MINISTRY OF EDUCATION AND TRAINING
QUY NHON UNIVERSITY

NGUYEN THI NGAN

SIMULTANEOUS DIAGONALIZATIONS OF MATRICES AND APPLICATIONS FOR SOME CLASSES OF OPTIMIZATION

DOCTORAL DISSERTATION IN MATHEMATICS

Binh Dinh - 2024

MINISTRY OF EDUCATION AND TRAINING
QUY NHON UNIVERSITY

NGUYEN THI NGAN

SIMULTANEOUS DIAGONALIZATIONS OF MATRICES AND APPLICATIONS FOR SOME CLASSES OF OPTIMIZATION

Speciality: Algebra and number theory
Code: 9 46 01 04

Reviewer 1: Assoc. Prof. Dr. Vu The Khoi

Reviewer 2: Assoc. Prof. Dr. Mai Hoang Bien

Reviewer 3: Assoc. Prof. Dr. Phan Thanh Nam

Board of Supervisors:
Dr. Thanh-Hieu Le
Prof. Ruey-Lin Sheu

Binh Dinh - 2024

Declaration

This dissertation was completed at the Department of Mathematics and Statistics, Quy Nhon University under the guidance of Dr. Le Thanh Hieu and Prof. Ruey-Lin Sheu. I hereby declare that the results presented in here are new and original. All of them were published in peer-reviewed journals and conferences. For using results from joint papers I have gotten permissions from my co-authors.

Binh Dinh, 2024

PhD. student

Nguyen Thi Ngan

Acknowledgments

This dissertation has been completed with the help that I am lucky to receive from my mentors, family and friends.

On this occasion, first and foremost, I would like to express my deepest gratitude to my advisors, Dr. Thanh-Hieu Le and Prof. Ruey-Lin Sheu, for their kindly help, encouragement, patient guidance and support during my studying time at Quy Nhon University. This work would not have been possible without their professional guidance and tireless enthusiasm. I am very fortunate to have had the opportunity to work with them.

I am grateful to Quy Nhon University and Tay Nguyen University for providing me with the opportunity to conduct my research and for all of the resources and support they provided. I would like to thank my mentors at the department of Mathematics and Statistics, Quy Nhon University, especially, Assoc. Prof. Le Cong Trinh, the department head, for his help during my study time at the department. I would also like to thank my colleagues at the department of Mathematics, faculty of Natural science and Technology, Tay Nguyen University, for their support and collaboration during my research.

Special thanks go to the staff of the Graduate Division, Quy Nhon University, especially, Assoc. Prof. Ho Xuan Quang, the Dean of the Graduate Division, and Ms. Huynh Thi Phuong Nga for their kindly help.

This is also an excellent opportunity for me to give thanks to my close friends for their friendship, their help and useful discussions in our weekly seminars.

Lastly, I would be remiss in not mentioning my family, especially my parents, my husband, and my children. Their belief in me has kept my spirits and motivation high during this process, their love inspires my life.

Contents

Table of Notations

$\mathbb{R}$	the field of real numbers
$\mathbb{R}^n$	the real vector space of real $n-$ vectors
$\mathbb{C}$	the field of complex numbers
$\mathbb{C}^n$	the complex vector space of complex $n-$ vectors
$\mathbb{F}$	a field (usually $\mathbb{R}$ or $\mathbb{C}$)
A, B, C, etc.	matrices
$\mathbb{F}^{m\times n}$	the set of all $m \times n$ matrices with entries in $\mathbb{F}$.
$\mathbb{R}^n_+$	the set of all n-dimensional real nonnegative vectors
$\mathbb{R}^n_{++}$	the set of all n-dimensional real positive vectors
$\mathbb{H}^n$	the set of $n \times n$ Hermitian matrices
$\mathcal{S}^n$	the set of $n \times n$ real symmetric matrices
$\mathcal{S}^n(\mathbb{C})$	the set of $n \times n$ complex symmetric matrices
x, y, z etc.	column vector; $x = (x_i) \in \mathbb{F}^n$
I_n	the identity matrix in $\mathbb{F}^{n\times n}$
0	zero scalar, vector, or matrix
$\bar{A}$	the matrix of complex conjugates of entries of $A \in \mathbb{C}^{m\times n}$
A^T	the transpose of $A \in \mathbb{C}^{m\times n}$
A^*	the conjugate transpose of $A \in \mathbb{C}^{m\times n}$, $A^* = \bar{A}^T$
A^{-1}	the inverse of a nonsingular $A \in \mathbb{F}^{n\times n}$
$(A)_p$	the $p \times p$ matrix
$A_{p\times q}$	the $p \times q$ matrix
0_p	the $p \times p$ zero matrix
$\mathrm{rank}A$	the rank of $A \in \mathbb{F}^{m\times n}$
$\mathrm{Ker}A$	the kernel of $A \in \mathbb{F}^{m\times n}$
$A \succeq 0$	matrix A is positive semidefinite
$A \succ 0$	matrix A is positive definite
$\dim_{\mathbb{F}} \ker C_t$	the dimension of $\mathbb{F}$-vector space $\ker C_t$
SDC	"simultaneously diagonalizable via congruence" or "simultaneous diagonalization via congruence"
SDS	"simultaneously diagonalizable via similarity"
`diag.`	diagonal
sym.	symmetric
invert.	invertible
dim	dimension

Introduction

Let $\mathcal{C} = \{C_1, C_2, \dots, C_m\}$ be a collection of $n \times n$ matrices with elements in $\mathbb{F}$, where $\mathbb{F}$ is the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers. If there is a nonsingular matrix R such that R^*C_iR are all diagonal, the collection $\mathcal{C}$ is then said to be *simultaneously diagonalizable via congruence*, where R^* is the conjugate transpose of R if C_i are Hermitian and simply the transpose of R if C_i are either complex or real symmetric matrices. Moreover, if there exists a nonsingular matrix S such that $S^{-1}C_iS$ is diagonal for every $i = 1, 2, \dots, m$ then $\mathcal{C}$ is called *simultaneously diagonalizable via similarity*, shortly SDS. For convenience, throughout the dissertation we use "SDC" to stand for either "simultaneously diagonalizable via congruence" or "simultaneous diagonalization via congruence" if no confusion will arise. The SDS problem is well-known and is completely solved. But the SDC problem is still open in some senses. The SDC of $\mathcal{C}$ implies that a single change of basis $x = Ry$ makes all the quadratic forms x^*C_ix simultaneously become the canonical forms. Specifically, if $R^*C_iR = \texttt{diag}(\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in})$ is the diagonal matrix with diagonal elements $\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{in}$, then x^*C_ix is transformed to the sum of squares $y^*(R^*C_iR)y = \sum_{j=1}^n \alpha_{ij}|y_j|^2$, for $i = 1, 2, \dots, m$. This is one of the properties connecting the SDC of matrices with many applications such as variational analysis [31], signal processing [14, 52, 62], quantum mechanics [57], medical imaging analysis [2, 13, 67] and many others, please see references therein. Especially, the SDC suggests a promising approach for solving quadratically constrained quadratic programming (QCQP) [17, 74, 5]. In recent studies by Ben-Tal and Hertog [6], Jiang and Li [37], Alizadeh [4], Taati [54], Adachi and Nakatsukasa [1], the SDC of two or three real symmetric matrices has been efficiently applied for solving QCQP with one or two constraints. Ben-Tal and Hertog [6] showed that if the matrices in the objective and constraint functions are SDC, the QCQP with one constraint can be recast as a convex second-order cone programming (SOCP) problem; the QCQP with two constraints can also be transformed into an equivalent SOCP under the SDC together with additional appropriate assumptions. We know that the convex SOCP is solvable efficiently in polynomial time [4]. Jiang and Li [37] applied the SDC for some classes of QCQP including the generalized trust region subproblem (GTRS), which is exactly the QCQP with one constraint, and its variants. Especially the homogeneous version of QCQP, i.e., when the linear terms in the objective and constraint functions are all zero, is reduced to a linear program if the matrices are SDC. Salahi and Taati [54] derived an efficient algorithm for solving GTRS under the SDC condition. Also under the SDC assumption, Adachi and Nakatsukasa [1] compute the positive definite interval $I_{\succ}(C_0, C_1) = \{\mu \in \mathbb{R} : C_0 + \mu C_1 \succ 0\}$ of the matrix pencil and propose an eigenvalue-based algorithm for a definite feasible

GTRS, i.e., the GTRS satisfies the Slater condition and $I_{\succ}(C_0, C_1) \neq \emptyset$.

Those important applications stimulate various studies on the problem, that we call *the SDC problem* in this dissertation. It is to find conditions on $\{C_1, C_2, \ldots, C_m\}$ ensuring the existence of a congruence matrix R for the SDC problem of real symmetric matrices [70, 27, 41, 65, 37], the SDC problem of complex symmetric matrices [34, 11] and the SDC problem of Hermitian matrices [74, 7, 34]. However, for the real setting, the best SDC results so far can only solve the case of two matrices while the case of more than two matrices is solved under the assumption of a positive semidefinite matrix pencil [37]. On the other hand, for the SDC problem of complex matrices, including the complex symmetric and Hermitian matrices, can be equivalently rephrased as a simultaneous diagonalization via similarity (SDS) problem [74, 7, 8, 11]. More importanly, the obtained results do not include algorithms for finding a congruence matrix R, except for the case of two real symmetric matrices by Jiang and Li [37]. Those unsolved issues inspire us to investigate, in this dissertation, algorithms for determining whether a class $\mathcal{C}$ is SDC and compute a congruence matrix R if it indeed is.

The SDC problem was first developed by Weierstrass [70] in 1868. He obtained sufficient SDC conditions for a pair of real symmetric matrices. Since then, several authors have extended those results, including Muth 1905 [45], Finsler 1937 [18], Albert 1938 [3], Hestenes 1940 [28], and various others. See, for example, [12, 27, 29, 30, 34, 44, 65]. The results for two matrices obtained so far can be shortly reviewed as follows. If at least one of the matrices C_1, C_2 is nonsingular, referred to as a nonsingular pair, suppose it is C_1, then C_1, C_2 are SDC if and only if $C_1^{-1}C_2$ is similarly diagonalizable [27], see also [64, 65]. If the non-singularity is not assumed, the obtained SDC results of C_1, C_2 were only sufficient. Specifically,

a) if there exist scalars $\mu_1, \mu_2 \in \mathbb{R}$ such that $\mu_1 C_1 + \mu_2 C_2 \succ 0$, then C_1, C_2 are SDC [30, 65];

b) if $\{x \in \mathbb{R}^n : x^T C_1 x = 0\} \cap \{x \in \mathbb{R}^n : x^T C_2 x = 0\} = \{0\}$ then C_1, C_2 are SDC [44, 59, 65].

Actually, the classical Finsler theorem [18] in 1937 indicated that these two conditions a) and b) are equivalent whenever $n \geq 3$. It has to wait until Hoi [74] in 1970 and independently Becker [5] in 1980 for a necessary and sufficient SDC condition for a pair of Hermitian matrices. Unfortunately, when more than two matrices are involved, none of those aforementioned results remains true. In 1990 and 1991, Binding [7, 8] provided some equivalent conditions, which link to the generalized eigenvalue problem and numerical range of Hermitian matrices or to the generalized eigenvalue problem, for a finite collection of Hermitian matrices to be SDC by a unitary matrix. However, there is still lack of algorithms for finding a congruence matrix R. In 2002, Hiriart-Urruty and M. Torki [29] and then, in 2007, Hiriart-Urruty [30] proposed an open problem to *find sensible and "palpable" conditions on* $C_1, C_2, \ldots, C_m$ *ensuring they are simultaneously diagonalizable via congruence*. In 2016 Jiang and Li [37] obtained a necessary and sufficient SDC condition for a pair of real symmetric matrices and proposed an algorithm for finding a congruence matrix R if it exists. Nevertheless, we find that the result of Jiang and Li [37] is not complete. A missing case not considered in their paper is now added to make it up in this dissertation. For more than two matrices, Jiang and Li [37] proposed a necessary and sufficient SDC condition under the existence assumption of a semidefinite matrix pencil. After this result, an open question still remains to be investigated: *solving the SDC problem of more than two real symmetric matrices without semidefinite matrix pencil assumption?* In 2020, Bustamante et al. [11] proposed a necessary and sufficient condition for a set of complex symmetric matrices to be SDC by equivalently rephrasing the SDC problem as the classical problem of simultaneous diagonalization via similarity (SDS) of a new related set of matrices. A procedure to determine in a finite number of steps whether or not a set of complex symmetric matrices is SDC was also proposed. However, the SDC results of complex symmetric matrices may not hold for the real setting. That is, even the given matrices $C_1, C_2, \ldots, C_m$ are all real, the resulting matrices R and $R^T C_i R$ may have to be complex, please see [11, Example 16], and also in Example 2.1.7. Apparently, the SDC of complex symmetric matrices does also not hold for the Hermitian matrices, please see [34, Theorem 4.5.15], Example 2.1.7.

The dissertation presents several new results on the SDC of Hermitian matrices and of real symmetric matrices. Specially, the results include algorithms for answering whether the matrices are SDC and returning a congruence matrix if it exists. We also present some applications of the SDC of $\mathcal{C}$ to some related problems including computing the positive semidefinite interval of matrix pencil; solving QCQP, GTRS in particular; and maximizing a sum of generalized Rayleigh quotients.

The dissertation is organized as follows. In Chapter 1 we present some related concepts and obtained results so far of the SDC problem including the SDC of real symmetric matrices, complex symmetric matrices and Hermitian matrices. In Chapter 2 we first focus on solving the SDC problem of Hermitian matrices, i.e., when C_i are all Hermitian. This part is based on the results in [42]. The main contributions of this part are as follows.

- We develop sufficient and necessary conditions (see Theorems 2.1.4 and 2.1.5) for a collection of finitely many Hermitian matrices to be simultaneously diagonalizable via $*$-congruence. The proofs use only matrix computation techniques;

- Interestingly, one of the conditions shown in Theorem 2.1.5 requires the existence of a positive definite solution of a system of linear equations over Hermitian matrices. This leads to the use of the SDP solvers (for example, SDPT3 [63]) for checking the simultaneous diagonalizability of the initial Hermitian matrices. In case the matrices are SDC, i.e., such a positive definite solution exists, we apply the existing Jacobi-like method in [10, 43] to simultaneously diagonalize the commuting Hermitian matrices that are the images of the initial ones under the congruence defined by the square root of the above positive definite solution. The Hermitian SDC problem is hence completely solved. As a consequence, this solves the long-standing SDC problem for real symmetric matrices mentioned as an open problem in [30], and for arbitrary square matrices since any square matrix is a summation of its Hermitian and skew Hermitian parts (see Theorem 2.1.6);

- In line with giving the equivalent condition that requires the maximum rank of Hermitian pencils (Theorem 2.1.2), we suggest a Schmüdgen-like algorithm for finding such the maximum rank in Algorithm 2. This methodology may also be applied in some other simultaneous diagonalizations, for example, that in [11];

- Finally, we propose corresponding algorithms the most important one of which is Algorithm 6 for solving the Hermitian SDC problem. These are implemented in MATLAB. The main algorithm consists of two stages which are summarized as follows: For $C_1, \ldots, C_m \in \mathbb{H}^n$,

  STAGE 1: Checking if there is a positive definite matrix P solving an appropriate semidefinite program based on Theorem 2.1.5 iii). Our main contribution stays in this part.

  STAGE 2: If such a P exists, apply Algorithm 5 [10, 43] to find a unitary matrix V that simultaneously diagonalizes the new commuting Hermitian matrices $\sqrt{P}C_i\sqrt{P}$, $i = 1, \ldots, m$.

The second part of Chapter 2 is based on [49], which focuses on the SDC problem of the real symmetric matrices, i.e., when C_i are all real symmetric. Although, in Theorem 2.1.5, our results (i)-(iii) on the Hermitian matrices can also apply to the real setting, get we find that the decomposition techniques for two matrices in [37] can be generalized to construct an inductive procedure for the SDC problem of $\mathcal{C}$ with $m \geq 3$. The approach based on [37] may be better than the SDP one, please see Example 2.2.2. To this end, the collection $\mathcal{C}$ is divided into two cases: the *nonsingular* *collection,* denoted by $\mathcal{C}_{ns}$, when at least one $C_i \in \mathcal{C}$ is non-singular. Without loss of generality, we always assume that C_1 is non-singular. On the other hand, the *singular collection,* denoted by $\mathcal{C}_s$, when all $C_i's$ in $\mathcal{C}$ are non-zero but singular. For the non-singular collection $\mathcal{C}_{ns}$, the arguments first apply to $\{C_1, C_2\}$; if C_1, C_2 are SDC then a matrix $Q^{(1)}$ is constructed at the first iteration such that $C_2^{(1)} := (Q^{(1)})^T C_2 Q^{(1)}$ is a non-homogeneous dilation of $C_1^{(1)} := (Q^{(1)})^T C_1 Q^{(1)}$, while $C_j^{(1)} := (Q^{(1)})^T C_j Q^{(1)}, j \geq 3$ share the same block diagonal structure of $C_1^{(1)}$, please see Lemma 2.2.2 and Remark 2.2.1 below. At the second iteration, $\{C_1^{(1)}, C_3^{(1)}\}$ are checked. If $C_1^{(1)}, C_3^{(1)}$ are SDC, then $Q^{(2)}$ is constructed such that $C_3^{(2)} := (Q^{(2)})^T C_3^{(1)} Q^{(2)}$ and $C_2^{(2)} := (Q^{(2)})^T C_2^{(1)} Q^{(2)}$ are non-homogeneous dilations of $C_1^{(2)} := (Q^{(2)})^T C_1^{(1)} Q^{(2)}$. Next, $\{C_1^{(2)}, C_4^{(2)}\}$ are considered at the third step; and so forth. These results are presented in Sect. 2.2.1. For the singular collection $\mathcal{C}_s$, we also begin with $\{C_1, C_2\}$. If the matrices C_1 and C_2 are SDC, we find a nonsingular matrix U_1 to get

$$\begin{aligned}
\hat{C}_1 &:= U_1^T C_1 U_1 = \mathtt{diag}((C_{11})_{p_1}, 0_{n-p_1}), p_1 < n, \\
\hat{C}_2 &:= U_1^T C_2 U_1 = \mathtt{diag}((C_{21})_{p_1}, 0_{n-p_1})
\end{aligned}$$

such that $(C_{11})_{p_1}, (C_{21})_{p_1}$ are SDC and $(C_{21})_{p_1}$ is nonsingular. At the second step, we consider the SDC of $\hat{C}_1, \hat{C}_2$ and $\hat{C}_3 = U_1^T C_3 U_1$. If they are SDC, we find a nonsingular matrix U_2 to get

$$\begin{aligned}
\breve{C}_1 &:= U_2^T \hat{C}_1 U_2 = \mathtt{diag}((C_{11})_{p_2}, 0_{n-p_2}), p_1 \leq p_2, \\
\breve{C}_2 &:= U_2^T \hat{C}_2 U_2 = \mathtt{diag}((C_{21})_{p_2}, 0_{n-p_2}), \\
\breve{C}_3 &:= U_2^T \hat{C}_3 U_2 = \mathtt{diag}((C_{31})_{p_2}, 0_{n-p_2})
\end{aligned}$$

such that $(C_{11})_{p_2}, (C_{21})_{p_2}, (C_{31})_{p_2}$ are SDC and $(C_{31})_{p_2}$ is nonsingular; and so forth. By this way, we show that if $\mathcal{C}_s$ is SDC, we can create a new collection $\tilde{\mathcal{C}}_s = \{\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m\}$ such that $\tilde{C}_i = \mathtt{diag}((C_{i1})_p, 0_{n-p}), p \leq n$, and $(C_{(m-1)1})_p$ is nonsingular. Importantly, the given collection $\mathcal{C}_s$ is SDC if and only if $(C_{11})_p, (C_{21})_p, \ldots, (C_{(m-1)1})_p, (C_{m1})_p$ are SDC. Therefore, we move from the SDC of a singular collection to the SDC of a non-singular collection; please see Theorem 2.2.3 in Sect. 2.2.3.

Chapter 3 is devoted to presenting some applications of the SDC results. We first show how to explore the SDC properties of two real symmetric matrices C_1 and C_2 to compute the positive semidefinite interval $I_{\succeq}(C_1, C_2) = \{\mu \in \mathbb{R} : C_1 + \mu C_2 \succeq 0\}$ of matrix pencil $C_1 + \mu C_2$. Indeed, we show that if C_1, C_2 are not SDC, then $I_{\succeq}(C_1, C_2)$ has at most one value μ, while if C_1, C_2 are SDC, $I_{\succeq}(C_1, C_2)$ could be empty, a singleton set or an interval. Each case helps to analyze when the GTRS is unbounded from below, has a unique Lagrange multiplier or has an optimal Lagrange multiplier μ^* in a given closed interval. Such a μ^* can be computed by a bisection algorithm. This results follow from [47]. The next application will be for QCQP which takes the following format

$$
\text{(QCQP)} \qquad \begin{array}{ll} \min & x^T C_1 x + 2a_1^T x \\ \text{s.t.} & x^T C_i x + 2a_i^T x + b_i \leq 0, \ i = 2, \ldots, m, \end{array}
$$

where $a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$. We show that if the matrices C_i in the objective and constraint fucntions are SDC, the QCQP can be relaxed to a convex SOCP problem. In general, the ralaxation admits a positive gap. That is, the optimal value of the relaxed SOCP is strictly less than that of the primal QCQP. The cases with a tight ralaxation will be presented in that chapter. Especially, if the matrices C_i are SDC and the QCQP is homogeneous, i.e., $a_i = 0$ for $i = 1, 2, \ldots, m$, then QCQP is reduced to a linear programming after two times of changing variables. A special case of the homogeneous QCQP, which minimizes a quadratic form subjective to two homogeneous quadratic constraints over the unit sphere [46], is reduced to a linear programming problem on a simplex if the matrices are SDC. Finally, we show the applications for solving a generalized Rayleigh quotient problem which maximizes a sum of generalized Rayleigh quotients.

Chapter 1

Preliminaries

The main purpose of this chapter is to provide basic concepts and existing results for matrices such as similarity diagonalization, spectral decomposition and others. For completeness, some results are accompanied by a short proof. In addition, most of SDC results of two matrices, including of real symmetric matrices, complex symmetric matrices and Hermitian matrices, will be presented in this chapter. We also present our new result on decomposition of two real singular symmetric matrices into blocks, which is a missing case in Jiang and Li's study [37] and now dealt with in this dissertation. Please see Lemma 1.2.8 and Theorem 1.2.1 below.

1.1 Some prepared concepts for the SDC problems

Let us begin with some notations, $\mathbb{F}$ denotes the field of real numbers $\mathbb{R}$ or complex ones $\mathbb{C}$, and $\mathbb{F}^{n\times n}$ is the set of all $n \times n$ matrices with entries in $\mathbb{F}$; $\mathbb{H}^n$ denotes the set of $n \times n$ Hermitian matrices, $\mathcal{S}^n$ denotes the set of $n \times n$ real symmetric matrices and $\mathcal{S}^n(\mathbb{C})$ denotes the set of $n \times n$ complex symmetric matrices. In addition,

- The matrices $C_1, C_2, \ldots, C_m \in \mathbb{F}^{n\times n}$ are said to be SDS on $\mathbb{F}$, shortly written as $\mathbb{F}$-*SDS* or shorter *SDS*, if there exists a nonsingular matrix $P \in \mathbb{F}^{n\times n}$ such that every $P^{-1}C_iP$ is diagonal in $\mathbb{F}^{n\times n}$.

  When $m = 1$, we will say "C_1 is *similar* to a diagonal matrix" or "C_1 is diagonalizable (via similarity)" as usual;

- The matrices $C_1, C_2, \ldots, C_m \in \mathbb{H}^n$ are said to be SDC on $\mathbb{C}$, shortly written as $*$-*SDC,* if there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that every P^*C_iP is diagonal in $\mathbb{R}^{n\times n}$. Here we emphasize that P^*C_iP must be real (if diagonal) due to the hemitianian of C_i and P^*C_iP.

When $m = 1$, we will say "C_1 is congruent to a diagonal matrix" as usual;

- The matrices $C_1, C_2, \ldots, C_m \in \mathcal{S}^n$ are said to be SDC on $\mathbb{R}$, shortly written as $\mathbb{R}$-*SDC,* if there exists a nonsingular matrix $P \in \mathbb{R}^{n\times n}$ such that every P^TC_iP is diagonal in $\mathbb{R}^{n\times n}$.

  When $m = 1$, we will also say "C_1 is congruent to a diagonal matrix" as usual;

- Matrices $C_1, C_2, \ldots, C_m \in \mathcal{S}^n(\mathbb{C})$ are said to be SDC on $\mathbb{C}$ if there exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that every P^TC_iP is diagonal in $\mathbb{C}^{n\times n}$. We also abbreviate this as $\mathbb{C}$-*SDC*.

  When $m = 1$, we will also say "C_1 is congruent to a diagonal matrix" as usual.

Some important properties of matrices which will be used later in the dissertation.

Lemma 1.1.1 ([34], Lemma 1.3.10)**.** *Let $A \in \mathbb{F}^{n\times n}$, $B \in \mathbb{F}^{m\times m}$. The matrix $M = \texttt{diag}(A, B)$ is diagonalizable via similarity if and only if so are both A and B.*

Lemma 1.1.2 ([34], Problem 15, Section 1.3)**.** *Let $A, B \in \mathbb{F}^{n\times n}$ and*

$$A = \texttt{diag}(\alpha_1 I_{n_1}, \ldots, \alpha_k I_{n_k})$$

with distinct scalars α_i's. If $AB = BA$, then $B = \texttt{diag}(B_1, \ldots, B_k)$ with $B_i \in \mathbb{F}^{n_i\times n_i}$ for every $i = 1, \ldots, k$. Furthermore, B is Hermitian (resp., symmetric) if and only if so are all B_i's.

Proof. Partition B as $B = (B_{ij})_{i,j=1,2,\ldots,k}$, where each B_{ii} is a square submatrix of size $n_i \times n_i$, $i = 1, 2, \ldots, k$ and off-diagonal blocks $B_{ij}, i \neq j$, are of appropriate sizes. It then follows from

$$\begin{pmatrix} \alpha_1 B_{11} & \ldots & \alpha_1 B_{1k} \\ \vdots & \ddots & \vdots \\ \alpha_k B_{k1} & \ldots & \alpha_k B_{kk} \end{pmatrix} = AB = BA = \begin{pmatrix} \alpha_1 B_{11} & \ldots & \alpha_k B_{1k} \\ \vdots & \ddots & \vdots \\ \alpha_1 B_{k1} & \ldots & \alpha_k B_{kk} \end{pmatrix}$$

that $\alpha_i B_{ij} = \alpha_j B_{ij}, \forall i \neq j$. Thus $B_{ij} = 0$ for every $i \neq j$.

□

Lemma 1.1.3 ([34], Theorem 4.1.5)**.** *(***The spectral theorem of Hermitian matrices***) Every $A \in \mathbb{H}^n$ can be diagonalized via similarity by a unitary matrix. That is, it can be written as $A = U\Lambda U^*$, where U is unitary and Λ is real diagonal and is uniquely defined up to a permutation of diagonal elements.*

Moreover, if $A \in \mathcal{S}^n$ then U can be picked to be real.

We now present some preliminary result on the rank of a matrix pencil, which is the main ingredient in our study on Hermitian matrices in Chapter 2.

Lemma 1.1.4. *Let $C_1, \ldots, C_m \in \mathbb{H}^n$ and denote $\mathfrak{C}(\lambda) = \lambda_1 C_1 + \cdots + \lambda_m C_m$, $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$. Then the following hold*

(i) $\bigcap_{\lambda\in\mathbb{R}^m} \ker \mathfrak{C}(\lambda) = \bigcap_{i=1}^m \ker C_i = \ker C$, *where* $C = \begin{pmatrix} C_1 & \ldots & C_m \end{pmatrix}^*$.

(ii) $\max\{\text{rank}\mathfrak{C}(\lambda) |\ \lambda \in \mathbb{R}^m\} \le \text{rank}C$.

(iii) *Suppose* $\dim_{\mathbb{F}}(\bigcap_{i=1}^m \ker C_i) = k$. *Then* $\bigcap_{i=1}^m \ker C_i = \ker \mathfrak{C}(\underline{\lambda})$ *for some* $\underline{\lambda} \in \mathbb{R}^m$ *if and only if* $\text{rank}\mathfrak{C}(\underline{\lambda}) = \max_{\lambda\in\mathbb{R}^m} \text{rank}\mathfrak{C}(\lambda) = \text{rank}C = n - k$.

Proof.

(i) We have $\bigcap_{i=1}^m \ker C_i \subseteq \bigcap_{\lambda\in\mathbb{R}^m} \ker \mathfrak{C}(\lambda)$.

On the other hand, for any $x \in \bigcap_{\lambda\in\mathbb{R}^m} \ker \mathfrak{C}(\lambda)$, we have $\mathfrak{C}(\lambda)x = \sum_{i=1}^m \lambda_i C_i x = 0, \forall \lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$. Implying $\sum_{i=1}^m \lambda_i C_i x = 0, \forall \lambda = (0, \ldots, 0, 1, 0, \ldots, 0) \in \mathbb{R}^m$. Then, $C_i x = 0, \forall i = 1, 2, \ldots, m$, and $\bigcap_{\lambda\in\mathbb{R}^m} \ker \mathfrak{C}(\lambda) \subseteq \bigcap_{i=1}^m \ker C_i$.

Similarly, we also have $\bigcap_{i=1}^m \ker C_i = \ker C$.

(ii) The part (ii) follows from the fact that

$$\text{rank}\mathfrak{C}(\lambda) = \text{rank}\left[\begin{pmatrix} \lambda_1 I & \ldots & \lambda_m I \end{pmatrix} \begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix}\right] \le \text{rank}\begin{pmatrix} C_1 \\ \vdots \\ C_m \end{pmatrix} = \text{rank}C,$$

for all $\lambda \in \mathbb{R}^m$.

(iii) Using the part (i), we have $\ker C = \bigcap_{i=1}^m \ker C_i \subseteq \ker \mathfrak{C}(\underline{\lambda})$. Then by the part (ii),

$$\begin{aligned} \bigcap_{i=1}^m \ker C_i = \ker \mathfrak{C}(\underline{\lambda}) &\Longleftrightarrow \dim_{\mathbb{F}}(\ker \mathfrak{C}(\underline{\lambda})) = \dim_{\mathbb{F}}\left(\bigcap_{i=1}^m \ker C_i\right) = n - \text{rank}C \\ &\Longleftrightarrow \text{rank}\mathfrak{C}(\underline{\lambda}) = \text{rank}C = n - k \ge \text{rank}\mathfrak{C}(\lambda), \forall \lambda \in \mathbb{R}^m. \end{aligned}$$

This is certainly equivalent to $n-k = \text{rank}\mathfrak{C}(\underline{\lambda}) = \max_{\lambda\in\mathbb{R}^m}\text{rank}\mathfrak{C}(\lambda)$.

□

Compared with the SDC, which has existed for a long time in literature, the SDS seems to be solved much earlier as shown in [34].

Lemma 1.1.5 ([34], Theorem 1.3.19)**.** *Let $C_1, \ldots, C_m \in \mathbb{F}^{n\times n}$ be such that each of them is similar to a diagonal matrix in $\mathbb{F}^{n\times n}$. Then $C_1, \ldots, C_m$ are $\mathbb{F}$-SDS if and only if C_i commutes with C_j for $i < j$.*

The following result is simple but important to Lemma 1.2.14 below and Theorem 2.1.4 in Chapter 2.

Lemma 1.1.6. *Let $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m \in \mathbb{H}^n$ be singular and $C_1, C_2, \ldots, C_m \in \mathbb{H}^p$, $p < n$ so that*

$$\tilde{C}_i = \textbf{\textit{diag}}((C_i)_p, 0_{n-p}). \tag{1.1}$$

Then $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $*$-SDC if and only if $C_1, C_2, \ldots, C_m$ are $*$-SDC.

Moreover, the lemma is also true for the real symmetric setting: $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m \in \mathcal{S}^n$ are $\mathbb{R}$-SDC if and only if $C_1, C_2, \ldots, C_m \in \mathcal{S}^p$ are $\mathbb{R}$-SDC.

Proof. If $C_1, C_2, \ldots, C_m$ are $*$-SDC by a nonsingular matrix Q then $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $*$-SDC by the nonsingular matrix $\tilde{Q} = \texttt{diag}(Q, I_{n-p})$ with I_{n-p} being the $(n-p)\times(n-p)$ unit matrix.

Conversely, suppose $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $*$-SDC by a nonsingular matix U. Partition

$$U = \begin{pmatrix} U_1 & U_2 \\ U_3 & U_4 \end{pmatrix},$$

where $U_1 \in \mathbb{C}^{p\times p}, U_4 \in \mathbb{C}^{(n-p)\times(n-p)}$.

For every $i = 1, 2, \ldots, m$, the matrix

$$U^* \begin{pmatrix} C_i & 0 \\ 0 & 0_p \end{pmatrix} U = \begin{pmatrix} U_1^*\hat{C}_iU_1 & U_1^*\hat{C}_iU_2 \\ U_2^*\hat{C}_iU_1 & U_2^*\hat{C}_iU_2 \end{pmatrix}$$

is diagonal. This implies $U_1^*C_iU_1$ and $U_2^*C_iU_2$ are diagonal. Since U is nonsingular, we can assume U_1 is nonsingular after multiplying on the right of U by an appropriate permutation matrix. This means U_1 simultaneously diagonalizes $\tilde{C}_i$'s.

The case $\tilde{C}_i \in \mathcal{S}^n, C_i \in \mathcal{S}^p, i = 1, 2, \ldots, m$, is proved similarly.

□

1.2 Existing SDC results

In this section we recall the obtained SDC results so far. The simplest case is of two matrices.

Lemma 1.2.1 ([27], p.255). *Two real symmetric matrices C_1, C_2, with C_1 nonsingular, are $\mathbb{R}$-SDC if and only if $C_1^{-1}C_2$ is real similarly diagonalizable.*

A similar result but for Hermitian matrices was presented in [34, Theorem 4.5.15]. That is, if $C_1, C_2 \in \mathbb{H}^n, C_1$ is nonsingular, then C_1 and C_2 are $*$-SDC if and only if $C_1^{-1}C_2$ is real similarly diagonalizable. This conclusion also holds for complex symmetric matrices as presented in Lemma 1.2.2 below. However, the resulting diagonals in Lemma 1.2.2 may not be real.

Lemma 1.2.2 ([34], Theorem 4.5.15). *Let $C_1, C_2 \in \mathcal{S}^n(\mathbb{C})$ and C_1 is a nonsingular matrix. Then, the following conditions are equivalent:*

(i) The matrices C_1 and C_2 are $\mathbb{C}$-SDC.

(ii) There is a nonsingular $P \in \mathbb{C}^{n\times n}$ such that $P^{-1}C_1^{-1}C_2P$ is diagonal.

If the non-singularity is not assumed, the results were only sufficient.

Lemma 1.2.3 ([65], p.221). *Let $C_1, C_2 \in \mathcal{S}^n$. If $\{x \in \mathbb{R}^n : x^TC_1x = 0\} \cap \{x \in \mathbb{R}^n : x^TC_2x = 0\} = \{0\}$ then C_1 and C_2 can be diagonalized simultaneously by a real congruence transformation, provided $n \geq 3$.*

Lemma 1.2.4 ([65], p.230). *Let $C_1, C_2 \in \mathcal{S}^n$. If there exist scalars $\mu_1, \mu_2 \in \mathbb{R}$ such that $\mu_1C_1 + \mu_2C_2 \succ 0$ then C_1 and C_2 are simultaneously diagonalizable over $\mathbb{R}$ by congruence.*

This result holds also for the Hermitian matrices as presented in [34, Theorem 7.6.4]. In fact, the two Lemmas 1.2.3 and 1.2.4 are equivalent when $n \geq 3$, which is exactly Finsler's Theorem [18]. If the positive definiteness is relaxed to positive semidefiniteness, the result is as follows.

Lemma 1.2.5 ([41], Theorem 10.1). *Let $C_1, C_2 \in \mathbb{H}^n$. Suppose that there exists a positive semidefinite linear combination of C_1 and C_2, i.e., $\alpha C_1 + \beta C_2 \succeq 0$ for some $\alpha, \beta \in \mathbb{R}$, and $\ker(\alpha C_1 + \beta C_2) \subseteq \ker C_1 \cap \ker C_2$. Then C_1 and C_2 are simultaneously diagonalizable via congruence (i.e $*$-SDC), or if C_1 and C_2 are real symmetric then they are $\mathbb{R}$-SDC.*

For a singular pair of real symmetric matrices, a necessary and sufficient SDC condition, however, has to wait until 2016 when Jiang and Li [37] obtained not only theoretical SDC results but also an algorithm. The results are based on the following lemma.

Lemma 1.2.6 ([37], Lemma 5)**.** *For any two $n \times n$ singular real symmetric matrices C_1 and C_2, there always exists a nonsingular matrix U such that*

$$\tilde{A} := U^T C_1 U = \begin{pmatrix} A_1 & 0_{p\times(n-p)} \\ 0_{(n-p)\times p} & 0_{n-p} \end{pmatrix} \tag{1.2}$$

and

$$\tilde{B} := U^T C_2 U = \begin{pmatrix} B_1 & 0_{p\times q} & B_2 \\ 0_{q\times p} & B_3 & 0_{q\times r} \\ B_2^T & 0_{r\times q} & 0_r \end{pmatrix} \tag{1.3}$$

where $p, q, r \geq 0, p+q+r = n$, A_1 is a nonsingular diagonal matrix, A_1 and B_1 have the same dimension of $p \times p$, B_2 is a $p \times r$ matrix, and B_3 is a $q \times q$ nonsingular diagonal matrix.

We observe that in Lemma 1.2.6, B_3 is confirmed to be a nonsingular $q \times q$ diagonal matrix. However, we will see that the singular pair $C_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ and $C_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ cannot be converted to the forms (1.2) and (1.3). Indeed, in general we have the following result.

Lemma 1.2.7. *If $C_1 = \begin{pmatrix} \underbrace{(\hat{A}_1)_p}_{\text{invert. \& diag.}} & 0 \\ 0 & 0_{n-p} \end{pmatrix}$; $C_2 = \begin{pmatrix} (\hat{B}_1)_p & \hat{B}_2 \\ \hat{B}_2^T & 0_{n-p} \end{pmatrix} \in \mathcal{S}^n$ such that $\hat{A}_1$ is a $p \times p$ nonsingular diagonal matrix, $\hat{B}_1$ is a $p \times p$ symmetric matrix and $\hat{B}_2$ is a $p \times (n-p)$ nonzero matrix, $p < n$ then C_1 and C_2 cannot be transformed into the forms* (1.2) *and* (1.3)*, respectively.*

Proof. We suppose in contrary that C_1 and C_2 can be transformed into the forms (1.2) and (1.3), respectively. That is there exists a nonsingular U such that

$$U^T C_1 U = \begin{pmatrix} \underbrace{(A_1)_p}_{\text{invert. \& diag.}} & 0 & 0 \\ 0 & 0_{s_1} & 0 \\ 0 & 0 & 0_{n-p-s_1} \end{pmatrix}, \tag{1.4}$$

and

$$U^T C_2 U = \begin{pmatrix} (B_1)_p & 0 & B_2 \\ 0 & \underbrace{(B_3)_{s_1}}_{invert.\ \&\ diag.} & 0 \\ B_2^T & 0 & 0_{n-p-s_1} \end{pmatrix}. \tag{1.5}$$

where $(A_1)_p$ is a $p \times p$ nonsingular diagonal matrix and B_3 is a $s_1 \times s_1$ nonsingular diagonal matrix, $s_1 \leq n-p$.

We write $\hat{B}_2 = (\hat{B}_3\ \hat{B}_4)$ such that $\hat{B}_3$ is a $p \times s_1$ matrix and $\hat{B}_4$ is of size $p \times (n-p-s_1)$. Then C_1, C_2 are rewritten as

$$C_1 = \begin{pmatrix} \underbrace{(\hat{A}_1)_p}_{invert.\ \&\ diag.} & 0 & 0 \\ 0 & 0_{s_1} & 0 \\ 0 & 0 & 0_{n-p-s_1} \end{pmatrix}, \tag{1.6}$$

$$C_2 = \begin{pmatrix} (\hat{B}_1)_p & \hat{B}_3 & \hat{B}_4 \\ \hat{B}_3^T & 0_{s_1} & 0 \\ \hat{B}_4^T & 0 & 0_{n-p-s_1} \end{pmatrix} \tag{1.7}$$

and U is partitioned to have the same block structure as C_1, C_2 :

$$U = \begin{pmatrix} (U_1)_p & U_2 & U_3 \\ U_4 & (U_5)_{s_1} & U_6 \\ U_7 & U_8 & (U_9)_{n-p-s_1} \end{pmatrix}$$

Then

$$U^T C_1 U = \begin{pmatrix} U_1^T \hat{A}_1 U_1 & U_1^T \hat{A}_1 U_2 & U_1^T \hat{A}_1 U_3 \\ U_2^T \hat{A}_1 U_1 & U_2^T \hat{A}_1 U_2 & U_2^T \hat{A}_1 U_3 \\ U_3^T \hat{A}_1 U_1 & U_3^T \hat{A}_1 U_1 & U_3^T \hat{A}_1 U_3 \end{pmatrix}. \tag{1.8}$$

From (1.4) and (1.8), we have $U_1^T \hat{A}_1 U_1 = A_1$. Since $\hat{A}_1, A_1$ are nonsingular, U_1 must be nonsingular. On the other hand, $U_1^T \hat{A}_1 U_2 = U_1^T \hat{A}_1 U_3 = 0$ with U_1 and $\hat{A}_1$ nonsingular, there must be $U_2 = U_3 = 0$. The matrix U is then

$$U = \begin{pmatrix} (U_1)_p & 0 & 0 \\ U_4 & (U_5)_{s_1} & U_6 \\ U_7 & U_8 & (U_9)_{n-p-s_1} \end{pmatrix}$$

and

$$U^T C_2 U = \begin{pmatrix} \bar{B}_1 & \bar{B}_2 & \bar{B}_3 \\ \bar{B}_2^T & 0 & 0 \\ \bar{B}_3^T & 0 & 0 \end{pmatrix}, \tag{1.9}$$

where $\bar{B}_1 = U_1^T \hat{B}_1 U_1 + U_4^T \hat{B}_3^T U_1 + U_7^T \hat{B}_4^T U_1 + U_1^T \hat{B}_3 U_4 + U_1^T \hat{B}_4 U_7$; $\bar{B}_2 = U_1^T \hat{B}_3 U_5 + U_1^T \hat{B}_4^T U_8$ and $\bar{B}_3 = U_1^T \hat{B}_3 U_6 + U_1^T \hat{B}_4^T U_9$. Both (1.9) and (1.5) imply that $B_3 = 0$. This is a contradition since B_3 is nonsingular. We complete the proof. □

Lemma 1.2.7 shows that the case $q = 0$ was not considered in Jiang and Li's study, and it is now included in our Lemma 1.2.8 below. The proof is almost similar to that of Lemma 1.2.6. However, for the sake of completeness, we also show it concisely here.

