New notions of simultaneous diagonalizability of quadratic forms with applications to QCQPs

Alex L. Wang * Rujun Jiang †

January 28, 2021

Abstract

A set of quadratic forms is simultaneously diagonalizable via congruence (SDC) if there exists a basis under which each of the quadratic forms is diagonal. This property appears naturally when analyzing quadratically constrained quadratic programs (QCQPs) and has important implications in this context. This paper extends the reach of the SDC property by studying two new related but weaker notions of simultaneous diagonalizability. Specifically, we say that a set of quadratic forms is almost SDC (ASDC) if it is the limit of SDC sets and d-restricted SDC (d-RSDC) if it is the restriction of an SDC set in up to d-many additional dimensions. Our main contributions are a complete characterization of the ASDC pairs and the nonsingular ASDC triples, as well as a sufficient condition for the 1-RSDC property for pairs of quadratic forms. Surprisingly, we show that *every* singular pair is ASDC and that *almost every* pair is 1-RSDC.

We accompany our theoretical results with preliminary numerical experiments applying the RSDC property to QCQPs with a single quadratic constraint.

1 Introduction

This paper investigates two new notions of *simultaneous diagonalizability* of quadratic forms over both $\mathbb{C}^n$ and $\mathbb{R}^n$.[1]

Let $\mathbb{H}^n$ denote the real vector space of $n \times n$ Hermitian matrices. Recall that a set of matrices $\mathcal{A} \subseteq \mathbb{H}^n$ is said to be *simultaneously diagonalizable via congruence* (SDC) if there exists an invertible $P \in \mathbb{C}^{n\times n}$ such that P^*AP is diagonal for every $A \in \mathcal{A}$. Here, P^* is the conjugate transpose of P.

The SDC property has attracted significant interest in recent years in the context of solving specific classes of quadratically constrained quadratic programs (QCQPs) and their relaxations [14, 17, 19, 21, 31, 33, 34]. We will refer to QCQPs in which the involved quadratic forms are SDC as *diagonalizable QCQPs*. While such diagonalizable QCQPs are not easier to solve in any broad complexity-theoretic sense (indeed, binary integer programs can be cast naturally as QCQPs even only using *diagonal* quadratic constraints), they do benefit from a number of advantages over more general QCQPs: It is well known that the standard Shor semidefinite program (SDP) relaxation of a diagonalizable QCQP is equivalent to a second order cone program [31]. Consequently, the standard SDP relaxation can be computed substantially faster for diagonalizable QCQPs than for

*Carnegie Mellon University, Pittsburgh, PA, USA, alw1@cs.cmu.edu

†Corresponding author. School of Data Science, Fudan University, Shanghai, China, rjjiang@fudan.edu.cn

[1] While all of our results hold with only minor modifications over both $\mathbb{C}^n$ and $\mathbb{R}^n$, where quadratic forms correspond to Hermitian matrices and real symmetric matrices respectively, we will simplify our presentation by only discussing the Hermitian setting; see Appendix C for a discussion of our results in the real symmetric setting.

arbitrary QCQPs. This idea has been used effectively to build cheap convex relaxations within branch-and-bound frameworks [33, 34]; see also [19] for an application to portfolio optimization. Additionally, qualitative properties of the standard SDP relaxation are often easier to analyze in the context of diagonalizable QCQPs. For example, a long line of work has investigated when the SDP relaxations of certain diagonalizable QCQPs are *exact* (for various definitions of exact) and have given sufficient conditions for these properties [2–4, 8, 10, 13, 14, 18, 29]. Often, such arguments rely on conditions (such as convexity[2] or polyhedrality) of the quadratic image [23] or the set of convex Lagrange multipliers [31]. In this context, the SDC property ensures that both of these sets are polyhedral. While such conditions have been generalized beyond only diagonalizable QCQPs, the sufficient conditions often become much more difficult to verify [30, 31].

Besides its implications in the context of QCQPs, the SDC property also finds applications in areas such as signal processing, multivariate statistics, medical imaging analysis, and genetics; see [5, 28] and references therein.

In this paper, we take a step towards increasing the applicability of the important consequences of the SDC property by investigating two weaker notions of simultaneous diagonalizability of quadratic forms, which we will refer to as *almost SDC* (ASDC) and *d-restricted SDC* (d-RSDC).

