ = \bigoplus_{j=1}^{\ell} \bigoplus_{i=1}^{\bar{m}_j} \rho_{ij}(A), \quad A \in \mathcal{T}, \tag{A.2}$$

for some ℓ *and* $\bar{m}_1, \ldots, \bar{m}_\ell$*, where each* ρ_{ij} *is an irreducible representation of* $\mathcal{T}$*, and* ρ_{ij} *and* $\rho_{i'j'}$ *are equivalent (as representations) if and only if* $j = j'$*.*

Proof. Let W be a $\mathcal{T}$-invariant subspace, and $W^\perp$ be the orthogonal complement of W. For any $x \in W$, $y \in W^\perp$ and $A \in \mathcal{T}$ we have $A^\top x \in W$ and hence $x^\top(Ay) = (A^\top x)^\top y = 0$, which shows $Ay \in W^\perp$. Hence $W^\perp$ is also a $\mathcal{T}$-invariant subspace. If W (or $W^\perp$) is not irreducible, we can decompose W (or $W^\perp$) into orthogonal $\mathcal{T}$-invariant subspaces. Repeating this we can arrive at a decomposition of $\mathbb{R}^n$ into mutually orthogonal irreducible subspaces. An orthonormal basis compatible with this decomposition gives the desired matrix Q, and the diagonal blocks of the block-diagonal matrix $Q^\top AQ$ give the irreducible representations $\rho_{ij}(A)$. □

Equation (A.2) shows that, by partitioning the column set of Q appropriately as $Q = (Q_{ij} \mid i = 1, \ldots, \bar{m}_j; j = 1, \ldots, \ell)$, we have

$$\rho_{ij}(A) = Q_{ij}^\top AQ_{ij}, \quad A \in \mathcal{T}. \tag{A.3}$$

Lemma A.6. $\mathcal{T}$ *is a finite-dimensional semisimple* $\mathbb{R}$*-algebra.*

Proof. For each (i, j) in the decomposition (A.2) in Lemma A.5, $\{\rho_{ij}(A) \mid A \in \mathcal{T}\}$ is an irreducible $*$-algebra, which is simple by Lemma A.4. This means that $\mathcal{T}$ is semisimple. □

Lemma A.7. *If two irreducible representations* ρ *and* $\tilde{\rho}$ *of* $\mathcal{T}$ *are equivalent, there exists an orthogonal matrix* S *such that* $\rho(A) = S^\top \tilde{\rho}(A)S$ *for all* $A \in \mathcal{T}$*.*

Proof. By the equivalence of ρ and $\tilde{\rho}$ there exists a nonsingular S such that $S\rho(A) = \tilde{\rho}(A)S$ for all $A \in \mathcal{T}$. This means also that $\rho(A)S^\top = S^\top \tilde{\rho}(A)$ for all $A \in \mathcal{T}$ (Proof: Since $\mathcal{T}$ is a $*$-algebra, we may replace A with $A^\top$ in the first equation to obtain $S\rho(A^\top) = \tilde{\rho}(A^\top)S$, which is equivalent to $S\rho(A)^\top = \tilde{\rho}(A)^\top S$. The transposition of this expression yields the desired equation). It then follows that

$$\tilde{\rho}(A)(SS^\top) = (SS^\top)\tilde{\rho}(A), \qquad A \in \mathcal{T}.$$

Let α be an eigenvalue of $SS^\top$, where $\alpha > 0$ since $SS^\top$ is positive-definite. Then

$$\tilde{\rho}(A)(SS^\top - \alpha I) = (SS^\top - \alpha I)\tilde{\rho}(A), \qquad A \in \mathcal{T}.$$

By Schur's lemma (or directly, since the kernel of $SS^\top - \alpha I$ is a nonzero subspace and $\tilde{\rho}$ is irreducible), we must have $SS^\top - \alpha I = O$. This shows that $S/\sqrt{\alpha}$ serves as the desired orthogonal matrix. □

We now start the proof of Theorem 3.1. By Lemma A.6 we can apply the Wedderburn–Artin theorem (Theorem A.1) to $\mathcal{T}$ to obtain an algebra-isomorphism

$$\mathcal{T} \simeq \mathcal{M}_{n_1}(D_1) \times \cdots \times \mathcal{M}_{n_\ell}(D_\ell). \tag{A.4}$$

Note that the last statement in (1) of Theorem A.1 allows us to assume that r in (A.1) for $R = \mathcal{T}$ is equal to ℓ in (A.2).

By Frobenius' theorem (Theorem A.2) we have $D_j = \mathbb{R}, \mathbb{C}$, or $\mathbb{H}$ for each $j = 1, \dots, \ell$. Depending on the cases we define a representation $\tilde{\rho}_j$ of $\mathcal{M}_{n_j}(D_j)$ over $\mathbb{R}$ as follows. Recall notations $C(v, w)$ and $H(v, w, x, y)$ in Theorem 3.1.