Lemma 1.2.8. *Let both $C_1, C_2 \in \mathcal{S}^n$ be non-zero singular with rank$(C_1) = p < n$. There exists a nonsingular matrix U_1, which diagonalizes C_1 and rearrange its non-zero eigenvalues as*

$$\tilde{C}_1 = U_1^T C_1 U_1 = \begin{pmatrix} \underbrace{(C_{11})_p}_{invert.\ \&\ diag.} & 0 \\ 0 & 0_{n-p} \end{pmatrix}, \tag{1.10}$$

while the same congruence U_1 puts $\tilde{C}_2 = U_1^T C_2 U_1$ into two possible forms: either

$$\tilde{C}_2 = U_1^T C_2 U_1 = \begin{pmatrix} (C_{21})_p & C_{22} \\ C_{22}^T & 0_{n-p} \end{pmatrix}, \tag{1.11}$$

or

$$\tilde{C}_2 = U_1^T C_2 U_1 = \begin{pmatrix} (C_{21})_p & 0 & C_{25} \\ 0 & \underbrace{(C_{26})_{s_1}}_{invert.\ \&\ diag.} & 0 \\ C_{25}^T & 0 & 0_{n-p-s_1} \end{pmatrix}. \tag{1.12}$$

where C_{11} is a nonsingular diagonal matrix, C_{11} and C_{21} have the same dimension of $p \times p$, C_{26} is a $s_1 \times s_1$ nonsingular diagonal matrix, $s_1 \leq n - p$. If $s_1 = n - p$ then C_{25} does not exist.

Proof. One first finds an orthogonal matrix Q_1 such that

$$\tilde{C}_1 = Q_1^T C_1 Q_1 = \mathtt{diag}(\underbrace{\mathtt{diag}(\alpha_1, \alpha_2, \ldots, \alpha_p)}_{=(C_{11})_p,\ invert.}, 0_{n-p}); \tag{1.13}$$

$$Q_1^T C_2 Q_1 = \begin{pmatrix} (M_{21})_p & M_{22} \\ M_{22}^T & \underbrace{(M_{23})_{n-p}}_{sym.} \end{pmatrix}. \tag{1.14}$$

We see that (1.13) is already in the form of (1.10). If $M_{23} = 0$ in (1.14),

$$\tilde{C}_2 = Q_1^T C_2 Q_1 = \begin{pmatrix} (M_{21})_p & M_{22} \\ (M_{22})^T & 0_{n-p} \end{pmatrix},$$

which is (1.11).

Otherwise, $\text{rank} M_{23} := s_1 \geq 1$. Let P_1 be an orthogonal matrix to diagonalize the symmetric M_{23} as

$$P_1^T M_{23} P_1 = \texttt{diag}(\underbrace{(C_{26})_{s_1}}_{invert.\ \&\ diag.}, 0_{n-p-s_1}).$$

Define $H_1 = \texttt{diag}(I_p, (P_1)_{n-p})$ and compute

$$H_1^T Q_1^T C_2 Q_1 H_1 = \begin{pmatrix} (M_{21})_p & C_{24} & C_{25} \\ C_{24}^T & (C_{26})_{s_1} & 0 \\ C_{25}^T & 0 & 0_{n-p-s_1} \end{pmatrix},$$

where $(C_{24}, C_{25})_{p\times(n-p)} = M_{22} P_1$. Define further that

$$V_1 = \begin{pmatrix} I_p & 0 & 0 \\ -C_{26}^{-1} C_{24}^T & I_{s_1} & 0 \\ 0 & 0 & I_{n-p-s_1} \end{pmatrix}; \text{ and } U_1 = Q_1 H_1 V_1.$$

Note that the matrix $H_1 V_1$ does not change $Q_1^T C_1 Q_1$ that we have

$$\begin{aligned} \tilde{C}_1 = U_1^T C_1 U_1 &= V_1^T H_1^T Q_1^T C_1 Q_1 H_1 V_1 = Q_1^T C_1 Q_1 \text{ (as in (1.13))} \\ \tilde{C}_2 = U_1^T C_2 U_1 &= V_1^T H_1^T Q_1^T C_2 Q_1 H_1 V_1 \\ &= \begin{pmatrix} \underbrace{M_{21} - C_{24} C_{26}^{-1} (C_{24})^T}_{=(C_{21})_p} & 0 & C_{25} \\ 0 & (C_{26})_{s_1} & 0 \\ C_{25}^T & 0 & 0_{n-p-s_1} \end{pmatrix}. \end{aligned}$$

These are what we need in (1.10) and (1.12).

□

Using Lemma 1.2.6, Jiang and Li proposed the following result and algorithm.

Lemma 1.2.9 ([37], Theorem 6)**.** *Two singular matrices C_1 and C_2, which take the forms (1.2) and (1.3), respectively, are $\mathbb{R}$-SDC if and only if A_1 and B_1 are $\mathbb{R}$-SDC and B_2 is a zero matrix or $r = n - p - s_1 = 0$ (B_2 does not exist).*

Algorithm 1 Procedure to check whether two matrices C_1 and C_2 are $\mathbb{R}$- SDC

INPUT: Matrices $C_1, C_2 \in \mathcal{S}^n$

1: Apply the spectral decomposition to C_1 such that $\mathbf{A} := Q_1^T C_1 Q_1 = \mathtt{diag}(A_1, 0)$, where A_1 is a nonsingular diagonal matrix, and express $\mathbf{B} := Q_1^T C_2 Q_1 = \begin{pmatrix} B_1 & B_2 \\ B_2^T & B_3 \end{pmatrix}$.

2: Apply the spectral decomposition to B_3 such that $V_1^T B_3 V_1 = \begin{pmatrix} B_6 & 0 \\ 0 & 0 \end{pmatrix}$, where B_6 is a nonsingular diagonal matrix; define $Q_2 := \mathtt{diag}(I, V_1)$ and set $\hat{A} := Q_2^T \mathbf{A} Q_2 = \mathbf{A}$ and

$$\hat{B} := Q_2^T \mathbf{B} Q_2 = \begin{pmatrix} B_1 & B_4 & B_5 \\ B_4^T & B_6 & 0 \\ B_5^T & 0 & 0 \end{pmatrix}$$

3: **If** B_5 exists and $B_5 \neq 0$ **then**

4: **return** "not $\mathbb{R}$-SDC," **else**

5: Define

$$Q_3 := \begin{pmatrix} I_p & 0_{p\times q} & 0_{p\times(n-p-q)} \\ -B_6^{-1}B_4^T & I_q & 0_{q\times(n-p-q)} \\ 0_{(n-p-q)\times p} & 0_{(n-p-q)\times q} & I_{(n-p-q)} \end{pmatrix};$$

further define $\tilde{A} := Q_3^T \hat{A} Q_3 = \hat{A} = \mathbf{A}$,

$$\tilde{B} := Q_3^T \hat{B} Q_3 = \begin{pmatrix} B_1 - B_4 B_6^{-1} B_4^T & 0 & 0 \\ 0 & B_6 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

6: **If** there exists a nonsingular matrix V_2 such that $V_2^{-1} A_1^{-1} (B_1 - B_4 B_6^{-1} B_4^T) V_2 = \mathtt{diag}(\lambda_1 I_{n_1}, \ldots, \lambda_t I_{n_t})$, **then**

7: Find $R_k, k = 1, 2, \ldots, t$, which is a spectral decomposition matrix of the k^{th} diagonal block of $V_2^T A_1 V_2$; Define $R := \mathtt{diag}(R_1, R_2, \ldots, R_k)$, $Q_4 := \mathtt{diag}(V_2 R, I)$, and $P := Q_1 Q_2 Q_3 Q_4$

8: **return** two diagonal matrices $Q_4^T \tilde{A} Q_4$ and $Q_4^T \tilde{B} Q_4$ and the corresponding congruent matrix P, **else**

9: **return** "not $\mathbb{R}$-SDC"

10: **end if**

11: **end if**

As mentioned, the case $q = 0$ was not considered in Lemma 1.2.6, Lemma 1.2.9 thus does not completely characterize the SDC of C_1 and C_2. We now apply Lemma 1.2.8 to completely characterize the SDC of C_1 and C_2. Note that if $\tilde{C}_1 = U_1^T C_1 U_1$ and $\tilde{C}_2 = U_1^T C_2 U_1$ are put into (1.10) and (1.12), the SDC of C_1 and C_2 is solved by Lemma 1.2.9. Here, we would like to add an additional result to supplement Lemma 1.2.9: *Suppose $\tilde{C}_1$ and $\tilde{C}_2$ are put into (1.10) and (1.11). Then $\tilde{C}_1$ and $\tilde{C}_2$ are $\mathbb{R}$-SDC if and only if C_{11} (in (1.10)) and C_{21} (in (1.11)) are $\mathbb{R}$-SDC; and $C_{22} = 0$ (in (1.11)).* The new result needs to accomplish a couple of lemmas below.

Lemma 1.2.10. *Suppose that $A, B \in \mathcal{S}^n$ of the following forms are $\mathbb{R}$-SDC*

$$A = \boldsymbol{diag}(\underbrace{(A_1)_p}_{invert.}, 0_{n-p}), \;\; B = \begin{pmatrix} (B_1)_p & (B_2)_{p\times(n-p)} \\ B_2^T & 0_{n-p} \end{pmatrix} \tag{1.15}$$

with A_1 nonsingular and $p < n$. Then, the congruence P can be chosen to be

$$P = \begin{pmatrix} \underbrace{(P_1)_p}_{invert.} & 0 \\ P_3 & P_4 \end{pmatrix} \quad \text{such that} \quad P^TAP = \begin{pmatrix} \underbrace{(P_1^TA_1P_1)_p}_{invert.\&diag.} & 0 \\ 0 & 0_{n-p} \end{pmatrix}$$

and

$$P^TBP = \begin{pmatrix} \underbrace{P_1^TB_1P_1 + P_1^TB_2P_3 + P_3^TB_2^TP_1}_{diag.} & P_1^TB_2P_4 \\ \underbrace{P_4^TB_2^TP_1}_{=0} & 0_{n-p} \end{pmatrix}$$

and thus B must be singular. In other words, if A and B take the form (1.15) and B is nonsingular, then $\{A, B\}$ cannot be $\mathbb{R}$-SDC.

Proof. Since A, B are $\mathbb{R}$-SDC and $\operatorname{rank}(A) = p$ by the assumption, we can choose the congruence P so that the p non-zero diagonal elements of P^TAP are arranged to the north-western corner, while P^TBP is still diagonal. That is,

$$P = \begin{pmatrix} (P_1)_p & P_2 \\ P_3 & (P_4)_{n-p} \end{pmatrix} \Longrightarrow P^TAP = \begin{pmatrix} \underbrace{(P_1^TA_1P_1)_p}_{invert.\ \&\ diag.} & \underbrace{(P_1^TA_1P_2)_{p\times(n-p)}}_{=\,0} \\ \underbrace{P_2^TA_1P_1}_{=0} & \underbrace{(P_2^TA_1P_2)_{n-p}}_{=\,0} \end{pmatrix}.$$

Since $P_1^TA_1P_1$ is nonsingular diagonal and A_1 is nonsingular, P_1 must be invertible. Then, the off-diagonal $P_1^TA_1P_2 = 0$ implies that $P_2 = 0_{p\times(n-p)}$. Consequently, P and P^TBP are of the following forms

$$P = \begin{pmatrix} P_1 & 0 \\ P_3 & P_4 \end{pmatrix} \text{ and } P^TBP = \begin{pmatrix} \underbrace{P_1^TB_1P_1 + P_1^TB_2P_3 + P_3^TB_2^TP_1}_{diag.} & P_1^TB_2P_4 \\ \underbrace{P_4^TB_2^TP_1}_{=0} & 0_{n-p} \end{pmatrix}.$$

Notice that P^TBP is singular, and thus B must be singular, too. The proof is thus complete.

□

Lemma 1.2.11. *Let $A, B \in \mathcal{S}^n$ take the following formats*

$$A = \textbf{\textit{diag}}((A_1)_p, 0_{n-p}),\ B = \begin{pmatrix} (B_1)_p & (B_2)_{p\times(n-p)} \\ B_2^T & 0_{n-p} \end{pmatrix},$$

with A_1 nonsingular and B_2 of full column rank. Then, $\ker A \bigcap \ker B = \{0\}$.

Lemma 1.2.12. *Let $A, B \in \mathcal{S}^n$ with* $\ker A \bigcap \ker B = \{0\}$. *If $\alpha A + \beta B$ is singular for all real couples $(\alpha, \beta) \in \mathbb{R}^2$, then A and B are not $\mathbb{R}$-SDC.*

Proof. Suppose contrarily that A and B were $\mathbb{R}$-SDC by a congruence P such that

$$P^TAP = D_1 = \mathtt{diag}(a_1, a_2, \ldots, a_n);\ P^TBP = D_2 = \mathtt{diag}(b_1, b_2, \ldots, b_n).$$

Then, $P^T(\alpha A + \beta B)P = \mathtt{diag}(\alpha a_1 + \beta b_1, \alpha a_2 + \beta b_2, \ldots, \alpha a_n + \beta b_n)$. By assumption, $\alpha A + \beta B$ is singular for all $(\alpha, \beta) \in \mathbb{R}^2$ so that at least one of $\alpha a_i + \beta b_i = 0, \forall(\alpha, \beta) \in \mathbb{R}^2$. Let us say $\alpha a_1 + \beta b_1 = 0, \forall(\alpha, \beta) \in \mathbb{R}^2$. It implies that $a_1 = b_1 = 0$. Let $e_1 = (1, 0, \ldots, 0)^T$ be the first unit vector and notice that $Pe_1 \neq 0$ since P is nonsingular. Then,

$$P^TAPe_1 = D_1e_1 = 0;\ P^TBPe_1 = D_2e_1 = 0 \Longrightarrow 0 \neq Pe_1 \in \ker A \bigcap \ker B,$$

which is a contradiction.

□

Lemma 1.2.13. *Let $A, B \in \mathcal{S}^n$ be both singular taking the following formats*

$$A = \textbf{\textit{diag}}(\underbrace{(A_1)_p}_{invert.}, 0_{n-p}); B = \begin{pmatrix} (B_1)_p & B_2 \\ B_2^T & 0_{n-p} \end{pmatrix},$$

with A_1 nonsingular and B_2 of full column-rank. Then A and B are not $\mathbb{R}$-SDC.

Proof. From Lemma 1.2.11, we know that $\ker A \cap \ker B = \{0\}$. If $\alpha A + \beta B$ is singular for all $(\alpha, \beta) \in \mathbb{R}^2$, Lemma 1.2.12 asserts that A and B are not SDC. Otherwise, there is $(\tilde{\alpha}, \tilde{\beta}) \in \mathbb{R}^2$ such that $\tilde{\alpha}A + \tilde{\beta}B$ is nonsingular. Surely, $\tilde{\alpha} \neq 0, \tilde{\beta} \neq 0$. Then,

$$C = \frac{\tilde{\alpha}}{\tilde{\beta}}A + B = \begin{pmatrix} (\frac{\tilde{\alpha}}{\tilde{\beta}}A_1 + B_1)_p & B_2 \\ B_2^T & 0 \end{pmatrix} \text{ is nonsingular .}$$

By Lemma 1.2.10, A and C are not $\mathbb{R}$-SDC. So, A and B are not $\mathbb{R}$-SDC, either.

□

Lemma 1.2.14. *Let $C_1, C_2 \in \mathcal{S}^n$ be both singular and U_1 be nonsingular that puts $\tilde{C}_1 = U_1^T C_1 U_1$ and $\tilde{C}_2 = U_1^T C_2 U_1$ into (1.10) and (1.11) in Lemma 1.2.8. If C_{22} is nonzero, $\tilde{C}_1$ and $\tilde{C}_2$ are not $\mathbb{R}$-SDC.*

Proof. By Lemma 1.2.13, if C_{22} is of full column-rank, $\tilde{C}_1$ and $\tilde{C}_2$ are not $\mathbb{R}$-SDC. So we suppose that C_{22} has its column rank $q < n - p$ and set $s = n - p - q > 0$. There is a $(n-p) \times (n-p)$ nonsingular matrix U such that $C_{22}U = \begin{pmatrix} \hat{C}_{22} & 0_{p\times s} \end{pmatrix}$, where $\hat{C}_{22}$ is a $p \times q$ full column-rank matrix. Let $Q = \mathtt{diag}(I_p, U)$. Then,

$$\begin{aligned}\hat{C}_2 = Q^T \tilde{C}_2 Q &= \begin{pmatrix} I_p & 0_{p\times(n-p)} \\ 0_{(n-p)\times p} & U^T \end{pmatrix} \begin{pmatrix} C_{21} & C_{22} \\ C_{22}^T & 0 \end{pmatrix} \begin{pmatrix} I_p & 0_{p\times(n-p)} \\ 0_{(n-p)\times p} & U \end{pmatrix} \\ &= \begin{pmatrix} C_{21} & \hat{C}_{22} & 0_{p\times s} \\ \hat{C}_{22}^T & 0_q & 0_{q\times s} \\ 0_{s\times p} & 0_{s\times q} & 0_s \end{pmatrix};\end{aligned}$$

and

$$\hat{C}_1 = Q^T \tilde{C}_1 Q = \begin{pmatrix} C_{11} & 0_{p\times q} & 0_{p\times s} \\ 0_{q\times p} & 0_q & 0_{q\times s} \\ 0_{s\times p} & 0_{s\times q} & 0_s \end{pmatrix}.$$

Observe that, by Lemma 1.2.13, the two leading principal submatrices

$$A = \begin{pmatrix} C_{11} & 0_{p\times q} \\ 0_{q\times p} & 0_q \end{pmatrix}, B = \begin{pmatrix} C_{21} & \hat{C}_{22} \\ \hat{C}_{22}^T & 0_q \end{pmatrix}$$

of $\hat{C}_1$ and $\hat{C}_2$, respectively, are not $\mathbb{R}$-SDC since C_{11} is nonsingular (due to (1.10)) and $\hat{C}_{22}$ is of full column rank. By Lemma 1.1.6, $\hat{C}_1$ and $\hat{C}_2$ cannot be $\mathbb{R}$-SDC. Then, $\tilde{C}_1$ and $\tilde{C}_2$ cannot be $\mathbb{R}$-SDC, either. The proof is complete. □

Now, Theorem 1.2.1 comes as a conclusion.

Theorem 1.2.1. *Let C_1 and C_2 be two symmetric singular matrices of $n \times n$. Let U_1 be the nonsingular matrix that puts $\tilde{C}_1 = U_1^T C_1 U_1$ and $\tilde{C}_2 = U_1^T C_2 U_1$ into the format of (1.10) and (1.11) in Lemma 1.2.8. Then, $\tilde{C}_1$ and $\tilde{C}_2$ are $\mathbb{R}$-SDC if and only if C_{11}, C_{21} are $\mathbb{R}$-SDC and $C_{22} = 0_{p\times r}$, where $r = n - p$.*

When more than two matrices involved, the aforementioned results no longer hold true. Specifically, for more than two real symmetric matrices, Jiang and Li [37] need a positive semidefiniteness assumption of the matrix pencil. Their results can be shortly reviewd as follows.

Theorem 1.2.2 ([37], Theorem 10)**.** *If there exists* $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ *such that* $\lambda_1 C_1 + \ldots + \lambda_m C_m \succ 0$*, where, without loss of generality,* λ_m *is assumed not to be zero, then* $C_1, \ldots, C_m$ *are* $\mathbb{R}$*-SDC if and only if* $P^T C_i P$ *commute with* $P^T C_j P, \forall i \neq j$*,* $1 \leq i, j \leq m-1$*, where* P *is any nonsingular matrix that makes*

$$P^T(\lambda_1 C_1 + \ldots + \lambda_m C_m)P = I.$$

If $\lambda_1 C_1 + \ldots + \lambda_m C_m \succeq 0$, but there does not exist $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ such that $\lambda_1 C_1 + \ldots + \lambda_m C_m \succ 0$ and suppose $\lambda_m \neq 0$, then a nonsingular matrix Q_1 and the corresponding $\lambda \in \mathbb{R}^m$ are found such that

$$\mathfrak{C}_m := {Q_1}^T(\lambda_1 C_1 + \lambda_2 C_2 + ... + \lambda_m C_m)Q_1 = \mathtt{diag}(I_p, 0),$$

and

$$\mathfrak{C}_i = {Q_1}^T C_i Q_1 = \begin{pmatrix} {\mathfrak{C}_i}^1 & {\mathfrak{C}_i}^2 \\ ({\mathfrak{C}_i}^2)^T & {\mathfrak{C}_i}^3 \end{pmatrix} \tag{1.16}$$

where $\dim {\mathfrak{C}_i}^1 = dim I_p < n$. If all ${\mathfrak{C}_i}^3, i = 1, 2, \ldots, m$, are $\mathbb{R}$-SDC, then, by rearranging the common 0's to the lower right corner of the matrix, there exists a nonsingular matrix $Q_2 = \mathtt{diag}(I_p, V)$ such that

$$A_m = {Q_2}^T \mathfrak{C}_m Q_2 = \mathtt{diag}(I_p, 0) \tag{1.17}$$

and

$$A_i = {Q_2}^T \mathfrak{C}_i Q_2 = \begin{pmatrix} {A_i}^1 & {A_i}^2 & {A_i}^4 \\ ({A_i}^2)^T & {A_i}^3 & 0 \\ ({A_i}^4)^T & 0 & 0 \end{pmatrix} \tag{1.18}$$

where ${A_i}^1 = \mathfrak{C}_i^1, {A_i}^3, i = 1, 2, \ldots, m-1$, are all diagonal matrices and do not have common 0's in the same positions.

For any diagonal matrices D and E, define $supp(D) := \{i | D_{ii} \neq 0\}$ and $supp(D) \cup supp(E) := \{i | D_{ii} \neq 0 \text{ or } \ E_{ii} \neq 0\}$.

Lemma 1.2.15 ([37], Lemma 12)**.** *For* k $(k \geq 2)$ $n \times n$ *nonzero diagonal matrices* $D^1, D^2, \ldots, D^k$*, if there exists no common 0's in the same position, then the following procedure will find* $\mu_i \in \mathbb{R}, i = 1, 2, \ldots, k$*, such that* $\sum_{i=1}^k \mu_i D^i$ *is nonsingular.*

Step 1. Let $D = D^1, \mu_1 = 1$ *and* $\mu_i = 0$*, for* $i = 1, 2, \ldots, n, j = 1$*.*

Step 2. Let $D^* = D + \mu_{j+1} D^{j+1}$ *where* $\mu_{j+1} = \dfrac{s}{n}, s \in \{0, 1, 2, \ldots, n\}$ *with* s *being chosen such that* $D^* = D + \mu_{j+1} D^{j+1}$ *and* $supp(D^*) = supp(D) \cup supp(D^{j+1})$*;*

Step 3. Let $D = D^*, j = j+1$*; if* D *is nonsingular or* $j = n$*, STOP and output* D*; else, go to Step 2,* Define

$$D = \sum_{i=1}^{m-1} \lambda_i A_i = \begin{pmatrix} D_1 & D_2 & D_4 \\ D_2^T & D_3 & 0 \\ D_4^T & 0 & 0 \end{pmatrix} \tag{1.19}$$

where $\mu_i, i = 1, 2, \ldots, m-1$, are chosen, via the procedure in Lemma 1.2.15, such that D_3 is nonsingular.

Theorem 1.2.3 ([37], Theorem 13)**.** *If $\mathfrak{C}(\lambda) = \lambda_1 C_1 + \ldots + \lambda_m C_m \succeq 0$, but there does not exist $\lambda \in \mathbb{R}^m$ such that $\mathfrak{C}(\lambda) = \lambda_1 C_1 + \ldots + \lambda_m C_m \succ 0$ and suppose $\lambda_m \neq 0$, then $C_1, C_2, \ldots, C_m$ are $\mathbb{R}$-SDC if and only if $C_1, \ldots, C_{m-1}$ and $\mathfrak{C}(\lambda) = \lambda_1 C_1 + \ldots + \lambda_m C_m \succeq 0$ are $\mathbb{R}$-SDC if and only if A_i^3 (defined in (1.16)), $i = 1, 2, \ldots, m$ are $\mathbb{R}$-SDC, and the following conditions are also satisfied:*

1. *$D_4 = 0$ and $A_i^4 = 0, i = 1, 2, \ldots, m-1$.*
2. *$A_i^2 = D_2 D_3^{-1} A_i^3, i = 1, 2, \ldots, m-1$.*
3. *$A_i^1 - A_i^2 D_3^{-1} D_2^T, i = 1, 2, \ldots, m-1$, mutually commute, where A_i^1, A_i^2, A_i^3 and A_i^4 are defined in (1.18) and D is defined in (1.19).*

We notice that the assumption for the positive semidefiniteness of a matrix pencil is very restrictive. It is not difficult to find a counter example. Let

$$C_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}; C_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & -4 \end{pmatrix};$$

$$C_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}.$$

We see that C_1, C_2, C_3 are $\mathbb{R}$-SDC by a nonsingular matrix

$$P = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ \sqrt{2}-1 & \sqrt{2}+1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

However, we can check that there exists no positive semidefinite linear combination of C_1, C_2, C_3 because the inequality $\lambda_1 C_1 + \lambda_2 C_2 + \lambda_3 C_3 \succeq 0$ has no solution $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3, \lambda \neq 0$.

For a set of more than two Hermitian matrices, Binding [7] showed that the SDC problem can be equivalently transformed to the SDS type under the assumption that there exists a nonsingular linear combination of the matrices.

Lemma 1.2.16 ([7], Corollary 1.3)**.** *Let $C_1, C_2, \ldots, C_m$ be Hermitian matrices. If $\mathfrak{C}(\lambda) = \lambda_1 C_1 + \ldots + \lambda_m C_m$ is nonsingular for some $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m$. Then $C_1, C_2, \ldots, C_m$ are $*$-SDC if and only if $\mathfrak{C}(\lambda)^{-1}C_1, \mathfrak{C}(\lambda)^{-1}C_2, \ldots, \mathfrak{C}(\lambda)^{-1}C_m$ are SDS.*

As noted in Lemma 1.1.5, $\mathfrak{C}(\lambda)^{-1}C_1, \mathfrak{C}(\lambda)^{-1}C_2, \ldots, \mathfrak{C}(\lambda)^{-1}C_m$ are SDS if and only if each of which is diagonalizable and $\mathfrak{C}(\lambda)^{-1}C_i$ commutes with $\mathfrak{C}(\lambda)^{-1}C_j, i < j$.

The unsolved case when $\mathfrak{C}(\lambda) = \lambda_1 C_1 + \ldots + \lambda_m C_m$ is singular for all $\lambda \in \mathbb{R}^m$ is now solved in this dissertation. Please see Theorem 2.1.4 in Chapter 2.

A similar result but for complex symmetric matrices has been developed by Bustamante et al. [11]. Specifically, the authors showed that the SDC problem of complex symmetric matrices can always be equivalently rephrased as an SDS problem.

Lemma 1.2.17 ([11], Theorem 7)**.** *Let $C_1, C_2, \ldots, C_m \in \mathcal{S}^n(\mathbb{C})$ have maximum pencil rank n. For any $\lambda_0 = (\lambda_1, \ldots, \lambda_m) \in \mathbb{C}^m$, $\mathfrak{C}(\lambda_0) = \sum_{i=1}^m \lambda_i C_i$ with $\operatorname{rank}\mathfrak{C}(\lambda_0) = n$ then $C_1, C_2, \ldots, C_m$ are $\mathbb{C}$-SDC if and only if, $\mathfrak{C}(\lambda_0)^{-1}C_1, \ldots, \mathfrak{C}(\lambda_0)^{-1}C_m$ are SDS.*

When $\max_{\lambda \in \mathbb{C}^m} \operatorname{rank}\mathfrak{C}(\lambda) = r < n$ and $\dim \bigcap_{j=1}^m \operatorname{Ker}C_j = n - r$, there must exist a nonsingular $Q \in \mathbb{C}^{n\times n}$ such that $Q^T C_i Q = \texttt{diag}(\tilde{C}_i, 0_{n-r})$. Fix $\lambda_0 \in S^{2m-1}$, where $S^{2m-1} := \{x \in \mathbb{C}^m, \|x\| = 1\}$, $\|.\|$ denotes the usual Euclidean norm, such that $r = \operatorname{rank}\mathfrak{C}(\lambda_0)$. Reduced pencil $\tilde{C}_i$ then has nonsingular $\tilde{\mathfrak{C}}(\lambda_0)$.

Let $L_j := \tilde{\mathfrak{C}}(\lambda_0)^{-1}\tilde{C}_j, j = 1, 2, \ldots, m$, be $r \times r$ matrices, the SDC problem is now rephrased into an SDS one as follows.

Lemma 1.2.18 ([11], Theorem 14)**.** *Let* $C_1, C_2, \ldots, C_m \in \mathcal{S}^n(\mathbb{C})$ *have maximum pencil rank* $r < n$*. Then* $C_1, C_2, \ldots, C_m \in S^n(\mathbb{C})$ *are* $\mathbb{C}$*-SDC if and only if* $\dim \bigcap_{j=1}^m \mathrm{Ker} C_j = n - r$ *and* $L_1, L_2, \ldots, L_m$ *are SDS.*

Chapter 2

Solving the SDC problems of Hermitian matrices and real symmetric matrices

This chapter is devoted to presenting the SDC results first for a collection of Hermitian matrices and later for a collection of real symmetric matrices. In Section 2.1 we show the SDC results of Hermitian matrices, i.e., all matrices $C_i \in \mathcal{C}$ are Hermitian. We first provide some equivalent conditions for $\mathcal{C}$ to be SDC. Interestingly, one of these conditions requires a positive definite solution to an appropriate system of linear equations over Hermitian matrices. Based on this theoretical result, we propose a polynomial-time algorithm for numerically solving the Hermitian SDC problem. The proposed algorithm is a combination of (i) detecting whether the initial matrix collection is simultaneously diagonalizable via congruence by solving an appropriate semidefinite program and (ii) using an Jacobi-like algorithm for simultaneously diagonalizing (via congruence) the new collection of commuting Hermitian matrices derived from the previous stage. Illustrative examples and numerical tests with MATLAB are also presented. In Section 2.2 we present a constructive and inductive method for finding the SDC conditions of real symmetric matrices. Such a constructive approach helps conclude whether $\mathcal{C}$ is SDC or not and construct a congruence matrix R if it is.

2.1 The Hermitian SDC problem

This section present two methods for solving the Hermitian SDC problem: The max-rank method and the SDP method. The results are based on [42] by Le and

Nguyen.

2.1.1 The max-rank method

The max-rank method based on Theorem 2.1.4 below, in which it requires a max rank Hermitian pencil. To find this max rank we will apply the Schmüdgen's procedure [56], which is summaried as follows. Let $F \in \mathbb{H}^n$ partition as

$$F = \begin{pmatrix} \alpha & \beta \\ \beta^* & \hat{F} \end{pmatrix}, \alpha \in \mathbb{R}.$$

We then have the relations

$$X_+X_- = X_-X_+ = \alpha^2.I_n, \alpha^4 F = X_+\tilde{F}X_+^*, \tilde{F} = X_-FX_-^*, \tag{2.1}$$

where

$$X_\pm = \begin{pmatrix} \alpha & 0 \\ \pm\beta^* & \alpha I_{n-1} \end{pmatrix}, \tilde{F} = \begin{pmatrix} \alpha^3 & 0 \\ 0 & \alpha(\alpha\hat{F} - \beta^*\beta) \end{pmatrix} := \begin{pmatrix} \alpha^3 & 0 \\ 0 & F_1 \end{pmatrix} \in \mathbb{H}^n. \tag{2.2}$$

We now apply (2.1) and (2.2) to the pencil $F = \mathfrak{C}(\lambda) = \lambda_1 C_1 + \lambda_2 C_2 + \ldots + \lambda_m C_m$, where $C_i \in \mathbb{H}^n, \lambda \in \mathbb{R}^m$. In the situation of Hermitian matrices, we have a constructive proof for Theorem 2.1.1 that leads to a procedure for determining a maximum rank linear combination.

Fistly, we have the following lemma by direct computations.

Lemma 2.1.1. *Let $A = (a_{ij}) \in \mathbb{H}^n$ and P_{ik} be the $(1k)$-permutation matrix, i.e, that is obtained by interchaning the columns 1 and k of the identity matrix. The following hold true:*

(i) If $a_{11} = 0$ and $a_{kk} \neq 0$ (always real) for some $k = 1, 2, \ldots, n$, then

$$P_{1k}^* A P_{1k} = \begin{pmatrix} a_{kk} & \beta \\ \beta^* & B \end{pmatrix}, B^* = B.$$

(ii) Let $S = I_n + e_k e_t^*$, where e_k is the kth unit vector of $\mathbb{C}^n$. Then the (t,t)th entry of S^*AS is $\tilde{a} =: a_{kk} + a_{tt} + a_{kt} + a_{tk} \in \mathbb{R}$. Moreover,

$$P_{1t}^* S^* A S P_{1t} = \begin{pmatrix} \tilde{a} & \beta \\ \beta^* & B \end{pmatrix}, B^* = B.$$

As a consequence, if all diagonal entries of A are zero and a_{kt} has nonzero real part for some $1 \le k < t \le n$, then $\tilde{a} = a_{kt} + a_{tk} \neq 0$.

(iii) Let $T = I_n + ie_k e_t^*$, where $i^2 = -1$. Then the (t,t)th entry of T^*AT is $\tilde{a} =: a_{kk} + a_{tt} + i(a_{tk} - \bar{a}_{tk}) \in \mathbb{R}$. Moreover,

$$P_{1t}^* T^* A T P_{1t} = \begin{pmatrix} \tilde{a} & \beta \\ \beta^* & B \end{pmatrix}, B^* = B.$$

As a consequence, if all diagonal entries of A are zero and a_{kt} has nonzero image part for some $1 \le k < t \le n$, then $\tilde{a} = i(a_{tk} - \bar{a}_{tk})$.

Theorem 2.1.1. *Let $\mathfrak{C} = \mathfrak{C}(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ be a Hermitian pencil, i.e, $\mathfrak{C}(\lambda)^* = \mathfrak{C}(\lambda)$ for every $\lambda \in \mathbb{R}^m$. Then there exist polynomial matrices $\mathfrak{X}_+, \mathfrak{X}_- \in \mathbb{F}[\lambda]^{n\times n}$ and polynomials $b, d_j \in \mathbb{R}[\lambda], j = 1, 2, \dots, n$ (note that b, d_j are always real even when $\mathbb{F}$ is the complex field) such that*

$$\mathfrak{X}_+\mathfrak{X}_- = \mathfrak{X}_-\mathfrak{X}_+ = b^2 I_n \tag{2.3a}$$

$$b^4\mathfrak{C} = \mathfrak{X}_+ \boldsymbol{diag}(d_1, d_2, \dots, d_n)\mathfrak{X}_+^*, \tag{2.3b}$$

$$\mathfrak{X}_-\mathfrak{C}\mathfrak{X}_-^* = \boldsymbol{diag}(d_1, d_2, \dots, d_n). \tag{2.3c}$$

Proof. We apply Schmüdgen's procedure (2.1)-(2.2) step-by-step to $\mathfrak{C}_0 = \mathfrak{C}, \mathfrak{C}_1, \dots,$ where

$$\mathfrak{C}_{t-1} = \begin{pmatrix} \alpha_t & \beta_t \\ \beta_t^* & \hat{\mathfrak{C}}_t \end{pmatrix} = \mathfrak{C}_{t-1}^* \in \mathbb{H}^{n-t+1}, \mathfrak{C}_t = \alpha_t(\alpha_t\hat{\mathfrak{C}}_t - \beta^*\beta) \in \mathbb{H}^{n-t}, \alpha_t \in \mathbb{R}[\lambda],$$

for $t = 1, 2, \dots,$ until there exists a diagonal or zero matrix $\mathfrak{C}_k \in \mathbb{F}[\lambda]^{(n-k)\times(n-k)}$.

If the $(1,1)$st entry of $\mathfrak{C}_t$ is zero, by Lemma 2.1.1 we can find a nonsingular matrix $T \in \mathbb{F}^{n\times n}$ for that of $T^*\mathfrak{C}_t T$ being nonzero. Therefore, we can assume every matrix $\mathfrak{C}_t$ has a nonzero $(1,1)$st entry.

We now describe the process in more detail. At the first step, partition $\mathfrak{C}_0$ as

$$\mathfrak{C}_0 = \begin{pmatrix} \alpha_1 & \beta_1 \\ \beta_1^* & \hat{\mathfrak{C}}_1 \end{pmatrix}, \hat{\mathfrak{C}}_1^* = \hat{\mathfrak{C}}_1 \in \mathbb{F}[\lambda]^{(n-1)\times(n-1)}, 0 \neq \alpha_1 \in \mathbb{R}[\lambda].$$

Assign $\mathfrak{C}_1 = \alpha_1(\alpha_1\hat{\mathfrak{C}}_1 - \beta_1^*\beta_1) \in \mathbb{H}^{n-1}$ and

$$\mathfrak{X}_{1\pm} = X_{1\pm}(\lambda) = \begin{pmatrix} \alpha_1 & 0 \\ \pm\beta_1^* & \alpha_1 I_{n-1} \end{pmatrix}.$$

Then, by (2.2), we have

$$\begin{aligned} \mathfrak{X}_{1+}\mathfrak{X}_{1-} = \mathfrak{X}_{1-}\mathfrak{X}_{1+} = \alpha_1^2 I_n, \\ \mathfrak{X}_{1-}\mathfrak{C}\mathfrak{X}_{1-}^* = \begin{pmatrix} \alpha_1^3 & 0 \\ 0 & \mathfrak{C}_1 \end{pmatrix} := \tilde{\mathfrak{C}}_1, \alpha_1^4\mathfrak{C} = \mathfrak{X}_{1+}\tilde{\mathfrak{C}}_1\mathfrak{X}_{1+}^*. \end{aligned} \tag{2.4}$$

If $\mathfrak{C}_1$ is diagonal, stop. Otherwise, let's go to the second step by partitioning $\mathfrak{C}_1 = \begin{pmatrix} \alpha_2 & \beta_2 \\ \beta_2^* & \hat{\mathfrak{C}}_2 \end{pmatrix}$ and continue applying Schmüdgen's procedure (2.2) to $\mathfrak{C}_1$ in the second step

$$\mathfrak{Y}_{2\pm} = \begin{pmatrix} \alpha_2 & 0 \\ \pm\beta_2^* & \alpha_2 I_{n-2} \end{pmatrix}, \mathfrak{Y}_{2-}\mathfrak{C}_1\mathfrak{Y}_{2-}^* = \begin{pmatrix} \alpha_2^3 & 0 \\ 0 & \mathfrak{C}_2 \end{pmatrix}, \mathfrak{C}_2 = \alpha_2(\alpha_2\hat{\mathfrak{C}}_2 - \beta_2^*\beta_2) \in \mathbb{H}^{n-2}.$$

Accumulating

$$\mathfrak{X}_{2-} = \begin{pmatrix} \alpha_2 & 0 \\ 0 & \mathfrak{Y}_{2-} \end{pmatrix} \mathfrak{X}_{1-}, \mathfrak{X}_{2+} = \mathfrak{X}_{1+} \begin{pmatrix} \alpha_2 & 0 \\ 0 & \mathfrak{Y}_{2+} \end{pmatrix}$$

and

$$\mathfrak{X}_{2-}\mathfrak{C}\mathfrak{X}_{2-}^* = \begin{pmatrix} \alpha_1^3\alpha_2^3 & 0 & 0 \\ 0 & \alpha_2^3 & 0 \\ 0 & 0 & \mathfrak{C}_2 \end{pmatrix} = \begin{pmatrix} \alpha_2^3\mathtt{diag}(\alpha_1^3, \alpha_2) & 0 \\ 0 & \mathfrak{C}_2 \end{pmatrix} := \tilde{\mathfrak{C}}_2,$$

then $\mathfrak{X}_{2-}\mathfrak{X}_{2+} = \mathfrak{X}_{2+}\mathfrak{X}_{2-} = \alpha_1^2\alpha_2^2 I_n = b^2 I_n$. The second step completes.

Suppose now we have at the $(k-1)$th step that

$$\mathfrak{X}_{(k-1)-}\mathfrak{C}\mathfrak{X}_{(k-1)-}^* = \begin{pmatrix} \mathtt{diag}(d_1, d_2, \ldots, d_{k-1}) & 0 \\ 0 & \mathfrak{C}_{k-1} \end{pmatrix} := \tilde{\mathfrak{C}}_{k-1},$$

where $\mathfrak{C}_{k-1} = \mathfrak{C}_{k-1}^* \in \mathbb{F}[\lambda]^{(n-k+1)\times(n-k+1)}$, and $d_1, d_2, \ldots, d_{k-1}$ are all not identically zero. If $\mathfrak{C}_{k-1}$ is not diagonal (and suppose that its $(1,1)$st entry is nonzero), then partition $\mathfrak{C}_{k-1}$ and go to the kth step with the following updates:

$$\begin{aligned} \mathfrak{C}_{k-1} &= \begin{pmatrix} \alpha_k & \beta_k \\ \beta_k^* & \hat{\mathfrak{C}}_k \end{pmatrix}, \mathfrak{C}_k = \alpha_k(\alpha_k\hat{\mathfrak{C}}_k - \beta_k^*\beta_k), b = \prod_{t=1}^{k} \alpha_t, \\ \mathfrak{X}_{k+} &= \mathfrak{X}_{(k-1)+} \cdot \begin{pmatrix} \alpha_k I & 0 \\ 0 & \mathfrak{Y}_{k+} \end{pmatrix}, \mathfrak{X}_{k-} = \begin{pmatrix} \alpha_k I_{k-1} & 0 \\ 0 & \mathfrak{Y}_{k-} \end{pmatrix} . \mathfrak{X}_{(k-1)-}, \\ \mathfrak{X}_{k-}\mathfrak{C}\mathfrak{X}_{k-}^* &= \begin{pmatrix} \mathtt{diag}(d_1, d_2, \ldots, d_{k-1}, d_k) & 0 \\ 0 & \mathfrak{C}_k \end{pmatrix} := \tilde{\mathfrak{C}}_k, \end{aligned} \tag{2.5}$$

where $\mathfrak{Y}_{k\pm} = \begin{pmatrix} \alpha_k & 0 \\ \pm\beta_k^* & \alpha_k I_{n-k} \end{pmatrix}$ and

$$d_k = \alpha_k^3, d_j = \alpha_j^3 \prod_{t=j+1}^{k} \alpha_t^2, j = 1, 2, \ldots, k-1. \tag{2.6}$$

The procedure immediately stops if $\mathfrak{C}_k$ is diagonal, and $\mathfrak{X}_\pm$ in (2.3c) will be $\mathfrak{X}_{k\pm}$. □

The proof of Theorem 2.1.1 gives a comprehensive update according to Schmügen's procedure. However, we only need the diagonal elements of $\tilde{\mathfrak{C}}_k$ to determine the maximum rank of $\mathfrak{C}(\lambda)$ at the end. The following theorem allows us to determine such a maximum rank linear combination.