1.1 Related work

Canonical forms for pairs of quadratic forms. Weierstrass [32] and Kronecker (see [15]) proposed canonical forms for pairs of real quadratic or bilinear forms under simultaneous reductions.[3] These canonical forms were also subsequently extended to the complex case; see [16, 25–27] for historical accounts of these developments, as well as collected and simplified proofs.

While this line of work was not specifically developed to understand the SDC property, it nonetheless gives a complete characterization of the SDC property for pairs of quadratic forms. We will make extensive use of these canonical forms (see Proposition 3) in this paper.

The SDC property for sets of three or more quadratic forms and SDC algorithms. There has been much recent interest in understanding the SDC property for more general m-tuples of quadratic forms. In fact, the search for "sensible and "palpable" conditions" for this property appeared as an open question on a short list of 14 open questions in nonlinear analysis and optimization [7].

A recent line of work has given a complete characterization of the SDC property for general m-tuples of quadratic forms: In the real symmetric setting, Jiang and Li [14] gave a complete characterization of this property under a semidefiniteness assumption. This result was then improved upon by Nguyen et al. [21] who removed the semidefiniteness assumption. Le and Nguyen [17] additionally extend these characterizations to the case of Hermitian matrices. Bustamante et al. [5] gave a complete characterization of the simultaneous diagonalizability of an m-tuple of *symmetric complex* matrices under $^\intercal$-congruence.[4]

We further remark that this line of work is "algorithmic" and gives numerical procedures for deciding if a given set of quadratic forms is SDC. See [17] and references therein.

[2] The convexity of the quadratic image is sometimes referred to as "hidden convexity."

[3] Reduction refers to one of similarity, congruence, equivalence, or strict equivalence.

[4] We emphasize that Bustamante et al. [5] consider complex symmetric matrices and adopt $^\intercal$-congruence as their notion of congruence.

The almost SDS property. An analogous theory for the *almost* simultaneous diagonalizability of *linear operators* has been studied in the literature. In this setting, the congruence transformation is naturally replaced by a similarity transformation[5] and the SDC property is replaced by simultaneous diagonalizability *via similarity* (SDS). A widely cited theorem due to Motzkin and Taussky [20] shows that every pair of commuting linear operators, i.e., a pair of matrices in $\mathbb{C}^{n\times n}$, is almost SDS. This line of investigation was more recently picked up by O'meara and Vinsonhaler [22] who showed that triples of commuting linear operators are almost SDS under a regularity assumption on the dimensions of eigenspaces associated with the linear operators.

1.2 Main contributions and outline

In this paper, we explore two weaker notions of simultaneous diagonalizability which we will refer to as *almost SDC* (ASDC) and *d-restricted SDC* (d-RSDC); see Sections 2 and 5 for precise definitions. Informally, $\mathcal{A} \subseteq \mathbb{H}^n$ is ASDC if it is the limit of SDC sets and d-RSDC if it is the restriction of an SDC set in $\mathbb{H}^{n+d}$ to $\mathbb{H}^n$. While *a priori* the ASDC and d-RSDC properties (for d small) may seem almost as restrictive as the SDC property, we will see (Theorems 2 and 4) that these properties actually hold much more widely in certain settings.

A summary of our contributions, along with an outline of the paper, follows:

- In Section 2, we formally define the SDC and ASDC properties and review known characterizations of the SDC property. We additionally highlight a number of behaviors of the SDC property which will later contrast with those of the ASDC property.
- In Section 3, we give a complete characterization of the ASDC property for Hermitian pairs. In particular, Theorem 2 states that *every* singular[6] Hermitian pair $\{A, B\} \subseteq \mathbb{H}^n$ is ASDC. The proof of this statement relies on the canonical form for pairs of Hermitian matrices [16] under congruence transformations and the invertibility of a certain matrix related to the eigenvalues of an "arrowhead" matrix.
- In Section 4, we give a complete characterization of the ASDC property for *nonsingular* Hermitian triples. While our characterization is easy to state, its proof is less straightforward and relies on facts about block matrices with Toeplitz upper triangular blocks. We review the relevant properties of such matrices in Appendix B.
- In Section 5, we formally define the d-RSDC property and highlight its relation to the ASDC property. We then show in Theorem 4 and Corollary 2 that the 1-RSDC property holds for almost every Hermitian pair.
- In Section 6, we construct obstructions to *a priori* plausible generalizations of our developments in Sections 3 to 5. Section 6.1 shows that, in contrast to Theorem 2, there exist singular Hermitian triples which are *not* ASDC. The same construction can be interpreted as a Hermitian triple which is not d-RSDC for any $d < \lfloor n/2 \rfloor$; this contrasts with Theorem 4. Next, Section 6.2 shows that a natural generalization of our characterizations of the ASDC property for Hermitian pairs and triples cannot hold for general m-tuples; specifically this natural generalization fails for $m \geq 5$ in the Hermitian setting and $m \geq 7$ in the real symmetric setting.