(i) If $D_j = \mathbb{R}$, then $\tilde{\rho}_j(A) = A \in \mathcal{M}_{n_j}(\mathbb{R})$.

(ii) If $D_j = \mathbb{C}$ and $A = (a_{pq}) \in \mathcal{M}_{n_j}(\mathbb{C})$ with $a_{pq} = v_{pq} + \imath w_{pq} \in \mathbb{C}$ $(p, q = 1, \dots, n_j)$, then

$$\tilde{\rho}_j(A) = \begin{bmatrix} C(v_{11}, w_{11}) & \cdots & C(v_{1n_j}, w_{1n_j}) \\ \vdots & \ddots & \vdots \\ C(v_{n_j1} w_{n_j1}) & \cdots & C(v_{n_jn_j}, w_{n_jn_j}) \end{bmatrix} \in \mathcal{M}_{2n_j}(\mathbb{R}).$$

(iii) If $D_j = \mathbb{H}$ and $A = (a_{pq}) \in \mathcal{M}_{n_j}(\mathbb{H})$ with $a_{pq} = v_{pq} + \imath w_{pq} + \jmath x_{pq} + k y_{pq} \in \mathbb{H}$ $(p, q = 1, \dots, n_j)$, then

$$\tilde{\rho}_j(A) = \begin{bmatrix} H(v_{11}, w_{11}, x_{11}, y_{11}) & \cdots & H(v_{1n_j}, w_{1n_j}, x_{1n_j}, y_{1n_j}) \\ \vdots & \ddots & \vdots \\ H(v_{n_j1}, w_{n_j1}, x_{n_j1}, y_{n_j1}) & \cdots & H(v_{n_jn_j}, w_{n_jn_j}, x_{n_jn_j}, y_{n_jn_j}) \end{bmatrix} \in \mathcal{M}_{4n_j}(\mathbb{R}).$$

We may assume, by Theorem A.3 and renumbering the indices, that ρ_{ij} in (A.2) is equivalent to $\tilde{\rho}_j$ for $i = 1, \dots, \bar{m}_j$ and $j = 1, \dots, \ell$. Then for each (i, j) there exists an orthogonal matrix S_{ij} such that

$$\rho_{ij}(A) = S_{ij}^\top \tilde{\rho}_j(A) S_{ij}, \quad A \in \mathcal{T} \tag{A.5}$$

by Lemma A.7.

With S_{ij} in (A.5) and Q_{ij} in (A.3) we put $P_{ij} = Q_{ij} S_{ij}$ and define $P = (P_{ij} \mid i = 1, \dots, \bar{m}_j; j = 1, \dots, \ell)$, which is an $n \times n$ orthogonal matrix. Then (A.2) is rewritten as

$$P^\top A P = \bigoplus_{j=1}^{\ell} \bigoplus_{i=1}^{\bar{m}_j} \tilde{\rho}_j(A) = \bigoplus_{j=1}^{\ell} (I_{\bar{m}_j} \otimes \tilde{\rho}_j(A)), \quad A \in \mathcal{T}. \tag{A.6}$$

This is the formula (3.2) with $B_j = \tilde{\rho}_j(A)$. We have thus proven Theorem 3.1.

References

[1] Y. Bai, E. de Klerk, D. V. Pasechnik and R. Sotirov, Exploiting group symmetry in truss topology optimization, Optimization and Engineering, DOI 10.1007/s11081-008-9050-6. Published online: 17 June 2008

[2] G. P. Barker, L. Q. Eifler and T. P. Kezlan, A non-commutative spectral theorem, Linear Algebra and Its Applications, Vol. 20 (1978), pp. 95–100.

[3] B. Borchers, CSDP 2.3 user's guide, a C library for semidefinite programming, Optimization Methods and Software, Vol. 11 & 12 (1999), pp. 597–611.

[4] E. de Klerk, D.V. Pasechnik and A. Schrijver, Reduction of symmetric semidefinite programs using the regular $*$-representation, Mathematical Programming, Series B, Vol. 109 (2007), pp. 613–624.

[5] E. de Klerk and R. Sotirov, Exploiting group symmetry in semidefinite relaxations of the quadratic assignment, CentER Discussion paper 2007-44, Center for Economic Research, Tilburg University, June 2007.

[6] W. Eberly and M. Giesbrecht: Efficient decomposition of separable algebras, Journal of Symbolic Computation, Vol. 37 (2004), pp. 35–81.

[7] K. Gatermann and P.A. Parrilo, Symmetry groups, semidefinite programs, and sums of squares, Journal of Pure and Applied Algebra, Vol. 192 (2004), pp. 95–128.

[8] D. Gijswijt, Matrix Algebras and Semidefimite Programming Techniques for Codes, Ph.D. thesis, University of Amsterdam, 2005.

[9] M. Golubitsky, I. Stewart and D.G. Schaeffer, Singularities and Groups in Bifurcation Theory, Vol. 2, Springer-Verlag, New York, 1988.