Theorem 2.1.2. *Use notation as in Theorem 2.1.1, and suppose $\mathfrak{C}_k$ in (2.5) is diagonal but every $\mathfrak{C}_t$, $t = 0, 1, 2, \ldots, k-1$, is not so. Consider the modification of (2.5) as*

$$\begin{aligned}
\mathfrak{C}_{k-1} &= \begin{pmatrix} \alpha_k & \beta_k \\ \beta_k^* & \hat{\mathfrak{C}}_k \end{pmatrix}, & \mathfrak{C}_k &= \alpha_k(\alpha_k \hat{\mathfrak{C}}_k - \beta_k^* \beta_k), \\
\mathfrak{X}_{k-} &= \begin{pmatrix} I_{k-1} & 0 \\ 0 & \mathfrak{Y}_{k-} \end{pmatrix} .\mathfrak{X}_{(k-1)-}, & \mathfrak{Y}_{k\pm} &= \begin{pmatrix} \alpha_k & 0 \\ \pm\beta_k^* & \alpha_k I_{n-k} \end{pmatrix}, \\
\mathfrak{X}_{k-}\mathfrak{C}\mathfrak{X}_{k-}^* &= \begin{pmatrix} \boldsymbol{diag}(\alpha_1^3, \alpha_2^3, \ldots, \alpha_{k-1}^3, \alpha_k^3) & 0 \\ 0 & \mathfrak{C}_k \end{pmatrix} := \tilde{\mathfrak{C}}_k,
\end{aligned} \tag{2.7}$$

Moreover, let $d_i = \alpha_i^3, i = 1, 2, \ldots, k,$ and $\mathfrak{C}_k = \boldsymbol{diag}(d_{k+1}, d_{k+2}, \ldots, d_n), d_j \in \mathbb{R}[\lambda], j = 1, 2, \ldots, n,$ and some of $d_{k+1}, d_{k+2}, \ldots, d_n$ may be identically zero. The following hold true.

(i) α_t divides α_{t+1} (and therefore d_t divides d_{t+1}) for every $t \le k-1$, and if $k < n$, then α_k divides every $d_j, j > k$.

(ii) The pencil $\mathfrak{C}(\lambda)$ has the maximum rank r if and only if there exists a permutation such that $\tilde{\mathfrak{C}}(\lambda) = \boldsymbol{diag}(d_1, d_2, \ldots, d_r, 0, \ldots, 0)$, d_j is not identically zero for every $j = 1, 2, \ldots, r$. In addition, the maximum rank of $\mathfrak{C}(\lambda)$ achieves at $\hat{\lambda}$ if and only if $\alpha_k(\hat{\lambda}) \neq 0$ or $(\prod_{t=k+1}^{r} d_t(\hat{\lambda})) \neq 0$, respectively, depends upon $\mathfrak{C}_k$ being identically zero or not.

Proof.

(i) The construction of $\mathfrak{C}_1, \ldots, \mathfrak{C}_k$ imply that α_t divides $\alpha_{t+1}, t = 1, 2, \ldots, k-1$. In particular, α_k is divisible by $\alpha_t, \forall t = 1, 2, \ldots, k-1$. Moreover, if $k < n$, then α_k divides $d_j, \forall j = k+1, \ldots, n,$ (since $\mathfrak{C}_k = \alpha_k(\alpha_k \hat{\mathfrak{C}}_k - \beta_k^* \beta_k) = \mathtt{diag}(d_{k+1}, d_{k+2}, \ldots, d_n)$), provided by the formula of $\mathfrak{C}_k$ in (2.7).

(ii) We first note that after an appropriate number of permutations, $\tilde{\mathfrak{C}}_k$ must be of the form $\tilde{\mathfrak{C}}_k = \mathtt{diag}(d_1, d_2, \ldots, d_k, \ldots, d_r, 0, \ldots, 0)$, with $d_1, d_2, \ldots, d_r$ not identically zero. Moreover, $k \le r$, in which the equality occurs if and only if $\mathfrak{C}_k$ is zero because $\mathfrak{C}_t$ is determined only when $\alpha_t = \mathfrak{C}_{t-1}(1,1) \neq 0$.

Finally, since $d_k, \ldots, d_r$ are real polynomials, one can pick a $\hat{\lambda} \in \mathbb{R}^m$ such that $\prod_{t=k}^{r} d_t(\hat{\lambda}) \neq 0$. By i), $d_i(\hat{\lambda}) \neq 0$ for all $i = 1, \ldots, r$, and hence rank$\mathfrak{C}(\hat{\lambda}) = r$ is the maximum rank of the pencil $\mathfrak{C}(\lambda)$.

□

The updates of $\mathfrak{X}_{k-}$ and d_j as in (2.7) are really more simple than that in (2.3c). Therefore, we use (2.7) to propose the following algorithm.

Algorithm 2 Schmüdgen-like algorithm determining maximum rank of a pencil.

INPUT: Hermitian matrices $C_1, \ldots, C_m \in \mathbb{H}^n$.

OUTPUT: A real m-tuple $\hat{\lambda} \in \mathbb{R}^m$ that maximizes the rank of the pencil $\mathfrak{C} =: \mathfrak{C}(\lambda)$.

1: Set up $\mathfrak{C}_0 = \mathfrak{C}$ and α_1, $\tilde{\mathfrak{C}}_1$ (containing $\mathfrak{C}_1$), $\mathfrak{X}_{1\pm}$ as in (2.7).

2: $k \leftarrow 1$.

3: **While** $\mathfrak{C}_k$ is not diagonal **do**

4: $k \leftarrow k + 1$.

5: Do the computations as in (2.7) to obtain $\alpha_k, \mathfrak{X}_{k-}, \tilde{\mathfrak{C}}_k$ containing $\mathfrak{C}_k$.

6: **Endwhile**

7: Pick a $\hat{\lambda} \in \mathbb{R}^m$ that satisfies Theorem 2.1.2 (ii).

Let us consider the following example to see how the algorithm works.

Example 2.1.1. Given singular matrices: $C_1 = \begin{pmatrix} -1 & -2-2i & 0 \\ -2+2i & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix}$;

$$C_2 = \begin{pmatrix} 1 & i & -i \\ -i & 1 & -1 \\ i & -1 & 2 \end{pmatrix}; C_3 = \begin{pmatrix} 1 & 1+i & 2 \\ 1-i & 2 & 2(1-i) \\ 2 & 2(1+i) & 4 \end{pmatrix}.$$

$$\begin{aligned} \mathfrak{C} &= xC_1 + yC_2 + zC_3 \\ &= \begin{pmatrix} -x+y+z & -2x+z+(-2x+y+z)i & 2z-yi \\ -2x+z-(-2x+y+z)i & -3x+y+2z & -y+2z-2zi \\ 2z+yi & -y+2z+2zi & 2y+4z \end{pmatrix} \end{aligned}$$

and

$$\alpha_1 = -x+y+z; \qquad \beta_1 = (-2x+z+(-2x+y+z)i; 2z-yi);$$
$$\hat{\mathfrak{C}}_1 = \begin{pmatrix} -3x+y+2z & -y+2z-2zi \\ -y+2z+2zi & 2y+4z \end{pmatrix};$$
$$\mathfrak{C}_1 = \alpha_1(\alpha_1.\hat{\mathfrak{C}}_1 - \beta_1^*\beta_1)$$
$$= \alpha_1 \begin{pmatrix} -5x^2+yz+3xz & -xy+2xz+2yz+i(-2xy+yz-2xz) \\ -xy+2xz+2yz-i(-2xy+yz-2xz) & y^2-2xy-4xz+6yz \end{pmatrix}$$

We have

$$X_{1\pm} := Y_{1\pm} = \begin{pmatrix} \alpha_1 & 0 \\ \pm\beta_1^* & \alpha_1 I_2 \end{pmatrix}$$

and

$$X_{1-}.\mathfrak{C}.X_{1-}^* = \begin{pmatrix} \alpha_1^3 & 0 \\ 0 & \mathfrak{C}_1 \end{pmatrix},$$

$$\mathfrak{C}_1 = \begin{pmatrix} \alpha_2 & \beta_2 \\ \beta_2^* & \hat{\mathfrak{C}}_2 \end{pmatrix};$$
$$\mathfrak{C}_2 = \alpha_2(\alpha_2.\hat{C}_2 - \beta_2^*.\beta_2) := \gamma$$

where

$$\alpha_2 = \alpha_1(-5x^2+yz+3xz); \quad \beta_2 = \alpha_1(-xy+2xz+2yz+i(-2xy+yz-2xz));$$
$$\hat{\mathfrak{C}}_2 = \alpha_1(y^2-2xy-4xz+6yz);$$
$$\gamma = \alpha_1.\alpha_2^2(y^2-2xy-4xz+6yz)$$
$$-\alpha_1^2.\alpha_2(5x^2y^2+8x^2z^2+5y^2z^2+4x^2yz-8xy^2z+4xyz^2).$$

$$Y_{2-} = \begin{pmatrix} \alpha_2 & 0 \\ -\beta_2^* & \alpha_2 \end{pmatrix}; X_{2-} = \begin{pmatrix} 1 & 0 \\ 0 & Y_{2-} \end{pmatrix}.X_{1-},$$

$$X_{2-}.\mathfrak{C}.X_{2-}^* = \begin{pmatrix} \mathtt{diag}(\alpha_1^3, \alpha_2^3) & 0 \\ 0 & \gamma \end{pmatrix}.$$

We now choose $\alpha_1, \alpha_2, \gamma$ such that the matrix $X_{2-}.\mathfrak{C}.X_{2-}^*$ is nonsingular, for example $\alpha_1 = 1; \alpha_2 = -1$ and $\gamma = 19$, corresponding to $(x,y,z) = (1,1,1)$. Then

$$\mathfrak{C} = C_1 + C_2 + C_3 = \begin{pmatrix} 1 & -1 & 2-i \\ -1 & 0 & 1-2i \\ 2+i & 1+2i & 6 \end{pmatrix} \text{ with } \det\mathfrak{C} = -19.$$

Now, we revisit a link between the Hermitian-SDC and SDS problems: A finite collection of Hermitian matrices is $*$-SDC if and only if an appropriate collection of same size matrices is SDS.

First, we present the necessary and sufficient conditions for simultaneous diagonalization via congruence of commuting Hermite matrices. This result is given, e.g., in [34, Theorem 4.1.6] and [7, Corollary 2.5]. To show how Algorithm 3 performs and finds a nonsingular matrix simultaneously diagonalizing commuting matrices, we give a constructive proof using only a matrix computation technique. The idea of the proof follows from that of [37, Theorem 9] for real symmetric matrices.

Theorem 2.1.3. *The matrices $I, C_1, \ldots, C_m \in \mathbb{H}^n, m \geq 1$ are $*$-SDC if and only if they are commuting. Moreover, when this the case, there are $*$-SDC by a unitary matrix (resp., orthogonal one) if $C_1, C_2, \ldots, C_m$ are complex (resp., all real).*

Proof. If $I, C_1, \ldots, C_m \in \mathbb{H}^n, m \geq 1$ are $*$-SDC, then there exists a nonsingular matrix $U \in \mathbb{C}^{n\times n}$ such that $U^*IU, U^*C_1U, \ldots, U^*C_mU$ are diagonal. Note that,

$$U^*IU = \texttt{diag}(d_1, d_2, \ldots, d_n) \succ 0 \tag{2.8}$$

Let $D = \texttt{diag}(\frac{1}{\sqrt{d_1}}, \ldots, \frac{1}{\sqrt{d_m}})$ and $V = UD$. Then V must be unitary and $V^*C_iV = DU^*C_iUD$ is diagonal for every $i = 1, 2, \ldots, m$.

Thus $V^*C_iV.V^*C_jV = V^*C_jV.V^*C_iV$, $\forall i \neq j$, and hence $C_iC_j = C_jC_i$, $\forall i \neq j$. Moreover, each V^*C_iV is real since it is Hermitian.

On the contrary, we prove by induction on m.

In the case $m = 1$, the proposition is true since any Hermitian matrix can be diagonalized by a unitary matrix.

For $m \geq 2$, we suppose the proposition holds true for $m - 1$ matrices.

Now, we consider an arbitrary collection of Hermitian matrices $I, C_1, \ldots, C_m$. Let P be a unitary matrix that diagonalizes C_1 :

$$P^*P = I, \quad P^*C_1P = \texttt{diag}(\alpha_1 I_{n_1}, \ldots, \alpha_k I_{n_k}),$$

where α_i's are distinct and real eigenvalues of C_1. Since C_1 and C_i commute for all $i = 2, \ldots, m$, so do P^*C_1P and P^*C_iP. By Lemma 1.1.2, we have

$$P^*C_iP = \texttt{diag}(C_{i1}, C_{i2}, \ldots, C_{ik}), \quad i = 2, 3, \ldots, m,$$

where each C_{it} is Hermitian of size n_t.

Now, for each $t = 1, 2, \ldots, k$, since $C_{it}C_{jt} = C_{jt}C_{it}$, $\forall i, j = 2, 3, \ldots, m$, (by $C_iC_j = C_jC_i$,) the induction hypothesis leads to the fact that

$$I_{n_t}, C_{2t}, \ldots, C_{mt} \tag{2.9}$$

are $*$-SDC by a unitary matrix Q_t. Determine $U = P\ \mathtt{diag}(Q_1, Q_2, \ldots, Q_k)$. Then

$$\begin{aligned} U^*C_1U &= \mathtt{diag}(\alpha_1 I_{n1}, \ldots, \alpha_k I_{nk}), \\ U^*C_iU &= \mathtt{diag}(Q_1^*C_{i1}Q_1, \ldots, Q_k^*C_{ik}Q_k), i = 2, 3, \ldots, m, \end{aligned} \tag{2.10}$$

are all diagonal.

□

In the above proof, the fewer multiple eigenvalues the starting matrix C_1 has, the fewer number of collection as in (2.9) need to be solved. Algorithm 3 below takes this observation into account at the first step. To this end, the algorithm computes the eigenvalue decomposition of all matrices $C_1, C_2, \ldots, C_m$ for finding a matrix with the minimum number of multiple eigenvalues.

Algorithm 3 Solving the $*$-SDC problem of commuting Hermitian matrices

INPUT: Commuting matrices $C_1, C_2, \ldots, C_m$.

OUTPUT: Unitary matrix U making $U^*C_1U, \ldots, U^*C_mU$ be all diagonal.

1: Pick a matrix with the minimum number of multiple eigenvalues, say, C_1.

2: Find an eigenvalue decomposition of $C_1 : C_1 = P^*\mathtt{diag}(\alpha_1 I_{n_1}, \ldots, \alpha_k I_{n_k}), n_1 + n_2 + \ldots + n_k = n$, $\alpha_1, \ldots, \alpha_k$ are distinct real eigenvalues, and $P^*P = I$.

3: Compute the diagonal blocks of $P^*C_iP, i \geq 2 :$

$$P^*C_iP = \mathtt{diag}(C_{i1}, C_{i2}, \ldots, C_{ik}), C_{it} \in \mathbb{H}^{n_t}, \forall t = 1, 2, \ldots, k.$$

where $C_{2t}, \ldots, C_{mt}$ pairwise commute for every $t = 1, 2, \ldots, k$.

4: For each $t = 1, 2, \ldots, k$ simultaneously diagonalize the collection of matrices $I_{n_t}, C_{2t}, \ldots, C_{mt}$ by a unitary matrix Q_t.

5: Define $U = P\mathtt{diag}(Q_1, \ldots, Q_k)$.

In the example below, we see that when C_1 has no multiple eigenvalue, the algorithm 3 immediately gives the congruence matrix in one step.

Example 2.1.2. Let

$$C_1 = \begin{pmatrix} 1 & 1+i \\ 1-i & 2 \end{pmatrix}; C_2 = \begin{pmatrix} 2 & 3+3i \\ 3-3i & 5 \end{pmatrix}; C_3 = \begin{pmatrix} -1 & -2-2i \\ -2+2i & -3 \end{pmatrix}$$

be commuting matrices and C_1 has two distinct eigenvalues, then we immediately find a unitary matrix $P = \begin{pmatrix} \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3}(1-i) & \frac{\sqrt{6}}{6}(i-1) \end{pmatrix}$ such that $P^*C_1P = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$, $P^*C_2P = \begin{pmatrix} 8 & 0 \\ 0 & -\frac{\sqrt{6}+3}{9} \end{pmatrix}$; $P^*C_3P = \begin{pmatrix} -15 & 0 \\ 0 & -\frac{5}{3} \end{pmatrix}$ are all diagonals.

Using Theorem 2.1.3, we describe comprehensively the SDC property of a collection of Hermitian matrices in Theorem 2.1.4 below. Its results are combined from [7] and references therein, but we restate and give a constructive proof leading to Algorithm 4. It is worth mentioning that in Theorem 2.1.4 below, $\mathfrak{C}(\lambda)$ is a Hermitian pencil, i.e., the parameter λ appearing in the theorem is always real if $\mathbb{F}$ is the field of real or complex numbers.

Theorem 2.1.4. *Let $0 \neq C_1, C_2, \ldots, C_m \in \mathbb{H}^n$ with $\dim_{\mathbb{C}}(\bigcap_{t=1}^m kerC_t) = q$, (always $q < n$.)*

1. *If $q = 0$, then the following hold:*

   (i) If $det\mathfrak{C}(\lambda) = 0$, for all $\lambda \in \mathbb{R}^m$ (over only real m-tuple λ), then $C_1, \ldots, C_m$ are not $*$-SDC.

   (ii) Otherwise, there exists $\lambda \in \mathbb{R}^m$ such that $\mathfrak{C}(\lambda)$ is nonsingular. The matrices $C_1, \ldots, C_m$ are $*$-SDC if and only if $\mathfrak{C}(\lambda)^{-1}C_1, \ldots, \mathfrak{C}(\lambda)^{-1}C_m$ pairwise commute and every $\mathfrak{C}(\lambda)^{-1}C_i, i = 1, 2, \ldots, m$, is similar to a real diagonal matrix.

2. *If $q > 0$, then there exists a nonsingular matrix V such that*

$$V^*C_iV = \textbf{diag}(\hat{C}_i, 0_q), \forall i = 1, 2, \ldots, m, \tag{2.11}$$

   where 0_q is the $q \times q$ zero matrix and $\hat{C}_i \in \mathbb{H}^{n-q}$ with $\bigcap_{t=1}^m ker\hat{C}_t = 0$. Moreover, $C_1, \ldots, C_m$ are $*$-SDC if and only if $\hat{C}_1, \hat{C}_2, \ldots, \hat{C}_m$ are $*$-SDC.

Proof.

1. Suppose $q = 0$,

   (i) If $\det\mathfrak{C}(\lambda) = 0$, for all $\lambda \in \mathbb{R}^m$ (over only real m-tuple λ), we prove that $C_1, \dots, C_m$ are not $*$-SDC. Assume the opposite, $C_1, \dots, C_m$ were $*$-SDC by a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ and then

$$C_i = P^* D_i P, D_i = \mathtt{diag}(\alpha_{i1}, \quad \alpha_{i2}, \dots, \alpha_{in})$$

   where D_i is real matrix, forall $i = 1, 2, \dots, m$. Moreover,

$$\mathfrak{C}(\lambda) = \sum_{i=1}^{m} \lambda_i C_i = \sum_{i=1}^{m} \lambda_i P^* D_i P = P^* (\sum_{i=1}^{m} \lambda_i D_i) P.$$

   The real polynomial (with real variable λ)

$$\mathrm{det}\mathfrak{C}(\lambda) = (\mathrm{detP})^2 . \Pi_{\mathrm{j=1}}^{\mathrm{n}} (\sum_{\mathrm{i=1}}^{\mathrm{m}} \alpha_{\mathrm{ij}} \lambda_{\mathrm{i}}); \lambda_{\mathrm{i}} \in \mathbb{R}, \mathrm{i} = 1, 2, \dots, \mathrm{m},$$

   is hence identically zero because of the hypothesis. But $\mathbb{R}[\lambda_1, \lambda_2, \dots, \lambda_m]$ is an integer domain, and there must exist an identically zero factor, say, there exists $j \in \{1, 2, \dots, n\}$ such that $(\alpha_{1j}, \alpha_{2j}, \dots, \alpha_{mj}) = 0$.

   Picking the vector $0 \neq x$ with $Px = e_j$, where e_j is the jth unit vector in $\mathbb{C}^n$, one obtains

$$C_i x = P^* D_i P x = P^* D_i e_j = 0, \forall i = 1, 2, \dots, m.$$

   It implies that $0 \neq x \in \bigcap_{t=1}^{m} kerC_t$, contradicting the hypothesis. Part (i) is thus proved.

   (ii) Otherwise, there exists $\lambda \in \mathbb{R}^m$ such that $\mathfrak{C}(\lambda)$ is nonsingular.

   Firstly, suppose $C_1, \dots, C_m$ are $*$-SDC by a nonsingular matrix $P \in \mathbb{C}^{n\times n}$, then $P^* C_i P$ are all real diagonal. As a consequence,

$$P^{-1}\mathfrak{C}(\lambda)^{-1} C_i P = [P^* \mathfrak{C}(\lambda) P]^{-1} (P^* C_i P)$$

   is real diagonal for every $i = 1, 2, \dots, m$. This yieds the pairwise commutativity of $P^{-1}\mathfrak{C}(\lambda)^{-1}C_1P, P^{-1}\mathfrak{C}(\lambda)^{-1}C_2P, \dots, P^{-1}\mathfrak{C}(\lambda)^{-1}C_mP$ and hence that of $\mathfrak{C}(\lambda)^{-1}C_1, \mathfrak{C}(\lambda)^{-1}C_2, \dots, \mathfrak{C}(\lambda)^{-1}C_m$.

   Conversely, suppose $\mathfrak{C}(\lambda)^{-1}C_1, \mathfrak{C}(\lambda)^{-1}C_2, \dots, \mathfrak{C}(\lambda)^{-1}C_m$ pairwise commute and every $\mathfrak{C}(\lambda)^{-1}C_i, i = 1, 2, \dots, m$, is similar to a real diagonal matrix. Then there exists a nonsingular $Q \in \mathbb{C}^{n\times n}$ such that $Q^{-1}\mathfrak{C}(\lambda)^{-1}C_iQ = M_i$ are all real diagonal.

We have $Q^*\mathfrak{C}(\lambda)Q.M_i = Q^*C_iQ, i = 1, 2, \ldots, m$. Since C_i is Hermitian, so is Q^*C_iQ. Then

$$Q^*\mathfrak{C}(\lambda)Q.M_i = Q^*C_iQ = (Q^*C_iQ)^* = (Q^*\mathfrak{C}(\lambda)Q.M_i)^* = M_i.Q^*\mathfrak{C}(\lambda)Q.$$

Therefore, we have

$$\begin{aligned}
Q^*C_iQ.Q^*C_jQ =& Q^*\mathfrak{C}(\lambda)Q.M_i.Q^*\mathfrak{C}(\lambda)Q.M_j \\
=& Q^*\mathfrak{C}(\lambda)Q.M_i.M_j.Q^*\mathfrak{C}(\lambda)Q \\
=& Q^*\mathfrak{C}(\lambda).M_j.M_i.Q^*\mathfrak{C}(\lambda)Q \\
=& Q^*\mathfrak{C}(\lambda).M_j.Q^*\mathfrak{C}(\lambda)Q.M_i \\
=& Q^*C_jQ.Q^*C_iQ
\end{aligned}$$

or $Q^*C_1Q, Q^*C_2Q, \ldots, Q^*C_mQ$ pairwise commute. By the Theorem 2.1.3, $I, Q^*C_1Q, Q^*C_2Q, \ldots, Q^*C_mQ$ are $*$-SDC. Implying $C_1, C_2, \ldots, C_m$ are $*$-SDC.

2. Suppose $q > 0$, let $C \in \mathbb{C}^{mn\times n}$ be the matrix containing $C_1, C_2, \ldots, C_m$, and $C = UDV^*$ be a singular value decomposition. Since $\text{rank}C = n - q$, the last q columns of V are an orthonormal basis of $\text{Ker}C = \bigcap_{i=1}^{m} \text{Ker}C_i$. One then can check that V^*C_iV has the form (2.11) for every $i = 1, 2, \ldots, m$.

   Moreover, by Lemma 1.1.6, $C_1, \ldots, C_m$ are $*$-SDC if and only if $\hat{C}_1, \hat{C}_2, \ldots, \hat{C}_m$ are $*$-SDC.

□

The following algorithm checks that the Hermitian matrices $C_1, C_2, \ldots, C_m$ are $*$-SDC or not.

Algorithm 4 The SDC of Hermitian matrices in a link with SDS.

INPUT: Matrices $C_1, C_2, \ldots, C_m \in \mathbb{H}^n$

OUTPUT: Conclude whether $C_1, C_2, \ldots, C_m$ are $*$-SDC or not.

1: Compute a singular value decomposition $C = U\sum V^*$, of $C = (C_1^*, C_2^*, \ldots, C_m^*)^*$, $\sum = \mathtt{diag}(\sigma_1, \ldots, \sigma_{n-q}, 0, \ldots, 0)$, $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{n-q} > 0, 0 \leq q \leq n-1$. Then $\mathrm{dim}_{\mathbb{F}}(\cap_{\mathrm{t=1}}^{\mathrm{m}}\mathrm{kerC_t}) = \mathrm{q}$.

2: If $q = 0$:

Step 1: If $\det\mathfrak{C}(\lambda) = 0$, for all $\lambda \in \mathbb{R}^m$, then $C_1, C_2, \ldots, C_m$ are not $*$-SDC. Else, go to Step 2.

Step 2: Find a $\underline{\lambda} \in \mathbb{R}^m$ such that $\mathfrak{C} := \mathfrak{C}(\underline{\lambda})$ is nonsingular.
(a) If there exists $i \in \{1, 2, \ldots, m\}$ such that $\mathfrak{C}^{-1}C_i$ is not similar to a diagonally real matrix, then conclude the given matrices are not $*$-SDC. Else, go to (b).
(b) If $\mathfrak{C}^{-1}C_1, \ldots, \mathfrak{C}^{-1}C_m$ are not commuting, which is equivalent to that $C_i\mathfrak{C}^{-1}C_j$ is not Hermitian for some $i \neq j$, then conclude the given matrices are not $*$-SDC. Else, they are $*$-SDC.

3: Else $(q > 0)$:

Step 3: For the singular value decomposition $C = U\sum V^*$ determined at the beginning, the matrix V satisfies (2.11.) Pick the matrices $\hat{C}_i$ being the $(n-q) \times (n-q)$ top-left submatrix of C_i.

Step 4: Go to Step 1 with the resulting matrices $\hat{C}_1, \ldots, \hat{C}_m \in \mathbb{H}^{n-q}$.

In Algorithm 4, Step 1 checks whether the maximum rank of the pencil $\mathfrak{C}(\lambda)$ is strictly less than its size or not. This is because of the following equivalence:

$$\det\mathfrak{C}(\lambda) = 0, \forall\lambda \in \mathbb{R}^n \setminus \{0\} \Longleftrightarrow \max\{\mathrm{rank}\mathfrak{C}(\lambda) \mid \lambda \in \mathbb{R}^m\} < n.$$

The terminology "maximum rank linear combination" is due to this equivalence and Lemma 1.1.4.

We now consider some examples in which all given matrices are singular. We apply Theorem 2.1.2 and Theorem 2.1.4 to solve the Hermitian SDC problem.

Example 2.1.3. Given three matrices as in Example 2.1.1, we use Algorithm 4 to check whether the matrices are $*$-SDC.

Observe that $\mathfrak{C} = C_1 + C_2 + C_3 = \begin{pmatrix} 1 & -1 & 2-i \\ -1 & 0 & 1-2i \\ 2+i & 1+2i & 6 \end{pmatrix}$ is nonsingular and $\text{rank}(C_1^*, C_2^*, C_3^*)^* = 3$. So $\dim(\bigcap_{i=1}^3 \text{ker} C_i) = 0$.

$$\mathfrak{C}^{-1} = \frac{1}{19}\begin{pmatrix} 5 & -10-3i & 1-2i \\ -10+3i & -1 & 3-3i \\ 1+2i & 3+3i & 1 \end{pmatrix}$$

and

$$A := \mathfrak{C}^{-1}C_1 = \frac{1}{19}\begin{pmatrix} 21-14i & 20-i & 0 \\ 12-5i & 29+14i & 0 \\ -13-2i & -7-15i & 0 \end{pmatrix};$$

$$B := \mathfrak{C}^{-1}C_2 = \frac{1}{19}\begin{pmatrix} 4+11i & -11+4i & 12-6i \\ -7+7i & -7-7i & 10+4i \\ 4 & 4i & 1-4i \end{pmatrix}.$$

It is easy to check that $AB \neq BA$. Therefore, by Theorem 2.1.4 (case 1(ii)), C_1, C_2, C_3 are not $*$-SDC.

Example 2.1.4. The matrices

$$C_1 = \begin{pmatrix} 1 & 3 & -1 \\ 3 & 6 & 0 \\ -1 & 0 & -2 \end{pmatrix},\ C_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & 2 \\ 0 & 2 & -1 \end{pmatrix},\ C_3 = \begin{pmatrix} -1 & -3 & 2 \\ -3 & -5 & 4 \\ 2 & 4 & -3 \end{pmatrix}$$

are all singular since $\text{rank}(C_1) = \text{rank}(C_2) = \text{rank}(C_3) = 2$. We furthermore have $\dim(\bigcap_{i=1}^3 \ker C_i) = 0$ since $\text{rank}(C_1\ \ C_2\ \ C_3)^T = 3$. We will prove these matrices are not SDC by applying Theorem 2.1.4 (case 1 (ii)) as follows. Consider the linear combination

$$\mathfrak{C} = xC_1 + yC_2 + zC_3 = \begin{pmatrix} x-z & 3x-3z & 2z-x \\ 3x-3z & 6x-3y-5z & 2y+4z \\ 2z-x & 2y+4z & -2x-y-3z \end{pmatrix}$$

Applying Scmüdgen's procedure we have

$$\tilde{\mathfrak{C}}_1 = \mathfrak{X}_{1-}\mathfrak{C}\mathfrak{X}_{1-}^* = \begin{pmatrix} (x-z)^3 & 0 \\ 0 & \mathfrak{C}_1 \end{pmatrix},\ \mathfrak{X}_{1-} = \begin{pmatrix} x-z & 0 & 0 \\ 3z-3x & x-z & 0 \\ x-2z & 0 & x-z \end{pmatrix}$$

where

$$\mathfrak{C}_1 = (x-z)\begin{pmatrix} -(x-z)(3x+3y-4z) & (3x+2y-2z)(x-z) \\ (3x+2y-2z)(x-z) & -(x-2z)^2-(x-z)(2x+y+3z) \end{pmatrix}$$

$$=: \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}.$$

Determine

$$\mathfrak{X}_{2-} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & -\beta & \alpha \end{pmatrix} \mathfrak{X}_{1-}.$$

We then have

$$\mathfrak{X}_{2-}\mathfrak{C}\mathfrak{X}_{2-}^* = \begin{pmatrix} (x-z)^3 & 0 & 0 \\ 0 & \alpha^3 & 0 \\ 0 & 0 & \alpha(\alpha\gamma-\beta^2) \end{pmatrix}.$$

Notice that none of the diagonal elements $(x-z)^3$, α and $\alpha(\alpha\gamma-\beta^2)$ in the latest matrix are identically zero. By Theorem 2.1.2, we pick (x, y, z) such that all these elements do not vanish. For example, $(x, y, z) = (2, 0, 3)$ yields $\alpha = 6$, $\beta = 0$, $\gamma = 3$, and $\alpha(\alpha\gamma - \beta) = 108 \neq 0$. Then

$$\mathfrak{X}_- = \begin{pmatrix} -1 & 0 & 0 \\ 3 & -1 & 0 \\ -4 & 0 & -1 \end{pmatrix}, \quad \mathfrak{C} = 2C_1 + 3C_3 = \begin{pmatrix} -1 & -3 & 4 \\ -3 & -3 & 12 \\ 4 & 12 & -13 \end{pmatrix}, \quad \operatorname{rank}\mathfrak{C} = 3.$$

In this case, $(\mathfrak{C}^{-1}C_1)(\mathfrak{C}^{-1}C_2) \neq (\mathfrak{C}^{-1}C_2)(\mathfrak{C}^{-1}C_1)$ although every one of $\mathfrak{C}^{-1}C_1$, $\mathfrak{C}^{-1}C_2$, $\mathfrak{C}^{-1}C_3$ is similar to a real diagonal matrix.

Example 2.1.5. The matrices

$$C_1 = \begin{pmatrix} -1 & -4 & 4 \\ -4 & -16 & 16 \\ 4 & 16 & -16 \end{pmatrix}, \quad C_2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 2 \\ 0 & 2 & -4 \end{pmatrix}, \quad C_3 = \begin{pmatrix} -1 & -3 & 2 \\ -3 & -9 & 6 \\ 2 & 6 & -4 \end{pmatrix}$$

are all singular and $\dim(\ker C_1 \bigcap \ker C_2 \bigcap \ker C_3) = 1$. This intersection is spanned by, e.g., $x = (-4, 2, 1)$. Consider the linear combination

$$\mathfrak{C} = xC_1 + yC_2 + zC_3 = \begin{pmatrix} -x-z & -4x-3z & 4x+2z \\ -4x-3z & -16x-y-9z & 16x+2y+6z \\ 4x+2z & 16x+2y+6z & -16x-4y-4z \end{pmatrix}.$$

Applying Schmüdgen's procedure, we have

$$\mathfrak{X}_{1-}\mathfrak{C}\mathfrak{X}_{1-}^* = \begin{pmatrix} (-x-z)^3 & 0 \\ 0 & \mathfrak{C}_1 \end{pmatrix}, \quad \mathfrak{X}_{1-} = \begin{pmatrix} -x-z & 0 & 0 \\ -4x-3z & -x-z & 0 \\ 4x+2z & 0 & -x-z \end{pmatrix},$$

where

$$\mathfrak{C}_1 = (-x-z)\begin{pmatrix} xy+yz+zx & -2(xy+yz+zx) \\ -2(xy+yz+zx) & 4(xy+yz+zx) \end{pmatrix} =: \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}.$$

Let

$$\mathfrak{X}_{2-} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & -\beta & \alpha \end{pmatrix} \mathfrak{X}_{1-}.$$

We then have

$$\mathfrak{X}_{2-}\mathfrak{C}\mathfrak{X}_{2-}^* = \begin{pmatrix} (-x-z)^3 & 0 & 0 \\ 0 & \alpha^3 & 0 \\ 0 & 0 & \alpha(\alpha\gamma-\beta^2) \end{pmatrix},$$

where

$$\begin{aligned} \alpha &= (-x-z)(xy+yz+zx), \\ \beta &= 2(-x-z)(xy+yz+zx) = -2\alpha, \\ \gamma &= 4(-x-z)(xy+yz+zx) = 4\alpha. \end{aligned}$$

It is easy to check that $\alpha\gamma - \beta^2 = 0$ for all x, y, z. The procedure stops. We have $r = \text{rank}\mathfrak{C}(\lambda) = 2$. Since $\bigcap_{i=1}^3 \ker C_i = \{(-4a, 2a, a) \mid a \in \mathbb{R}\}$, $\dim(\bigcap_{i=1}^3 \ker C_i) = 1$.

We now apply Theorem 2.1.4 (case 2) to prove that these matrices are not $*$-SDC. Picking

$$Q = \begin{pmatrix} 1 & 0 & -4 \\ 0 & 1 & 2 \\ 4 & -2 & 1 \end{pmatrix},$$

then

$$Q^*C_1Q = \begin{pmatrix} -225 & 180 & 0 \\ 180 & -144 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \hat{C}_1 = \begin{pmatrix} -225 & 180 \\ 180 & -144 \end{pmatrix},$$

$$Q^*C_2Q = \begin{pmatrix} -64 & 40 & 0 \\ 40 & -25 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \hat{C}_2 = \begin{pmatrix} -64 & 40 \\ 40 & -25 \end{pmatrix},$$

$$Q^*C_3Q = \begin{pmatrix} -49 & 49 & 0 \\ 49 & -49 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \hat{C}_3 = \begin{pmatrix} -49 & 49 \\ 49 & -49 \end{pmatrix}.$$

We can check that $\det \hat{\mathfrak{C}} = -441 \neq 0$ with $\hat{\mathfrak{C}} = -\hat{C}_1 + \hat{C}_2$, and furthermore that $\hat{\mathfrak{C}}^{-1}\hat{C}_1$ and $\hat{\mathfrak{C}}^{-1}\hat{C}_3$ does not commute.

By Theorem 2.1.4 (case 1 (ii)), $\hat{C}_1, \hat{C}_2, \hat{C}_3$ are not $*$-SDC. Hence, neither C_1, C_2, C_3.

2.1.2 The SDP method

Now, we give some equivalent $*$-SDC conditions for Hermitian matrices in the following theorem.

Theorem 2.1.5. *The following conditions are equivalent:*

(i) The matrices $C_1, C_2, \ldots, C_m \in \mathbb{H}^n$ are $*$-SDC.

(ii) There exists a nonsingular matrix $P \in \mathbb{C}^{n\times n}$ such that $P^*C_1P, P^*C_2P, \ldots, P^*C_mP$ commute.

(iii) There exists a positive definite $X = X^* \in \mathbb{H}^n$ that solves the following systems:

$$C_iXC_j = C_jXC_i, \quad 1 \leq i < j \leq m. \tag{2.12}$$

We note that the theorem is also true for the real setting: If C_i's are all real then the corresponding matrices P, X in all conditions above can be all picked to be real.

Proof. (i) $\Rightarrow$ (ii) If the matrices $C_1, C_2, \ldots, C_m \in \mathbb{H}^n$ are $*$-SDC, then there is a nonsingular $P \in \mathbb{C}^{n\times n}$ such that $P^*C_1P, P^*C_2P, \ldots, P^*C_mP$ are diagonal. The latter matrices clearly commute.

(ii) $\Rightarrow$ (iii) Let $P = QU$ be a polar decomposition, $Q = Q^*$ be positive definite, and U be unitary. We have

$$\begin{aligned}(U^*Q^*C_iQU)(U^*Q^*C_jQU) &= (P^*C_iP)(P^*C_jP) = (P^*C_jP)(P^*C_iP) \\ &= (U^*Q^*C_jQU)(U^*Q^*C_iQU), \quad i \neq j.\end{aligned}$$

Consequently, QC_iQ and QC_jQ commute:

$$QC_iQ.QC_jQ = QC_jQ.QC_iQ, \quad \forall i \neq j.$$

Then $C_iQ^2C_j = C_jQ^2C_i, \forall i \neq j$. Therefore, (2.12) holds true for $X = Q^2$.

(iii) $\Rightarrow$ (i) If X is a positive definite matrix which satisfies (2.12), then Q can be picked as the square root of X. From (2.12), the matrices $QC_1Q, QC_2Q, \ldots, QC_mQ$ are $*$-SDC by Theorem 2.1.3. So are $C_1, C_2, \ldots, C_m$.

Finally, suppose C_i's are all real symmetric and let $X \in \mathbb{H}^n$ be a positive definite matrix satisfying (2.12). Let Y, Z be the real and imaginary parts, respectively, of X. Then $Y^T = Y$ and $Z^T = -Z$. It is well-known in the literature that Y is also positive definite. Substituting Y, Z into (2.12) and comparing the real and the imaginary parts, one obtains

$$C_iYC_j = C_jYC_i, \;\; C_iZC_j = C_jZC_i, \;\; 1 \le i < j \le m.$$

The matrices $\sqrt{Y}C_1\sqrt{Y}, \ldots, \sqrt{Y}C_m\sqrt{Y}$ are $\mathbb{R}$-SDC by an orthogonal matrix P, and $C_1, \ldots, C_m$ are $\mathbb{R}$-SDC by $\sqrt{Y}P$. The orthogonality of P is due to Theorem 2.1.3. $\square$

Based on the Theorems 2.1.3 and 2.1.5, we give Algorithm 6, consisting of two stages:

Stage 1: detect whether a collection of Hermitian matrices are SDC by solving a linear system of the form (2.12) and obtaining commuting Hermitian matrices. This stage is based on Theorem 2.1.5(iii), and it is the most significant contribution of this section. In this stage, an SDP solvers is used to find a positive definite matrix under the images of the initial Hermitian matrices under congruence (The image of a matrix X under the congruence matrix P is defined as P^*XP) are commuting; and

Stage 2: simultaneously diagonalize via congruence the resulting image matrices by a unitary matrix.

Algorithm 3 can be applied to the second stage. However, it requires to compute the eigenvalue decomposition of all matrices $C_1, \ldots, C_m$ in step 1, while simultaneously diagonalizing k collections of submatrices in step 4. This may cause high computational complexity. We hence prefer Algorithm 6 below to Algorithm 3 for this stage. Algorithm 5 exploits the works in [10, 43], where the work in [10] proposes a Jacobi-like algorithm for simultaneously diagonalizing two commuting normal matrices, and that in [43] extends to several normal ones together with MATLAB implementations. It is also worth mentioning that Algorithm 4 can be used for solving the Hermitian SDC problem. However, in doing so, it needs some auxiliary steps. For example, eigenvalue decomposition for matrices C_i's; a maximum-rank linear combination and its pseudoinverse; and determination of whether $\det\mathfrak{C}(\lambda)$ being identically zero on $\mathbb{R}^m$ or not (Algorithm 2). Algorithm 6 below does not do these things and does not care many cases as in Algorithm 4. It first reformulates and solves the problem of detecting the SDC property as a semidefinite program, which is a famous numerical method with many excellent toolboxes widely used in engineering and related areas, and then performs an ideal Jacobi-like approximation [10, 43] for the resulting matrices.

The Jacobi-like method in [10, 43] can be summarized as follows. Suppose $C_i = [c_{uv}^{(i)}] \in \mathbb{H}^n$ and let

$$\mathrm{off}_2 = \mathrm{off}_2(C_1, \dots, C_m) = \sum_{i=1}^{m} \sum_{u \neq v} |c_{uv}^{(i)}|^2, \tag{2.13a}$$

$$R(u, v, c, s) = I_n + (c-1)\mathbf{e}_u\mathbf{e}_u^T - \bar{s}\mathbf{e}_u\mathbf{e}_v^T + s\mathbf{e}_v\mathbf{e}_u^T + (\bar{c}-1)\mathbf{e}_v\mathbf{e}_v^T, \tag{2.13b}$$

where $c, s \in \mathrm{C}_{\mathrm{rd}}$ with $|c|^2 + |s|^2 = 1$. $R(u, v, c, s)$ is called a (u, v)-Givens or (u, v)-plane rotation matrix, $1 \leq u < v \leq n$.