[5] Recall that two matrices $A, B \in \mathbb{C}^{n\times n}$ are similar if there exists an invertible $P \in \mathbb{C}^{n\times n}$ such that $A = P^{-1}BP$.
[6] See Definition 3.

- In Section 7, we revisit one of the key motivations for studying the ASDC and d-RSDC properties—solving QCQPs more efficiently. In this context, we propose new diagonal reformulations of QCQPs when the quadratic forms involved are either ASDC or d-RSDC. We accompany these reformulations with preliminary numerical experiments which highlight interesting directions for future research.

Remark 1. In the main body of this paper, we will state and prove our results for only the Hermitian setting. Nevertheless, our results and proofs extend almost verbatim to the real symmetric setting by replacing the canonical form of a pair of Hermitian matrices (Proposition 3) by the notationally more involved canonical form for a pair of real symmetric matrices (see [16, Theorem 9.2]). As no new ideas or insights are required for handling the real symmetric setting, we defer formally stating our results in the real symmetric setting and discussing the necessary modifications to our proofs to Appendix C. $\square$

1.3 Notation

Let $\mathbb{N} = \{1, 2, \dots\}$ and $\mathbb{N}_0 = \{0, 1, \dots\}$. For $m, n \in \mathbb{N}_0$, let $[m, n] = \{m, m+1, \dots, n\}$ and $[n] = \{1, \dots, n\}$. By convention, if $m \geq n+1$ (respectively, $n \leq 0$), then $[m, n] = \varnothing$ (respectively, $[n] = \varnothing$). For $n \in \mathbb{N}$, let $\mathbb{H}^n$ denote the real vector space of $n \times n$ Hermitian matrices. For $\alpha \in \mathbb{C}$, $v \in \mathbb{C}^n$, and $A \in \mathbb{C}^{n\times n}$, let α^*, v^*, and A^* denote the conjugate of α, conjugate transpose of v, and conjugate transpose of A respectively. Let $\operatorname{span}(\cdot)$ and $\dim(\cdot)$ denote the real span and real dimension respectively. For $n, m \in \mathbb{N}$, let I_n, 0_n, and $0_{n\times m}$ denote the $n \times n$ identity matrix, $n \times n$ zero matrix, and $n \times m$ zero matrix respectively. When $n \in \mathbb{N}$ is clear from context, let $e_i \in \mathbb{C}^n$ denote the ith standard basis vector. Given a complex subspace $V \subseteq \mathbb{C}^n$ with $\mathbb{C}$-dimension k, a surjective map $U : \mathbb{C}^k \to V$, and $A \in \mathbb{H}^n$, let $A|_V \in \mathbb{H}^k$ denote the restriction of A to V, i.e., $A|_V = U^*AU$. When the map U is inconsequential, we will omit specifying U. For $\alpha_1, \dots, \alpha_k \in \mathbb{C}$, let $\operatorname{Diag}(\alpha_1, \dots, \alpha_k) \in \mathbb{C}^{k\times k}$ denote the diagonal matrix with ith entry α_i. For $A_1, \dots, A_k$ complex square matrices, let $\operatorname{Diag}(A_1, \dots, A_k)$ denote the block diagonal matrix with ith block A_i. Given $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$, let $A \oplus B \in \mathbb{C}^{(n+m)\times(n+m)}$ and $A \otimes B \in \mathbb{C}^{nm\times nm}$ denote the direct sum and Kronecker product of A and B respectively. Given $A, B \in \mathbb{C}^{n\times n}$, let $[A, B] = AB - BA$ denote the commutator of A and B. For $A \in \mathbb{C}^{n\times n}$, let $\|A\|$ denote the spectral norm of A. Given $\alpha \in \mathbb{C}$, let $\operatorname{Re}(\alpha)$ and $\operatorname{Im}(\alpha)$ denote the real and imaginary parts of α respectively. We will denote the imaginary unit by the symbol i in order to distinguish it from the variable i, which will often be used as an index.