[10] D. Henrion and J. B. Lasserre, Convergent relaxations of polynomial matrix inequalities and static output feedback, IEEE Transactions on Automatic Control, Vol. 51 (2006), pp. 192–202.

[11] C. W. J. Hol and C. W. Scherer, Sum of squares relaxations for polynomial semidefinite programming, Proc. Symp. on Mathematical Theory of Networks and Systems (MTNS), Leuven, Belgium, 2004.

[12] L. Jansson, J.B. Lasserre, C. Riener and T. Theobald, Exploiting symmetries in SDP-relaxations for polynomial optimization, LAAS-report, Toulouse, September 2006.

[13] Y. Kanno and M. Ohsaki, Eigenvalue optimization of structures via polynomial semidefinite programming, METR 2007-31, Department of Mathematical Informatics, University of Tokyo, May 2007.

[14] Y. Kanno, M. Ohsaki, K. Murota, and N. Katoh, Group symmetry in interior-point methods for semidefinite program, Optimization and Engineering, Vol. 2 (2001), pp. 293–320.

[15] M. Kojima, Sums of squares relaxations of polynomial semidefinite programs, Research Report B-397, Tokyo Institute of Technology, November 2003.

[16] M. Kojima, S. Kojima and S. Hara, Linear algebra for semidefinite programming, Research Report B-290, Tokyo Institute of Technology, October 1994; also in RIMS Kokyuroku 1004, Kyoto University, pp. 1–23, 1997.

[17] M. Kojima and M. Muramatsu, An extension of sums of squares relaxations to polynomial optimization problems over symmetric cones, Mathematical Programming, Vol. 110 (2007), pp. 315–336.

[18] M. Kojima and M. Muramatsu, A note on sparse SOS and SDP relaxations for polynomial optimization problems over symmetric cones, Computational Optimization and Applications, Vol. 42 (2009), pp. 31–41.

[19] T. Y. Lam, A First Course in Noncommutative Rings, 2nd ed., Springer-Verlag, New York, 2001.

[20] L. Lamberti and C. Pappalettere, An efficient sequential linear programming algorithm for engineering optimization, Journal of Engineering Design, Vol. 16 (2005), pp. 353–371.

[21] J. B. Lasserre, Global optimization with polynomials and the problems of moments, SIAM Journal on Optimization, Vol. 11 (2001), pp. 796–817.

[22] T. Maehara and K. Murota: A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras, Part II: general algorithm, submitted for publication. (Original version: "A numerical algorithm for block-diagonal decomposition of matrix $*$-algebras with general irreducible components," METR 2008-26, University of Tokyo, May 2008.)

[23] W. Miller, Jr., Symmetry Groups and Their Applications, Academic Press, New York, 1972.

[24] Yu. Nesterov and A. Nemirovskii, Interior-Point Polynomial Algorithms in Convex Programming, Stud. Appl. Math. Vol. 13, SIAM, Philadelphia, 1994.

[25] M. Ohsaki, K. Fujisawa, N. Katoh and Y. Kanno, Semi-definite programming for topology optimization of trusses under multiple eigenvalue constraints, Computer Methods in Applied Mechanics and Engineering, Vol. 180 (1999), pp. 203-217.

[26] I. Pólik: Addendum to the SeDuMi user guide: version 1.1, Advanced Optimization Laboratory, McMaster University, Ontario, Canada. `http://sedumi.mcmaster.ca`.

[27] J.-P. Serre, Linear Representations of Finite Groups, Springer-Verlag, Berlin, 1977.

[28] J.F. Sturm, Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, Optimization Methods and Software, Vol. 11 & 12 (1999), pp. 625–653.

[29] K.C. Toh, M.J. Todd and R.H. Tütüncü, SDPT3—a MATLAB software package for semidefinite programming, version 1.3, Optimization Methods and Software, Vol. 11 & 12 (1999), pp. 545–581.

[30] F. Vallentin, Symmetry in semidefinite programs, Linear Algebra and Its Applications, to appear.

[31] H. Waki, S. Kim, M. Kojima, M. Muramatsu and H. Sugimoto, SparsePOP: a sparse semidefinite programming relaxation of polynomial optimization problems, ACM Transactions on Mathematical Software, Vol. 35 (2008), Article No. 15.

[32] J.H.M. Wedderburn, Lectures on Matrices, American Mathematical Society, New York, 1934; Dover, Mineola, N.Y., 2005.

[33] H. Wolkowicz, R. Saigal and L. Vandenberghe, Handbook of Semidefinite Programming: Theory, Algorithms, and Applications, Kluwer Academic Publishers, Boston, 2000.

[34] M. Yamashita, K. Fujisawa and M. Kojima, Implementation and evaluation of SDPA6.0 (SemiDefinite Programming Algorithm 6.0), Optimization Methods and Software, Vol. 18 (2003), pp. 491–505.

[35] B. Yang, Stress, Strain, and Structural Dynamics, Academic Press, Burlington, 2005.