It can be verified that for a given pair (c, s) and every pair $(u, v) \in \{1, \dots, n\}^2$, the following holds true:

$$\begin{aligned}\mathrm{off}_2(RC_1R^*, \dots, RC_mR^*) = \mathrm{off}_2(C_1, \dots, C_m) &- \sum_{i=1}^{m} \left(|c_{uv}^{(i)}|^2 + |c_{vu}^{(i)}|^2\right) \\ &+ \sum_{i=1}^{m} \left|c^2\bar{c}_{uv}^{(i)} + cs(\bar{c}_{uu}^{(i)} - \bar{c}_{vv}^{(i)}) - s^2\bar{c}_{vu}^{(i)}\right|^2 \\ &+ \sum_{i=1}^{m} \left|c^2 c_{vu}^{(i)} + cs(c_{uu}^{(i)} - c_{vv}^{(i)}) - s^2 c_{uv}^{(i)}\right|^2.\end{aligned} \tag{2.14}$$

In the methods [10, 43], at the loop with respect to each (u, v), it tries to find c, s that makes $\mathrm{off}_2(RC_1R^*, \dots, RC_mR^*) < \mathrm{off}_2(C_1, \dots, C_m)$. The values of c, s can be looked for, as shown in, e.g., [24], so that the last sum on the right-hand side of (2.14) is minimized. This is equivalent to the minimization of the amount $\|M_{uv}z\|_2$ with

$$M_{uv} = \begin{bmatrix} \bar{c}_{uv}^{(1)} & (\bar{c}_{uu}^{(1)} - \bar{c}_{vv}^{(1)}) & -\bar{c}_{vu}^{(1)} \\ c_{vu}^{(1)} & (c_{uu}^{(1)} - c_{vv}^{(1)}) & -c_{vu}^{(1)} \\ \vdots & \vdots & \vdots \\ \bar{c}_{uv}^{(m)} & (\bar{c}_{uu}^{(m)} - \bar{c}_{vv}^{(m)}) & -\bar{c}_{vu}^{(m)} \\ c_{vu}^{(m)} & (c_{uu}^{(m)} - c_{vv}^{(m)}) & -c_{vu}^{(m)} \end{bmatrix}, \quad z = \begin{bmatrix} c^2 \\ sc \\ s^2 \end{bmatrix}.$$

Note that $\|z\|_2 = 1$ and c, s can be parameterized as $c = \cos(\theta_{uv})$, $s = e^{\mathbf{i}\phi_{uv}} \sin(\theta_{uv})$, $(\theta_{uv}, \phi_{uv}) \in [-\frac{\pi}{4}, \frac{\pi}{4}] \times [-\pi, \pi]$. As mentioned in [10], minimizing the amount $\|M_{uv}z\|_2$ may be "relatively complicated", and it suffices to approximate this minimization problem to one that minimizes

$$g_{uv}(c,s) = \sum_{i=1}^{m} \left(|cs(\bar{c}_{uu}^{(i)} - \bar{c}_{vv}^{(i)}) + c^2\bar{c}_{uv}^{(i)} - s^2\bar{c}_{vu}^{(i)}| + |cs(c_{uu}^{(i)} - c_{vv}^{(i)}) - s^2 c_{uv}^{(i)} + c^2 c_{vu}^{(i)}| \right) \tag{2.15}$$

in (c, s) defined above. The following algorithm is a pseudo-code of the work [43] we record it here for convenience.

Algorithm 5 SDC of commuting Hermitian matrices [10, 43].

INPUT: Commuting Hermitian matrices $C_1, \ldots, C_m \in \mathbb{C}^{n\times n}$, a tolerance $\epsilon > 0$.
OUTPUT: A unitary matrix U such that $\mathrm{off}_2 \leq \epsilon \sum_{i=1}^{m} \|C_i\|_F =: \nu_\epsilon$.

1: Accumulate $Q = I_n$.

2: **While** $\mathrm{off}_2 > \nu_\epsilon$ **do**

3: For every pair (u, v), $1 \leq u < v \leq n$, determine the rotation $R(u, v, c, s)$ such that $(c, s) = (\cos\theta_{uv}, e^{\mathbf{i}\phi}\sin\theta_{uv})$ minimizes the function g_{uv} in (2.15).

4: Accumulate $Q = QR(u, v, c, s)$, $C_i = R(u, v, c, s)^* C_i R(u, v, c, s)$, $i = 1, \ldots, m$.

5: **Endwhile**

For each pair (u, v), $1 \leq u < v \leq n$, Algorithm 5, which summarizes the work [43], requires: $O(m)$ flops for approximating the minimum of $g_{uv}^{(i)}$ in (2.15); $O(n)$ flops for updating $Q = QR(u, v, c, s)$; $O(mn)$ flops for updating m matrices

$$C_i = R(u, v, c, s)^* C_i R(u, v, c, s).$$

The whole algorithm hence needs $O(mn^3)$ complex flops.

For $m = 2$, it is shown in [10] that Algorithm 5 locally quadratically converges.

By analogous methodology, this rate of convergence is still valid for $m \geq 2$ matrices.

We now exploit Algorithm 5 to propose our main algorithm below.

Algorithm 6 Solving the Hermitian SDC problem.

INPUT: Hermitian matrices $C_1, C_2, \ldots, C_m \in \mathbb{H}^n$ (not necessarily commuting).
OUTPUT: A nonsingular matrix U such that $U^* C_i U$'s are diagonal (if exists).

```
Compute n − q = rank(C_1 ... C_m).
If q = 0 then
    Solve the system (2.12) by using a SDP solver.
    If ∃P ≻ 0 solving (2.12) then
        Compute the square root Q of P, Q^2 = P.
        Apply Algorithm 5 to the matrices QC_iQ's and obtain a unitary matrix V.
        Return U = QV.
    else: Conclude the given matrices are not SDC.
    endif
else
    Compute a SVD of (C_1 ... C_m) := UΛV^*.
    Obtain the matrices Ĉ_i as in (2.1) from V^*C_iV.
    Similarly proceed as the case q = 0 for the matrices Ĉ_i.
endif
```

To illustrate Algorithm 6, we consider the following examples.

Example 2.1.6. Let

$$C_1 = \begin{pmatrix} 1 & 3 & -2 \\ 3 & 16 & -10 \\ -2 & -10 & 6 \end{pmatrix}, C_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -3 & 2 \\ 0 & 2 & -1 \end{pmatrix}, C_3 = \begin{pmatrix} -1 & -3 & 2 \\ -3 & -5 & 4 \\ 2 & 4 & -3 \end{pmatrix}.$$

To apply Theorem 2.1.5, we need to find

$$X = \begin{pmatrix} x & y & z \\ y & t & u \\ z & u & v \end{pmatrix} \succ 0 \ \left(\Leftrightarrow x > 0, xt > y^2, \det(X) > 0\right) \tag{2.16}$$

such that

$$C_1XC_2 = (C_1XC_2)^*,\ \ C_1XC_3 = (C_1XC_3)^*,\ \ C_2XC_3 = (C_2XC_3)^*.$$

By directly computation,

$$C_1XC_2 = (C_1XC_2)^* \Leftrightarrow \begin{cases} 12u & -9t & -4v & -3y & +2z & = 0 \\ -7u & +6t & +2v & +2y & -z & = 0 \\ 2u & +2t & -2v & & +z & = 0, \end{cases}$$

$$C_1XC_3 = (C_1XC_3)^* \Leftrightarrow \begin{cases} 40u & -33t & -12v & -11y & +6z & = 0 \\ 7u & -6t & -2v & -2y & +z & = 0 \\ 18u & -14t & -6v & -4y & +3z & = 0, \end{cases}$$

$$C_2XC_3 = (C_2XC_3)^* \Leftrightarrow \begin{cases} -12u & +9t & +4v & +3y & -2z & = 0 \\ 4u & -2t & -2v & & +z & = 0 \\ 7u & -6t & -2v & -2y & +z & = 0. \end{cases}$$

Combining the linear equations above, we obtain

$$u = 2y,\ \ t = y,\ \ v = 3y + \frac{z}{2}.$$

Let us pick $y = 1, z = 4, x = 6$ with which $X = \begin{pmatrix} 6 & 1 & 4 \\ 1 & 1 & 2 \\ 4 & 2 & 5 \end{pmatrix} \succ 0$ satisfies $C_iXC_j = C_jXC_i$, $1 \leq i < j \leq 3$. Thus three initial matrices are $\mathbb{R}$-SDC on $\mathbb{R}$, and so are they on $\mathbb{C}$.

Example 2.1.7. As shown in [11], the matrices $C_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ are $\mathbb{C}$-SDC. However, they are not $*$-SDC by Theorem 2.1.4 since C_1 is nonsingular and

$$C_1^{-1}C_2 = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$$

has only complex eigenvalues $\frac{1 \pm i\sqrt{3}}{2}$. Similarly, they are not $\mathbb{R}$-SDC by Theorem 1.2.1.

We can also check this by applying Theorem 2.1.5 as follows. The matrices are $*$-SDC if and only if there is a positive definite matrix $X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \succ 0$, which is equivalent to $x > 0$ and $xz > y^2$ (it suffices to deal with the real world of X) such that

$$C_1XC_2 = C_2XC_1 \left(= (C_1XC_2)^*\right).$$

This is equivalent to

$$\begin{cases} x > 0, & xz > y^2 \\ x + y + z & = 0. \end{cases}$$

The last condition is impossible to satisfy since there do not exist $x, z > 0$ such that $xz > y^2 = (x+z)^2$. Thus C_1 and C_2 are not $*$-SDC on $\mathbb{R}$.

We finish this part by stating the relationship between the SDC problems for arbitrarily square and Hermitian matrices. The study of the Hermitian SDC problem is again confirmed to be meaningful in SDC theory. The theorem below refers to the notation of the Hermitian and skew-Hermitian parts, respectively, of a square matrix A as follows:

$$\mathcal{H}(A) = \frac{1}{2}(A + A^*) = \mathcal{H}(A)^*, \ \ \mathcal{S}(A) = \frac{1}{2}(A - A^*) = -\mathcal{S}(A)^*, \ \ \mathbf{i^2} = -1.$$

We further note that both $\mathcal{H}(A)$ and $\mathbf{i}\mathcal{S}(A)$ are Hermitian matrices.

Theorem 2.1.6. (see, e.g., in [35, Section 1.7, Problem 18]) *The square matrices $A_1, \ldots, A_m \in \mathbb{F}^{n\times n}$ are $*$-SDC if and only if so are $\mathcal{H}(A_t), \mathbf{i}\mathcal{S}(A_t)$, $t = 1, \ldots, m$.*

Proof. If $A_1, \ldots, A_m$ are SDC by a nonsingular matrix P then $P^*A_tP = D_t$ and $P^*A_t^*P = D_t^*$ are diagonal for every $t = 1, \ldots, m$. As a result, $P^*\mathcal{H}(A_t)P = \mathcal{H}(D_t)$, $P^*\mathbf{i}\mathcal{S}(A_t)P = \mathbf{i}\mathcal{S}(D_t)$ are real diagonal for every $t = 1, \ldots, m$.

The opposite direction is analogously proved by noticing that

$$A_t = \mathcal{H}(A_t) - \mathbf{i}[\mathbf{i}\mathcal{S}(A_t)], \ \ t = 1, \ldots, m.$$

□

Example 2.1.8. Given two square complex matrices

$$A_1 = \begin{pmatrix} 10 - 28\mathbf{i} & -2 + 16\mathbf{i} & -6 - 2\mathbf{i} \\ -6 + 12\mathbf{i} & 2 - 7\mathbf{i} & 2 + 2\mathbf{i} \\ 2 + 6\mathbf{i} & -2 - 2\mathbf{i} & 2 - 2\mathbf{i} \end{pmatrix}, \ A_2 = \begin{pmatrix} 21 - 4\mathbf{i} & -8 + 5\mathbf{i} & -3 - 6\mathbf{i} \\ -8 - \mathbf{i} & 4 - \mathbf{i} & 3\mathbf{i} \\ -3 + 6\mathbf{i} & -3\mathbf{i} & 3 \end{pmatrix}.$$

Their Hermitian and skew-Hermitian parts, respectively, are

$$\mathcal{H}(A_1) = \begin{pmatrix} 10 & -4 + 2\mathbf{i} & -2 - 4\mathbf{i} \\ -4 - 2\mathbf{i} & 2 & 2\mathbf{i} \\ -2 + 4\mathbf{i} & -2\mathbf{i} & 2 \end{pmatrix}, \ \mathbf{i}\mathcal{S}(A_1) = \begin{pmatrix} 28 & -14 + 2\mathbf{i} & -2 - 4\mathbf{i} \\ -14 - 2\mathbf{i} & 7 & 2\mathbf{i} \\ -2 + 4\mathbf{i} & -2\mathbf{i} & 2 \end{pmatrix},$$

$$\mathcal{H}(A_2) = \begin{pmatrix} 21 & -8 + 3\mathbf{i} & -3 - 6\mathbf{i} \\ -8 - 3\mathbf{i} & 4 & 3\mathbf{i} \\ -3 + 6\mathbf{i} & -3\mathbf{i} & 3 \end{pmatrix}, \ \mathbf{i}\mathcal{S}(A_2) = \begin{pmatrix} 4 & -2 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

For short, let C_1, C_2, C_3, C_4 be $\mathcal{H}(A_1), \mathbf{i}\mathcal{S}(A_1), \mathcal{H}(A_2), \mathbf{i}\mathcal{S}(A_2) \in \mathbb{H}^3$, respectively. By Theorem 2.1.6, it suffices to check whether the four latter Hermitian matrices are $*$-SDC or not. Once again we apply Theorem 2.1.5 to find a positive definite matrix

$X = \begin{pmatrix} x & y & z \\ \bar{y} & t & u \\ \bar{z} & \bar{u} & v \end{pmatrix} \in \mathbb{H}^3$. That is, we need $x > 0, xt > |y|^2$, $\det(X) = (uy\bar{z} + \overline{uy}z) + xtv - x|u|^2 - t|z|^2 - v|y|^2 > 0$, $x, t, v \in \mathbb{R}$, and $C_iXC_j = C_jXC_i$, $1 \le i < j \le 4$.

These equations are equivalent to

$$\begin{cases}
\begin{array}{lllllllll}
 & 9y_1 & -7y_2 & -18z_1 & +9z_2 & -5t & +10u_1 & -5u_2 & =0 \\
14x & -7y_1 & & -10z_1 & -2z_2 & & +5u_1 & & =0 \\
-18x & & +7y_2 & +38z_1 & -10z_2 & +5t & -20u_1 & +5u_2 & =0 \\
9x & -5y_1 & -y_2 & -9z_1 & & & +5u_1 & & =0 \\
18x & -19y_1 & +5y_2 & & -9z_2 & +5t & & +5u_2 & =0 \\
 & & & & & & & & \\
 & 3y_1 & +y_2 & -6z_1 & +3z_2 & -t & +2u_1 & -u_2 & =0 \\
2x & -y_1 & & +2z_1 & -2z_2 & & -u_1 & & =0 \\
6x & & +y_2 & -10z_1 & +2z_2 & -t & +4u_1 & -u_2 & =0 \\
3x & -y_1 & +y_2 & -3z_1 & & & +u_1 & & =0 \\
6x & -5y_1 & +y_2 & & -3z_2 & +t & & +u_2 & =0 \\
 & & & & & & & & \\
 & 2y_1 & -y_2 & -4z_1 & +2z_2 & -t & +2u_1 & -u_2 & =0 \\
4x & & -y_2 & -8z_1 & +2z_2 & -t & +4u_1 & -u_2 & =0 \\
 & & & & & & & & \\
 & -21y_1 & +35y_2 & +42z_1 & -21z_2 & +13t & -26u_1 & +13u_2 & =0 \\
70x & -35y_1 & & -26z_1 & -10z_2 & & +13u_1 & & =0 \\
42x & & -35y_2 & -94z_1 & +26z_2 & -13t & +52u_1 & -13u_2 & =0 \\
21x & -13y_1 & -5y_2 & -21z_1 & & & +13u_1 & & =0 \\
42x & -47y_1 & +13y_2 & & -21z_2 & +13t & & +13u_2 & =0 \\
 & & & & & & & & \\
 & -2y_1 & & +4z_1 & -2z_2 & +t & -2u_1 & +u_2 & =0 \\
 & & & 2z_1 & & & -u_1 & & =0 \\
4x & & & -8z_1 & +2z_2 & -t & +4u_1 & -u_2 & =0 \\
 & & & & & & & & \\
 & 6y_1 & -5y_2 & -12z_1 & +6z_2 & -3t & +6u_1 & -3u_2 & =0 \\
10x & -5y_1 & & -6z_1 & & & +3u_1 & & =0 \\
-12x & & +5y_2 & +24z_1 & -6z_2 & +3t & -12u_1 & +3u_2 & =0 \\
2x & -y_1 & & -2z_1 & & & +u_1 & & =0 \\
4x & -4y_1 & +y_2 & & -2z_2 & +t & & +u_2 & =0,
\end{array}
\end{cases}$$

and, is equivalent to

$$\begin{cases} 2x & -y_1 & & +2z_1 & & & -u_1 & & =0 \\ & y_1 & +2y_2 & +36z_1 & -18z_2 & & -19u1 & & =0 \\ & & 5y_2 & +76z_1 & -38z_2 & +t & -40u_1 & +u_2 & =0 \\ & & & 2z_1 & -z_2 & & -u_1 & & =0 \\ & & & & z_2 & & & & =0 \\ & & & & & t & -2u_1 & +u_2 & =0. \end{cases}$$

A solution X of these equations must be in the form

$$X = \begin{pmatrix} z_1 & 2z_1 & z_1 \\ 2z_1 & 4z_1 - u_2 & 2z_1 + \mathbf{i}u_2 \\ z_1 & 2z_1 - \mathbf{i}u_2 & v \end{pmatrix}, \tag{2.17}$$

where $z_1 \in \mathbb{R}$ is the real part of z, and $u_2 \in \mathbb{R}$ is the imaginary part of u. In addition, these parameters must satisfy $z_1 > 0$, $-z_1u_2 > 0$, $-z_1u_2(v+u_2-z_1) = \det(X) > 0$ to ensure the positive definiteness of X. For example, one can pick $X = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 2-\mathbf{i} \\ 1 & 2+\mathbf{i} & 3 \end{pmatrix}$ with respect to $z_1 = 1$, $u_2 = -1$, $v = 3$. This yields that C_1, C_2, C_3, C_4 are $*$-SDC.

Numerical experiment for this problem will be shown in Example 2.1.9 below.

2.1.3 Numerical tests

We now give some numerical tests illustrating Algorithm 6 implemented in MATLAB R2015a running on a PC with Intel Core i3 CPU 3.3GHz, 8GB RAM, Windows 10 x64 operating system. In each test, we set up a collection of Hermitian matrices that are surely SDC as follows: Fix a nonsingular matrix P whose entries are randomly taken from a uniform distribution on the interval $(0, 1)$ and pick m diagonal matrices D_i whose diagonal elements are in $(-1, 1)$, then construct $C_i = P^*D_iP$. These latter matrices $C_1, \ldots, C_m$ are clearly $*$-SDC by P. Note that the diagonal entries of D_i could be zero, making the matrices $C_1, \ldots, C_m$ so generated be either singular or not.

The first stage of Algorithm 6 is implemented with the `CVX` toolbox [26] calling `SDPT3` version 4.0 [63] that solves the following semidefinite program

$$\min\{s \mid X - sI_n \succeq 0, \;\; s \geq \epsilon, \;\; C_iXC_j = C_jXC_i, 1 \leq i < j \leq m\}, \tag{2.18}$$

where the tolerance $\epsilon > 0$ is given. We then exploit the MATLAB function `sqrtm.m`, which executes the algorithm proposed in [15], to compute the square root Q of X. For the second stage, thank to the works in [43] executing Algorithm 5. In our experiment, we pick ϵ to be the floating-point relative accuracy `eps` of MATLAB for the first stage, while keeping their tolerance for the second stage to be `eps` to the power of $\frac{3}{2}$[1][43].

Tables 2.1 and 2.2 show some numerical tests for real and complex Hermitian SDC problems, respectively. Each result in these tables is the average of five executions. Because the input matrices are randomly chosen, they should be linearly independent. We hence pick $m \leq \dim_{\mathbb{R}} \mathcal{S}^n = \frac{n(n+1)}{2}$ in Table 2.1 and $m \leq \dim_{\mathbb{R}} \mathbb{H}^n = n^2$ in Table 2.2.

The errors of the first stage are estimated by

$$\text{Err1} = \max_{1 \leq i < j \leq m} \|C_i X C_j - C_j X C_i\|_2 ,$$

while those of the whole algorithm are estimated as

$$\text{Err2} = \max_{1 \leq i \leq m} \|U^* C_i U - \texttt{diag}(\texttt{diag}(U^* C_i U))\|_2 ,$$

where $\texttt{diag}(\texttt{diag}(X))$ denotes the diagonal matrix whose diagonal is that of X.

Table 2.1: Numerical tests for the real Hermitian SDC problem.

m, number of matrices	n, size of matrices	Err1	Err2	CPU time (s)
6	3	4.71e-14	4,62e-14	2.37
10	10	5.55e-14	5.21e-12	9.67
15	10	6.50e-13	5.34e-12	17.56
20	10	2.98e-13	9.25e-11	24.49
30	10	7.92e-13	1.30e-11	86.05
10	15	2.55e-11	1.60e-10	59.69
30	15	2.31e-11	2.86e-10	632.23
10	20	8.80e-11	1.39e-10	337.37

Example 2.1.9. We now revisit the matrices in Example 2.1.8 with numerical performance based on Algorithm 6. The first stage gives a positive definite matrix

$$X \simeq \begin{pmatrix} 257.78 & 515.55 & 257.78 \\ 515.55 & 1457 & 515.55 - \mathbf{i}\,425.93 \\ 257.78 & 515.55 + \mathbf{i}\,425.93 & 1537.1 \end{pmatrix}.$$

[1] MATLAB codes of Algorithm 6, and `Julia` codes for the first stage are available at `https://sites.google.com/a/qnu.edu.vn/le-thanh-hieu/experiments`.

Table 2.2: Numerical tests for the complex Hermitian SDC problem.

m, number of matrices	n, size of matrices	Err1	Err2	CPU time (s)
9	3	1.97e-13	2.30e-13	4.08
10	10	2.63e-13	4.12e-13	17.61
15	10	2.97e-13	8.29e-13	30.25
20	10	3.33e-12	2.41e-12	51.31
10	15	2.92e-11	4.02e-11	144.08
10	20	2.86e-10	2.48e-10	742.38

It turns out that this is a special case of that in (2.17) with $z_1 \simeq 257.78$, $u_2 \simeq -425.93$, $v \simeq 1537.1$ and $t \simeq 1457$. Stage 2 of Algorithm 6 is performed for the matrices $\sqrt{X}C_i\sqrt{X}$, $i = 1, \ldots, 4$, and one obtains the unitary matrix

$$Q \simeq \begin{pmatrix} 0.70043 & -0.57468 & -0.42326 \\ 0.66343 + \mathbf{i}\ 0.043507 & 0.68597 - \mathbf{i}0.10828 & 0.16650 + \mathbf{i}\ 0.21902 \\ 0.24454 - \mathbf{i}\ 0.087015 & -0.42389 + \mathbf{i}0.43089 & 0.46223 - \mathbf{i}\ 0.72904 \end{pmatrix}.$$

The final nonsingular matrix simultaneously diagonalizes C_1, C_2, C_3, C_4, and hence A_1, A_2 is

$$U = \sqrt{X}Q \simeq \begin{pmatrix} 16.0554 & -0.0000 & 0.0000 \\ 32.1108 & 20.6021 - \mathbf{i}\ 1.2155 & 0.0000 \\ 16.0554 & 1.2155 + \mathbf{i}\ 20.6021 & 15.6427 - \mathbf{i}\ 24.6720 \end{pmatrix}.$$

Moreover,

$$U^*C_1U \simeq \mathtt{diag}(0, 0, 1706.8), \qquad U^*C_2U \simeq \mathtt{diag}(-515.55, 2129.63, 1706.80),$$
$$U^*C_3U \simeq \mathtt{diag}(515.55, 425.93, 2560.20), \qquad U^*C_4U \simeq \mathtt{diag}(0, 425.93, 0).$$

Example 2.1.10. Consider the two matrices

$$\begin{pmatrix} 45 & 10 & 0 & 5 & 0 & 0 \\ 10 & 45 & 5 & 0 & 0 & 0 \\ 0 & 5 & 45 & 10 & 0 & 0 \\ 5 & 0 & 10 & 45 & 0 & 0 \\ 0 & 0 & 0 & 0 & 16.4 & -4.8 \\ 0 & 0 & 0 & 0 & -4.8 & 13.6 \end{pmatrix}, \begin{pmatrix} 27.5 & -12.5 & -.5 & -4.5 & -2.04 & 3.72 \\ -12.5 & 27.5 & -4.5 & -.5 & 2.04 & -3.72 \\ -.5 & -4.5 & 24.5 & -9.5 & -3.72 & -2.04 \\ -4.5 & -.5 & -9.5 & 24.5 & 3.72 & 2.04 \\ -2.04 & 2.04 & -3.72 & 3.72 & 54.76 & -4.68 \\ 3.72 & -3.72 & -2.04 & 2.04 & -4.68 & 51.24 \end{pmatrix}.$$

which are proved to be positive definite in [19, 52]. Algorithm 6 gives Err1$\simeq 2.89e-13$ and Err2$\simeq 4.68e-14$.

2.2 An alternative solution method for the SDC problem of real symmetric matrices

As indicated in Theorem 2.1.5, equivalent conditions (i)-(iii) hold also for the real setting, i.e., when C_i are all real symmetric. Then R and R^*C_iR can be picked to be real. However, solving an SDP problem for a positive definite matrix X may not efficient, in particular when the dimension n or the number m of the matrices is large. In this section, we propose an alternative method for solving the real SDC problem of real symmetric matrices, i.e., $C_i \in \mathcal{C}$ are real symmetric and the congruence matrice R and R^TC_iR are also real. The method is iterative which begins with only two matrices C_1, C_2. If the two matrices C_1, C_2 are SDC, we include C_3 to consider the SDC of C_1, C_2, C_3, and so forth. We divide $\mathcal{C} = \{C_1, C_2, \ldots, C_m\} \subset \mathcal{S}^n$ into two cases. The first case is called the *nonsingular collection* (in Section 2.2.1), when at least one $C_i \in \mathcal{C}$ is nonsingular. The other case is called the *singular collection* (in Section 2.2.3), when all $C_i's$ in $\mathcal{C}$ are non-zero but singular. When $\mathcal{C}$ is a nonsingular collection, we always assume that C_1 is nonsingular. A nonsingular collection will be denoted by $\mathcal{C}_{ns}$, while $\mathcal{C}_s$ represents the singular collection. The results are based on [49].

2.2.1 The SDC problem of nonsingular collection

Consider a nonsingular collection $\mathcal{C}_{ns} = \{C_1, C_2, \ldots, C_m\} \subset \mathcal{S}^n$ and assume that C_1 is nonsingular. Let us outline the approach to determine the SDC of $\mathcal{C}_{ns}$. First, in below Lemmas 2.2.1 we show that if $\mathcal{C}_{ns}$ is $\mathbb{R}$-SDC, it is necessary that

(N1) $C_1^{-1}C_i,\ i = 2, 3, \ldots, m$ is real similarly diagonalizable;

(N2) $C_jC_1^{-1}C_i$ is symmetric, for every $i = 2, 3, \ldots, m$ and every $j \neq i$.

Conversely, for the sufficiency, we use (N1) and (N2) to decompose, iteratively, all matrices in $\mathcal{C}_{ns}$ into block diagonal forms of smaller and smaller size until all of them become the so-called non-homogeneous dilation of the same block structure (to be seen later) with certain scaling factors. Then, the $\mathbb{R}$-SDC of $\mathcal{C}_{ns}$ is readily achieved.

Firstly, we have following lemma.

Lemma 2.2.1. *If a nonsingular collection $\mathcal{C}_{ns}$ is $\mathbb{R}$-SDC, then*

(N1) $C_1^{-1}C_i,\ i = 2, 3, \ldots, m$ is real similarly diagonalizable;

(N2) $C_jC_1^{-1}C_i$ *is symmetric, for every* $i = 2, 3, \dots, m$ *and every* $j \neq i$.

Proof. If $C_1, C_2, \dots, C_m$ are SDC by a nonsingular real matrix P then

$$P^TC_iP = D_i, i = 1, 2, \dots, m,$$

are real diagonal. Since C_1 is nonsingular, D_1 is nonsingular and we have

$$C_i = (P^T)^{-1}D_iP^{-1}; i = 1, 2, \dots, m; C_1^{-1} = PD_1^{-1}P^T.$$

Then $P^{-1}C_1^{-1}C_iP = D_1^{-1}D_i$ are real diagonal. That is $C_1^{-1}C_i$ are real similarly diagonalizable, $i = 2, 3, \dots, m$. For $2 \leq i < j \leq m$, we have

$$\begin{aligned} C_jC_1^{-1}C_i =& ((P^T)^{-1}D_jP^{-1})(PD_1^{-1}P^T)((P^T)^{-1}D_iP^{-1}) \\ =& (P^T)^{-1}D_jD_1^{-1}D_iP^{-1}. \end{aligned}$$

The matrices $D_jD_1^{-1}D_i$ are symmetric, so are $C_jC_1^{-1}C_i$. □

By Theorem 2.2.1 and Theorem 2.2.2 below, we will show that (N1) and (N2) are indeed sufficient for $\mathcal{C}_{ns}$ to be SDC. Let us begin with Lemma 2.2.2.

Lemma 2.2.2. *Let* $\mathcal{C}_{ns} = \{C_1, C_2, \dots, C_m\} \subset \mathcal{S}^n$ *be a nonsingular collection with* C_1 *invertible. Suppose* $C_1^{-1}C_2$ *is real similarly diagonalized by invertible matrix* Q *to have* r *distinct eigenvalues* $\beta_1, \dots, \beta_r$*; each of multiplicity* m_t*,* $t = 1, 2, \dots, r$*, respectively. Then,*

$$Q^TC_1Q = \texttt{diag} \underbrace{((A_1)_{m_1}, (A_2)_{m_2}, \dots, (A_r)_{m_r})}_{m_1+\dots+m_r=n,\ \text{each } A_t\text{: sym. invert.}}; \tag{2.19}$$

$$Q^TC_2Q = \texttt{diag}(\beta_1A_1, \beta_2A_2 \dots, \beta_rA_r). \tag{2.20}$$

In addition, if $C_jC_1^{-1}C_2, j = 3, 4, \dots, m$*, are symmetric, we can further block diagonalize* $C_3, C_4, \dots, C_m$ *to adopt the same block structure as in (2.19), such that*

$$Q^TC_jQ = \texttt{diag}\underbrace{((C_{j1})_{m_1}, (C_{j2})_{m_2}, \dots, (C_{jr})_{m_r})}_{\text{each } C_{jt}\text{: sym.}}\ j = 3, 4, \dots, m. \tag{2.21}$$

Proof. Since $C_1^{-1}C_2$ is similarly diagonalizable by Q, by assumption, there is

$$J := Q^{-1}C_1^{-1}C_2Q = \texttt{diag}(\beta_1I_{m_1}, \dots, \beta_rI_{m_r}) \tag{2.22}$$

with $m_1 + m_2 + \dots + m_r = n$. From (2.22), we have, for $j = 1, 2, \dots, m$,

$$(Q^TC_jQ)J = (Q^TC_jQ)(Q^{-1}C_1^{-1}C_2Q) = Q^TC_jC_1^{-1}C_2Q. \tag{2.23}$$

When $j = 1$, by substituting (2.22) into (2.23), we have

$$(Q^T C_1 Q)J = (Q^T C_1 Q) \cdot \texttt{diag}(\beta_1 I_{m_1}, \ldots, \beta_r I_{m_r}) = Q^T C_2 Q. \tag{2.24}$$

Since $Q^T C_1 Q, Q^T C_2 Q$ are both real symmetric and J is diagonal, Lemma 1.1.2 asserts that $Q^T C_1 Q$ is a block diagonal matrix with the same partition as J. That is, we can write

$$Q^T C_1 Q = \texttt{diag}((A_1)_{m_1}, (A_2)_{m_2}, \ldots, (A_r)_{m_r}), \tag{2.25}$$

which proves (2.19). Plugging both (2.25) and (2.22) into (2.24), we obtain

$$\begin{aligned} & \texttt{diag}((A_1)_{m_1}, (A_2)_{m_2}, \ldots, (A_r)_{m_r})\texttt{diag}(\beta_1 I_{m_1}, \ldots, \beta_r I_{m_r}) \\ = \ & \texttt{diag}(\beta_1 A_1, \ldots, \beta_r A_r) = Q^T C_2 Q, \end{aligned}$$

which proves (2.20).

Finally, for $j = 3, 4, \ldots, m$ in (2.23), due to the assumption that $C_j C_1^{-1} C_2$ are symmetric, so are $Q^T C_j C_1^{-1} C_2 Q$. By Lemma 1.1.2 again, $Q^T C_j Q$ are all block diagonal matrices with the same partition as J, which is exactly (2.21). □

Remark 2.2.1. When there is a nonsingular Q that puts $Q^T C_1 Q$ and $Q^T C_2 Q$ to (2.19) and (2.20), we say that $Q^T C_2 Q$ is a non-homogeneous dilation of $Q^T C_1 Q$ with scaling factors $\{\beta_1, \beta_2, \ldots, \beta_r\}$. In this case, since $A_1, A_2, \ldots, A_r$ are symmetric, there exist orthogonal matrices $H_i, i = 1, 2, \ldots, r$ such that $H_i^T A_i H_i$ is diagonal. Let $H = \texttt{diag}(H_1, H_2, \ldots, H_r)$, $Q^T C_1 Q$ and $Q^T C_2 Q$ are $\mathbb{R}$-SDC by the congruence H. Then, C_1 and C_2 are $\mathbb{R}$-SDC by the congruence QH.

For $m = 2$, Remark 2.2.1 and (N1) together give Theorem 1.2.1.

Another special case of Lemma 2.2.2 is when $C_1^{-1} C_2$ has n distinct real eigenvalues.

Corollary 2.2.1. *Let $\mathcal{C}_{ns} = \{C_1, C_2, \ldots, C_m\} \subset \mathcal{S}^n$ be a nonsingular collection with C_1 invertible. Suppose $C_1^{-1} C_2$ has n distinct real eigenvalues, i.e., $r = n$ in Lemma 2.2.2. Then, $C_1, C_2, \ldots, C_m$ are SDC if and only if $C_i C_1^{-1} C_2$ are symmetric for every $i = 3, \ldots, m$.*

Proof. If $C_1, C_2, \ldots, C_m$ are $\mathbb{R}$-SDC, by (N_2), we have $C_i C_1^{-1} C_2$ are symmetric for every $i = 3, \ldots, m$.

For the converse, since $C_1^{-1} C_2$ has n distinct eigenvalues, it is similarly diagonalizable. By assumption, $C_i C_1^{-1} C_2$ are symmetric. Then, by Lemma 2.2.2, the matrices $C_1, C_2, \ldots, C_m$ can be decomposed into block diagonals, each block is of size one. So $C_1, C_2, \ldots, C_m$ are $\mathbb{R}$-SDC. □

It then comes with our first main result, Theorem 2.2.1, below.

Theorem 2.2.1. *Let $\mathcal{C}_{ns} = \{C_1, C_2, \dots, C_m\} \subset \mathcal{S}^n$, $m \geq 3$ be a nonsingular collection with C_1 invertible. Suppose for each i the matrix $C_1^{-1}C_i$ is real similarly diagonalizable. If $C_j C_1^{-1} C_i$ are symmetric for $2 \leq i < j \leq m$, then there exists a nonsingular real matrix R such that*

$$\begin{aligned} R^T C_1 R =& \boldsymbol{diag}(A_1, A_2, \dots, A_s), \\ R^T C_2 R =& \boldsymbol{diag}(\alpha_1^2 A_1, \alpha_2^2 A_2, \dots, \alpha_s^2 A_s), \\ \dots & \quad \dots \\ R^T C_m R =& \boldsymbol{diag}(\alpha_1^m A_1, \alpha_2^m A_2, \dots, \alpha_s^m A_s), \end{aligned} \tag{2.26}$$

where $A_t's$ are nonsingular and symmetric, $\alpha_t^i, t = 1, 2, \dots, s$, are real numbers. When the nonsingular collection $\mathcal{C}_{ns}$ is transformed into the form of (2.26) by a congruence R, the collection $\mathcal{C}_{ns}$ is indeed $\mathbb{R}$-SDC.

Proof. Suppose $C_1^{-1}C_2$ is diagonalized by a nonsingular $Q^{(1)}$ with distinct eigenvalues $\beta_1^{(1)}, \beta_2^{(1)}, \dots, \beta_{r^{(1)}}^{(1)}$ having multiplicity $m_1^{(1)}, m_2^{(1)}, \dots, m_{r^{(1)}}^{(1)}$, respectively. Here the superscript $^{(1)}$ denotes the first iteration. Since $C_j C_1^{-1} C_2$ is symmetric for $j = 3, 4, \dots, m$, Lemma 2.2.2 assures that

$$C_1^{(1)} = {Q^{(1)}}^T C_1 Q^{(1)} = \texttt{diag}\underbrace{((A_1^{(1)})_{m_1^{(1)}}, (A_2^{(1)})_{m_2^{(1)}}, \dots, (A_{r^{(1)}}^{(1)})_{m_{r^{(1)}}^{(1)}})}_{sym.\ \&\ invert.}, \tag{2.27}$$

$$C_2^{(1)} = {Q^{(1)}}^T C_2 Q^{(1)} = \texttt{diag}(\beta_1^{(1)} A_1^{(1)}, \beta_2^{(1)} A_2^{(1)}, \dots, \beta_{r^{(1)}}^{(1)} A_{r^{(1)}}^{(1)}), \tag{2.28}$$

$$C_j^{(1)} = {Q^{(1)}}^T C_j Q^{(1)} = \texttt{diag}\underbrace{(C_{j1}^{(1)}, C_{j2}^{(1)}, \dots, C_{jr^{(1)}}^{(1)})}_{sym.}, \quad j = 3, 4, \dots, m; \tag{2.29}$$

where all members in $\{C_1^{(1)}, C_2^{(1)}, C_3^{(1)}, \dots, C_m^{(1)}\}$ adopt the same block structure, each having $r^{(1)}$ diagonal blocks.

As for the second iteration, we use the assumption that $C_1^{-1}C_3$ is similarly diagonalizable. Then,

$${C_1^{(1)}}^{-1} C_3^{(1)} = \texttt{diag}\left({A_1^{(1)}}^{-1} C_{31}^{(1)}, \dots, {A_{r^{(1)}}^{(1)}}^{-1} C_{3r^{(1)}}^{(1)}\right) \tag{2.30}$$

is also similarly diagonalizable. Since a block diagonal matrix is diagonalizable if and only if each of its blocks is diagonalizable, (2.30) implies that each ${A_t^{(1)}}^{-1} C_{3t}^{(1)}$, $t = 1, 2, \dots, r^{(1)}$ is diagonalizable. Let $Q_t^{(2)}$ (the superscript $^{(2)}$ denotes the second iteration) diagonalize ${A_t^{(1)}}^{-1} C_{3t}^{(1)}$ into l_t distinct eigenvalues $\beta_{t1}^{(2)}, \beta_{t2}^{(2)}, \dots, \beta_{tl_t}^{(2)}$, each having multiplicity $m_{t1}^{(2)}, m_{t2}^{(2)}, \dots, m_{tl_t}^{(2)}$, respectively. Then,

$$Q^{(2)} = \texttt{diag}(Q_1^{(2)}, Q_2^{(2)}, \dots, Q_{r^{(1)}}^{(2)})$$

diagonalizes ${C_1^{(1)}}^{-1} C_3^{(1)}$.

Now, applying Lemma 2.2.2 to $\{A_t^{(1)}, C_{3t}^{(1)}\}$ for each $t = 1, 2, \ldots, r^{(1)}$, we have

$$Q_t^{(2)^T} A_t^{(1)} Q_t^{(2)} = \mathtt{diag}\underbrace{((A_{t1}^{(2)})_{m_{t1}^{(2)}}, (A_{t2}^{(2)})_{m_{t2}^{(2)}}, \ldots, (A_{tl_t}^{(2)})_{m_{tl_t}^{(2)}})}_{sym.\ \&\ invert.}; \tag{2.31}$$

$$Q_t^{(2)^T} C_{3t}^{(1)} Q_t^{(2)} = \mathtt{diag}(\beta_{t1}^{(2)} A_{t1}^{(2)}, \beta_{t2}^{(2)} A_{t2}^{(2)}, \ldots, \beta_{tl_t}^{(2)} A_{tl_t}^{(2)}). \tag{2.32}$$

Let us re-enumerate the indices of all sub-blocks into a sequence from $r^{(1)}$ to $r^{(2)}$:

$$\begin{aligned} & \{11, 12, \ldots, 1l_1\}; \{21, 22, \ldots, 2l_2\}; \cdots; \{r^{(1)}1, r^{(1)}2 \ldots, r^{(1)}l_{r^{(1)}}\} \\ \Longrightarrow\ & \{1, 2, \ldots, l_1; l_1+1, l_1+2, \ldots, l_1+l_2; \ldots; \sum_{k=1}^{r^{(1)}-1} l_k + 1, \ldots, r^{(2)}\} \end{aligned} \tag{2.33}$$

so that

$$A_{11}^{(2)} \to A_1^{(2)};\ A_{12}^{(2)} \to A_2^{(2)}; \cdots; A_{1l_1}^{(2)} \to A_{l_1}^{(2)};\ A_{21}^{(2)} \to A_{l_1+1}^{(2)};\ A_{22}^{(2)} \to A_{l_1+2}^{(2)} \text{ and so on.}$$

Assemble (2.31) and (2.32) for all $t = 1, 2, \ldots, r^{(1)}$ together and then use the re-index (2.33), there is

$$C_1^{(2)} = Q^{(2)^T} C_1^{(1)} Q^{(2)} = \mathtt{diag}(A_1^{(2)}, A_2^{(2)}, \ldots, A_{r^{(2)}}^{(2)}), \tag{2.34}$$

$$C_3^{(2)} = Q^{(2)^T} C_3^{(1)} Q^{(2)} = \mathtt{diag}(\beta_1^{(2)} A_1^{(2)}, \beta_2^{(2)} A_2^{(2)}, \ldots, \beta_{r^{(2)}}^{(2)} A_{r^{(2)}}^{(2)}). \tag{2.35}$$

In other words, at the first iteration, C_1 is congruent (via $Q^{(1)}$) to a block diagonal matrix $C_1^{(1)}$ of $r^{(1)}$ blocks as in (2.27), while at the second iteration, each of the $r^{(1)}$ blocks is further decomposed (via $Q^{(2)}$) into many more finer blocks ($r^{(2)}$ blocks) as in (2.34). Simultaneously, the same congruence matric $Q^{(1)}Q^{(2)}$ makes C_3 into $C_3^{(2)}$ in (2.35), which is a non-homogeneous dilation of $C_1^{(2)}$ with scaling factors $\{\beta_1^{(2)}, \beta_2^{(2)}, \ldots, \beta_{r^{(2)}}^{(2)}\}$.