2 Preliminaries

In this section, we define our main objects of study and recall some useful results from the literature.

Definition 1. A set of Hermitian matrices $\mathcal{A} \subseteq \mathbb{H}^n$ is *simultaneously diagonalizable via congruence* (SDC) if there exists an invertible $P \in \mathbb{C}^{n\times n}$ such that P^*AP is diagonal for all $A \in \mathcal{A}$. $\square$

Remark 2. The SDC property is the natural notion for simultaneous diagonalization in the context of quadratic forms. Indeed, suppose $\mathcal{A} \subseteq \mathbb{H}^n$ is SDC and let P denote the corresponding invertible matrix. Then, performing the change of variables $y = P^{-1}x$, we have that $x^*Ax = y^*(P^*AP)y$ is separable in y for every $A \in \mathcal{A}$. $\square$

Observation 1. *The SDC property is closed under taking spans and subsets. In particular, $\mathcal{A} \subseteq \mathbb{H}^n$ is SDC if and only if $\{A_1, \dots, A_m\}$ is SDC for some basis $\{A_1, \dots, A_m\}$ of* $\operatorname{span}(\mathcal{A})$.

We begin by studying the following relaxation of the SDC property.

Definition 2. A set of Hermitian matrices $\mathcal{A} \subseteq \mathbb{H}^n$ is *almost simultaneously diagonalizable via congruence* (ASDC) if there exists a mapping $f : \mathcal{A} \times \mathbb{N} \to \mathbb{H}^n$ such that

- for all $A \in \mathcal{A}$, the limit $\lim_{j\to\infty} f(A, j)$ exists and is equal to A, and
- for all $j \in \mathbb{N}$, the set $\{f(A, j) : A \in \mathcal{A}\}$ is SDC. $\square$

Observation 2. *The ASDC property is closed under taking spans and subsets. In particular, $\mathcal{A} \subseteq \mathbb{H}^n$ is ASDC if and only if $\{A_1, \dots, A_m\}$ is ASDC for some basis $\{A_1, \dots, A_m\}$ of* $\operatorname{span}(\mathcal{A})$.

When $|\mathcal{A}|$ is finite, we will use the following equivalent definition of ASDC.

Observation 3. *Let $A_1, \dots, A_m \in \mathbb{H}^n$. Then $\{A_1, \dots, A_m\}$ is ASDC if and only if for all $\epsilon > 0$, there exist $\tilde{A}_1, \dots, \tilde{A}_m \in \mathbb{H}^n$ such that*

- *for all $i \in [m]$, the spectral norm $\left\| A_i - \tilde{A}_i \right\| \leq \epsilon$, and*
- *$\left\{\tilde{A}_1, \dots, \tilde{A}_m\right\}$ is SDC.*

We will additionally need the following two definitions.

Definition 3. A set of Hermitian matrices $\mathcal{A} \subseteq \mathbb{H}^n$ is *nonsingular* if there exists a nonsingular $A \in \operatorname{span}(\mathcal{A})$. Else, it is *singular*. $\square$

Definition 4. Given a set of Hermitian matrices $\mathcal{A} \subseteq \mathbb{H}^n$, we will say that $S \in \mathcal{A}$ is a *max-rank element* of $\operatorname{span}(\mathcal{A})$ if $\operatorname{rank}(S) = \max_{A\in\mathcal{A}} \operatorname{rank}(A)$. $\square$

2.1 Characterization of SDC

A number of necessary and/or sufficient conditions for the SDC property have been given in the literature [5, 9, 16]. For our purposes, we will need the following two results. The first result gives a characterization of the SDC property for nonsingular sets of Hermitian matrices and is well-known (see [9, Theorem 4.5.17]). The second result, due to Bustamante et al. [5], gives a characterization of the SDC property for singular sets of Hermitian matrices by reducing to the nonsingular case. For completeness, we provide a short proof for each of these results in Appendix A.

Proposition 1. *Let $\mathcal{A} \subseteq \mathbb{H}^n$ and suppose $S \in \operatorname{span}(\mathcal{A})$ is nonsingular. Then, $\mathcal{A}$ is SDC if and only if $S^{-1}\mathcal{A}$ is a commuting set of diagonalizable matrices with real eigenvalues.*

Proposition 2. *Let $\mathcal{A} \subseteq \mathbb{H}^n$ and suppose $S \in \operatorname{span}(\mathcal{A})$ is a max-rank element of $\operatorname{span}(\mathcal{A})$. Then, $\mathcal{A}$ is SDC if and only if $\operatorname{range}(A) \subseteq \operatorname{range}(S)$ for every $A \in \mathcal{A}$ and $\left\{A|_{\operatorname{range}(S)} : A \in \mathcal{A}\right\}$ is SDC.*

We close this section with two lemmas highlighting consequences of the SDC property which we will compare and contrast with consequences of the ASDC property.

Lemma 1. *Let $\mathcal{A} \subseteq \mathbb{H}^n$ and suppose $S \in \operatorname{span}(\mathcal{A})$ is positive definite. Then, $\mathcal{A}$ is SDC if and only if $S^{-1/2}\mathcal{A}S^{-1/2}$ is a commuting set.*

Proof. This follows as an immediate corollary to Proposition 1 and the fact that $S^{-1}A$ has the same eigenvalues as the Hermitian matrix $S^{-1/2}AS^{-1/2}$. $\blacksquare$

In particular, when $\mathrm{span}(\mathcal{A})$ contains a positive definite matrix, the SDC and ASDC properties can be shown to be equivalent.

Corollary 1. *Let $\mathcal{A} \subseteq \mathbb{H}^n$ and suppose $S \in \mathrm{span}(\mathcal{A})$ is positive definite. Then, $\mathcal{A}$ is SDC if and only if $\mathcal{A}$ is ASDC.*

Despite Corollary 1, we will see soon that the ASDC property is qualitatively quite different to the SDC property in a number of settings (in particular, for singular Hermitian pairs; see Theorem 2). Specifically, we will contrast the following consequence of the SDC property.

Lemma 2 ([17, Lemma 9])**.** *Let $\mathcal{A} \subseteq \mathbb{H}^n$ and suppose there exists a common block decomposition*

$$A = \begin{pmatrix} \bar{A} & \\ & 0_d \end{pmatrix}$$

for all $A \in \mathcal{A}$. Then $\mathcal{A}$ is SDC if and only if $\left\{\bar{A} :\ A \in \mathcal{A}\right\} \subseteq \mathbb{H}^{n-d}$ is SDC.

3 The ASDC property of Hermitian pairs

In this section, we will give a complete characterization of the ASDC property for Hermitian pairs. We will switch the notation above and label our matrices $\mathcal{A} = \{A, B\}$. Our analysis will proceed in two cases: when $\{A, B\}$ is nonsingular and singular respectively.

3.1 A canonical form for a Hermitian pair

In this section and the next, we will make regular use of the canonical form for a Hermitian pair [16, 26].

We will need to define the following special matrices. For $n \geq 2$, let

$$F_n = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}, \qquad G_n = \begin{pmatrix} & & & 0 \\ & & \iddots & 1 \\ & 0 & \iddots & \\ 0 & 1 & & \end{pmatrix}, \quad \text{and} \quad H_n = \begin{pmatrix} & & 1 & 0 \\ & \iddots & 0 & \\ 1 & \iddots & & \\ 0 & & & \end{pmatrix}.$$

Set $F_1 = (1)$ and $G_1 = H_1 = (0)$.

The following proposition is adapted[7] from [16, Theorem 6.1].

Proposition 3. *Let $A, B \in \mathbb{H}^n$ and suppose A is a max-rank element of $\mathrm{span}(\{A, B\})$. Then, there exists an invertible $P \in \mathbb{C}^{n\times n}$ such that $P^*AP = \mathrm{Diag}(S_1, \dots, S_m)$ and $P^*BP = \mathrm{Diag}(T_1, \dots, T_m)$ are block diagonal matrices with compatible block structure. Here, $m = m_1+m_2+m_3+m_4$ corresponds to four different types of blocks where each $m_i \in \mathbb{N}_0$ may be zero. Additionally, $m_4 \in \{0, 1\}$.*