As for $C_2^{(1)}$ in (2.28), after the first iteration it has already become a non-homogeneous dilation of $C_1^{(1)}$ in (2.27) with scaling factors $\{\beta_1^{(1)}, \beta_2^{(1)}, \ldots, \beta_{r^{(1)}}^{(1)}\}$. Since $C_1^{(1)}$ continues to split into finer sub-blocks as in (2.34), $C_2^{(1)}$ will be synchronously decomposed, along with $C_1^{(1)}$, into a block diagonal matrix of $r^{(2)}$ blocks having the original scaling factors $\{\beta_1^{(1)}, \beta_2^{(1)}, \ldots, \beta_{r^{(1)}}^{(1)}\}$. Specifically, we can expand the scaling factors $\{\beta_1^{(1)}, \beta_2^{(1)}, \ldots, \beta_{r^{(1)}}^{(1)}\}$ to become a sequence of $r^{(2)}$ terms as follows:

$$\begin{aligned} & \{\underbrace{\beta_1^{(1)}, \beta_1^{(1)}, \ldots, \beta_1^{(1)}}_{l_1}; \underbrace{\beta_2^{(1)}, \beta_2^{(1)}, \ldots, \beta_2^{(1)}}_{l_2}; \cdots; \underbrace{\beta_{r^{(1)}}^{(1)}, \beta_{r^{(1)}}^{(1)}, \ldots, \beta_{r^{(1)}}^{(1)}}_{l_{r^{(1)}}}\} \\ \triangleq\ & \{[\beta_1^{(1)}], [\beta_2^{(1)}], \ldots, [\beta_{l_1}^{(1)}]; [\beta_{l_1+1}^{(1)}], \ldots, [\beta_{l_1+l_2}^{(1)}]; \ldots; [\beta_{\sum_{k=1}^{r^{(1)}-1} l_k+1}^{(1)}], \ldots, [\beta_{r^{(2)}}^{(1)}]\}. \end{aligned} \tag{2.36}$$

With this notation, we can express

$$C_2^{(2)} = Q^{(2)^T} C_2^{(1)} Q^{(2)} = \mathtt{diag}([\beta_1^{(1)}]A_1^{(2)}, [\beta_2^{(1)}]A_2^{(2)}, \dots, [\beta_{r^{(2)}}^{(1)}]A_{r^{(2)}}^{(1)}). \tag{2.37}$$

For $C_4^{(1)}$ up to $C_m^{(1)}$, let us take $C_4^{(1)}$ for example because all the others $C_5^{(1)}, C_6^{(1)}, \dots, C_m^{(1)}$ can be analogously taken care of. By the assumption that $C_4C_1^{-1}C_3$ is symmetric,we also have that

$$C_4^{(1)} C_1^{(1)^{-1}} C_3^{(1)} = \mathtt{diag}\left(C_{41}^{(1)} A_1^{(1)^{-1}} C_{31}^{(1)}, \dots, C_{4r^{(1)}}^{(1)} A_{r^{(1)}}^{(1)^{-1}} C_{3r^{(1)}}^{(1)}\right) \tag{2.38}$$

is symmetric. Since, for each $t = 1, 2, \dots, r^{(1)}$, $A_t^{(1)^{-1}} C_{3t}^{(1)}$ is similarly diagonalizable by $Q_t^{(2)}$; and $C_{4t}^{(1)} A_t^{(1)^{-1}} C_{3t}^{(1)}$ is symmetric, by Lemma 2.2.2, $C_{4t}^{(1)}$ can be further decomposed into finer blocks to become

$$Q_t^{(2)^T} C_{4t}^{(1)} Q_t^{(2)} = \mathtt{diag}\underbrace{(C_{4,t1}^{(2)}, C_{4,t2}^{(2)}, \dots, C_{4,tl_t}^{(2)})}_{sym.}. \tag{2.39}$$

Under the re-indexing formula (2.33) and (2.36), we have

$$C_4^{(2)} = Q^{(2)^T} C_4^{(1)} Q^{(2)} = \mathtt{diag}(C_{41}^{(2)}, C_{42}^{(2)}, \dots, C_{4r^{(2)}}^{(2)}). \tag{2.40}$$

Similarly, we have

$$C_j^{(2)} = Q^{(2)^T} C_j^{(1)} Q^{(2)} = \mathtt{diag}(C_{j1}^{(2)}, C_{j2}^{(2)}, \dots, C_{jr^{(2)}}^{(2)}); j = 5, 6, \dots, m. \tag{2.41}$$

As the process continues, at the third iteration we use the condition that $C_1^{-1}C_4$ is diagonalizable and $C_jC_1^{-1}C_4$, $5 \le j \le m$ symmetric to ensure the existence of a congruence $Q^{(3)}$, which puts $\{C_2^{(2)}, C_3^{(2)}, C_4^{(2)}\}$ as non-homogeneous dilation of the first matrix $C_1^{(2)}$, whereas from $C_5^{(2)}$ up to the last $C_m^{(2)}$ are all block diagonal matrices with the same pattern as the first matrix $C_1^{(2)}$. At the final iteration, there is a congruence matrix $Q^{(m-1)}$ that puts $\{C_2^{(m-1)}, C_3^{(m-1)}, \dots, C_m^{(m-1)}\}$ as non-homogeneous dilation of $C_1^{(m-1)}$. Define

$$R = Q^{(1)}Q^{(2)}Q^{(3)} \cdots Q^{(m-1)}.$$

Then, the nonsingular congruence matrix R transforms the collection $\{R^TC_iR : i = 1, 2, \dots, m\}$ into block diagonal forms of (2.26). By Remark 2.2.1, the collection $\mathcal{C}_{ns} = \{C_1, C_2, \dots, C_m\}$, $m \ge 3$ is $\mathbb{R}$-SDC and the proof is complete. □

With $(N_1), (N_2)$ and Theorem 2.2.1, we can now completely characterize the $\mathbb{R}$-SDC of a nonsingular collection $\mathcal{C}_{ns} = \{C_1, C_2, \dots, C_m\}$.

Theorem 2.2.2. *Let $\mathcal{C}_{ns} = \{C_1, C_2, \dots, C_m\} \subset \mathcal{S}^n$, $m \ge 3$ be a nonsingular collection with C_1 invertible. The collection $\mathcal{C}_{ns}$ is $\mathbb{R}$-SDC if and only if for each $2 \le i \le m$, the matrix $C_1^{-1}C_i$ is real similarly diagonalizable and $C_jC_1^{-1}C_i, 2 \le i < j \le m$ are all symmetric.*

2.2.2 Algorithm for the nonsingular collection

Return to (2.19), (2.20) and (2.21), in Lemma 2.2.2, where each C_i is decomposed into block diagonal form. Let us call *column* t to be the family of submatrices $\{C_{it}|i = 3, 4, \ldots, m\}$ of the t^{th} block. If each C_{it} in the family satisfies

$$C_{it} = \alpha_t^i A_t, \text{ for some } \alpha_t^i \in \mathbb{R}, i = 3, 4, \ldots, m, \tag{2.42}$$

we say that (2.42) holds for column t. Since A_t are symmetric, there are orthogonal matrices U_t such that $(U_t)^T A_t U_t$ are diagonal. Therefore, if (2.42) holds for all columns $t, t = 1, 2, \ldots, r$, the given matrices $C_1, C_2, \ldots, C_m$ are $\mathbb{R}$-SDC with the congruence matrix $P = Q \cdot \texttt{diag}(U_1, U_2, \ldots, U_r)$. Note that (2.42) always holds for column t with $m_t = 1$.

From the proof of Theorem 2.2.1, we indeed apply repeatedly Lemma 2.2.2 for nonsingular pairs. That idea suggests us to propose an algorithm for finding R as follows.

The procedure A below decompose the matrices into block diagonals.

Procedure A:

Step 1. Find a matrix R for $C_1, C_2, \ldots, C_m$ (by Lemma 2.2.2) such that

$$\begin{aligned} R^T C_1 R =& \texttt{diag}(C_{11}, C_{12}, \ldots, C_{1r}), \\ R^T C_2 R =& \texttt{diag}(\alpha_1^2 C_{11}, \alpha_2^2 C_{12}, \ldots, \alpha_r^2 C_{1r}), \\ R^T C_i R =& \texttt{diag}(C_{i1}, C_{i2}, \ldots, C_{ir}), 3 \le i \le m, \end{aligned}$$

If (2.42) holds for all columns t, $t = 1, 2, \ldots, r$, return R and stop. Else, set $j := 3$ and go to Step 2.

Step 2. While $j < m$ **do**

For $t = 1$ **to** r **do**

If (2.42) does not hold for column t, apply Lemma 2.2.2 for $C_{1t}, C_{jt}, \ldots, C_{mt}$ to find Q_t :

$$\begin{aligned} (Q_t)^T C_{1t} Q_t =& \texttt{diag}(C_{1t}^{(1)}, C_{1t}^{(2)}, \ldots, C_{1t}^{(l_t)}), \\ (Q_t)^T C_{jt} Q_t =& \texttt{diag}(\alpha_{t1}^j C_{1t}^{(1)}, \alpha_{t2}^j C_{1t}^{(2)}, \ldots, \alpha_{tl_t}^j C_{1t}^{(l_t)}), \\ (Q_t)^T C_{it} Q_t =& \texttt{diag}(C_{it}^{(1)}, C_{it}^{(2)}, \ldots, C_{it}^{(l_t)}), \;\; i = j+1, \ldots, m. \end{aligned}$$

Else set $Q_t := I_{m_t}$ and $l_t = 1$, here $m_t \times m_t$ is the size of C_{1t}.

EndFor

Update $R := R \cdot \texttt{diag}(Q_1, \ldots, Q_r)$.

- Reset the number of blocks: $r := l_1 + l_2 + \ldots + l_r$,
- Reset the blocks (use auxiliary variables if necessary)

$$C_{11} := C_{11}^{(1)}, \ldots, C_{1l_1} := C_{11}^{(l_1)}, C_{1(l_1+1)} := C_{12}^{(1)}, \ldots, C_{1r} := C_{1r}^{(l_r)},$$
$$C_{i1} := C_{i1}^{(1)}, \ldots, C_{il_1} := C_{i1}^{(l_1)}, C_{i(l_1+1)} := C_{i2}^{(1)}, \ldots, C_{ir} := C_{ir}^{(l_r)}, i = j+1, \ldots, m.$$

If (2.42) holds for all columns t, $t = 1, 2, \ldots, r$, return R and Stop.

Else, $j := j + 1$.

EndWhile

To see how the algorithm works, we consider the following example where the matrices given satisfy Theorem 2.2.1.

Example 2.2.1. We consider the following four 5×5 real symmetric matrices:

$$C_1 = \begin{pmatrix} 2 & 4 & -6 & -8 & -14 \\ 4 & 10 & -14 & -20 & -38 \\ -6 & -14 & 22 & 22 & 18 \\ -8 & -20 & 22 & 60 & 186 \\ -14 & -38 & 18 & 186 & 761 \end{pmatrix}, C_2 = \begin{pmatrix} 5 & 10 & -15 & -20 & -35 \\ 10 & 25 & -35 & -50 & -95 \\ -15 & -35 & 55 & 55 & 45 \\ -20 & -50 & 55 & 150 & 465 \\ -35 & -95 & 45 & 465 & 1900 \end{pmatrix}$$

$$C_3 = \begin{pmatrix} -1 & -2 & 3 & 4 & 7 \\ -2 & -5 & 7 & 10 & 19 \\ 3 & 7 & -11 & -11 & -9 \\ 4 & 10 & -11 & -25 & -73 \\ 7 & 19 & -9 & -73 & -295 \end{pmatrix}, C_4 = \begin{pmatrix} 1 & 2 & -3 & -4 & -7 \\ 2 & 5 & -7 & -10 & -19 \\ -3 & -7 & 17 & -7 & -93 \\ -4 & -10 & -7 & 83 & 395 \\ -7 & -19 & -93 & 395 & 2104 \end{pmatrix}.$$

Step 1. Apply Lemma 2.2.2 we have $R = \begin{pmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 0 & 5 \\ 1 & 0 & 0 & 0 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ such that

$$R^T C_1 R := \begin{pmatrix} 60 & 22 & -20 & -8 & 0 \\ 22 & 22 & -14 & -6 & 0 \\ -20 & -14 & 10 & 4 & 0 \\ -8 & -6 & 4 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} := \texttt{diag}(C_{11}, C_{12});$$

$$R^T C_2 R := \begin{pmatrix} 150 & 55 & -50 & -20 & 0 \\ 55 & 55 & -35 & -15 & 0 \\ -50 & -35 & 25 & 10 & 0 \\ -20 & -15 & 10 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} = \texttt{diag}\left(\frac{5}{2}C_{11}, \frac{5}{3}C_{12}\right),$$

where $C_{11} := \begin{pmatrix} 60 & 22 & -20 & -8 \\ 22 & 22 & -14 & -6 \\ -20 & -14 & 10 & 4 \\ -8 & -6 & 4 & 2 \end{pmatrix}; C_{12} := (3);$

$$R^T C_3 R := \begin{pmatrix} -25 & -11 & 10 & 4 & 0 \\ -11 & -11 & 7 & 3 & 0 \\ 10 & 7 & -5 & -2 & 0 \\ 4 & 3 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix} := \texttt{diag}(C_{31}, C_{32}),$$

where $C_{31} := \begin{pmatrix} -25 & -11 & 10 & 4 \\ -11 & -11 & 7 & 3 \\ 10 & 7 & -5 & -2 \\ 4 & 3 & -2 & -1 \end{pmatrix}, C_{32} := (4);$

$$R^T C_4 R := \begin{pmatrix} 83 & -7 & -10 & -4 & 0 \\ -7 & 17 & -7 & -3 & 0 \\ -10 & -7 & 5 & 2 & 0 \\ -4 & -3 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix} := \texttt{diag}(C_{41}, C_{42}),$$

where $C_{41} := \begin{pmatrix} 83 & -7 & -10 & -4 \\ -7 & 17 & -7 & -3 \\ -10 & -7 & 5 & 2 \\ -4 & -3 & 2 & 1 \end{pmatrix}; C_{42} := (7).$

Observe that (2.42) does not hold for column 1 which involves the sub-matrices C_{11}, C_{31}, C_{41}, (note that at this iteration we have only two columns: $r = 2$) we set $j := 3$ and go to Step 2.

Step 2. For $t = 1$ to 2 do

- $t = 1$: (2.42) does not hold for column 1, we apply Lemma 2.2.2 for column 1 including matrices C_{11}, C_{31}, C_{41} as follows. Find $Q_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 5 \\ 1 & 0 & 0 & 3 \end{pmatrix}$ such that

$$(Q_1)^T C_{11} Q_1 = \begin{pmatrix} 2 & 4 & -6 & 0 \\ 4 & 10 & -14 & 0 \\ -6 & -14 & 22 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} := \mathtt{diag}(C_{11}^{(1)}, C_{11}^{(2)})$$

$$(Q_1)^T C_{31} Q_1 = \begin{pmatrix} -1 & -2 & 3 & 0 \\ -2 & -5 & 7 & 0 \\ 3 & -7 & -11 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} := \mathtt{diag}\left(-\frac{1}{2}C_{11}^{(1)}, 2C_{11}^{(2)}\right),$$

where $C_{11}^{(1)} = \begin{pmatrix} 2 & 4 & -6 \\ 4 & 10 & -14 \\ -6 & -14 & 22 \end{pmatrix}; C_{11}^{(2)} = (2);$

$$(Q_1)^T C_{41} Q_1 = \begin{pmatrix} 1 & 2 & -3 & 0 \\ 2 & 5 & -7 & 0 \\ -3 & -7 & 17 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} := \mathtt{diag}(C_{41}^{(1)}, C_{41}^{(2)})$$

where $C_{41}^{(1)} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ -3 & -7 & 17 \end{pmatrix}; C_{41}^{(2)} := (0).$

- $t = 2$: (2.42) holds for column 2, set $Q_2 = 1, l_2 = 1$.

  Update $R := R \cdot \mathtt{diag}(Q_1, Q_2)$, here $\mathtt{diag}(Q_1, Q_2) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 1 & 0 & 5 & 0 \\ 1 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$

  Reset the following:

  The number of blocks: $r = l_1 + l_2 = 2 + 1 = 3$,

  The blocks: Use auxiliary variables:

$$M_{11} := C_{11}^{(1)}, M_{12} := C_{11}^{(2)}, M_{13} := C_{12};$$
$$M_{41} := C_{41}^{(1)}, M_{42} := C_{41}^{(2)}, M_{43} := C_{42}.$$

Now, reset $C_{1t} := M_{1t}, C_{4t} := M_{4t},\ \ t = 1, 2, 3$. We have

$$C_{11} = \begin{pmatrix} 2 & 4 & -6 \\ 4 & 10 & -14 \\ -6 & -14 & 22 \end{pmatrix}, C_{12} = (2), C_{13} = (3)$$

and

$$C_{41} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 5 & -7 \\ -3 & -7 & 17 \end{pmatrix}, C_{42} = (0), C_{43} = (7).$$

Observe that (2.42) does not hold for column 1. We set $j := j+1 = 4$ and repeat Step 2.

For $t = 1$ to 3 do

• $t = 1$: (2.42) does not hold for column 1. We apply Lemma 2.2.2 for C_{11}, C_{41} as follows: Find $Q_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$ such that

$$(Q_1)^T C_{11} Q_1 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 4 & 10 \end{pmatrix} = \texttt{diag}(C_{11}^{(1)}, C_{11}^{(2)},$$

$$(Q_1)^T C_{41} Q_1 = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 5 \end{pmatrix} := \texttt{diag}\left(\frac{7}{2}C_{11}^{(1)}, \frac{1}{2}C_{11}^{(2)}\right)$$

where $C_{11}^{(1)} = (2), C_{11}^{(2)} = \begin{pmatrix} 2 & 4 \\ 4 & 10 \end{pmatrix}$.

• $t = 2, 3$: (2.42) holds for columns 2, 3, we set $Q_2 = 1, Q_3 = 1$.

At this iteration we already have $j = m$, so we return $R := R \cdot \texttt{diag}(Q_1, Q_2, Q_3)$ $= \begin{pmatrix} 1 & 1 & 0 & 3 & 2 \\ 1 & 0 & 1 & 5 & 2 \\ 1 & 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$. It is not difficult to check that R is the desired matrix:

$$R^T C_1 R = \texttt{diag}(A_1, A_2, A_3, A_4),$$
$$R^T C_2 R = \texttt{diag}\left(\frac{5}{2}A_1, \frac{5}{2}A_2, \frac{5}{2}A_3, \frac{5}{3}A_4\right),$$

$$R^T C_3 R = \texttt{diag}\left(-\frac{1}{2}A_1, -\frac{1}{2}A_2, 2A_3, \frac{4}{3}A_4\right),$$
$$R^T C_4 R = \texttt{diag}\left(\frac{7}{2}A_1, \frac{1}{2}A_2, 0A_3, \frac{7}{3}A_4\right),$$

where $A_1 := (2); A_2 := \begin{pmatrix} 2 & 4 \\ 4 & 10 \end{pmatrix}; A_3 := (2); A_4 := (3).$

The algorithm for solving the SDC problem of a nonsingular collection $\mathcal{C}_{ns}$ is now stated as follows.

Algorithm 7 Solving the SDC problem for a nonsingular collection

---

INPUT: Real symmetric matrices $C_1, C_2, \ldots, C_m$; C_1 is nonsingular.

OUTPUT: NOT $\mathbb{R}$-SDC or a nonsingular real matrix P that simultaneously diagonalizes $C_1, C_2, \ldots, C_m$

Step 1. (Checking $\mathbb{R}$-SDC)

If $C_1^{-1}C_i$ is not real similarly diagonalizable for some i or $C_j C_1^{-1} C_i$ is not symmetric for some $i < j$ then NOT $\mathbb{R}$-SDC and STOP.

Else, go to Step 2.

Step 2.
- Apply **Procedure A** to find R, which satisfies (2.26);
- Let $U_t, t = 1, 2, \ldots, r$, be orthogonal matrices such that $U_t^T A_t U_t$ are diagonal, define $U = \texttt{diag}(U_1, U_2, \ldots, U_r)$.

Return $P = RU$.

Example 2.2.2. We consider again the three matrices given in Example 2.1.6. Recall that Algorithm 6 requires three steps: (1) finding X; (2) computing the square root Q of $X : Q^2 = X$; and (3) applying Algorithm 5 to the matrices QC_1Q, QC_2Q, QC_3Q to obtain a unitary matrix V and returning the congruence matrix $P = QV$. Here, Algorithm 7 requires only one step as follows. The matrix $C_1^{-1}C_2 = \begin{pmatrix} 0 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & -1 & -\frac{1}{2} \end{pmatrix}$ is real similarly diagonalizable by $P = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix}$. Since $C_1^{-1}C_2$ has three distinct eigenvalues, which are $0, -1, -\frac{1}{2}$, the matrices C_1, C_2, C_3 are $\mathbb{R}$-SDC via congruence P.

2.2.3 The SDC problem of singular collection

Let $\mathcal{C}_s = \{C_1, C_2, \ldots, C_m\} \subset \mathcal{S}^n$ be a singular collection in which every $C_i \neq 0$ is singular. Consider the first two matrices C_1, C_2. If they are not $\mathbb{R}$-SDC so is not $\mathcal{C}_s$. Otherwise, by Lemmas 1.2.8, Theorem 1.2.1 and Lemma 1.2.9, there is a nonsingular U_1 that converts C_1, C_2 to block diagonal matrices

$$\tilde{C}_1 = \texttt{diag}(\underbrace{(C_{11})_p}_{invert.\ \&\ diag.}, 0_{n-p});\ \tilde{C}_2 = \texttt{diag}((C_{21})_p, \underbrace{(C_{26})_{s_1}}_{invert.\ \&\ diag.}, 0_{n-p-s_1}) \tag{2.43}$$

where C_{11} and C_{26} are both nonsingular diagonal, $p > 0$, $s_1 \geq 0$; and 0_{n-p} denotes the zero matrix of size $(n-p) \times (n-p)$. We emphasize that $s_1 = 0$ corresponds to (1.11) in Lemma 1.2.8, while $s_1 > 0$ to (1.12) in Lemma 1.2.8. Also by Theorem 1.2.1 and Lemma 1.2.9, the $\mathbb{R}$-SDC of $\{C_1, C_2\}$ implies the $\mathbb{R}$-SDC of $\{(C_{11})_p, (C_{21})_p\}$, the latter of which is a nonsingular collection of smaller matrix size $p < n$.

Suppose $\{C_{11}, C_{21}\}$ are $\mathbb{R}$-SDC, say, by $(W)_p$. Let $Q_1 = \texttt{diag}((W)_p, I_{n-p})$, where I_{n-p} is the identity matrix of dimension $n-p$. Then,

$$\begin{aligned}
\tilde{C}_1' = Q_1^T \tilde{C}_1 Q_1 &= \texttt{diag}(\underbrace{\underbrace{(W^T C_{11} W)_p}_{\triangleq \tilde{C}_{11}':\ invert.\ \&\ diag.}, \underbrace{0_{s_1}}_{s_1 \geq 0}}_{\triangleq \hat{C}_{11}}, 0_{n-p-s_1});\\
\tilde{C}_2' = Q_1^T \tilde{C}_2 Q_1 &= \texttt{diag}(\underbrace{(W^T C_{21} W)_p}_{\triangleq \tilde{C}_{21}':\ diag.}, \underbrace{(C_{26})_{s_1}}_{\triangleq \tilde{C}_{26}':invert.\ \&\ diag.}, 0_{n-p-s_1}).
\end{aligned}$$

It allows us to choose a large enough μ_1 such that $\mu_1 \tilde{C}_{11}' + \tilde{C}_{21}'$ is invertible (where $\tilde{C}_{21}' = W^T C_{21} W$). Then,

$$\begin{aligned}
\mu_1 \tilde{C}_1' + \tilde{C}_2' &= Q_1^T(\mu_1 \tilde{C}_1 + \tilde{C}_2) Q_1\\
&= \texttt{diag}(\underbrace{\underbrace{(\mu_1 \tilde{C}_{11}' + \tilde{C}_{21}')_p}_{invert.\ \&\ diag.}, \underbrace{(\tilde{C}_{26}')_{s_1}}_{invert.\ \&\ diag.}}_{\triangleq \hat{C}_{21}:invert.\ \&\ diag.}, 0_{n-p-s_1}).
\end{aligned}$$

Now include C_3 for determining the $\mathbb{R}$-SDC of $\{C_1, C_2, C_3\}$. We first transform C_3 by U_1, followed by Q_1, to obtain $\tilde{C}_3' = Q_1^T(U_1^T C_3 U_1) Q_1$. The idea is to apply Lemma 1.2.8 again to convert $\mu_1 \tilde{C}_1' + \tilde{C}_2'$ and $\tilde{C}_3'$ into the form (2.43), where, with the help of a sufficiently large $\mu_1 > 0$, the subblock $(\hat{C}_{21})_{p+s_1}$ in $\mu_1 \tilde{C}_1' + \tilde{C}_2'$ is nonsingular and diagonal and thus can be used to determine the $\mathbb{R}$-SDC of $\{\mu_1 \tilde{C}_1' + \tilde{C}_2', \tilde{C}_3'\}$. The entire Section is devoted to proving that the idea does indeed work. The main result, Theorem 2.2.3, states that, suppose that the first $m-1$ matrices are $\mathbb{R}$-SDC (otherwise, it is end of the story), there always exist a sequence of congruences matrices and a sequence of large enough constants which can reduce the $\mathbb{R}$-SDC of the *entire* singular collection $\mathcal{C}_s$ to become the $\mathbb{R}$-SDC of another nonsingular collection $\mathcal{C}_{ns}$ having a smaller matrix size.

Suppose C_1, C_2 are $\mathbb{R}$-SDC and we now include C_3 to determine the $\mathbb{R}$-SDC of $\{C_1, C_2, C_3\}$. By Theorem 1.2.1 and Lemma 1.2.9, there is a U_1 that converts C_1, C_2 to block diagonal matrices $\tilde{C}_1 = \texttt{diag}((C_{11})_p, 0_{n-p})$ and $\tilde{C}_2 = \texttt{diag}((C_{21})_p, (C_{26})_{s_1}, 0_{n-p-s_1})$ where C_{11} and C_{26} are both nonsingular diagonal, but $s_1 \geq 0$ could be 0. Moreover, $\mathbb{R}$-SDC of C_1, C_2 implies that C_{11}, C_{21} are $\mathbb{R}$-SDC, say, by the congruence $(W)_p$. Let $Q_1 = \texttt{diag}((W)_p, I_{n-p})$. Then,

$$\tilde{C}_1' = Q_1^T \tilde{C}_1 Q_1 = \texttt{diag}(\underbrace{\underbrace{(W^T C_{11} W)_p}_{\triangleq \tilde{C}_{11}':\ invert.\ \&\ diag.}, \underbrace{0_{s_1}}_{s_1 \geq 0}}_{\triangleq \hat{C}_{11}}, 0_{n-p-s_1}); \tag{2.44}$$

$$\tilde{C}_2' = Q_1^T \tilde{C}_2 Q_1 = \texttt{diag}(\underbrace{(W^T C_{21} W)_p}_{\triangleq \tilde{C}_{21}':\ diag.}, \underbrace{(C_{26})_{s_1}}_{\triangleq \tilde{C}_{26}':invert.\ \&\ diag.}, 0_{n-p-s_1}). \tag{2.45}$$

Synchronically, C_3 is first transformed to $\tilde{C}_3$ by U_1, followed by another transformation by Q_1 to become

$$\tilde{C}_3' = Q_1^T \underbrace{U_1^T C_3 U_1}_{\tilde{C}_3} Q_1 = \begin{pmatrix} \underbrace{(M_{31})_{p+s_1}}_{sym.,\ s_1 \geq 0} & M_{32} \\ M_{32}^T & \underbrace{(M_{33})_{n-p-s_1}}_{sym.} \end{pmatrix} \tag{2.46}$$

Note that, in (2.44), $\tilde{C}_{11}' = W^T C_{11} W$ is invertible due to C_{11} being invertible and $\operatorname{rank}(C_{11}) = \operatorname{rank}(\tilde{C}_{11}')$. It allows us to choose a large enough μ_1 such that $\mu_1 \tilde{C}_{11}' + \tilde{C}_{21}'$ is invertible (where $\tilde{C}_{21}' = W^T C_{21} W$). Then,

$$\begin{aligned} \mu_1 \tilde{C}_1' + \tilde{C}_2' &= Q_1^T (\mu_1 \tilde{C}_1 + \tilde{C}_2) Q_1 \\ &= \texttt{diag}(\underbrace{\underbrace{(\mu \tilde{C}_{11}' + \tilde{C}_{21}')_p}_{invert.\ \&\ diag.}, \underbrace{(\tilde{C}_{26}')_{s_1}}_{invert.\ \&\ diag.}}_{\triangleq \hat{C}_{21}:invert.\ \&\ diag.}, 0_{n-p-s_1}). \end{aligned} \tag{2.47}$$

Next, we are going to convert the pair $\mu_1 \tilde{C}_1' + \tilde{C}_2'$ and $\tilde{C}_3'$ into the form (1.10) and (1.11); or the form of (1.10) and (1.12) in Lemma 1.2.8, respectively. Notice that $\mu_1 \tilde{C}_1' + \tilde{C}_2' = \texttt{diag}((\hat{C}_{21})_{p+s_1}, 0_{n-p-s_1})$ is already in the form of (1.10).

- If, in (2.46), $M_{33} = 0$, $\tilde{C}_3'$ is thus in the form of (1.11). Let us rename

$$\hat{C}_1 = \tilde{C}_1' \text{ (in (2.44))};\ \hat{C}_2 = \mu_1 \tilde{C}_1' + \tilde{C}_2' \text{ (in (2.47))};\ \hat{C}_3 = \tilde{C}_3' = \begin{pmatrix} M_{31} & M_{32} \\ M_{32}^T & 0 \end{pmatrix} \tag{2.48}$$

and denote their north-west subblocks as in (2.44) and in (2.47)

$$\hat{C}_{11} = \mathtt{diag}((\tilde{C}'_{11})_p, 0_{s_1});\ \hat{C}_{21} = \mathtt{diag}(\underbrace{(\mu\tilde{C}'_{11} + \tilde{C}'_{21})_p, (\tilde{C}'_{26})_{s_1}}_{invert.\ \&\ diag.});\ \hat{C}_{31} = (M_{31})_{p+s_1}. \quad (2.49)$$

It is easy to see the following result.

Lemma 2.2.3. *Let $\{\hat{C}_1, \hat{C}_2, \hat{C}_3\}$ be singular matrices of the form (2.48). Then, $\{\hat{C}_1, \hat{C}_2, \hat{C}_3\}$ are $\mathbb{R}$-SDC if and only if the north-western sub-blocks of them,$\{\hat{C}_{11}, \hat{C}_{21}, \hat{C}_{31}\}$, as specified by (2.49) are $\mathbb{R}$-SDC; and $M_{32} = 0$.*

Proof. If $M_{32} = 0$ and the northwest sub-blocks $\hat{C}_{11}, \hat{C}_{21}, \hat{C}_{31}$ in (2.49) are $\mathbb{R}$-SDC by $(L_1)_{p+s_1}$, then the matrix $L = \mathtt{diag}(L_1, I_{n-p-s_1})$ simultaneously diagonalizes $\hat{C}_1, \hat{C}_2, \hat{C}_3$ via congruence.

Conversely, suppose $\hat{C}_1, \hat{C}_2, \hat{C}_3$ are $\mathbb{R}$-SDC. In particular, $\hat{C}_2, \hat{C}_3$ are $\mathbb{R}$-SDC. Since $\hat{C}_{21}$ is nonsingular and diagonal whereas $\hat{C}_3$ is in the form of (1.11), by Theorem 1.2.1, M_{32} must be 0. It implies that $\hat{C}_1, \hat{C}_2, \hat{C}_3$ have the same block structure. Specifically, $\hat{C}_1 = \mathtt{diag}((\hat{C}_{11})_{p+s_1}, 0)$, $\hat{C}_2 = \mathtt{diag}((\hat{C}_{21})_{p+s_1}, 0)$, and $\hat{C}_3 = \mathtt{diag}((M_{31})_{p+s_1}, 0)$. By Lemma 1.1.6, $\{\hat{C}_{11}, \hat{C}_{21}, \hat{C}_{31}\}$ are $\mathbb{R}$-SDC and the proof is complete. □

- Suppose, in (2.46), $M_{33} \neq 0$. Let an orthogonal $(P_2)_{n-p-s_1}$ be such that

$$P_2^T M_{33} P_2 = \mathtt{diag}(\underbrace{(C_{36})_{s_2}}_{invert.\ \&\ diag.,\ s_2>0}, 0_{n-p-s_1-s_2}),$$

with which we can form $H_2 = \mathtt{diag}(I_{p+s_1}, P_2)$ and compute

$$H_2^T \tilde{C}'_3 H_2 = \begin{pmatrix} (M_{31})_{p+s_1} & C_{34} & C_{35} \\ C_{34}^T & (C_{36})_{s_2} & 0 \\ C_{35}^T & 0 & 0_{n-p-s_1-s_2} \end{pmatrix}, \quad (2.50)$$

where $(C_{34}, C_{35})_{p\times(n-p)} = M_{32}P_2$. Define further that

$$V_2 = \begin{pmatrix} I_{p+s_1} & 0 & 0 \\ -C_{36}^{-1}C_{34}^T & I_{s_2} & 0 \\ 0 & 0 & I_{n-p-s_1-s_2} \end{pmatrix}, \text{ and } U_2 = H_2V_2 \quad (2.51)$$

so that

$$\breve{C}_3 \triangleq U_2^T \tilde{C}'_3 U_2 = \begin{pmatrix} \underbrace{M_{31} - C_{34}C_{36}^{-1}C_{34}^T}_{\triangleq(\breve{C}_{31})_{p+s_1},\ sym.} & 0 & C_{35} \\ 0 & \underbrace{(C_{36})_{s_2}}_{invert.\ \&\ diag.} & 0 \\ C_{35}^T & 0 & 0_{n-p-s_1-s_2} \end{pmatrix}. \quad (2.52)$$

More importantly, the transformation U_2,

$$U_2 = H_2V_2 = \begin{pmatrix} I_{p+s_1} & 0 \\ -P_2 \begin{bmatrix} C_{36}^{-1}C_{34}^T \\ 0 \end{bmatrix} & (P_2)_{n-p-s_1} \end{pmatrix}, \tag{2.53}$$

does not change $\tilde{C}_1'$ in (2.44) and $\mu_1\tilde{C}_1' + \tilde{C}_2'$ in (2.47), in the sense that

$$\breve{C}_1 \triangleq U_2^T\tilde{C}_1'U_2 = \tilde{C}_1' = \texttt{diag}(\underbrace{(W^TC_{11}W)_p, 0_{s_1}}_{\triangleq\breve{C}_{11}=\hat{C}_{11}:\ diagonal}, 0_{n-p-s_1}) \tag{2.54}$$

$$\begin{aligned} \breve{C}_2 &= U_2^T(\mu_1\tilde{C}_1' + \tilde{C}_2')U_2 = \mu_1\tilde{C}_1' + \tilde{C}_2' \\ &= \texttt{diag}(\underbrace{(\mu_1\tilde{C}_{11}' + \tilde{C}_{21}')_p, (\tilde{C}_{26}')_{s_1}}_{\triangleq\breve{C}_{21}=\hat{C}_{21}:\ invert.\ \&\ diag.}, 0_{n-p-s_1}). \end{aligned} \tag{2.55}$$

Notice that, in (2.54) and (2.55), $\hat{C}_{11}$ is renamed as $\breve{C}_{11}$, while $\hat{C}_{21}$ becomes $\breve{C}_{21}$. Then, we have the following main result.

Lemma 2.2.4. *The singular collection $\{\breve{C}_1, \breve{C}_2, \breve{C}_3\}$ in (2.54), (2.55), in (2.52) are $\mathbb{R}$-SDC if and only if the north-western sub-blocks of them, i.e. $\{\breve{C}_{11}, \breve{C}_{21}, \breve{C}_{31}\}$, are $\mathbb{R}$-SDC; and C_{35} in (2.52) is a zero matrix or does not exist.*

Proof. The sufficiency of Lemma 2.2.4 is easy. If C_{35} in (2.52) is a zero matrix or does not exist, and if the northwest sub-blocks $\breve{C}_{11}, \breve{C}_{21}, \breve{C}_{31}$ in (2.54), (2.55) and (2.52) are $\mathbb{R}$-SDC by $(L_1)_{p+s_1}$, then the matrix $L = \texttt{diag}(L_1, I_{s_2}, I_{n-p-p_2-s_2})$ simultaneously diagonalizes $\{\breve{C}_1, \breve{C}_2, \breve{C}_3\}$ via congruence.

To prove the necessity, suppose that $\breve{C}_1, \breve{C}_2, \breve{C}_3$ are $\mathbb{R}$-SDC by a congruence matrix Q. In particular, $\breve{C}_2, \breve{C}_3$ are $\mathbb{R}$-SDC in which

$$\breve{C}_{21} = \texttt{diag}(\underbrace{(\mu\tilde{C}_{11}' + \tilde{C}_{21}')_p, (\tilde{C}_{26}')_{s_1}}_{invert.\ \&\ diag.})\ (\text{in } \breve{C}_2), \quad \underbrace{(C_{36})_{s_2}}_{invert.\ \&\ diag.}\ (\text{in } \breve{C}_3)$$

are nonsingular diagonal. By Lemma 1.2.9, two matrices (here they are $\breve{C}_2, \breve{C}_3$) in the form of (1.10) and (1.12) are $\mathbb{R}$-SDC, there must be $C_{35} = 0$ in $\breve{C}_3$ (2.52) or C_{35} does not exist. Let us assume that $C_{35} = 0$. Then,

$$\begin{aligned} \breve{C}_2 &= \texttt{diag}(\underbrace{(\breve{C}_{21})_{p+s_1}}_{invert.\ \&\ diag.}, 0_{s_2}, 0_{n-p-s_1-s_2}) \\ \breve{C}_3 &= \texttt{diag}((\breve{C}_{31})_{p+s_1}, \underbrace{(C_{36})_{s_2}}_{invert.\ \&\ diag.}, 0_{n-p-s_1-s_2}) \end{aligned}$$

where $\breve{C}_{31} = M_{31} - C_{34}C_{36}^{-1}C_{34}^T$ has been defined in (2.52).

By Lemma 1.1.6, two matrices (which are $\breve{C}_2, \breve{C}_3$ with $p = p + s_1 + s_2$,) of form (1.1) are $\mathbb{R}$-SDC, the congruence Q that diagonalizes $\breve{C}_2, \breve{C}_3$ can be chosen to be

$$Q = \begin{pmatrix} (Q_1)_{p+s_1} & (Q_2)_{(p+s_1)\times s_2} & 0 \\ (Q_3)_{s_2\times(p+s_1)} & (Q_4)_{s_2} & 0 \\ 0 & 0 & I_{n-p-s_1-s_2} \end{pmatrix} \tag{2.56}$$

such that the first $p + s_1$ diagonal entries of the diagonal matrix $Q^T\breve{C}_2Q$ are all non-zero. We shall show that $Q_2 = 0_{(p+s_1)\times s_2}$ and $Q_3 = 0_{s_2\times(p+s_1)}$ so that $Q = \texttt{diag}\left((Q_1)_{p+s_1}, (Q_4)_{s_2}, I_{n-p-s_1-s_2}\right)$.

By Q in (2.56), $\breve{C}_2$ is congruent to the diagonal matrix

$$Q^T\breve{C}_2Q = \begin{pmatrix} (Q_1^T\breve{C}_{21}Q_1)_{p+s_1} & Q_1^T\breve{C}_{21}Q_2 & 0 \\ Q_2^T\breve{C}_{21}Q_1 & (Q_2^T\breve{C}_{21}Q_2)_{s_2} & 0 \\ 0 & 0 & 0_{n-p-s_1-s_2} \end{pmatrix}$$

in which $Q_1^T\breve{C}_{21}Q_1$ is nonsingular diagonal. Since $\breve{C}_{21}$ is also nonsingular, it implies that Q_1 must be nonsingular. Then, due to the off-diagonal block $Q_1^T\hat{C}_{21}Q_2 = 0$, we see that $Q_2 = 0$. Then,

$$Q^T\breve{C}_2Q = \texttt{diag}(\underbrace{(Q_1^T\breve{C}_{21}Q_1)_{p+s_1}}_{invert.\ \&\ diag.}, 0_{s_2}, 0_{n-p-s_1-s_2}). \tag{2.57}$$

Since $\breve{C}_1$ in (2.54) and $\breve{C}_2$ in (2.55) adopt the same block structure, there also is

$$Q^T\breve{C}_1Q = \texttt{diag}(\underbrace{(Q_1^T\breve{C}_{11}Q_1)_{p+s_1}}_{diag.}, 0_{s_2}, 0_{n-p-s_1-s_2}). \tag{2.58}$$

The same congruence Q also diagonalizes $\hat{C}_3$. Since $Q_2 = 0$ in (2.56),

$$Q^T\breve{C}_3Q = \begin{pmatrix} Q_1^T\breve{C}_{31}Q_1 + Q_3^TC_{36}Q_3 & Q_3^TC_{36}Q_4 & 0 \\ Q_4^TC_{36}Q_3 & Q_4^TC_{36}Q_4 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ is diagonal} \tag{2.59}$$

so that $Q_4^TC_{36}Q_3 = 0$. From (2.56), since Q is nonsingular and we have known that $Q_2 = 0$, there must be Q_4 nonsingular. From (2.52), we also know that C_{36} is nonsingular. Then, $Q_4^TC_{36}Q_3 = 0$ implies that $Q_3 = 0$, which proves that the congruence $Q = \texttt{diag}\left((Q_1)_{p+s_1}, (Q_4)_{s_2}, I_{n-p-s_1-s_2}\right)$ and (2.63) becomes

$$Q^T\breve{C}_3Q = \texttt{diag}(\underbrace{(Q_1^T\breve{C}_{31}Q_1)_{p+s_1}}_{diag.}, \underbrace{(Q_4^TC_{36}Q_4)_{s_2}}_{diag.}, 0_{n-p-s_1-s_2}). \tag{2.60}$$

Combining (2.58),(2.57),(2.60), we see that if $\breve{C}_1, \breve{C}_2, \breve{C}_3$ are $\mathbb{R}$-SDC by Q, then the north-western blocks $\breve{C}_{11}, \breve{C}_{21}, \breve{C}_{31}$ are $\mathbb{R}$-SDC by Q_1. The proof is complete. $\square$

In summary, when C_1 and C_2 are $\mathbb{R}$-SDC, there is

$$\breve{C}_1 = U_2^T \overbrace{Q_1^T \quad \underbrace{U_1^T(C_1)U_1}_{=\tilde{C}_1 \text{in Lemma 1.2.8}} \quad Q_1}^{=\tilde{C}_1' \text{ in (2.44)}} U_2$$

where U_1 is from Lemma 1.2.8 that puts C_1, C_2 in the form of $\tilde{C}_1 = U_1^T C_1 U_1$ (1.10) and $\tilde{C}_2 = U_1^T C_2 U_1$ (1.12); while Q_1 from (2.44)-(2.45) diagonalizes simultaneously $\tilde{C}_1$ and $\tilde{C}_2$; finally U_2 from (2.53) puts $\tilde{C}_3' = Q_1^T U_1^T C_3 U_1 Q_1$ in the form of (2.52). In addition,

$$\breve{C}_2 = U_2^T Q_1^T U_1^T(\mu C_1 + C_2)U_1 Q_1 U_2; \ \breve{C}_3 = U_2^T Q_1^T U_1^T(C_3)U_1 Q_1 U_2.$$

It is obvious that $\breve{C}_1, \breve{C}_2, \breve{C}_3$ are $\mathbb{R}$-SDC if and only if $C_1, \mu C_1 + C_2, C_3$ are $\mathbb{R}$-SDC; and, if and only if C_1, C_2, C_3 are $\mathbb{R}$-SDC. Therefore, from Lemma 2.2.4, we have the following result.