The first m_1-many blocks of P^*AP and P^*BP have the form

$$S_i = \sigma_i F_{n_i}, \qquad T_i = \sigma_i(\lambda_i F_{n_i} + G_{n_i}),$$

where $n_i \in \mathbb{N}$, $\sigma_i \in \{\pm 1\}$, and $\lambda_i \in \mathbb{R}$. The next m_2-many blocks of P^*AP and P^*BP have the form

$$S_i = \begin{pmatrix} & F_{n_i} \\ F_{n_i} & \end{pmatrix}, \qquad T_i = \begin{pmatrix} & \lambda_i F_{n_i} + G_{n_i} \\ \lambda_i^* F_{n_i} + G_{n_i} & \end{pmatrix}, \tag{1}$$

[7] The original statement of [16, Theorem 6.1] contains one additional type of block: those corresponding to the eigenvalues at infinity. These blocks do not exist in our setting by the assumption that A is a max-rank element of $\mathrm{span}(\{A, B\})$.

where $n_i \in \mathbb{N}$ and $\lambda_i \in \mathbb{C} \setminus \mathbb{R}$. The next m_3-many blocks of P^*AP and P^*BP have the form

$$S_i = \begin{pmatrix} & & F_{n_i} \\ & 0 & \\ F_{n_i} & & \end{pmatrix}, \qquad T_i = G_{2n_i+1},$$

where $n_i \in \mathbb{N}$. If $m_4 = 1$, then the last block of P^*AP and P^*BP has the form $S_m = T_m = 0_{n_m}$ for some $n_m \in \mathbb{N}$.

3.2 The nonsingular case

In this section, we will show that if A is invertible, then $\{A, B\}$ is ASDC if and only if $A^{-1}B$ has real eigenvalues. We begin by examining two examples that are representative of the situation when A is invertible.

Example 1. Let

$$A = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Noting that $A^{-1}B$ is not diagonalizable, we conclude via Proposition 1 that $\{A, B\}$ is not SDC. On the other hand, let $\epsilon > 0$ and define

$$\tilde{B} = \begin{pmatrix} \epsilon & 1 \\ 1 & 1 \end{pmatrix}.$$

Now, $A^{-1}\tilde{B}$ has eigenvalues $1 \pm \sqrt{\epsilon}$, whence by Proposition 1 $\left\{A, \tilde{B}\right\}$ is SDC. □

Example 2. Let

$$A = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$$

Noting that $A^{-1}B$ has non-real eigenvalues, we conclude via Proposition 1 (and the fact that eigenvalues vary continuously) that $\{A, B\}$ is not ASDC. □

While the set of diagonalizable matrices is dense in $\mathbb{C}^{n \times n}$, it is not immediately clear that the pairs $(A, B) \in \mathbb{H}^n \times \mathbb{H}^n$ such that $A^{-1}B$ exists and is diagonalizable is dense in $\mathbb{H}^n \times \mathbb{H}^n$. The following lemma shows that this is indeed the case.

Lemma 3. *Let $\{A, B\} \subseteq \mathbb{H}^n$ and suppose A is invertible. For all $\epsilon > 0$, there exists $\tilde{B}$ such that*

- $\left\|B - \tilde{B}\right\| \leq \epsilon$,
- *$A^{-1}\tilde{B}$ has simple eigenvalues (whence in particular, $A^{-1}\tilde{B}$ is diagonalizable), and*
- *$A^{-1}\tilde{B}$ and $A^{-1}B$ have the same number of real eigenvalues counted with multiplicity.*

Proof. We will apply Proposition 3 to $\{A, B\}$. Note that as A is invertible, we will have $m_3 = m_4 = 0$ in Proposition 3. For notational convenience, let $r = m_1$ and let $P, \sigma_1, \ldots, \sigma_r, n_1 \ldots, n_m, \lambda_1, \ldots, \lambda_m$ denote the quantities furnished by Proposition 3.

Fix $\epsilon > 0$ and let $\eta \in [-\epsilon, \epsilon]^m$ denote a vector to be chosen later. Define the blocks $\tilde{T}_i$ as

$$\begin{aligned}
&\tilde{T}_i \coloneqq \sigma_i((\lambda_i + \eta_i)F_{n_i} + G_{n_i} + \epsilon H_{n_i}),\ \forall i \in [r],\\
&\tilde{T}_i \coloneqq \begin{pmatrix} & (\lambda_i + \eta_i)F_{n_i} + G_{n_i} + \epsilon H_{n_i} \\ (\lambda_i^* + \eta_i)F_{n_i} + G_{n_i} + \epsilon H_{n_i} & \end{pmatrix},\ \forall i \in [r+1, m],
\end{aligned} \tag{2}$$