Lemma 2.2.5. *Let $\{C_1, C_2, C_3\} \subset \mathcal{S}^n$ be a singular collection and assume that C_1, C_2 are $\mathbb{R}$-SDC. Then, there is a nonsingular U and a constant μ such that $\breve{C}_1 = U^T C_1 U$, $\breve{C}_2 = U^T(\mu C_1 + C_2)U$, $\breve{C}_3 = U^T C_3 U$ be singular matrices of the forms (2.54), (2.55) and (2.52), respectively. Moreover, the collection $\{C_1, C_2, C_3\}$ is $\mathbb{R}$-SDC if and only if the northwestern nonsingular subblocks of them, $\{\breve{C}_{11}, \breve{C}_{21}, \breve{C}_{31}\}$, are $\mathbb{R}$-SDC; and C_{35} in (2.52) is either zero or does not exist.*

Lemmas 2.2.3 and 2.2.5 can be easily extended to more than three matrices. Theorem 2.2.3 below can be proved by induction. Firstly, we need the following lemma.

Lemma 2.2.6. *Suppose $\tilde{C}_1, \tilde{C}_2, \dots, \tilde{C}_m$ are singular matrices of the forms:*

for $i = 1, 2, \dots, m-1$,

$$\tilde{C}_i = \boldsymbol{diag}((C_{i1})_p, 0_s, 0_{r-s}) = \boldsymbol{diag}((\hat{C}_{i1})_{p+s}, 0_{r-s}) \tag{2.61}$$

for $i = m$,

$$\tilde{C}_m = \boldsymbol{diag}((C_{m1})_p, (C_{m6})_s, 0_{r-s}) = \boldsymbol{diag}((\hat{C}_{m1})_{p+s}, 0_{r-s}) \tag{2.62}$$

where $p, r \geq 1, s \geq 0$; in (2.61), $C_{i1}, i = 1, 2, \dots, m-1$, are real diagonal of size $p \times p$, $(\hat{C}_{i1})_{p+s} = \boldsymbol{diag}((C_{i1})_p, 0_s)$, and $C_{(m-1)1}$ is nonsingular; $(\hat{C}_{m1})_{p+s} = \boldsymbol{diag}((C_{m1})_p, (C_{m6})_s)$ in (2.62), C_{m1} is real symmetric of size $p \times p$, C_{m6} is real nonsingular diagonal of size $s \times s$. Then $\tilde{C}_1, \tilde{C}_2, \dots, \tilde{C}_{m-1}, \tilde{C}_m$ are $\mathbb{R}$-SDC if and only if their north-west blocks $C_{11}, C_{21}, \dots, C_{(m-1)1}, C_{m1}$ are $\mathbb{R}$-SDC.

Proof. Suppose first that $C_{11}, C_{21}, \ldots, C_{m1}$ are $\mathbb{R}$-SDC by Q_1. We define the matrix Q as follows: if $r > s$ then $Q = \texttt{diag}(Q_1, I_s, I_{r-s})$; if $r = s$ then $Q = \texttt{diag}(Q_1, I_s)$. Then $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $\mathbb{R}$-SDC by Q.

For the converse, suppose that $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $\mathbb{R}$-SDC and $r > s$. The case $r = s$ is proved similarly. By Lemma 1.1.6, $(\hat{C}_{11})_{p+s}, (\hat{C}_{(m-1)1})_{p+s}, \ldots, (\hat{C}_{m1})_{p+s}$ are $\mathbb{R}$-SDC by

$$Q = \begin{pmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{pmatrix},$$

where Q_1 and Q_4 are square matrices of size $p \times p$ and $s \times s$, respectively, such that the p nonzero elements of the diagonal matrix $Q^T\hat{C}_{(m-1)1}Q$ are put in the first p positions of the diagonal. Specifically for $\hat{C}_{(m-1)1}$, it is congruent to the diagonal matrix

$$Q^T\hat{C}_{(m-1)1}Q = \begin{pmatrix} Q_1^T C_{(m-1)1} Q_1 & Q_1^T C_{(m-1)1} Q_2 \\ Q_2^T C_{(m-1)1} Q_1 & Q_2^T C_{(m-1)1} Q_2 \end{pmatrix}$$

with $Q_1^T C_{(m-1)1} Q_1$ being nonsingular diagonal of $p \times p$, $Q_2^T C_{(m-1)1} Q_2$ being diagonal and $Q_1^T C_{(m-1)1} Q_2 = 0$. Since both $C_{(m-1)1}$ and $Q_1^T C_{(m-1)1} Q_1$ are nonsingular, the submatrix Q_1 must be nonsingular. The equation $Q_1^T C_{(m-1)1} Q_2 = 0$ thus implies that $Q_2 = 0$. Then we have

$$Q^T\hat{C}_{(m-1)1}Q = \begin{pmatrix} Q_1^T C_{(m-1)1} Q_1 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$Q^T\hat{C}_{i1}Q = \begin{pmatrix} Q_1^T C_{i1} Q_1 & 0 \\ 0 & 0 \end{pmatrix}, \quad i = 1, 2, \ldots, m-2,$$

such that $Q_1^T C_{i1} Q_1, i = 1, 2, \ldots, m-1$, are all diagonal.

Finally, for $i = m$,

$$Q^T\hat{C}_{m1}Q = \begin{pmatrix} Q_1^T C_{m1} Q_1 + Q_3^T C_{m6} Q_3 & Q_3^T C_{m6} Q_4 \\ Q_4^T C_{m6} Q_3 & Q_4^T C_{m6} Q_4 \end{pmatrix} \tag{2.63}$$

is diagonal. Then, $Q_1^T C_{m1} Q_1 + Q_3^T C_{m6} Q_3$ and $Q_4^T C_{m6} Q_4$ are diagonal and $Q_4^T C_{m6} Q_3 = 0$. We note that Q is nonsingular and $Q_2 = 0$, the matrix Q_4 must be nonsingular. Moreover, by assumption, C_{m6} is nonsingular so that $Q_4^T C_{m6} Q_3 = 0$ implies that $Q_3 = 0$. As a consequence, $\hat{C}_{m1}$ is reduced to

$$Q^T\hat{C}_{m1}Q = \begin{pmatrix} Q_1^T C_{m1} Q_1 & 0 \\ 0 & Q_4^T C_{m6} Q_4 \end{pmatrix},$$

so that $Q_1^T C_{m1} Q_1$ is diagonal.

Those arguments have shown that $C_{11}, C_{21}, \ldots, C_{(m-1)1}, C_{m1}$ are $\mathbb{R}$-SDC by the nonsingular matrix Q_1. $\square$

Theorem 2.2.3. *Let $\mathcal{C}_s = \{C_1, C_2, \ldots, C_m\} \subset \mathcal{S}^n$, $m \geq 3$ be a singular collection in which none is zero. If $C_1, C_2, \ldots, C_{m-1}$ are $\mathbb{R}$-SDC, then there exist a nonsingular real matrix Q and a positive vector $\mu = (\mu_1, \mu_2, \ldots, \mu_{m-2}, 1) \in \mathbb{R}^{m-1}_{++}$ such that*

$$\begin{aligned}
\tilde{C}_1 &= Q^T C_1 Q = \textbf{\textit{diag}}((C_{11})_p, 0_{n-p}), \ p < n; \\
\tilde{C}_2 &= Q^T(\mu_1 C_1 + C_2)Q = \textbf{\textit{diag}}((C_{21})_p, 0_{n-p}); \\
\tilde{C}_3 &= Q^T(\mu_2(\mu_1 C_1 + C_2) + C_3)Q = \textbf{\textit{diag}}((C_{31})_p, 0_{n-p}); \\
&\vdots \\
\tilde{C}_{m-1} &= Q^T(\mu_{m-2}(\cdots \mu_3(\mu_2(\mu_1 C_1 + C_2) + C_3) + C_4) + \cdots + C_{m-2}) + C_{m-1})Q \\
&= \textbf{\textit{diag}}((C_{(m-1)1})_p, 0_{n-p}); \qquad (2.64)
\end{aligned}$$

and either

$$\tilde{C}_m = Q^T C_m Q = \begin{pmatrix} (C_{m1})_p & C_{m2} \\ C_{m2}^T & 0_{n-p} \end{pmatrix}; \qquad (2.65)$$

or

$$\tilde{C}_m = Q^T C_m Q = \begin{pmatrix} (C_{m1})_p & 0 & C_{m5} \\ 0 & (C_{m6})_s & 0 \\ C_{m5}^T & 0 & 0_{n-p-s} \end{pmatrix}, \ s \leq n-p, \qquad (2.66)$$

where

- *the sub-matrices $(C_{i1})_p, i = 1, 2, \ldots, m-1$, are all diagonal of the same size. In particular, $(C_{(m-1)1})_p$ in (2.64) is nonsingular;*
- *in (2.65), $(C_{m1})_p$ is symmetric;*
- *in (2.66), $(C_{m1})_p$ is symmetric, $(C_{m6})_s$ is nonsingular diagonal; C_{m5} is either a $p \times (n-p-s)$ matrix if $s < n-p$ or does not exist if $s = n-p$.*

Moreover, the following three statements are equivalent.

(i) all matrices in the collection $\mathcal{C}_s$ are $\mathbb{R}$-SDC;

(ii) all matrices in the collection $\tilde{\mathcal{C}}_s = \{\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m\}$ are $\mathbb{R}$-SDC;

(iii) either sub-blocks $C_{11}, C_{21}, \ldots, C_{m1}$ with C_{m1} coming from (2.65) are $\mathbb{R}$-SDC and $C_{m2} = 0$; or sub-blocks $C_{11}, C_{21}, \ldots, C_{m1}$ with C_{m1} coming from (2.66) are $\mathbb{R}$-SDC and either $C_{m5} = 0$ or C_{m5} does not exist.

Proof. Proof for the initial step of mathematical induction:

1. Suppose C_1 and C_2 are $\mathbb{R}$-SDC. By Lemma 2.2.5, the theorem is true for $m = 3$.

2. Proof for the induction step on $m \geq 4$: Suppose (2.64) and (2.65) or (2.64) and (2.66) hold for $m-1$ matrices $C_1, C_2, \dots, C_{m-1}$, i.e., there exist a nonsingular matrix Q_1 and a vector $\hat{\mu} = (\mu_1, \mu_2, \dots, \mu_{m-3}, 1) \in \mathbb{R}^{m-2}_{++}$ such that

$$\begin{aligned}
\hat{C}_1 =& Q_1^T C_1 Q_1 = \texttt{diag}((\hat{C}_{11})_{p_1}, 0_{r_1}), \\
\hat{C}_2 =& Q_1^T (\mu_1 C_1 + C_2) Q_1 = \texttt{diag}((\hat{C}_{21})_{p_1}, 0_{r_1}), \\
&\vdots \\
\hat{C}_{m-2} =& Q_1^T (\mu_{m-3}(\dots \mu_2(\mu_1 C_1 + C_2) + C_3) + \dots + C_{m-2}) Q_1 \\
=& \texttt{diag}((\hat{C}_{(m-2)1})_{p_1}, 0_{r_1}),
\end{aligned} \tag{2.67}$$

and either

$$\hat{C}_{m-1} \quad = \quad Q_1^T C_{m-1} Q_1 = \begin{pmatrix} (\hat{C}_{(m-1)1})_{p_1} & \hat{C}_{(m-1)2} \\ \hat{C}^T_{(m-1)2} & 0_{n-p_1} \end{pmatrix}; \tag{2.68}$$

or

$$\hat{C}_{m-1} = \quad Q_1^T C_{m-1} Q_1 = \begin{pmatrix} (\hat{C}_{(m-1)1})_{p_1} & 0_{p_1 \times s_1} & \hat{C}_{(m-1)5} \\ 0_{s_1 \times p_1} & \hat{C}_{(m-1)6} & 0_{s_1 \times (r_1 - s_1)} \\ (\hat{C}_{(m-1)5})^T & 0_{(r_1 - s_1) \times s_1} & 0_{r_1 - s_1} \end{pmatrix}, \tag{2.69}$$

where

- the sub-matrices $(\hat{C}_{i1})_{p_1}, i = 1, 2, \dots, m-2$, are all diagonal of the same size. In particular, $(\hat{C}_{(m-2)1})_{p_1}$ is nonsingular;
- in (2.68), $(\hat{C}_{(m-1)1})_{p_1}$ is symmetric; $(\hat{C}_{(m-1)2})$ is a $p_1 \times (n - p_1)$.
- in (2.69), $(\hat{C}_{(m-1)1})_{p_1}$ is symmetric, $(\hat{C}_{(m-1)6})_s$ is nonsingular diagonal; $\hat{C}_{(m-1)5}$ is either a $p_1 \times (n - p_1 - s)$ matrix if $s < n - p_1$ or does not exist if $s = n - p_1$.

Since $C_1, C_2, \dots, C_{m-1}$ are $\mathbb{R}$-SDC, the collection $\hat{C}_1, \hat{C}_2, \dots, \hat{C}_{m-1}$ are $\mathbb{R}$-SDC. This implies that $\hat{C}_{m-2}, \hat{C}_{m-1}$ are $\mathbb{R}$-SDC.

• If $\hat{C}_{m-1}$ takes the form (2.68), $\hat{C}_{(m-1)2} = 0$ (by Lemma 1.2.1). Then, the matrices $\hat{C}_1, \hat{C}_2, \dots, \hat{C}_{m-1}$ have the form of (1.1). By Lemma 1.1.6, the submatrices $\hat{C}_{11}, \hat{C}_{21}, \dots, \hat{C}_{(m-1)1}$ are $\mathbb{R}$-SDC.

• If $\hat{C}_{m-1}$ takes the form (2.69), $\hat{C}_{(m-1)5}$ is zero or does not exist (by Lemma 1.2.9). Then, the matrices $\hat{C}_1, \hat{C}_2, \dots, \hat{C}_{m-1}$ have the form of (2.61), (2.62). By Lemma 2.2.6, the submatrices $\hat{C}_{11}, \hat{C}_{21}, \dots, \hat{C}_{(m-1)1}$ are $\mathbb{R}$-SDC.

Therefore, there is a nonsingular matrix P_1 such that

$$P_1^T \hat{C}_{i1} P_1 = \tilde{C}_{i1}, i = 1, 2, \ldots, m-1$$

are all diagonal. Set $Q_2 = \begin{pmatrix} P_1 & 0 \\ 0 & I_{r_1} \end{pmatrix}$ and $Q_3 = Q_1 Q_2$, then from (2.67),

$$\begin{aligned}
\tilde{C}_1 =& Q_3^T C_1 Q_3 = Q_2^T \hat{C}_1 Q_2 = \mathtt{diag}(\tilde{C}_{11}, 0_{r_1}), \\
\tilde{C}_2 =& Q_3^T (\mu_1 C_1 + C_2) Q_3 = Q_2^T \hat{C}_2 Q_2 = \mathtt{diag}(\tilde{C}_{21}, 0_{r_1}), \\
& \ldots \\
\tilde{C}_{m-2} =& Q_3^T (\mu_{m-3}(\ldots \mu_3(\mu_2(\mu_1 C_1 + C_2) + C_3) + C_4) + \ldots + C_{m-3}) + C_{m-2}) Q_3 \\
=& Q_2^T \hat{C}_{m-2} Q_2 = \mathtt{diag}(\tilde{C}_{(m-2)1}, 0_{r_1}) \\
\tilde{C}_{m-1} =& Q_3^T C_{m-1} Q_3 = Q_2^T \hat{C}_{m-1} Q_2 = \mathtt{diag}(\tilde{C}_{(m-1)1}, (\hat{C}_{(m-1)6})_{s_1}, 0_{r_1 - s_1}), s_1 \geq 0,
\end{aligned}$$

where all $\tilde{C}_{i1}$ are diagonal, $i = 1, 2, \ldots, m-1$; $\tilde{C}_{(m-2)1}, \hat{C}_{(m-1)6}$ are nonsingular diagonal. Notice that if $s_1 = 0$ then $\hat{C}_{(m-1)6}$ does not exist.

Suppose

$$\tilde{C}_{(m-2)1} = \mathtt{diag}(\eta_1, \eta_2, \ldots, \eta_{p_1}) \text{ and } \tilde{C}_{(m-1)1} = \mathtt{diag}(\gamma_1, \gamma_2, \ldots, \gamma_{p_1})$$

where $\eta_j \neq 0, j = 1, 2, \ldots, p_1$. We now define

$$\mu_{m-2} = \max_{1 \leq j \leq p_1} \left\{ \left| \frac{\gamma_j}{\eta_j} \right| + 1 \right\}.$$

Then, the matrix

$$\mu_{m-2} \tilde{C}_{(m-2)1} + \tilde{C}_{(m-1)1} = \mathtt{diag}(\mu_{m-2}\eta_1 + \gamma_1, \ldots, \mu_{m-2}\eta_{p_1} + \gamma_{p_1})$$

is nonsingular diagonal of size $p_1 \times p_1$. Let $r_2 = r_1 - s_1, p_2 = p_1 + s_1$ and

$$\begin{aligned}
C_{i1} =& \mathtt{diag}(\tilde{C}_{i1}, 0_{s_1}), i = 1, 2, \ldots, m-2, \\
C_{(m-1)1} =& \mathtt{diag}(\mu_{m-2} \tilde{C}_{(m-2)1} + \tilde{C}_{(m-1)1}, \hat{C}_{(m-1)6})
\end{aligned}$$

we will have

$$\tilde{C}_i = \mathtt{diag}((C_{i1})_{p_2}, 0_{r_2}), i = 1, 2, \ldots, m-1,$$

such that $C_{(m-1)1}$ is nonsingular diagonal and $\mu = (\mu_1, \ldots, \mu_{m-2}, 1) \in \mathbb{R}_{++}^{m-1}$.

Now for $Q_3^T C_m Q_3$ we make a partition as

$$\hat{C}_m = Q_3^T C_m Q_3 = \begin{pmatrix} N_{m1} & N_{m2} \\ N_{m3} & N_{m4} \end{pmatrix}$$

such that N_{m1} and N_{m4} are symmetric matrices of size $p_2 \times p_2$ and $r_2 \times r_2$, respectively. Using the same arguments as in (2.48), (2.52), there will be nonsingular matrices U such that $Q = Q_3 U$ satisfying $\tilde{C}_i = Q^T \tilde{C}_i Q$ for all $i = 1, 2, \ldots, m-1$ and $\tilde{C}_m = Q^T \hat{C}_m Q$ is of the form (2.65) or (2.66). Then the matrix Q will be the one we need to find and $\mu = (\mu_1, \ldots, \mu_{m-2}, 1) \in \mathbb{R}^{m-1}_{++}$.

Moreover, we have:

$(i) \Leftrightarrow (ii)$. If all matrices in the collection $\mathcal{C}_s$ are $\mathbb{R}$-SDC by P, then all matrices in the collection $\tilde{\mathcal{C}}_s = \{\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m\}$ are $\mathbb{R}$-SDC by $Q^{-1}P$;

Conversely, if $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $\mathbb{R}$-SDC by R, then $C_1, C_2, \ldots, C_m$ are $\mathbb{R}$-SDC by QR.

$(ii) \Leftrightarrow (iii)$ If $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $\mathbb{R}$-SDC, $\tilde{C}_{m-1}, \tilde{C}_m$ are $\mathbb{R}$-SDC. By Lemma 1.2.1, $C_{m2} = 0$ if $\tilde{C}_m$ is in the form of (2.65). Then, by Lemma 1.1.6, sub-blocks $C_{11}, C_{21}, \ldots, C_{m1}$ are $\mathbb{R}$-SDC. And by Lemma 1.2.9, $C_{m5} = 0$ or does not exist if $\tilde{C}_m$ is in the form of (2.66). Then, by Lemma 2.2.6, sub-blocks $C_{11}, C_{21}, \ldots, C_{m1}$ are $\mathbb{R}$-SDC.

Conversely, if sub-blocks $C_{11}, C_{21}, \ldots, C_{m1}$ with C_{m1} coming from (2.65) are $\mathbb{R}$-SDC and $C_{m2} = 0$, then $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $\mathbb{R}$-SDC (since Lemma 1.1.6). Or, if sub-blocks $C_{11}, C_{21}, \ldots, C_{m1}$ with C_{m1} coming from (2.66) are $\mathbb{R}$-SDC and either $C_{m5} = 0$ or C_{m5} does not exist, then $\tilde{C}_1, \tilde{C}_2, \ldots, \tilde{C}_m$ are $\mathbb{R}$-SDC (by Lemma 2.2.6). □

2.2.4 Algorithm for the singular collection

The following algorithm helps to solve the SDC problem of a singular collection.

Algorithm 8 Solving the SDC problem for a singular collection.

INPUT: A singular collection of real symmetric matrices $C_1, C_2, \ldots, C_m$

OUTPUT: NOT $\mathbb{R}$-SDC or a nonsingular real matrix Q that simultaneously diagonalizes $C_1, C_2, \ldots, C_m$

Step 1. Find a matrix Q such that $Q^T C_1 Q = \mathtt{diag}(\alpha_1, \alpha_2, \ldots, \alpha_p, 0_r) := \mathtt{diag}(C_{11}, 0_r), \alpha_i \neq 0.$

Step 2. For $i = 2$ **to** m **do**

Using Lemma 1.2.8 to find Q_i such that

$$Q_i^T Q^T C_i Q Q_i = \begin{pmatrix} (C_{i1})_p & (C_{i2})_{p\times(n-p)} \\ C_{i2}^T & 0_{n-p} \end{pmatrix}.$$

or

$$Q_i^T Q^T C_i Q Q_i = \begin{pmatrix} C_{i1} & 0_{p\times s} & C_{i5} \\ 0_{s\times p} & C_{i6} & 0_{s\times(r-s)} \\ (C_{i5})^T & 0_{(r-s)\times s} & 0_{r-s} \end{pmatrix}.$$

If $C_{i2} \neq 0$, or $C_{i5} \neq 0$ then NOT $\mathbb{R}$-SDC and STOP.

Else, apply Algorithm 7 for $C_{11}, \ldots, C_{i1}$.

If $C_{11}, \ldots, C_{i1}$ are not $\mathbb{R}$-SDC then $C_1, C_2, \ldots, C_m$ are NOT $\mathbb{R}$-SDC and STOP,

Else let P_i be the matrix returned when applying Algorithm 7 for

$C_{11}, \ldots, C_{i1}$, set $M_i := \mathtt{diag}(P_i, I_r), Q := QQ_iM_i$.

If $i = m$ then Stop. Else compute

$Q^T C_i Q := \mathtt{diag}(\beta_1, \beta_2, \ldots, \beta_p, \beta_{p+1}, \ldots, \beta_{p+s}, 0_{r-s}),$

$\mu = \max_{1\le j\le p} \left\{ \left| \frac{\beta_j}{\alpha_j} \right| + 1 \right\}.$

Set $\alpha_1 := \mu\alpha_1 + \beta_1, \ldots, \alpha_p := \mu\alpha_p + \beta_p, \alpha_{p+1} := \beta_{p+1}, \ldots, \alpha_{p+s} := \beta_{p+s}$

$p := p + s; r := n - p,$

EndIf

EndIf

EndFor

Return Q.

To end the section we consider the following simple example to see how the algorithm works. We suppose the first two matrices were diagonalized by Lemma 1.2.9 [37, Theorem 6].

Example 2.2.3. We consider the following singular collection of three matrices

$$C_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, C_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, C_3 = \begin{pmatrix} 5 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Step 1. Since C_1 are already diagonal: $C_1 = \mathtt{diag}(1, 0_3) = \mathtt{diag}(C_{11}, 0_3)$, so $Q := I$.

Step 2. For $i = 2$ to 3 do

• $i = 2$: find $Q_2 = I$ such that $Q_2^T Q^T C_2 Q Q_2 = \mathtt{diag}(0, 1, \frac{1}{4}, 0) = (C_{21}, C_{26}, 0_1)$, $C_{21} = \mathtt{diag}(0), C_{26} = \mathtt{diag}(1, \frac{1}{4})$.

Since C_{11}, C_{21} are already diagonal, let $P_2 := I_1, M_2 := \mathtt{diag}(I_1, I_3)$, update $Q := QQ_2M_2 = I$.

We find $\mu = 1$ then $\alpha_1 = 1, \alpha_2 = 1, \alpha_3 = \frac{1}{4}; p = 1 + 2 = 3, r = n - p = 4 - 3 = 1$.
• $i = 3$: Applying Lemma 1.2.8 to find Q_3 as follows: we have $\hat{C}_3 = Q^T C_3 Q = C_3 = \begin{pmatrix} M_{31} & M_{32} \\ (M_{32})^T & M_{33} \end{pmatrix}$, here $M_{31} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, M_{32} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $M_{33} = (1)$. Let $Q_3 = \begin{pmatrix} I_3 & 0 \\ -(M_{33})^{-1}(M_{32})^T & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$ we have $\tilde{C}_3 = Q_3^T \hat{C}_3 Q_3 = Q_3^T Q^T C_3 Q Q_3 = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, then $C_{31} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, C_{36} = (1)$, C_{35} does not exist. Apply Algorithm 7 for C_{11}, C_{21}, C_{31} we find $P_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 4 \end{pmatrix}$. Set $H_3 = \mathtt{diag}(P_3, I_1)$ and update $Q = QQ_3H_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 4 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$. We can check that C_1, C_2, C_3 are $\mathbb{R}$-SDC by Q.

Conclusion of Chapter 2

In the first part of this chapter we presented two different methods for solving the SDC problem of Hermitian matrices, which are the max-rank method shown in Theorem 2.1.4 and the Algorithm 4, and the SDP method, please see Theorem 2.1.5 and the Algorithm 6. In the second part of the chapter, we proposed a contructive and inductive algorithm for solving the SDC problem of the real symmetric matrices, which are Theorems 2.2.2, 2.2.3 and the Algorithms 7, 8. We also presented numerical experiments to show the efficiency of the algorithms.

Chapter 3

Some applications of the SDC results

In this chaper we show how the SDC of matrices can help to solve some problems. In Section 3.1, we show that the SDC of two real symmetric matrices can help to completely evaluate the positive semidefinite interval of matrix pencil. In Section 3.2 we use the SDC of matrices $C_1, C_2, \ldots, C_m$ to relax a QCQP to a convex SOCP, which is then a lower bound of such a QCQP. In some special cases, for example QCQP with one or two constraints, homogeneous QCQP, the relaxation is tight, and the QCQP is then equivalently transformed to a convex SOCP. Especially, also in this section, we present how to use the positive semidefinite interval of matrix pencil to completely solve an important case of the QCQP-the GTRS. Finally, an application of the SDC to maximizing a sum of generalized Rayleigh quotients is mentioned. The results of Section 3.1 and Subsection 3.2.1 are taken from [47]. The results of Subsection 3.2.2 are taken from [46].

3.1 Computing the positive semidefinite interval

Let C_1 and C_2 be real symmetric matrices. In this section we are concerned with finding the set $I_{\succeq}(C_1, C_2) = \{\mu \in \mathbb{R} : C_1 + \mu C_2 \succeq 0\}$ of real values μ such that the matrix pencil $C_1 + \mu C_2$ is positive semidefinite. If C_1, C_2 are not $\mathbb{R}$-SDC, $I_{\succeq}(C_1, C_2)$ either is empty or has only one value μ. When C_1, C_2 are $\mathbb{R}$-SDC, $I_{\succeq}(C_1, C_2)$, if not empty, can be a singleton or an interval. Especially, if $I_{\succeq}(C_1, C_2)$ is an interval and at least one of the matrices is nonsingular then its interior is the positive definite interval $I_{\succ}(C_1, C_2)$. If C_1, C_2 are both singular, then even $I_{\succeq}(C_1, C_2)$ is an interval, its interior may not be $I_{\succ}(C_1, C_2)$, but C_1, C_2 are then decomposed to block diagonals of submatrices A_1, B_1 with B_1 nonsingular such that $I_{\succeq}(C_1, C_2) = I_{\succeq}(A_1, B_1)$.

In this section, we show computing $I_{\succeq}(C_1, C_2)$ in two separate cases: C_1, C_2 are $\mathbb{R}$-SDC and C_1, C_2 are not $\mathbb{R}$-SDC.

3.1.1 Computing $I_{\succeq}(C_1, C_2)$ when C_1, C_2 are $\mathbb{R}$-SDC

Now, if C_1, C_2 are $\mathbb{R}$-SDC and C_2 is nonsingular, by Lemma 1.2.1, there is a nonsingular matrix P such that

$$J := P^{-1}C_2^{-1}C_1P = \texttt{diag}(\lambda_1 I_{m_1}, \ldots, \lambda_k I_{m_k}), \tag{3.1}$$

is a diagonal matrix, where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the k distinct eigenvalues of $C_2^{-1}C_1$, I_{m_t} is the identity matrix of size $m_t \times m_t$ and $m_1 + m_2 + \ldots + m_k = n$. We can suppose without loss of generality that $\lambda_1 > \lambda_2 > \ldots > \lambda_k$.

Observe that $P^TC_2P.J = P^TC_1P$ and P^TC_1P is symmetric. Lemma 1.1.2 indicates that P^TC_2P is a block diagonal matrix with the same partition as J. That is

$$P^TC_2P = \texttt{diag}(B_1, B_2 \ldots, B_k), \tag{3.2}$$

where B_t is real symmetric matrices of size $m_t \times m_t$ for every $t = 1, 2, \ldots, k$. We now have

$$P^TC_1P = P^TC_2P.J = \texttt{diag}(\lambda_1 B_1, \lambda_2 B_2 \ldots, \lambda_k B_k). \tag{3.3}$$

Both (3.2) and (3.3) show that C_1, C_2 are now decomposed into the same block structure and the matrix pencil $C_1 + \mu C_2$ now becomes

$$P^T(C_1 + \mu C_2)P = \texttt{diag}((\lambda_1 + \mu)B_1, (\lambda_2 + \mu)B_2 \ldots, (\lambda_k + \mu)B_k). \tag{3.4}$$

The requirement $C_1 + \mu C_2 \succeq 0$ is then equivalent to

$$(\lambda_i + \mu)B_i \succeq 0, i = 1, 2, \ldots, k. \tag{3.5}$$

Using (3.5) we compute $I_{\succeq}(C_1, C_2)$ as follows.

Theorem 3.1.1. *Suppose $C_1, C_2 \in \mathcal{S}^n$ are $\mathbb{R}$-SDC and C_2 is nonsingular.*

1. *If $C_2 \succ 0$ then $I_{\succeq}(C_1, C_2) = [-\lambda_k, +\infty)$;*

2. If $C_2 \prec 0$ *then* $I_{\succeq}(C_1, C_2) = (-\infty, -\lambda_1]$*;*

3. If C_2 *is indefinite then*

(i) if $B_1, B_2, \ldots, B_t \succ 0$ *and* $B_{t+1}, B_{t+2}, \ldots, B_k \prec 0$ *for some* $t \in \{1, 2, \ldots, k\}$*, then* $I_{\succeq}(C_1, C_2) = [-\lambda_t, -\lambda_{t+1}]$*.*

(ii) if $B_1, B_2, \ldots, B_{t-1} \succ 0$*,* B_t *is indefinite and* $B_{t+1}, B_{t+2}, \ldots, B_k \prec 0$*, then* $I_{\succeq}(C_1, C_2) = \{-\lambda_t\}$*,*

(iii) in other cases, that is either B_i, B_j *are indefinite for some* $i \neq j$ *or* $B_i \prec 0, B_j \succ 0$ *for some* $i < j$ *or* B_i *is indefinite and* $B_j \succ 0$ *for some* $i < j$*, then* $I_{\succeq}(C_1, C_2) = \emptyset$*.*

Proof. 1. If $C_2 \succ 0$ then $B_i \succ 0 \ \forall i = 1, 2, \ldots, k$. The inequality (3.5) is then equivalent to $\lambda_i + \mu \geq 0 \ \forall i = 1, 2, \ldots, k$. Since $\lambda_1 > \lambda_2 > \ldots > \lambda_k$, we need only $\mu \geq -\lambda_k$. This shows $I_{\succeq}(C_1, C_2) = [-\lambda_k, +\infty)$.

2. Similarly, if $C_2 \prec 0$ then $B_i \prec 0 \ \forall i = 1, 2, \ldots, k$. The inequality (3.5) is then equivalent to $\lambda_i + \mu \leq 0 \ \forall i = 1, 2, \ldots, k$. Then $I_{\succeq}(C_1, C_2) = (-\infty, -\lambda_1]$.

3. The case C_2 is indefinite:

(i) if $B_1, B_2, \ldots, B_t \succ 0$ and $B_{t+1}, B_{t+2}, \ldots, B_k \prec 0$ for some $t \in \{1, 2, \ldots, k\}$, the inequality (3.5) then implies

$$\begin{cases} \lambda_i + \mu \geq 0, \forall i = 1, 2, \ldots, t, \\ \lambda_i + \mu \leq 0, \forall i = t+1, \ldots, k. \end{cases}$$

Since $\lambda_1 > \lambda_2 > \ldots > \lambda_k$, we have $I_{\succeq}(C_1, C_2) = [-\lambda_t, -\lambda_{t+1}]$.

(ii) if $B_1, B_2, \ldots, B_{t-1} \succ 0$, B_t is indefinite and $B_{t+1}, B_{t+2}, \ldots, B_k \prec 0$ for some $t \in \{1, 2, \ldots, k\}$. The inequality (3.5) then implies

$$\begin{cases} \lambda_i + \mu \geq 0, \forall i = 1, 2, \ldots, t-1 \\ \lambda_t + \mu = 0 \\ \lambda_i + \mu \leq 0, \forall i = t+1, \ldots, k. \end{cases}$$

Since $\lambda_1 > \lambda_2 > \ldots > \lambda_k$, we have $I_{\succeq}(C_1, C_2) = \{-\lambda_t\}$.

(iii) if B_i, B_j are indefinite, (3.5) implies $\lambda_i + \mu = 0$ and $\lambda_j + \mu = 0$. This cannot happen since $\lambda_i \neq \lambda_j$. If $B_i \prec 0$ and $B_j \succ 0$ for some $i < j$, then

$$\begin{cases} \lambda_i + \mu \leq 0 \\ \lambda_j + \mu \geq 0 \end{cases}$$

implying $-\lambda_j \le \mu \le -\lambda_i$. This also cannot happen since $\lambda_i > \lambda_j$. Finally, if B_i is indefinite and $B_j \succ 0$ for some $i < j$. Again, by (3.5),

$$\begin{cases} \lambda_i + \mu = 0 \\ \lambda_j + \mu \ge 0 \end{cases}$$

implying $\lambda_i \le \lambda_j$. This also cannot happen. So $I_{\succeq}(C_1, C_2) = \emptyset$ in these all three cases.

□

The proof of Theorem 3.1.1 indicates that if C_1, C_2 are $\mathbb{R}$-SDC, C_2 is nonsingular and $I_{\succeq}(C_1, C_2)$ is an interval then $I_{\succ}(C_1, C_2)$ is nonempty. In that case we have $I_{\succ}(C_1, C_2) = \text{int}(I_{\succeq}(C_1, C_2))$, please see [44]. If C_2 is singular and C_1 is nonsingular, we have the following result.

Theorem 3.1.2. *Suppose $C_1, C_2 \in \mathcal{S}^n$ are $\mathbb{R}$-SDC, C_2 is singular and C_1 is nonsingular. Then*

(i) there always exists a nonsingular matrix U such that

$$U^T C_2 U = \textbf{\textit{diag}}(B_1, 0),$$

$$U^T C_1 U = \textbf{\textit{diag}}(A_1, A_3),$$

where B_1, A_1 are symmetric of the same size, B_1 is nonsingular;

(ii) if $A_3 \succ 0$ then $I_{\succeq}(C_1, C_2) = I_{\succeq}(A_1, B_1)$. Otherwise, $I_{\succeq}(C_1, C_2) = \emptyset$.

Proof. (i) Since C_2 is symmetric and singular, there is an orthogonal matrix Q_1 that puts C_2 into the form

$$\hat{C}_2 = Q_1^T C_2 Q_1 = \mathtt{diag}(B_1, 0)$$

such that B_1 is a nonsingular symmetric matrix of size $p \times p$, where $p = \text{rank}(B)$. Let $\hat{C}_1 := Q_1^T C_1 Q_1$. Since C_1, C_2 are $\mathbb{R}$-SDC, $\hat{C}_1, \hat{C}_2$ are $\mathbb{R}$-SDC too (the converse also holds true). We can write $\hat{C}_1$ in the following form

$$\hat{C}_1 = Q_1^T C_1 Q_1 = \begin{pmatrix} M_1 & M_2 \\ M_2^T & M_3 \end{pmatrix} \tag{3.6}$$

such that M_1 is a symmetric matrix of size $p \times p$, M_2 is a $p \times (n-p)$ matrix, M_3 is symmetric of size $(n-p) \times (n-p)$ and, importantly, $M_3 \ne 0$. Indeed, if $M_3 = 0$ then

$\hat{C}_1 = Q_1^T C_1 Q_1 = \begin{pmatrix} M_1 & M_2 \\ M_2^T & 0 \end{pmatrix}$. Then we can choose a nonsingular matrix H written in the same partition as $\hat{C}_1 : H = \begin{pmatrix} H_1 & H_2 \\ H_3 & H_4 \end{pmatrix}$ such that both $H^T\hat{C}_2H, H^T\hat{C}_1H$ are diagonal and $H^T\hat{C}_2H$ is of the form

$$H^T\hat{C}_2H = \begin{pmatrix} H_1^TB_1H_1 & H_1^TB_1H_2 \\ H_2^TB_1H_1 & H_2^TB_1H_2 \end{pmatrix} = \begin{pmatrix} H_1^TB_1H_1 & 0 \\ 0 & 0 \end{pmatrix},$$

where $H_1^TB_1H_1$ is nonsingular. This implies $H_2 = 0$. On the other hand,

$$H^T\hat{C}_1H = \begin{pmatrix} H_1^TM_1H_1 + H_3^TM_2^TH_1 + H_1^TM_2H_3 & H_1^TM_2H_4 \\ H_4^TM_2^TH_1 & 0 \end{pmatrix}$$

is diagonal implying that $H_1^TM_2H_4 = 0$, and so

$$H^T\hat{C}_1H = \begin{pmatrix} H_1^TM_1H_1 + H_3^TM_2^TH_1 + H_1^TM_2H_3 & 0 \\ 0 & 0 \end{pmatrix}.$$

This cannot happen since $\hat{C}_1$ is nonsingular.

Let P be an orthogonal matrix such that $P^TM_3P = \texttt{diag}(A_3, 0_{q-r})$, where A_3 is a nonsingular diagonal matrix of size $r\times r, r \leq q$ and $p+q = n$, and set $U_1 = \texttt{diag}(I_p, P)$. We then have

$$\tilde{C}_1 := U_1^T\hat{C}_1U_1 = \begin{pmatrix} M_1 & M_2P \\ (M_2P)^T & P^TM_3P \end{pmatrix} = \begin{pmatrix} M_1 & A_4 & A_5 \\ A_4^T & A_3 & 0 \\ A_5^T & 0 & 0 \end{pmatrix}, \tag{3.7}$$

where $\begin{pmatrix} A_4 & A_5 \end{pmatrix} = M_2P$, A_4 and A_5 are of size $p\times r$ and $p\times(q-r), r \leq q$, respectively. Let

$$U_2 = \begin{pmatrix} I_p & 0 & 0 \\ -A_3^{-1}A_4^T & I_r & 0 \\ 0 & 0 & I_{q-r} \end{pmatrix} \text{ and } U = Q_1U_1U_2.$$

We can verify that

$$U^TC_2U = U_2^TU_1^T(Q_1^TC_2Q_1))U_1U_2 = \hat{C}_2,$$

and, by (3.7),

$$U^TC_1U = U_2^T\tilde{C}_1U_2 = \begin{pmatrix} M_1 - A_4A_3^{-1}A_4^T & 0 & A_5 \\ 0 & A_3 & 0 \\ A_5^T & 0 & 0 \end{pmatrix}.$$

We denote $A_1 := M_1 - A_4A_3^{-1}A_4^T$ and rewrite the matrices as follows

$$U^TC_2U = \texttt{diag}(B_1, 0), U^TC_1U = \begin{pmatrix} A_1 & 0 & A_5 \\ 0 & A_3 & 0 \\ A_5^T & 0 & 0 \end{pmatrix}.$$

We now consider whether it can happen that $r < q$. We note that U^TC_1U, U^TC_2U are $\mathbb{R}$-SDC. We can choose a nonsingular congruence matrix K written in the form

$$K = \begin{pmatrix} K_1 & K_2 & K_3 \\ K_4 & K_5 & K_6 \\ K_7 & K_8 & K_9 \end{pmatrix}$$

such that not only the matrices $K^TU^TC_1UK, K^TU^TC_2UK$ are diagonal but also the matrix $K^TU^TC_2UK$ is remained a $p\times p$ nonsingular submatrix at the northwest corner. That is

$$K^TU^TC_2UK = \begin{pmatrix} K_1^TB_1K_1 & K_1^TB_1K_2 & K_1^TB_1K_3 \\ K_2^TB_1K_1 & K_2^TB_1K_2 & K_2^TB_1K_3 \\ K_3^TB_1K_1 & K_3^TB_1K_2 & K_3^TB_1K_3 \end{pmatrix} = \begin{pmatrix} K_1^TB_1K_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

is diagonal and $K_1^TB_1K_1$ is nonsingular diagonal of size $p \times p$. This implies that $K_2 = K_3 = 0$. Then

$$\begin{aligned} &K^TU^TC_1UK = \\ &= \begin{pmatrix} \substack{K_1^TA_1K_1+K_1^TA_2K_7 \\ +K_4^TA_3K_4+K_7^TA_2^TK_1} & K_1^TA_2K_8 + K_4^TA_3K_5 & K_1^TA_2K_9 + K_4^TA_3K_6 \\ K_8^TA_2^TK_1 + K_5^TA_3^TK_4 & K_5^TA_3K_5 & K_5^TA_3K_6 \\ K_9^TA_2^TK_1 + K_6^TA_3^TK_4 & K_6^TA_3K_5 & K_6^TA_3K_6 \end{pmatrix} \\ &= \begin{pmatrix} K_1^TA_1K_1 + K_1^TA_2K_7 + K_4^TA_3K_4 + K_7^TA_2^TK_1 & 0 & 0 \\ 0 & K_5^TA_3K_5 & 0 \\ 0 & 0 & K_6^TA_3K_6 \end{pmatrix} \end{aligned}$$

is diagonal implying that

$$K_1^TA_1K_1 + K_1^TA_2K_7 + K_4^TA_3K_4 + K_7^TA_2^TK_1, K_5^TA_3K_5, K_6^TA_3K_6$$

are diagonal. Note that U^TC_1U is nonsingular, $K_5^TA_3K_5, K_6^TA_3K_6$ must be nonsingular. But then $K_5^TA_3K_6 = 0$ with A_3 nonsingular is a contradiction. It therefore holds that $q = r$. Then

$$U^TC_2U = \texttt{diag}(B_1, 0), U^TC_1U = \texttt{diag}(A_1, A_3)$$

with B_1, A_1, A_3 as desired.

(ii) We note first that C_1 is nonsingular so is A_3. If $A_3 \succ 0$, then $C_1 + \mu C_2 \succeq 0$ if and only if $A_1 + \mu B_1 \succeq 0$. So it holds in that case $I_{\succeq}(C_1, C_2) = I_{\succeq}(A_1, B_1)$. Otherwise, A_3 is either indefinite or negative definite then $I_{\succeq}(C_1, C_2) = \emptyset$. □

The proofs of Theorems 3.1.1 and 3.1.2 reveal the following important result.

Corollary 3.1.1. *Suppose $C_1, C_2 \in \mathcal{S}^n$ are $\mathbb{R}$-SDC and either C_1 or C_2 is nonsingular. Then $I_{\succ}(C_1, C_2)$ is nonempty if and only if $I_{\succeq}(C_1, C_2)$ has more than one point.*

If C_1, C_2 are both singular, by Lemma 1.2.8, they can be decomposed in one of the following forms.

For any $C_1, C_2 \in \mathcal{S}^n$, there always exists a nonsingular matrix U that puts C_2 to

$$\tilde{C}_2 = U^T C_2 U = \begin{pmatrix} B_1 & 0_{p\times r} \\ 0_{r\times p} & 0_{r\times r} \end{pmatrix}$$

such that B_1 is nonsingular diagonal of size $p \times p$, and puts A to $\tilde{A}$ of either form

$$\tilde{C}_1 = U^T C_1 U = \begin{pmatrix} A_1 & A_2 \\ A_2^T & 0_{r\times r} \end{pmatrix} \tag{3.8}$$

where A_1 is symmetric of dimension $p \times p$ and A_2 is a $p \times r$ matrix, or

$$\tilde{C}_1 = U^T C_1 U = \begin{pmatrix} A_1 & 0_{p\times s} & A_2 \\ 0_{s\times p} & A_3 & 0_{s\times(r-s)} \\ A_2^T & 0_{(r-s)\times s} & 0_{(r-s)\times(r-s)} \end{pmatrix}, \tag{3.9}$$

where A_1 is symmetric of dimension $p \times p$, A_2 is a $p \times (r-s)$ matrix, and A_3 is a nonsingular diagonal matrix of dimension $s \times s$; $p, r, s \geq 0, p + r = n$.

It is easy to verify that C_1, C_2 are $\mathbb{R}$-SDC if and only if $\tilde{C}_1, \tilde{C}_2$ are $\mathbb{R}$-SDC. And we have:

i) If $\tilde{C}_1$ takes the form (3.8) then $\tilde{C}_2, \tilde{C}_1$ are $\mathbb{R}$-SDC if and only if B_1, A_1 are $\mathbb{R}$-SDC and $A_2 = 0$;

ii) If $\tilde{C}_1$ takes the form (3.9) then $\tilde{C}_2, \tilde{C}_1$ are $\mathbb{R}$-SDC if and only if B_1, A_1 are $\mathbb{R}$-SDC and $A_2 = 0$ or does not exist, i.e., $s = r$.

Now suppose that $\{C_1, C_2\}$ are $\mathbb{R}$-SDC, without loss of generality we always assume that $\tilde{C}_2, \tilde{C}_1$ are already $\mathbb{R}$-SDC. That is

$$\tilde{C}_2 = U^T C_2 U = \texttt{diag}(B_1, 0), \tilde{C}_1 = U^T C_1 U = \texttt{diag}(A_1, 0) \tag{3.10}$$

or

$$\tilde{C}_2 = U^T C_2 U = \texttt{diag}(B_1, 0), \tilde{C}_1 = U^T C_1 U = \texttt{diag}(A_1, A_4), \tag{3.11}$$

where A_1, B_1 are of the same size and diagonal, B_1 is nonsingular and if $\tilde{C}_1$ takes the form (3.8) or (3.9) and $A_2 = 0$ then $A_4 = \texttt{diag}(A_3, 0)$ or if $\tilde{C}_1$ takes the form (3.9) and A_2 does not exist then $A_4 = A_3$. Now we can compute $I_{\succeq}(C_1, C_2)$ as follows.

Theorem 3.1.3. *(i) If $\tilde{C}_2, \tilde{C}_1$ take the form (3.10), then $I_{\succeq}(C_1, C_2) = I_{\succeq}(A_1, B_1)$;*

(ii) If $\tilde{C}_2, \tilde{C}_1$ take the form (3.11), then $I_{\succeq}(C_1, C_2) = I_{\succeq}(A_1, B_1)$ if $A_4 \succeq 0$ and $I_{\succeq}(C_1, C_2) = \emptyset$ otherwise.

We note that B_1 is nonsingular, $I_{\succeq}(A_1, B_1)$ is therefore computed by Theorem 3.1.1. Especially, if $I_{\succeq}(A_1, B_1)$ has more than one point, then $I_{\succ}(A_1, B_1) \neq \emptyset$, see Corollary 3.1.1.

3.1.2 Computing $I_{\succeq}(C_1, C_2)$ when C_1, C_2 are not $\mathbb{R}$-SDC

In this section we consider $I_{\succeq}(C_1, C_2)$ when C_1, C_2 are not $\mathbb{R}$-SDC. We need first to show that if C_1, C_2 are not $\mathbb{R}$-SDC, then $I_{\succeq}(C_1, C_2)$ either is empty or has only one point.

Lemma 3.1.1. *If $C_1, C_2 \in \mathcal{S}^n$ are positive semidefinite then C_1 and C_2 are $\mathbb{R}$-SDC.*

Proof. Since C_1, C_2 are positive semidefinite, $C_1 + C_2 \succeq 0; C_1 + 2C_2 \succeq 0$ and $C_1 + 3C_2 \succeq 0$.

We show that $\text{Ker}(C_1 + 2C_2) \subseteq \text{Ker}C_1 \bigcap \text{Ker}C_2$. Let $x \in \text{Ker}(C_1 + 2C_2)$, we have $(C_1 + 2C_2)x = 0$. Implying $x^T(C_1 + 2C_2)x = 0$. Then, $x \in \mathbb{R}^n$

$$\begin{aligned} 0 \leq & x^T(C_1 + C_2)x = x^T(C_1 + 2C_2)x - x^T C_2 x = -x^T C_2 x \\ \text{and} \quad & x^T C_2 x \geq 0 \end{aligned}$$

which implies that $x^T C_2 x = 0$.

Since $x^T(C_1 + 2C_2)x = 0, x^T C_2 x = 0$, we have $x^T C_1 x = 0$ and $x^T(C_1 + 3C_2)x = 0$.

By $C_1 + 2C_2 \succeq 0; C_1 + 3C_2 \succeq 0$, and $x^T(C_1 + 2C_2)x = 0, x^T(C_1 + 3C_2)x = 0$, we have $(C_1 + 2C_2)x = 0, (C_1 + 3C_2)x = 0$. Implying $C_2 x = 0, C_1 x = 0$. Then $x \in \text{Ker}C_1 \bigcap \text{Ker}C_2$.

By Lemma 1.2.5, C_1 and C_2 are $\mathbb{R}$-SDC. □

Lemma 3.1.2. *If $C_1, C_2 \in \mathcal{S}^n$ are not $\mathbb{R}$-SDC then $I_{\succeq}(C_1, C_2)$ either is empty or has only one element.*

Proof. Suppose on the contrary that $I_{\succeq}(C_1, C_2)$ has more than one elements, then we can choose $\mu_1, \mu_2 \in I_{\succeq}(C_1, C_2), \mu_1 \neq \mu_2$ such that $C := C_1 + \mu_1 C_2 \succeq 0$ and $D := C_1 + \mu_2 C_2 \succeq 0$. By Lemma 3.1.1, C, D are $\mathbb{R}$-SDC, i.e., there is a nonsingular matrix P such that $P^T CP, P^T DP$ are diagonal. Then $P^T C_2 P$ is diagonal because $P^T CP - P^T DP = (\mu_1 - \mu_2) P^T C_2 P$ and $\mu_1 \neq \mu_2$. Since $P^T C_1 P = P^T CP - \mu_1 P^T C_2 P$, $P^T C_1 P$ is also diagonal. That is C_1, C_2 are $\mathbb{R}$-SDC and we get a contradiction. □

To know when $I_{\succeq}(C_1, C_2)$ is empty or has one element, we need the following result.

Lemma 3.1.3 (Theorem 1, [64]). *Let $C_1, C_2 \in \mathcal{S}^n$, C_2 be nonsingular. Let $C_2^{-1} C_1$ have the real Jordan normal form $\textbf{diag}(J_1, \dots J_r, J_{r+1}, \dots, J_m)$, where $J_1, \dots, J_r$ are Jordan blocks corresponding to real eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_r$ of $C_2^{-1} C_1$ and $J_{r+1}, \dots, J_m$ are Jordan blocks for pairs of complex conjugate roots $\lambda_i = a_i \pm \mathbf{i} b_i, a_i, b_i \in \mathbb{R}, i = r+1, r+2, \dots, m$ of $C_2^{-1} C_1$. Then there exists a nonsingular matrix U such that*

$$U^T C_2 U = \textbf{diag}(\epsilon_1 E_1, \epsilon_2 E_2, \dots, \epsilon_r E_r, E_{r+1}, \dots, E_m) \tag{3.12}$$

$$U^T C_1 U = \textbf{diag}(\epsilon_1 E_1 J_1, \epsilon_2 E_2 J_2, \dots, \epsilon_r E_r J_r, E_{r+1} J_{r+1}, \dots, E_m J_m) \tag{3.13}$$

where $\epsilon_i = \pm 1, E_i = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}$; $\dim E_i = \dim J_i = n_i; n_1 + n_2 + \dots + n_m = n$.

Theorem 3.1.4. *Let $C_1, C_2 \in \mathcal{S}^n$ be as in Lemma 3.1.3 and C_1, C_2 are not $\mathbb{R}$-SDC. The followings hold.*

(i) if $C_1 \succeq 0$ then $I_{\succeq}(C_1, C_2) = \{0\}$;

(ii) if $C_1 \not\succeq 0$ and there is a real eigenvalue λ_l of $C_2^{-1} C_1$ such that $C_1 + (-\lambda_l) C_2 \succeq 0$ then

$$I_{\succeq}(C_1, C_2) = \{-\lambda_l\};$$

(iii) if (i) and (ii) do not occur then $I_{\succeq}(C_1, C_2) = \emptyset$.

Proof. It is sufficient to prove only (iii). Lemma 3.1.3 allows us to decompose C_1 and C_2 as the forms (3.13) and (3.12), respectively. Since C_1, C_2 are not $\mathbb{R}$-SDC, at least one of the following cases must occur.

Case 1 *There is a Jordan block J_i such that $n_i \geq 2$ and $\lambda_i \in \mathbb{R}$.* We then consider the following principal minor of $C_1 + \mu C_2$:

$$Y = \epsilon_i(E_iJ_i + \mu E_i) = \epsilon_i \begin{pmatrix} 0 & 0 & \dots & 0 & \lambda_i + \mu \\ 0 & 0 & \dots & \lambda_i + \mu & 1 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \lambda_i + \mu & 1 & \dots & 0 & 0 \end{pmatrix}_{n_i \times n_i}.$$

If $n_i = 2$ then $Y = \epsilon_i \begin{pmatrix} 0 & \lambda_i + \mu \\ \lambda_i + \mu & 1 \end{pmatrix}$. Since $\mu \neq -\lambda_i$, $Y \not\succeq 0$ so $A + \mu B \not\succeq 0$. If $n_i > 2$ then Y always contains the following not positive semidefinite principal minor of size $(n_i - 1) \times (n_i - 1)$:

$$\epsilon_i \begin{pmatrix} 0 & 0 & \dots & \lambda_i + \mu & 1 \\ 0 & 0 & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \lambda_i + \mu & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}_{(n_i-1)\times(n_i-1)}.$$

So $A + \mu B \not\succeq 0$.

Case 2 *There is a Jordan block J_i such that $n_i \geq 4$ and $\lambda_i = a_i \pm \mathbf{i}b_i \notin \mathbb{R}$.* We then consider

$$Y = \epsilon_i(E_iJ_i + \mu E_i) = \epsilon_i \begin{pmatrix} 0 & 0 & \dots & b_i & a_i + \mu \\ 0 & 0 & \dots & a_i + \mu & -b_i \\ \dots & \dots & \dots & \dots & \dots \\ b_i & a_i + \mu & \dots & 0 & 0 \\ a_i + \mu & -b_i & \dots & 0 & 0 \end{pmatrix}_{n_i \times n_i}.$$

This matrix always contains either a principal minor of size 2×2 : $\epsilon_i \begin{pmatrix} b_i & a_i + \mu \\ a_i + \mu & -b_i \end{pmatrix}$ or a principal minor of size 4×4 :

$$\epsilon_i \begin{pmatrix} 0 & 0 & b_i & a_i + \mu \\ 0 & 0 & a_i + \mu & -b_i \\ b_i & a_i + \mu & 0 & 0 \\ a_i + \mu & -b_i & 0 & 0 \end{pmatrix}.$$

Both are not positive semidefinite for any $\mu \in \mathbb{R}$. □

Similarly, we have the following result.

Theorem 3.1.5. *Let $C_1, C_2 \in \mathcal{S}^n$ be not $\mathbb{R}$-SDC. Suppose C_1 is nonsingular and $C_1^{-1}C_2$ has real Jordan normal form* $\textbf{diag}(J_1, \ldots J_r, J_{r+1}, \ldots, J_m)$*, where $J_1, \ldots, J_r$ are Jordan blocks corresponding to real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ of $C_1^{-1}C_2$ and $J_{r+1}, \ldots, J_m$ are Jordan blocks for pairs of complex conjugate roots $\lambda_i = a_i \pm \mathbf{i}b_i, a_i, b_i \in \mathbb{R}, i = r+1, r+2, \ldots, m$ of $C_1^{-1}C_2$.*

(i) If $C_1 \succeq 0$ then $I_{\succeq}(C_1, C_2) = \{0\}$;

(ii) If $C_1 \not\succeq 0$ and there is a real eigenvalue $\lambda_l \neq 0$ of $C_1^{-1}C_2$ such that $C_1 + \left(-\frac{1}{\lambda_l}\right) C_2 \succeq 0$ then $I_{\succeq}(C_1, C_2) = \left\{-\frac{1}{\lambda_l}\right\}$;

(iii) If cases (i) and (ii) do not occur then $I_{\succeq}(C_1, C_2) = \emptyset$.

Finally, if C_1 and C_2 are not $\mathbb{R}$-SDC and both singular. Lemma 1.2.8 indicates that C_1 and C_2 can be simultaneously decomposed as $\tilde{C}_1$ and $\tilde{C}_2$ in either (3.8) or (3.9). If $\tilde{C}_1$ and $\tilde{C}_2$ take the forms (3.8) and $A_2 = 0$ then $I_{\succeq}(C_1, C_2) = I_{\succeq}(A_1, B_1)$, where A_1, B_1 are not $\mathbb{R}$-SDC and B_1 is nonsingular. In this case we apply Theorem 3.1.4 to compute $I_{\succeq}(A_1, B_1)$. If $\tilde{C}_1$ and $\tilde{C}_2$ take the forms (3.9) and $A_2 = 0$. In this case, if A_3 is not positive definite then $I_{\succeq}(C_1, C_2) = \emptyset$. Otherwise, $I_{\succeq}(C_1, C_2) = I_{\succeq}(A_1, B_1)$, where A_1, B_1 are not $\mathbb{R}$-SDC and B_1 is nonsingular, again we can apply Theorem 3.1.4. Therefore we need only to consider the case $A_2 \neq 0$ with noting that $I_{\succeq}(C_1, C_2) \subset I_{\succeq}(A_1, B_1)$.

Theorem 3.1.6. *Given $C_1, C_2 \in \mathcal{S}^n$ are not $\mathbb{R}$-SDC and singular such that $\tilde{C}_1$ and $\tilde{C}_2$ take the forms in either (3.8) or (3.9) with $A_2 \neq 0$. Suppose that $I_{\succeq}(A_1, B_1) = [a, b], a < b$. Then, if $a \notin I_{\succeq}(C_1, C_2)$ and $b \notin I_{\succeq}(C_1, C_2)$ then $I_{\succeq}(C_1, C_2) = \emptyset$.*

Proof. We consider $\tilde{C}_1$ and $\tilde{C}_2$ in (3.9), the form in (3.8) is considered similarly. Suppose in contrary that $I_{\succeq}(C_1, C_2) = \{\mu_0\}$ and $a < \mu_0 < b$. Since $I_{\succeq}(A_1, B_1)$ has more than one point, by Lemma 3.1.2, A_1 and B_1 are $\mathbb{R}$-SDC. Let Q_1 be a $p \times p$ nonsingular matrix such that $Q_1^T A_1 Q_1, Q_1^T B_1 Q_1$ are diagonal, then $Q_1^T(A_1 + \mu_0 B_1)Q_1 := \texttt{diag}(\gamma_1, \gamma_2, \ldots, \gamma_p)$ is a diagonal matrix. Moreover, B_1 is nonsingular, we have $I_{\succ}(A_1, B_1) = (a, b)$, please see Corollary 3.1.1. Then $\gamma_i > 0$ for $i = 1, 2, \ldots, p$ because $\mu_0 \in I_{\succ}(A_1, B_1)$. Let $Q := \texttt{diag}(Q_1, I_s, I_{r-s})$ we then have

$$Q^T(\tilde{C}_1 + \mu_0 \tilde{C}_2)Q = \begin{pmatrix} Q_1^T(A_1 + \mu_0 B_1)Q_1 & 0_{p\times s} & Q_1^T A_2 \\ 0_{s\times p} & A_3 & 0_{s\times(r-s)} \\ A_2^T Q_1 & 0_{(r-s)\times s} & 0_{(r-s)\times(r-s)} \end{pmatrix}.$$

We note that $I_{\succeq}(C_1, C_2) = \{\mu_0\}$ is singleton implying $\det(C_1 + \mu_0 C_2) = 0$ and so $\det(Q^T(\tilde{C}_1 + \mu_0\tilde{C}_2)Q) = 0$. On the other hand, since A_3 is nonsingular diagonal and $A_1 + \mu_0 B_1 \succ 0$, the first $p + s$ columns of the matrix $Q^T(\tilde{C}_1 + \mu_0\tilde{C}_2)Q$ are linearly independent. One of the following cases must occur: i) the columns of the right side submatrix $\begin{pmatrix} Q_1^T A_2 \\ 0_{s\times(r-s)} \\ 0_{(r-s)\times(r-s)} \end{pmatrix}$ are linearly independent and at least one column, suppose $(c_1, c_2, \ldots, c_p, 0, 0, \ldots, 0)^T$, is a linear combination of the columns of the matrix

$$\begin{pmatrix} Q_1^T(A_1 + \mu_0.B_1)Q_1 \\ 0_{s\times p} \\ A_2^T Q_1 \end{pmatrix} := (\mathrm{column_1}|\mathrm{column_2}|\ldots|\mathrm{column_p}),$$

where $\mathrm{column_i}$ is the ith column of the matrix or ii) the columns of the right side submatrix $\begin{pmatrix} Q_1^T A_2 \\ 0_{s\times(r-s)} \\ 0_{(r-s)\times(r-s)} \end{pmatrix}$ are linearly dependent. If the case i) occurs then there are scalars $a_1, a_2, \ldots, a_p$ which are not all zero such that

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_1\mathrm{column_1} + a_2\mathrm{column_2} + \ldots + a_p\mathrm{column_p}. \tag{3.14}$$

Equation (3.14) implies that $\begin{cases} c_1 = a_1\gamma_1 \\ c_2 = a_2\gamma_2 \\ \ldots \\ c_p = a_p\gamma_p \\ 0 = a_1c_1 + a_2c_2 + \ldots + a_pc_p \end{cases}$ which further implies

$$0 = (a_1)^2\gamma_1 + (a_1)^2\gamma_2 + \ldots + (a_p)^2\gamma_p.$$

This cannot happen with $\gamma_i > 0$ and $(a_1)^2 + (a_2)^2 + \ldots + (a_p)^2 \neq 0$. This contradiction shows that $I_{\succeq}(C_1, C_2) = \emptyset$. If the case ii) happens then there always exists a nonsingular matrix H such that

$$H^T Q^T(\tilde{C}_1 + \mu_0 \tilde{C}_2)QH = \begin{pmatrix} Q_1^T(A_1+\mu_0 B_1)Q_1 & 0_{p\times s} & \hat{A}_2 & 0 \\ 0_{s\times p} & A_3 & 0 & 0 \\ \hat{A}_2^T & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

where $\hat{A}_2$ is a full column-rank matrix. Let

$$\hat{C}_1 = \begin{pmatrix} Q_1^T A_1 Q_1 & 0_{p\times s} & \hat{A}_2 \\ 0_{s\times p} & A_3 & 0 \\ \hat{A}_2^T & 0 & 0 \end{pmatrix}, \hat{C}_2 = \begin{pmatrix} Q_1^T B_1 Q_1 & 0_{p\times s} & 0 \\ 0_{s\times p} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

we have $I_{\succeq}(C_1, C_2) = I_{\succeq}(\tilde{C}_1, \tilde{C}_2) = I_{\succeq}(\hat{C}_1, \hat{C}_2)$ and so $I_{\succeq}(\hat{C}_1, \hat{C}_2) = \{\mu_0\}$. This implies $\det(\hat{C}_1 + \mu_0 \hat{C}_2) = 0$, and the right side submatrix $\begin{pmatrix} \hat{A}_2 \\ 0 \\ 0 \end{pmatrix}$ is full column-rank. We return to the case i). □

3.2 Solving the quadratically constrained quadratic programming

We consider the following QCQP problem with m constraints:

$$(\mathrm{P_m}) \quad \begin{array}{ll} \min & f_0(x) = x^T C_0 x + a_0^T x \\ \text{s.t.} & f_i(x) = x^T C_i x + a_i^T x + b_i \le 0, \ i = 1, 2, \ldots, m, \end{array}$$

where $C_i \in \mathcal{S}^n$, $x, a_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$. When C_i are all positive semidefinite, $(\mathrm{P_m})$ is a convex problem, for which efficient algorithms are available such as the interior method [9, Chapter 11]. However, if convexity is not assumed, $(\mathrm{P_m})$ is in general very difficult, even its special form when all constraints are affine, i.e., $C_i = 0$ for $i = 1, 2, \ldots, m$, and C_0 is indefinite, is already NP-hard [66, 51].

If $C_0, C_1, \ldots, C_m$ are $\mathbb{R}$-SDC, a congruence matrix R is obtained so that

$$R^T C_i R = \texttt{diag}(\alpha_1^i, \ldots, \alpha_n^i).$$

By change of variables $x = Ry$, the quadratic forms $x^T C_i x$ become the sums of squares in y. That is,

$$x^T C_i x = y^T R^T C_i R y = \sum_{j=1}^{n} \alpha_j^i y_j^2.$$

Set $\alpha_i = (\alpha_1^i, \ldots, \alpha_n^i)^T$, $\xi_i = R^T a_i$ and $z_j = y_j^2, j = 1, 2, \ldots, n$, $(\mathrm{P_m})$ is then rewritten as follows.

$$(\mathrm{P_m}) \qquad \begin{array}{ll} \min & f_0(y,z) = \alpha_0^T z + \xi_0^T y \\ \text{s.t.} & f_i(y,z) = \alpha_i^T z + \xi_i^T y + b_i \leq 0, \ i = 1, 2, \ldots, m, \\ & y_j^2 = z_j, j = 1, 2, \ldots, n. \end{array} \tag{3.15}$$

The constraints $y_j^2 = z_j$ are not convex. By relaxing $y_j^2 \leq z_j$ for $j = 1, 2, \ldots, n$, we get the following relaxation of $(\mathrm{P_m})$:

$$(\mathrm{SP_m}) \qquad \begin{array}{ll} \min & f_0(y,z) = \alpha_0^T z + \xi_0^T y \\ \text{s.t.} & f_i(y,z) = \alpha_i^T z + \xi_i^T y + b_i \leq 0, \ i = 1, 2, \ldots, m, \\ & y_j^2 \leq z_j, j = 1, 2, \ldots, n. \end{array} \tag{3.16}$$

The problem $(\mathrm{SP_m})$ is a convex second-order cone programming (SOCP) problem and it can be solved in polynomial time by the interior algorithm [21].

Because of the relaxation $y_j^2 \leq z_j$, the optimal value of $(\mathrm{SP_m})$ is less than that of $(\mathrm{P_m})$. That is $v((\mathrm{SP_m})) \leq v((\mathrm{P_m}))$, here $v(\cdot)$ is the optimal value of the problem $(\cdot)$. In other words, the convex SOCP problem $(\mathrm{SP_m})$ is a lower bound of $(\mathrm{P_m})$. The relaxation is said to be tight, or exact, if $v((\mathrm{SP_m})) = v((\mathrm{P_m}))$, and in that case, the nonconvex problem $(\mathrm{P_m})$ is equivalently transformed to a convex problem $(\mathrm{SP_m})$. In 2014, Ben-Tal and Hertog [6] showed that $v((\mathrm{SP_1})) = v((\mathrm{P_1}))$ under the Slater condition, i.e., there is $\bar{x} \in \mathbb{R}^n$ such that $f_1(\bar{x}) < 0$, and $v((\mathrm{SP_2})) = v((\mathrm{P_2}))$ under some additional appropriate assumptions. In 2019, Adachi and Nakatsukasa [1] proposed an eigenvalue-based algorithm for a definite feasible $(\mathrm{P_1})$, i.e., the Slater condition is satisfied and the positive definite interval $I_{\succ}(C_0, C_1) = \{\mu \in \mathbb{R} : C_0 + \mu C_1 \succ 0\}$ is nonempty. It should be noticed that $I_{\succ}(C_0, C_1)$ can be empty even if $I_{\succeq}(C_0, C_1)$ is an interval and $(\mathrm{P_1})$ has optimal solutions. In the following, we explore the SDC of $C_i's$ to apply for some special cases of $(\mathrm{P_m})$.

3.2.1 Application for the GTRS

We write $(\mathrm{P_1})$ specifically as follows.

$$(\mathrm{P_1}) \qquad \begin{array}{ll} \min & f_0(x) = x^T C_0 x + a_0^T x \\ \text{s.t.} & f_1(x) = x^T C_1 x + a_1^T x + b_1 \leq 0. \end{array}$$

Problem $(\mathrm{P_1})$ itself arises from many applications such as time of arrival problems [32], double well potential problems [17], subproblems of consensus ADMM in solving quadratically constrained quadratic programming in signal processing [36]. In particular, it includes the trust-region subproblem (TRS) as a special case, in which $C_1 = I$ is the identity matrix, $a_1 = 0$ and $b_1 = -1$. In literature, it is thus often referred to as the generalized trust region subproblem (GTRS).

Without loss of generality, we only solve problem (P_1) under the Slater condition, i.e., there exists $\bar{x} \in \mathbb{R}^n$ such that $f_1(\bar{x}) < 0$. Because, if the Slater condition is violated, then $f_1(x) \geq 0$ for all $x \in \mathbb{R}^n$. Problem (P_1) is then either infeasible or reduced to an unconstrained quadratic problem, which can be solved efficiently [72].

In 1993, Moré [44] obtained the following important results for (P_1).

Lemma 3.2.1 ([44], Theorem 3.4)**.** *Suppose the Slater condition is satisfied. A vector $x^* \in \mathbb{R}^n$ is an optimal solution to (P_1) if and only if there exists $\mu^* \geq 0$ such that*

$$(C_0 + \mu^* C_1)x^* + a_0 + \mu^* a_1 = 0, \tag{3.17}$$

$$f_1(x^*) \leq 0, \tag{3.18}$$

$$\mu^* f_1(x^*) = 0, \tag{3.19}$$

$$C_0 + \mu^* C_1 \succeq 0. \tag{3.20}$$

Recall that $I_{\succ}(C_0, C_1) = \{\mu \in \mathbb{R} : C_0 + \mu C_1 \succ 0\}$.

Lemma 3.2.2 ([44])**.** *If $I_{\succ}(C_0, C_1)$ is nonempty, it is an open interval. Moreover, if μ is a finite endpoint of $I_{\succ}(C_0, C_1)$ then $C_0 + \mu C_1$ is positive semidefinite but not positive definite.*

Suppose $I_{\succ}(C_0, C_1) \neq \emptyset$, let $\varphi(\mu) := f_1[x(\mu)]$, where $x(\mu)$ is solved from the linear equation (3.17) and $\mu \in I_{\succ}(C_0, C_1)$.

Lemma 3.2.3 ([44], Theorem 5.2)**.** *Suppose $I_{\succ}(C_0, C_1) \neq \emptyset$. The function $\varphi(\mu)$ is strictly decreasing on $I_{\succ}(C_0, C_1)$, unless $x(\mu)$ is constant on $I_{\succ}(C_0, C_1)$ with $C_0 x(\mu) + a_0 = 0$ and $C_1 x(\mu) + a_1 = 0$ for all $\mu \in I_{\succ}(C_0, C_1)$.*

Lemmas 3.2.1, 3.2.2 and 3.2.3 together indicate that the optimal Lagrange multiplier μ^* of (P_1) can be found efficiently whenever $I_{\succeq}(C_0, C_1)$ is computed. Using the results in Subsection 3.2.1, we present algorithms for finding μ^* and $x^* = x(\mu^*)$ satisfying (3.17)-(3.20) as follows. Let $I = I_{\succeq}(C_0, C_1) \cap [0, \infty)$ denote the set of Lagrange multipliers of (P_1).

1. If $I = \emptyset$, then (P_1) has no optimal solution, it is in fact unbounded from below in this case [72].

2. If I has only one value μ, we solve the linear equation (3.17) for a corresponding solution $x(\mu)$. If μ and $x(\mu)$ satisfy (3.18)-(3.19), then $\mu^* = \mu$ and $x^* = x(\mu)$. Otherwise, (P_1) has no optimal solution.

3. If I is an interval, we need to detect whether there exist a $\mu \in I$ and a corresponding $x(\mu)$ satisfying (3.18)-(3.19). This case raises two questions: 1) how to test whether μ and $x(\mu)$ satisfy (3.18)-(3.19)? and 2) how to pick another $\mu \in I$ to continue the process if the current μ and $x(\mu)$ do not satisfy (3.18)-(3.19)? For question 1), if $\mu = 0$ we need to test whether $f_1(x(\mu)) \leq 0$; if $\mu > 0$ we need to test $f_1(x(\mu)) = 0$. Below, we present only checking the case $f_1(x(\mu)) = 0$ since checking $f_1(x(\mu)) \leq 0$ is done similarly. For question 2), we need to use Lemma 3.2.2 but not only for the case $I_{\succ}(C_0, C_1) \neq \emptyset$ but also $I_{\succ}(C_0, C_1) = \emptyset$. The details are as below.

Theorem 3.2.1. *If $\mu^* > 0$, then an optimal solution x^* of (P_1) is found by solving a quadratic equation.*

Proof. Since $\mu^* > 0$, x^* is an optimal solution of (P_1) if and only if x^* satisfies (3.17) and $f_1(x^*) = 0$. From the equation (3.17), x^* is of the form

$$x^* = x^0 + Ny, \tag{3.21}$$

where $x^0 = -(C_0 + \mu^* C_1)^+(a + \mu^* b)$, $(C_0 + \mu^* C_1)^+$ is the Moore-Penrose generalized inverse of the matrix $C_0 + \mu^* C_1$, $N \in \mathbb{R}^{n\times r}$ is a basic matrix for the null space of $C_0 + \mu^* C_1$ with $r = n - \mathrm{rank}(C_0 + \mu^* C_1), y \in \mathbb{R}^r$. Notice that the Moore-Penrose generalized inverse of a matrix $A \in \mathbb{F}^{m\times n}$ is defined as a matrix $A^+ \in \mathbb{F}^{n\times m}$ satisfying all of the following four criteria: 1) $AA^+A = A$; 2)$A^+AA^+ = A^+$; 3)$(AA^+)^* = AA^+$; 4)$(A^+A)^* = A^+A$. If $r = 0$ then $x^* = x^0 = (C_0 + \mu^* C_1)^{-1}(a + \mu^* b)$ is the unique solution of (3.17), checking if $f_1(x^*) = 0$ is then simply substituting x^* into $f_1(x)$. If $r > 0$, $f_1(x^*)$ is then a quadratic function of y as follows:

$$\begin{aligned} f_1(x^*) &= f_1(x^0 + Ny) \\ &= y^T(N^T C_1 N)y + 2(N^T(C_1x^0 + b))^T y + {x^0}^T C_1 x^0 + 2b^T x^0 + c \\ &:= y^T \tilde{C}_1 y + 2\tilde{b}^T y + \tilde{c} := \tilde{g}(y), \end{aligned}$$

where $\tilde{C}_1 = N^T C_1 N, \tilde{b} = N^T(C_1x^0 + b)$ and $\tilde{c} = {x^0}^T C_1 x^0 + 2b^T x^0 + c$. Checking whether $f_1(x^*) = 0$ is now equivalent to finding a solution y^* of the quadratic equation $\tilde{g}(y) = 0$. Making diagonal if necessary, we can suppose that $\tilde{C}_1 = \mathtt{diag}(\lambda_1, \ldots, \lambda_r)$ is already diagonal. The equation $\tilde{g}(y) = 0$ is then simply of the form

$$\sum_{i=1}^{r} \lambda_i y_i^2 + 2\sum_{i=1}^{r} \tilde{b}_i y_i + \tilde{c} = 0, \tag{3.22}$$

here $\tilde{b} = (\tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_r)^T$ and $y = (y_1, y_2, \ldots, y_r)^T$. Solving a solution y^* of this equation is as follows.

1. If there is an index i such that $\lambda_i = 0$ and $\tilde{b}_i \neq 0$, then

$$y^* = \underbrace{(0, \ldots, 0, -\frac{\tilde{c}}{2\tilde{b}_i}}_{i\text{th position}}, 0, \ldots, 0)^T$$

   is a solution of (3.22), and $x^* = x^0 + Ny^*$ is then an optimal solution to (P$_1$). Note that if $\lambda_i = 0$ and $\tilde{b}_i = 0$, then y_i does not play any role in $\tilde{g}(y) = 0$.

2. If $\lambda_t > 0$ and $\lambda_j < 0$ for some indexes t, j, suppose $t < j$, we then set $y_i = 0$ for all $i \neq t, i \neq j$, such that the equation (3.22) is reduced to

$$\lambda_t y_t^2 + \lambda_j y_j^2 + 2\bar{b}_t y_t + 2\bar{b}_j y_j + \bar{c} = 0.$$

   We write this equation in term of a quadratic equation of y_t with parameter y_j :

$$\lambda_t y_t^2 + 2\bar{b}_t y_t + \lambda_j y_j^2 + 2\bar{b}_j y_j + \bar{c} = 0. \tag{3.23}$$

   Let $\triangle(y_j) = \bar{b}_t^2 - \lambda_t(\lambda_j y_j^2 + 2\bar{b}_j y_j + \bar{c}) = -\lambda_t\lambda_j y_j^2 - 2\bar{b}_j\lambda_t y_j - \bar{c}\lambda_t + \bar{b}_t^2$. Since $-\lambda_t\lambda_j > 0$, $\triangle(y_j) \geq 0$ when $|y_j|$ is large enough. So we can choose y_j^* such that $\triangle(y_j^*) \geq 0$ and $y_t^* = \frac{-\bar{b}_t + \sqrt{\triangle(y_j^*)}}{\lambda_t}$. Then (y_t^*, y_j^*) is a solution of (3.23) and

$$y^* = (0, \ldots, 0, y_t^*, 0, \ldots, 0, y_j^*, 0 \ldots, 0)^T$$

   is a solution of (3.22). So $x^* = x^0 + Ny^*$ is optimal to (P_1).

3. If $\lambda_i > 0$ for all $i = 1, 2, \ldots, r$, the equation (3.22) can be rewritten as follows

$$\sum_{i=1}^{r} \lambda_i \left(y_i + \frac{\tilde{b}_i}{\lambda_i} \right)^2 + \beta = 0, \tag{3.24}$$

   where $\beta = \tilde{c} - \sum_{i=1}^{r} \frac{\tilde{b}_i^2}{\lambda_i}$. Now

   - if $\beta > 0$ then the equation $\tilde{g}(y) = 0$ has no solution so does the equation $f_1(x^*) = 0$. (P$_1$) has no optimal solution.
   - if $\beta = 0$, let $y^* = \left(-\frac{\tilde{b}_1}{\lambda_1}, -\frac{\tilde{b}_2}{\lambda_2}, \ldots, -\frac{\tilde{b}_r}{\lambda_r} \right)^T$, then $x^* = x^0 + Ny^*$ is an optimal solution of (P$_1$).
   - if $\beta < 0$, then $y^* = \left(-\frac{\tilde{b}_1}{\lambda_1}, -\frac{\tilde{b}_2}{\lambda_2}, \ldots, -\frac{\tilde{b}_{r-1}}{\lambda_{r-1}}, \sqrt{\frac{-\beta}{\lambda_r}} - \frac{\tilde{b}_r}{\lambda_r} \right)$ is a solution of (3.24). Then $x^* = x^0 + Ny^*$ is optimal to (P$_1$).

□

We emphasize that if C_0, C_1 are $\mathbb{R}$-SDC, the linear equation (3.17) can be transformed to having a simple form for solving. Indeed, without loss of generality we assume that C_0, C_1 are already diagonal:

$$C_0 = \texttt{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n), C_1 = \texttt{diag}(\beta_1, \beta_2, \ldots, \beta_n). \tag{3.25}$$

The linear equation (3.17) is then of the following simple form

$$(\alpha_i + \mu\beta_i)x_i = -(a_i + \mu b_i), i = 1, 2, \ldots, n. \tag{3.26}$$

If I has only one element μ, testing whether $\mu^* = \mu$ has been presented in the previous subsection. If I is an interval of the form $I = [\mu_1, \mu_2]$, where $\mu_1 \geq 0$ and μ_2 may be ∞, we need to test whether there is an optimal Lagrange multiplier $\mu^* \in I$ satisfying $\varphi(\mu^*) = 0$. We note that in this case C_0, C_1 are $\mathbb{R}$-SDC, see Lemma 3.1.2. For simplicity in presentation, we assume without loss of generality that C_0, C_1 are diagonal taking the form (3.25). The testing strategy is considered in the following two separate cases: $I_{PD} \neq \emptyset$ and $I_{PD} = \emptyset$, where $I_{PD} = I_{\succ}(C_0, C_1) \cap [0, +\infty)$.

Definition 3.2.1 ([1])**.** A GTRS satisfying the following two conditions is said to be *definite feasible.*

1. It is strictly feasible: there exists $\bar{x} \in \mathbb{R}^n$ such that $f_1(\bar{x}) < 0$, and
2. $I_{PD} \neq \emptyset$

Case 1: $I_{PD} \neq \emptyset$. Then (P_1) is definite feasible and it has a unique optimal solution x^* [44, Theorem 4.1] and, importantly, I is then the closure of $I_{PD} : I = \text{closure}(I_{PD})$, please see [44, Theorem 5.3]. By Lemma 3.2.3, the function $\varphi(\mu) = f_1[x(\mu)]$ is strictly decreasing on I_{PD}, unless $x(\mu)$ is constant on I_{PD}. Using this property of $\varphi(\mu)$, Adachi et al. [1] obtain the following result.

Lemma 3.2.4 ([1])**.** *Suppose the Slater condition holds for the* (P_1)*, i.e., there exists* $\tilde{x} \in \mathbb{R}^n$ *such that* $f_1(\tilde{x}) < 0$*, and* $I_{PD} \neq \emptyset$*.*

(a) If $\varphi(\mu) > 0$ *on* I_{PD} *and* $\mu_2 < \infty$*, then* $\mu^* = \mu_2$*;*

(b) If $\varphi(\mu) < 0$ *on* I_{PD} *then* $\mu^* = \mu_1$*;*

(c) If $\varphi(\mu)$ *changes its sign on* I_{PD} *then* $\mu^* \in I_{PD}$*;*

(d) If $\varphi(\mu_1) > 0$ *and* $\mu_2 = \infty$*, then* $\mu_1 < \mu^* < \mu_2$*.*

Lemma 3.2.4 suggests a strategy for finding μ^* as follows: If μ_2 is finite, we compute $\varphi(\mu)$ at endpoints: if $\varphi(\mu_1) = 0$ then $\mu^* = \mu_1$, if $\varphi(\mu_2) = 0$ then $\mu^* = \mu_2$. Otherwise, $\mu^* \in I_{PD}$. Then, we use a bisection algorithm for finding μ^* : let $\tilde{\mu} := \dfrac{\mu_1 + \mu_2}{2}$. If $\varphi(\mu_1)\varphi(\tilde{\mu}) < 0$ then set $\mu_2 := \tilde{\mu}$, else set $\mu_1 := \tilde{\mu}$ and continue the process with new μ_1 and μ_2. If $\mu_2 = \infty$ and $\varphi(\mu_1) > 0$ depending on how large the value $\varphi(\mu_1)$, we choose a positive number l, for example $l = \varphi(\mu_1)$, and set $\mu = \mu_1 + l$. If $\varphi(\mu) < 0$, we apply a bisection algorithm as mentioned above to find μ^* in $[\mu_1, \mu]$. If $\varphi(\mu) > 0$, we choose other $\mu := \mu_1 + 2l$ and continue the process.

Case 2: $I_{PD} = \emptyset$. As mentioned, (P_1) with $I_{PD} = \emptyset$ is referred to as the *hard case* [44, 33]. We now deal with this case as follows.

Theorem 3.2.2. *If I is an interval and $I_{PD} = \emptyset$ then (P_1) either is reduced to a definite feasible GTRS of smaller dimension or has no optimal solution.*

Proof. Since $I_{PD} = \emptyset$, by Corollary 3.1.1, C_0, C_1 are singular and decomposable in one of the forms (3.10) and (3.11) such that

$$I_{\succeq}(C_0, C_1) = I_{\succeq}(A_1, B_1) = \text{closure}\left(I_{\succ}(A_1, B_1)\right),$$

where B_1 is nonsingular. C_0, C_1 are assumed to be diagonal, the forms (3.10) and (3.11) are written as

$$C_1 = \texttt{diag}(\beta_1, \ldots, \beta_p, 0, \ldots, 0), C_0 = \texttt{diag}(\alpha_1, \ldots, \alpha_p, 0, \ldots, 0) \tag{3.27}$$

and

$$C_1 = \texttt{diag}(\beta_1, \ldots, \beta_p, 0, \ldots, 0), C_0 = \texttt{diag}(\alpha_1, \ldots, \alpha_p, \alpha_{p+1}, \ldots, \alpha_{p+s}, 0, \ldots, 0), \tag{3.28}$$

where $B_1 = \texttt{diag}(\beta_1, \beta_2, \ldots, \beta_p)$, $A_1 = \texttt{diag}(\alpha_1, \alpha_2, \ldots, \alpha_p)$ and

$$A_4 = \texttt{diag}(\alpha_{p+1}, \ldots, \alpha_{p+s}, 0, \ldots, 0).$$

Since B_1 is nonsingular $\beta_1, \beta_2, \ldots, \beta_p$ are nonzero.

If C_0, C_1 take the form (3.27), the equations (3.26) become

$$\begin{aligned}(\alpha_i + \mu\beta_i)x_i &= -(a_i + \mu b_i), i = 1, 2, \ldots, p;\\ 0 &= -(a_i + \mu b_i), i = p+1, \ldots, n.\end{aligned} \tag{3.29}$$

Observe now that if $a_i = b_i = 0$ for $i = p+1, \ldots, n$, then the (P_1) is reduced to a definite feasible GTRS of p variables with matrices A_1, B_1 such that $I_{\succ}(A_1, B_1) \neq$

$\emptyset$. Otherwise, if there are indexes $p+1 \le i, j \le n$ such that $b_i \neq 0, b_j \neq 0$ and $\frac{a_i}{b_i} \neq \frac{a_j}{b_j}$, then (3.29) has no solution x for all $\mu \in I$, if $b_i \neq 0$ and $\mu = -\frac{a_i}{b_i} \in I$ for some $p+1 \le i \le n$ then (3.29) may have solutions at only one $\mu \in I$. Checking whether $\mu^* = \mu$ has been discussed in the previous section.

Similarly, if C_0, C_1 take the form (3.28), the equations (3.26) become

$$\begin{aligned}(\alpha_i + \mu\beta_i)x_i &= -(a_i + \mu b_i), i = 1, 2, \dots, p; \\ \alpha_i x_i &= -(a_i + \mu b_i), i = p+1, p+2, \dots, p+s; \\ 0 &= -(a_i + \mu b_i), i = p+s+1, \dots, n.\end{aligned} \tag{3.30}$$

(P_1) either is reduced to a definte feasible GTRS of $p+s$ variables with matrices

$$\begin{aligned}\tilde{A}_1 &= \mathtt{diag}(A_1, \alpha_{p+1}, \dots, \alpha_{p+s}) = \mathtt{diag}(\alpha_1, \dots, \alpha_p, \alpha_{p+1}, \dots, \alpha_{p+s}), \\ \tilde{B}_1 &= \mathtt{diag}(B_1, \underbrace{0, \dots, 0}_{s \text{ zeros}}) = \mathtt{diag}(\beta_1, \dots, \beta_p, \underbrace{0, \dots, 0}_{s \text{ zeros}})\end{aligned}$$

such that $I_{\succ}(\tilde{A}_1, \tilde{B}_1) \neq \emptyset$, or has no solution x for all $\mu \in I$ or has only one Lagrange multiplier $\mu \in I$. $\square$

Example 3.2.1. Consider the following problem:

$$\begin{aligned}\min \quad & f(x) = x^T C_0 x + 2a^T x \\ \text{s.t.} \quad & g(x) = x^T C_1 x + 2b^T x + c \le 0,\end{aligned} \tag{3.31}$$

where

$$C_0 = \begin{pmatrix} 2 & -12 & -12 \\ -12 & -10 & 4 \\ -12 & 4 & 20 \end{pmatrix}, C_1 = \begin{pmatrix} 3 & 4 & -1 \\ 4 & 13 & 5 \\ -1 & 5 & 4 \end{pmatrix}, a = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, b = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}, c = 5.$$

We have

$$C_1^{-1} C_0 = \begin{pmatrix} \frac{5}{2} & -\frac{1}{2} & -7 \\ -\frac{3}{2} & -\frac{11}{6} & \frac{7}{3} \\ -\frac{1}{2} & \frac{19}{6} & \frac{1}{3} \end{pmatrix}$$

is not similar to a diagonally real matrix, C_0 and C_1 are not $\mathbb{R}$-SDC. By Theorem 3.1.4, we have $I_{\succeq}(C_0, C_1) = \{2\}$.

Now, solving $x(\mu)$, where $\mu = 2$ and checking if $g(x(\mu)) = 0$.

Firstly, we solve the linear equation $(C_0 + 2C_1)x = -(a + 2b)$. This equation is equivalent to

$$\begin{cases} 8x_1 - 4x_2 - 14x_3 & = 1 \\ -4x_1 + 16x_2 + 14x_3 & = 3 \\ 8x_1 - 4x_2 - 14x_3 & = 1 \end{cases} \Leftrightarrow \begin{cases} x_1 & = y_1 \\ x_2 & = \dfrac{1}{3} - \dfrac{y_1}{3} \\ x_3 & = -\dfrac{1}{6} + \dfrac{2y_1}{3} \end{cases},$$

where $y_1 \in \mathbb{R}$.

Put $x(\mu) = (x_1, x_2, x_3)^T := x^0 + N.y$, where $x^0 = \left(0, \frac{1}{3}, -\frac{1}{6}\right)^T$, $N = \begin{pmatrix} 1 \\ -\frac{1}{3} \\ \frac{2}{3} \end{pmatrix}$, and $y = y_1$.

Now, substituting $x(\mu)$ into $g(x(\mu))$, we get $\bar{g}(y) = -\frac{2}{3}y_1 + \frac{17}{3}$. Solving the equation $\bar{g}(y) = 0$, we have $y^* = y_1 = \frac{17}{2}$. And $x^* = x^0 + N.y = \left(\frac{17}{2}, -\frac{5}{2}, \frac{33}{6}\right)^T$ is then an optimal solution to the GTRS (3.31).

Example 3.2.2. Consider the following problem:

$$\begin{array}{ll} \min & f(x) = x^T C_0 x + 2a^T z \\ \text{s.t.} & g(x) = x^T C_1 x + 2b^T x + c \leq 0, \end{array} \tag{3.32}$$

where

$$C_0 = \begin{pmatrix} 4 & 4 & 0 & 2 \\ 4 & 8 & 4 & 4 \\ 0 & 4 & 4 & 2 \\ 2 & 4 & 2 & 2 \end{pmatrix}, C_1 = \begin{pmatrix} 2 & 4 & 2 & 2 \\ 4 & 18 & 4 & 34 \\ 2 & 4 & 2 & 2 \\ 2 & 34 & 2 & 92 \end{pmatrix}, a = \begin{pmatrix} -2 \\ -8 \\ -6 \\ -4 \end{pmatrix}, b = \begin{pmatrix} 4 \\ -8 \\ 8 \\ -54 \end{pmatrix}, c = 4.$$

We have C_0, C_1 are $\mathbb{R}$-SDC by $U = \begin{pmatrix} 3 & -1 & -3 & -5 \\ -3 & 1 & 3 & 6 \\ 3 & -1 & -2 & -5 \\ 1 & 0 & -1 & -2 \end{pmatrix}$ and

$$\tilde{C}_1 = U^T C_1 U = \texttt{diag}(2, 10, 0, 0)$$
$$\tilde{C}_0 = U^T C_0 U = \texttt{diag}(2, 0, 2, 0)$$

Put $x = Uy$, then the problem (3.32) is equivalent to the following problem:

$$\begin{array}{ll} \min & f(y) = y^T \tilde{C}_0 y + 2\bar{a}^T y \\ \text{s.t.} & g(y) = y^T \tilde{C}_1 y + 2\bar{b}^T y + c \leq 0, \end{array} \tag{3.33}$$

where

$$\bar{a} = (-4, 0, -2, 0)^T := (\bar{a}_1, \bar{a}_2, \bar{a}_3, \bar{a}_4)^T, \bar{b} \ = (6, -20, 2, 0)^T := \left(\bar{b}_1, \bar{b}_2, \bar{b}_3, \bar{b}_4\right)^T, c = 4.$$

Since $\bar{a}_4 = \bar{b}_4 = 0$, the problem (3.33) is reduced to a GTRS of 3 variables:

$$\begin{array}{ll} \min & f(y) = y^T A_1 y + 2a_1^T y \\ \text{s.t.} & g(y) = y^T B_1 y + 2b_1^T y + c \leq 0, \end{array} \tag{3.34}$$

where

$$\begin{array}{l} A_1 = \texttt{diag}(2, 0, 2); B_1 = \texttt{diag}(2, 10, 0) \\ a_1 = (-4, 0, -2)^T, b_1 = (6, -20, 2)^T, c = 4. \end{array}$$

By Theorem 3.1.3, we have $I_{\succeq}(A_1, B_1) = [0, +\infty)$.

For $\mu > 0$, we solve the linear equation $(A_1 + \mu B_1)y = -(a_1 + \mu b_1)$. The solution of this equation is $y(\mu) = \left(\dfrac{2-3\mu}{\mu+1}, 2, 1-\mu\right)^T$. And $\varphi(\mu) = g(y(\mu)) = -2\mu\left(\dfrac{25(\mu+2)}{(\mu+1)^2} + 2\right) < 0, \forall \mu > 0$. By Lemma 3.2.4, $\mu^* = 0$.

Now, substituting $\mu^* = 0$ into the linear equation $(A_1 + \mu^* B_1)y = -(a_1 + \mu^* b_1)$, we get

$$y(\mu^*) = (2, z_1, 1)^T := y^0 + N.z$$

where $y^0 = (2, 0, 1)^T, N = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and $z = z_1$.

Next, substituting $y(\mu^*)$ into $g(y(\mu^*))$, we get $\bar{g}(y) = 10y_1^2 - 40y_1 + 40$. Solving the equation $\bar{g}(y) = 0$, we have $z^* = z_1 = 2$. And $y^* = y^0 + N.z = (2, 2, 1)^T$ is an optimal solution to the GTRS (3.34). Implying $x^* = U(2, 2, 1, 0) = (1, -1, 2, 1)^T$ is then an optimal solution to the GTRS (3.32).

3.2.2 Applications for the homogeneous QCQP

If $(\mathrm{P_m})$ is homogeneous, i.e., $a_i = 0, i = 0, 1, \ldots, m$ and $C_0, C_1, \ldots, C_m$ are $\mathbb{R}$-SDC, then we do not need relax the constraints $z_j = y_j^2$ to $z_j \leq y_j^2$ but we can directly convert (3.15) to a linear programming in non-negative variables z_j as follows.

$$(\mathrm{LP_m}) \quad \begin{array}{lll} \lambda^* = & \min & \sum_{j=1}^n \alpha_j^0 z_j \\ & \text{s.t.} & \sum_{j=1}^n \alpha_j^i z_j + b_i \leq 0, \ i = 1, 2, \ldots, m, \\ & & z_j \geq 0, \ j = 1, 2, \ldots, n. \end{array}$$

The simplex algorithm is now applied for solving ($\mathrm{LP_m}$). Suppose $z^* = (z_1^*, z_2^*, \ldots, z_n^*)^T$ is an optimal solution of ($\mathrm{LP_m}$), then we define $y^* = (\sqrt{z_1^*}, \sqrt{z_2^*}, \ldots, \sqrt{z_n^*})^T$ and obtain an optimal solution x^* of the homogeneous ($\mathrm{P_m}$) as $x^* = Ry^*$.

We revisit the following special case of the homogeneous ($\mathrm{P_m}$) :

$$(\mathrm{Q}) \qquad \begin{array}{ll} \min & f_0(x) = x^T C_0 x \\ \text{s.t.} & f_1(x) = x^T C_1 x + b_1 \le 0, \\ & f_2(x) = x^T C_2 x + b_2 \le 0, \\ & f_3(x) = \|x\| = 1. \end{array}$$

It was shown in [46] that if the *Property J* fails, then (Q) is converted to an SDP problem, please see [46, Definition 1] for details on Property J. However, as mentioned, when n is large, the SDP problem is not solved efficiently. The following result can help to deal with such case if the SDC conditions hold.

Theorem 3.2.3. *If C_0, C_1, C_2 are $\mathbb{R}$-SDC by an orthogonal congruence matrix then* (Q) *is reduced to a linear programing problem over the unit simplex.*

Proof. Suppose C_0, C_1, C_2 are $\mathbb{R}$-SDC by an orthogonal congruence matrix R :

$$R^T C_i R = \mathtt{diag}(\alpha_1^i, \ldots, \alpha_n^i), i = 0, 1, 2.$$

We note that the constraint $\|x\| = 1$ is equivalently written as $\|x\|^2 = 1$ which is further written $x^T x = 1$. We make a change of coordinates $x = Ry$ and notice that $x^T x = y^T (R^T R) y = y^T y$. Then (Q) is rewritten as follows

$$(\mathrm{Q}) \qquad \begin{array}{ll} \min & \sum_{j=1}^n \alpha_j^0 y_j^2 \\ \text{s.t.} & \sum_{j=1}^n \alpha_j^1 y_j^2 + b_1 \le 0, \\ & \sum_{j=1}^n \alpha_j^2 y_j^2 + b_2 \le 0, \\ & \sum_{j=1}^n y_j^2 = 1. \end{array}$$

Let $z_j = y_j^2$, problem (Q) is then reduced to a linear programming problem over the unit simplex as follows.

$$(\mathrm{Q_1}) \qquad \begin{array}{ll} \min & \sum_{j=1}^n \alpha_j^0 z_j \\ \text{s.t.} & \sum_{j=1}^n \alpha_j^1 z_j + b_1 \le 0, \\ & \sum_{j=1}^n \alpha_j^2 z_j + b_2 \le 0, \\ & \sum_{j=1}^n z_j = 1, z_j \ge 0 \;\; j = 1, 2, \ldots, n. \end{array}$$

□

We should note that if the SDC conditions of $C_0, C_1, \ldots, C_m$ fail, even ($\mathrm{P_m}$) is homogeneous, it is still very hard to solve. Only some special cases have been discovered to be solved in polynomial time but by SDP relaxation, see for example [73].

3.3 Applications for maximizing a sum of generalized Rayleigh quotients

Given $n \times n$ matrices A, B. The ratio $R(A;x) := \frac{x^TAx}{x^Tx}, x \neq 0$, is called the Rayleigh quotient of the matrix A and $R(A,B;x) = \frac{x^TAx}{x^TBx}, B \succ 0$, is known as the generalized Rayleigh quotient of (A,B). We know that

$$\min_{x\neq 0} R(A;x) = \lambda_{\min}(A) \leq R(A;x) \leq \lambda_{\max}(A) = \max_{x\neq 0} R(A;x),$$

where $\lambda_{\min}(A), \lambda_{\max}(A)$ are the smallest and largest eigenvalues of A, respectively. Similarly,

$$\min_{x\neq 0} R(A,B;x) = \lambda_{\min}(A,B) \leq R(A,B;x) \leq \lambda_{\max}(A,B) = \max_{x\neq 0} R(A,B;x),$$

where $\lambda_{\min}(A,B), \lambda_{\max}(A,B)$ are the smallest and largest generalized eigenvalues of (A,B), respectively [34].

Due to the homogeneity: $R(A;x) = R(A;cx), R(A,B;x) = R(A,B;cx)$, for any non-zero scalar c, it holds that

$$\min(\max)_{x\neq 0} R(A;x) = \min(\max)_{\|x\|=1} R(A;x); \tag{3.35}$$

$$\min(\max)_{x\neq 0} R(A,B;x) = \min(\max)_{\|x\|=1} R(A,B;x). \tag{3.36}$$

Both (3.35) and (3.36) do not admit local non-global solution [22, 23] and they can be solved efficiently. However, difficulty will arise when we attemp to optimize a sum.

We consider the following simplest case of the sum:

$$\max_{x\neq 0} \frac{x^TA_1x}{x^TB_1x} + \frac{x^TA_2x}{x^TB_2x}, \tag{3.37}$$

where $B_1 \succ 0, B_2 \succ 0$. This problem has various applications such as for the downlink of a multi-user MIMO system [53], for the sparse Fisher discriminant analysis in pattern recognition and many others, please see [16, 20, 71, 75, 76, 48, 60, 69]. Zhang [75] showed that (3.37) admit many local-non global optima, please see [75, Example 3.1]. It is thus very hard to solve. Many studies later [75, 76, 46, 69] proposed different approximate methods for it. However, if the SDC conditions hold for (3.37), it can be equivalently reduced to a linear programming on the simplex [69]. We present in detail this conclusion as follows. Since $B_1 \succ 0$, there is a nonsingular matrix P such that $B_1 = P^TP$. Substitute $y = Px$ into (3.37), set $D = {P^{-1}}^TA_1P^{-1}, A = {P^{-1}}^TA_2P^{-1}, B = {P^{-1}}^TB_2P^{-1}$ and use the homogeneity, problem (3.37) is rewritten as follows.

$$\max_{\|y\|=1} \ y^TDy + \frac{y^TAy}{y^TBy}, \ B \succ 0. \tag{3.38}$$

Theorem 3.3.1 ([72]). *If A, B, D are $\mathbb{R}$-SDC by an orthogonal congruence matrix then (3.38) is reduced to a one-dimensional maximization problem over a closed interval.*

Proof. Suppose A, B, D are $\mathbb{R}$-SDC by an orthogonal matrix R :

$$R^TAR = \mathtt{diag}(a_1, a_2, \ldots, a_n), R^TBR = \mathtt{diag}(b_1, b_2, \ldots, b_n),$$

$$R^TDR = \mathtt{diag}(d_1, d_2, \ldots, d_n).$$

Making a change of variables $\eta = Ry$, problem (3.38) becomes

$$\begin{array}{ll} \max & \sum_{i=1}^n d_i\eta_i^2 + \dfrac{\sum_{i=1}^n a_i\eta_i^2}{\sum_{i=1}^n b_i\eta_i^2} \\ \text{s.t.} & \sum_{i=1}^n \eta_i^2 = 1. \end{array} \tag{3.39}$$

Let $z_i = \eta_i^2$, problem (3.39) becomes

$$\begin{array}{ll} \max & \sum_{i=1}^n d_iz_i + \dfrac{\sum_{i=1}^n a_iz_i}{\sum_{i=1}^n b_iz_i} \\ \text{s.t.} & z \in \triangle = \left\{z : \sum_{i=1}^n z_i = 1, z_i \geq 0,\ i = 1, 2, \ldots, n\right\}. \end{array} \tag{3.40}$$

Suppose $z^* = (z_1^*, z_2^*, \ldots, z_n^*)$ is an optimal solution to (3.40), we set $\alpha = \sum_{i=1}^n b_iz_i^*$. Problem (3.40) then shares the same optimal solution set with the following linear programming problem

$$\begin{array}{ll} \max & \sum_{i=1}^n d_iz_i + \dfrac{\sum_{i=1}^n a_iz_i}{\alpha} \\ \text{s.t.} & \sum_{i=1}^n b_iz_i = \alpha, z \in \triangle. \end{array} \tag{3.41}$$

We note now that (3.41) is a linear programming problem and its optimal solutions can only be the extreme points of $\triangle$. An extreme point of $\triangle$ has at most two nonzero elements. There is no loss of generality, suppose $(z_1, z_2, 0, \ldots, 0)^T \in \triangle$ is a candidate of the optimal solutions of (3.41). We have $z_2 = 1 - z_1$ and problem (3.41) becomes:

$$\begin{array}{ll} \max & d_1z_1 + d_2(1 - z_1) + \dfrac{a_1z_1 + a_2(1 - z_1)}{\alpha} \\ \text{s.t.} & b_1z_1 + b_2(1 - z_1) = \alpha; \\ & 0 \leq z_1 \leq 1. \end{array} \tag{3.42}$$

This is a one-dimensional maximization problem as desired. ◻

Now, we extend problem (3.37) to a sum of a finite number of ratios taking the following format

$$(\mathrm{R_m}) \qquad \max_{x\in\mathbb{R}^n\setminus\{0\}} \left\{\frac{x^TA_1x}{x^TB_1x} + \frac{x^TA_2x}{x^TB_2x} + \ldots + \frac{x^TA_mx}{x^TB_mx}\right\}$$

where $A_i, B_i \in S^n$ and $B_i \succ 0$. When $A_1, A_2, \ldots, A_m;\ B_1, B_2, \ldots, B_m$ are $\mathbb{R}$-SDC, problem ($\mathrm{R_m}$) is reduced to maximizing the sum-of-linear-ratios

$$(\mathrm{SLR_m}) \qquad \max_{z\geq 0, z\neq 0} \sum_{i=1}^{m} \frac{\alpha_i^T z}{\beta_i^T z}.$$

Even though both ($\mathrm{R_m}$) and ($\mathrm{SLR_m}$) are NP-hard, the latter can be better approximated by some methods, such as an interior algorithm in [21], a range-space approach in [58] and a branch-and-bound algorithm in [40, 38]. Please see a good survey on sum-of-ratios problems in [55].

Conclusion of Chapter 3

We computed the positive semidefinite interval $I_{\succeq}(C_1, C_2)$ of matrix pencil $C_1 + \mu C_2$ by exploring the SDC properties of C_1 and C_2. Specifically, if C_1 and C_2 are $\mathbb{R}$-SDC, $I_{\succeq}(C_1, C_2)$ can be an empty set or a single point or an interval as shown in Theorems 3.1.1, 3.1.2, 3.1.3. If C_1 and C_2 are not $\mathbb{R}$-SDC, $I_{\succeq}(C_1, C_2)$ can only be empty or singleton. Theorems 3.1.4, 3.1.5 and 3.1.6 present these situations. $I_{\succeq}(C_1, C_2)$ is then applied to solve the generalized trust region subproblems by only solving linear equations, please see Theorems 3.2.1, 3.2.2. We also showed that if the matrices in the quadratic terms of a QCQP problem are $\mathbb{R}$-SDC, the QCQP can be relaxed to a convex SOCP. A lower bound of QCQP is thus found by solving a convex problem. At the end of the chaper we presented the applications of the SDC for reducing a sum-of-generalized Rayleigh quotients to a sum-of-linear ratios.

Conclusions

In this dissertation, the SDC problem of Hermitian matrices and real symmetric matrices has been dealt with. The results obtained in the dissertation are not only theoretical but also algorithmic. On one hand, we proposed necessary and sufficient SDC conditions for a set of arbitrary number of either Hermitian matrices or real symmetric matrices. We also proposed a polynomial time algorithm for solving the Hermitian SDC problem, together with some numerical tests in MATLAB to illustrate for the main algorithm. The results in this part immediately hold for real Hermitian matrices, which is known as a long-standing problem posed in [30, Problem 12]. In addition, the main algorithm in this part can be applied to solve the SDC problem for arbitrarily square matrices by splitting the square matrices up into Hermitian and skew-Hermitian parts. On the other hand, we developed Jiang and Li' technique [37] for two real symmetric matrices to apply for a set of arbitrary number of real symmetric matrices.

1. Results on the SDC problem of Hermitian matrices.
   - Proposed an algorithm for solving the SDC problem of commuting Hermitian matrices (Algorithm 3);
   - Solved the SDC problem of Hermitian matrices by max-rank method (please see Theorem 2.1.4 and Algorithm 4);
   - Proposed a Schmüdgen-like method to find the maximum rank of a Hermitian matrix-pencil (please see Theorem 2.1.2 and Algorithm 2);
   - Proposed equivalent SDC conditions of Hermitian matrices linked with the existence of a positive definite matrix satisfying a system of linear equations (Theorem 2.1.5);
   - Proposed an algorithm for completely solving the SDC problem of complex or real Hermitian matrices (please see Algorithm 6).
2. Results on the SDC problem of real symmetric matrices.
   - Proposed necessary and sufficient SDC conditions for a collection of real symmetric matrices to be SDC (please see Theorem 2.2.2 for nonsingular collection and Theorem 2.2.3 for singular collection). These results are completeness and generalizations of Jiang and Li's method for two matrices [37].
   - Proposed an inductive method for solving the SDC problem of a singular collection. This method helps to move from study the SDC of a singular

collection to study the SDC of a nonsingular collection of smaller dimension as shown in Theorem 2.2.3. Moreover, we realize that a result by Jiang and Li [37] is not complete. A missing case not considered in their paper is now added to make it up in the dissertation, please see Lemma 1.2.8 and Theorem 1.2.1.

- Proposed algorithms for solving the SDC problems of nonsingular and singular collection (Algorithm 7 and Algorithm 8, respectively).

3. We apply above SDC results for dealing with the following problems.

    - Computed the positive semidefinite interval of matrix pencil $C_1+\mu C_2$ (please see Theorems 3.1.1, 3.1.2, 3.1.3, 3.1.4, 3.1.5 and 3.1.6);
    - Applied the positive semidefinite interval of matrix pencil for completely solving the GTRS (please see Theorems 3.2.1, 3.2.2);
    - Solved the homogeneous QCQP problems, the maximization of a sum of generalized Rayleigh quotients under the SDC of involved matrices.

Future research

The SDC problem has been completely solved on the field of real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$. A natural question to aks is whether the obtained SDC results are remained true on a finite field? on a commutative ring with unit? Moreover, as seen, the SDC conditions seem to be very strict. That is, not too many collections can satisfy the SDC conditions. This raises a question that how much disturbance on the matrices such that a not SDC collection becomes SDC? Those unsloved problems suggest our future research as follows.

1. Studying the SDC problems on a finite field, on a commutative ring with unit;

2. Studying the approximately simultaneous diagonalization via congruence of matrices. This problem can be stated as follows: Suppose the matrices $C_1, C_2, \ldots, C_m$, are not SDC. Given $\epsilon > 0$, whether there are matrices E_i with $\|E_i\| < \epsilon$ such that $C_1 + E_1, C_2 + E_2, \ldots, C_m + E_m$ are SDC?

   Some results on approximately simultaneously diagonalizable matrices for two real matrices and for three complex matrices can be found in [50, 68, 61].

3. Explore applications of the SDC results.

List of Author's Related Publication

1. V. B. Nguyen, **T. N. Nguyen,** R.L. Sheu (2020), " Strong duality in minimizing a quadratic form subject to two homogeneous quadratic inequalities over the unit sphere", J. Glob. Optim., 76, pp. 121-135.

2. T. H. Le, **T. N. Nguyen** (2022) , "Simultaneous Diagonalization via Congruence of Hermitian Matrices: Some Equivalent Conditions and a Numerical Solution", SIAM J. Matrix Anal. Appl., 43, Iss. 2, pp. 882-911.

3. V. B. Nguyen, **T. N. Nguyen** (2024), "Positive semidefinite interval of matrix pencil and its applications to the generalized trust region subproblems", Linear Algebra Appl., 680, pp. 371-390.

4. **T. N. Nguyen,** V. B. Nguyen, T. H. Le, R. L. Sheu, "Simultaneous Diagonalization via Congruence of m Real Symmetric Matrices and Its Implications in Quadratic Optimization", Preprint.

Bibliography

[1] S. Adachi, Y. Nakatsukasa (2019), "Eigenvalue-based algorithm and analysis for nonconvex QCQP with one constraint", *Math. Program.*, Ser. A 173, pp. 79-116.

[2] B. Afsari (2008), "Sensitivity Analysis for the Problem of Matrix Joint Diagonalisation", *SIAM J. Matrix Anal. Appl.*, 30, pp. 1148-1171.

[3] A. A. Albert (1938), "A quadratic form problem in the calculus of variations", *Bull. Amer. Math. Soc*, 44, pp. 250-253.

[4] F. Alizadeh, D. Goldfarb (2003), "Second-order cone programming", *Math. Program., Ser. B*, 95, pp. 3-51.

[5] R. I. Becker (1980), "Necessary and sufficient conditions for the simultaneous diagonability of two quadratic forms", *Linear Algebra Appl.*, 30, pp. 129-139.

[6] A. Ben-Tal, D. Hertog (2014), "Hidden conic quadratic representation of some nonconvex quadratic optimization problems", *Math. Program.*, 143, pp. 1-29.

[7] P. Binding (1990), "Simultaneous diagonalisation of several Hermitian matrices", *SIAM J. Matrix Anal. Appl.*, 11, pp. 531-536.

[8] P. Binding, C. K. Li (1991), "Joint ranges of Hermitian matrices and simultaneous diagonalization", *Linear Algebra Appl.*, 151, pp. 157-167.

[9] S. Boyd, L. Vandenberghe (2004), *Convex Optimization*, Cambridge University Press, Cambridge.

[10] A. Bunse-Gerstner, R. Byers, V.Mehrmann (1993), "Numerical methods for simultaneous diagonalization", *SIAM J. Matrix Anal. Appl*, 14, pp. 927-949.

[11] M. D. Bustamante, P. Mellon, M. V. Velasco (2020), "Solving the problem of simultaneous diagonalisation via congruence",*SIAM J. Matrix Anal. Appl*, 41, No. 4, pp. 1616-1629 .

[12] E. Calabi (1964), " Linear systems of real quadratic forms", *Proc. Amer. Math. Soc*, 15, pp. 844-846.

[13] J. F. Cardoso , A. Souloumiac (1993), "Blind beamforming for non-Gaussian signals", *IEE Proc. F Radar and Signal Process.*, 140, pp. 362-370.

[14] L. De Lathauwer (2006), "A link betwen the canonical decomposition in multilinear algebra and simultaneous matrix diagonalisation", *SIAM J. Matrix Anal. Appl.*, 28, pp. 642-666.

[15] E. Deadman, N. J. Higham, R. Ralha (2013), *Blocked Schur algorithms for computing the matrix square root*, in Proceedings of the International Workshop on Applied Parallel Computing, Lecture Notes in Comput. Sci. 7782, P. Manninen and P. Oster, eds., Springer, New York, pp. 171-182.

[16] M. M. Dundar, G. Fung, J. Bi, S. Sandilya, B. Rao (2005), *Sparse Fisher discriminant analysis for computer aided detection*, Proceedings of SIAM International Conference on Data Mining.

[17] J. M. Feng, G. X. Lin, R. L. Sheu, Y. Xia (2012), "Duality and solutions for quadratic programming over single non-homogeneous quadratic constraint", *J. Glob. Optim.*, 54(2), pp. 275-293.

[18] P. Finsler (1937), " Über das vorkommen definiter und semidefiniter formen in scharen quadratischer formen", *Comment. Math. Helv*, 9, pp. 188-192.

[19] B. N. Flury, W. Gautschi (1986), "An algorithm for simultaneous orthogonal transformation of several positive definite symmetric matrices to nearly diagonal form", *SIAM J. Sci. Stat. Comput.*, 7, pp. 169-184.

[20] E. Fung, K. Ng. Michael (2007), "On sparse Fisher discriminant method for microarray data analysis", *Bio.*, 2, pp. 230-234.

[21] R.W. Freund, F. Jarre (2001), "Solving the sum-of-ratios problem by an interior-point method", *J. Glob. Optim.*, 19, pp. 83-102.

[22] X. B. Gao, G. H. Golub, L. Z. Liao (2008), "Continuous methods for symmetric generalized eigenvalue problems," *Linear Algebra Appl.*, 428, pp. 676-696.

[23] G. H. Golub, L. Z. Liao (2006), "Continuous methods for extreme and interior eigenvalue problems," *Linear Algebra Appl.*, 415, pp. 31-51.

[24] H. H. Goldstine, L. P. Horwitz (1959), "A procedure for the diagonalization of normal matrices", *J.ACM*, 6, pp. 176-195.

[25] G. H. Golub, C. F. Van Loan (1996), *Matrix Computations*, 3rd edn. Johns Hopkins University Press, Baltimore.

[26] M. Grant, S. P. Boyd, CVX (2011), "MATLAB Software for Disciplined Convex Programming", *Version 1.21, http://cvxr.com/cvx.*

[27] W. Greub (1958), *Linear Algebra*, 1st ed., Springer-Verlag, p.255. ; Heidelberger Taschenbücker (1976), Bd. 179.

[28] M. R. Hestenes, E. J. McShane (1940), "A theorem on quadratic forms and its application in the calculus of variations", *Transactions of the AMS*, 47, pp. 501-512.

[29] J. B. Hiriart-Urruty, M. Torki (2002), " Permanently going back and forth between the "quadratic world" and "convexity world" in optimization", *Appl. Math. Optim.*, 45, pp. 169-184.

[30] J. B. Hiriart-Urruty (2007), "Potpourri of conjectures and open questions in Nonlinear analysis and Optimization", *SIAM Rev.*, 49, pp. 255-273.

[31] J. B. Hiriart-Urruty, J. Malick (2012), "A fresh variational-analysis look at the positive semidefinite matrices world", *J. Optim. Theory Appl.*, 153, pp. 551-577.

[32] H. Hmam (2010), *Quadratic Optimization with One Quadratic Equality Constraint.*, Tech. report, DSTO-TR-2416, Warfare and Radar Division DSTO Defence Science and Technology Or- ganisation, Australia.

[33] Y. Hsia, G. X. Lin, R. L. Sheu (2014), "A revisit to quadratic programming with one inequality quadratic constraint via matrix pencil", *Pac J Optim.*, 10(3), pp. 461-481.

[34] R. A. Horn, C. R. Johnson (1985), *Matrix analysis*, Cambridge University Press, Cambridge.

[35] R. A. Horn, C. R. Johnson (1991), *Topics in Matrix Analysis*, Cambridge University Press, Cambridge.

[36] K. Huang, N. D. Sidiropoulos (2016), "Consensus-ADMM for general quadratically constrained quadratic programming, *IEEE Trans. Signal Process.*, 64, pp. 5297-5310.

[37] R. Jiang, D. Li (2016), "Simultaneous Diagonalization of Matrices and Its Applications in Quadratically Constrained Quadratic Programming", *SIAM J. Optim.*, 26, pp. 1649-1668.

[38] H. W. Jiao, S. Y. Liu (2015), "A practicable branch and bound algorithm for sum of linear ratios problem", *Eur. J. Oper. Res.*, 243, Issue 3, 16, pp. 723-730.

[39] L. Kronecker (1874), “Monatsber”, *Akad. Wiss. Berl.*, pp. 397.

[40] T. Kuno (2002), “A branch-and-bound algorithm for maximizing the sum of several linear ratios”, *J. Glob. Optim.*, 22, pp. 155-174

[41] P. Lancaster, L. Rodman (2005), “ Canonical forms for Hermitian matrix pairs un- der strict equivalence and congruence”, *SIAM Rev.*, 47, pp. 407-443.

[42] T. H. Le, T. N. Nguyen (2022), “Simultaneous Diagonalization via Congruence of Hermitian Matrices: Some Equivalent Conditions and a Numerical Solution”, *SIAM J. Matrix Anal. Appl. ,* 43, Iss. 2, pp. 882-911.

[43] C. Mendl (2020), “simdiag.m”, Matlab Central File Exchange, *http://www.mathworks.com/matlabcentral/fileexchange/46794-simdiag-m”.*

[44] J. J. Moré (1993), “ Generalization of the trust region problem”, *Optim. Methods Softw.*, 2, pp. 189-209.

[45] P. Muth (1905), “ Über relle Äquivalenz von Scharen reeller quadratischer Formen”, *J. Reine Angew. Math*, 128, pp. 302-321.

[46] V. B. Nguyen, T. N. Nguyen, R.L. Sheu (2020), “ Strong duality in minimizing a quadratic form subject to two homogeneous quadratic inequalities over the unit sphere”, *J. Glob. Optim.*, 76, pp. 121-135.

[47] V. B. Nguyen, T. N. Nguyen (2024), “Positive semidefinite interval of matrix pencil and its applications to the generalized trust region subproblems”, *Linear Algebra Appl.*, 680, pp. 371-390.

[48] V. B. Nguyen, R. L. Sheu, Y. Xia (2016), “ Maximizing the sum of a generalized Rayleigh quotient and another Rayleigh quotient on the unit sphere via semidefinite programming”, *J. of Glob. Optim.*, 64, pp. 399-416.

[49] T. N. Nguyen, V. B. Nguyen, T. H. Le R. L. Sheu, “Simultaneous Diagonalization via Congruence of m Real Symmetric Matrices and Its Implications in Quadratic Optimization”, Preprint.

[50] K. C. O’Meara, C. Vinsonhaler (2006), “ On approximately simultaneously diagonalizable matrices”, *Linear Algebra Appl.*, 412, pp. 39-74.

[51] P. M. Pardalos, S. A. Vavasis (1991), “Quadratic programming with one negative eigenvalue is NP-Hard”, *J. Global Optim.*, 1, pp. 15-22.

[52] D. T. Pham (2001), "Joint approximate diagonalisation of positive definite matrices", *SIAM. J. Matrix Anal. Appl.* , 22, pp. 1136-1152.

[53] G. Primolevo, O. Simeone, U. Spagnolini (2006), *Towards a joint optimization of scheduling and beamforming for MIMO downlink*, IEEE Ninth International Symposium on Spread Spectrum Techniques and Applications, pp. 493-497.

[54] M. Salahi, A. Taati (2018), "An efficient algorithm for solving the generalized trust region subproblem", *Comput. Appl. Math.*, 37 , pp. 395-413.

[55] S. Schaible, J. Shi (2003), "Fractional programming: The sum-of-ratios cases", *Optim. Methods Softw.*, 18, No. 2, pp. 219-229.

[56] K. Schmüdgen (2009), "Noncommutative real algebraic geometry-some basic concepts and first ideas, in Emerging Applications of Algebraic Geometry ", *The IMA Volumes in Mathematics and its Applications, M. Putinar and S. Sullivant, eds.*, 149, Springer New York, pp. 325-350.

[57] R. Shankar (1994), *Principles of quantum mechanics*, Plenum Press, New York.

[58] R. L. Sheu, W. I. Wu, I. Birble (2008), "Solving the sum-of-ratios problem by stochastic search algorithm", *J. Glob. Optim.*, 42(1), pp. 91-109.

[59] R. J. Stern, H. Wolkowicz (1995), "Indefinite trust region subproblems and nonsymmetric eigen- value perturbations", *SIAM J. Optim.*, 5, pp. 286-313.

[60] J. G. Sun (1991), "Eigenvalues of Rayleigh quotient matrices", *Numerische Math.*, 59, Issue 1, pp. 603-614.

[61] B. D. Sutton (2023), "Simultaneous diagonalization of nearly commuting Hermitian matrices: do-one-then-do-the-other", *IMA J. Numerical Anal.*, pp. 1-29.

[62] P. Tichavsky, A. Yeredor (2009), "Fast approximate joint diagonalisation incorporating weight matrices", *IEEE Trans. Signal Process.*, 57, pp. 878-891.

[63] K. C. Toh, M. J. Todd, R. H. Tut ünc ü(1999), "SDPT3—A MATLAB software package for semidefinite programming", *Optim. Methods Softw.*, 11, pp. 545-581.

[64] F. Uhlig (1976), "A canonical form for a pair of real symmetric matrices that generate a nonsingular pencil", *Linear Algebra Appl.*, 14, pp. 189-209.

[65] F. Uhlig (1979), " A recurring theorem about pairs of quadratic forms and extensions: A survey", *Linear Algebra Appl.*, 25, pp. 219-237.

[66] S. A. Vavasis (1990), "Quadratic programming is in NP", *Inf. Process. Lett.*, 36(2), pp. 73-77.

[67] L. Wang, L. Albera, A. Kachenoura, H. Z. Shu, L. Senhadji (2013), "Nonnegative joint diagonalisation by congruence based on LU matrix factorization",*IEEE Signal Process. Lett.*, 20, pp. 807-810.

[68] A. L. Wang, R. Jiang (2021), "New notions of simultaneous diagonalizability of quadratic forms with applications to QCQPs", arXiv preprint arXiv:2101.12141.

[69] L. F. Wang, Y. Xia (2019), "A Linear-Time Algorithm for Globally Maximizing the Sum of a Generalized Rayleigh Quotient and a Quadratic Form on the Unit Sphere", *SIAM J. Optim.*, 29(3), pp. 1844-1869.

[70] K. Weierstrass (1868), *Zur Theorie der bilinearen und quadratischen Formen*, Monatsb. Berliner Akad. Wiss., pp. 310-338.

[71] M. C. Wu, L. S. Zhang, Z. X. Wang, D. C. Christiani, X. H. Lin (2009), "Sparse linear discriminant analysis for simultaneous testing for the significance of a gene set/pathway and gene selection", *Bio.*, 25, pp. 1145-1151.

[72] Y. Xia, S. Wang, R. L. Sheu (2016), "S-Lemma with Equality and Its Applications", *Math. Program., Ser. A*, 156, Issue 1-2, pp. 513-547.

[73] Y. Ye, S. Z. Zhang (2003), " New results on quadratic minimization", *SIAM J. Optim.*, 14, No. 1, pp. 245-267.

[74] A.Y. Yik-Hoi (1970), "A necessary and sufficient condition for simultaneously diagonalisation of two Hermitian matrices and its applications", *Glasg. Math. J.*, 11, pp. 81-83.

[75] L. H. Zhang (2013), "On optimizing the sum of the Rayleigh quotient and the generalized Rayleigh quotient on the unit sphere", *Comput. Optim. Appl.*, 54, pp. 111-139.

[76] L. H. Zhang (2014), "On a self-consistent-field-like iteration for maximizing the sum of the Rayleigh quotients", *J. Comput. Appl. Math.*, 257, pp.14-28.

Index