and set $\tilde{B} \coloneqq P^{-*} \operatorname{Diag}(\tilde{T}_1, \dots, \tilde{T}_m) P^{-1}$. Then, $P^{-1}A^{-1}\tilde{B}P = \operatorname{Diag}(S_1^{-1}\tilde{T}_1, \dots, S_k^{-1}\tilde{T}_k)$ is again a block diagonal matrix. Note that for $i \in [r]$, the block

$$S_i^{-1}\tilde{T}_i = (\lambda_i + \eta_i)I_{n_i} + F_{n_i}G_{n_i} + \epsilon F_{n_i}H_{n_i}$$

is a Toeplitz tridiagonal matrix. Similarly, for $i \in [r+1, m]$, the block

$$S_i^{-1}\tilde{T}_i = \begin{pmatrix} (\lambda_i^* + \eta_i)I_{n_i} + F_{n_i}G_{n_i} + \epsilon F_{n_i}H_{n_i} & \\ & (\lambda_i + \eta_i)I_{n_i} + F_{n_i}G_{n_i} + \epsilon F_{n_i}H_{n_i} \end{pmatrix} \tag{3}$$

is a direct sum of Toeplitz tridiagonal matrices. Then, as the closed form of eigenvalues of Toeplitz tridiagonal matrices are known [9], $A^{-1}\tilde{B}$ has eigenvalues

$$\begin{aligned}
&\bigcup_{i=1}^{r} \left\{ \lambda_i + \eta_i + 2\sqrt{\epsilon}\cos\left(\frac{\pi j}{n_i + 1}\right) : j \in [n_i] \right\}\\
&\qquad \cup \bigcup_{i=r+1}^{m} \left\{ \lambda + \eta_i + 2\sqrt{\epsilon}\cos\left(\frac{\pi j}{n_i + 1}\right) : j \in [n_i],\ \lambda \in \{\lambda_i, \lambda_i^*\} \right\}.
\end{aligned}$$

It is clear then that $\eta \in [-\epsilon, \epsilon]^m$ can be picked so that $A^{-1}\tilde{B}$ has only simple eigenvalues. Note also that the quantity $\eta_i + 2\sqrt{\epsilon}\cos\left(\frac{\pi j}{n_i+1}\right)$ is real so that $A^{-1}B$ and $A^{-1}\tilde{B}$ have the same number of real eigenvalues counted with multiplicity. ■

The following theorem follows as a simple corollary to our developments thus far.

Theorem 1. *Let $A, B \in \mathbb{H}^n$ and suppose A is invertible. Then, $\{A, B\}$ is ASDC if and only if $A^{-1}B$ has real eigenvalues.*

Proof. ($\Rightarrow$) This direction holds trivially by continuity of eigenvalues and the assumption that A is invertible.

($\Leftarrow$) Let $\epsilon > 0$. Then, applying Lemma 3 to $\{A, B\}$, we get $\tilde{B}$ such that $\left\|B - \tilde{B}\right\| \leq \epsilon$ and $A^{-1}\tilde{B}$ is a matrix with real simple eigenvalues. We deduce by Proposition 1 that $\left\{A, \tilde{B}\right\}$ is SDC. ■

3.3 The singular case

In the remainder of this section, we investigate the ASDC property when $\{A, B\}$ is singular. We will show, surprisingly, that *every* singular Hermitian pair is ASDC. We begin with an example and some intuition.

Example 3. In contrast to the SDC property (cf. Lemma 2), the ASDC property of a pair $\{A, B\}$ in the singular case does not reduce to the ASDC property of $\left\{\bar{A}, \bar{B}\right\}$, where $\bar{A}$ and $\bar{B}$ are the restrictions of A and B to the joint range of A and B. For example, let

$$A = \begin{pmatrix} & 1 & \\ 1 & & \\ & & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix},$$

and let $\bar{A}$ and $\bar{B}$ denote the respective 2×2 leading principal submatrices.

By Theorem 1, $\left\{\bar{A}, \bar{B}\right\}$ is not ASDC (and in particular not SDC). On the other hand, we claim that $\{A, B\}$ is ASDC: For $\epsilon > 0$, consider the matrices

$$\tilde{A} = \begin{pmatrix} & 1 & \\ 1 & & \\ & & \epsilon \end{pmatrix}, \qquad \tilde{B} = \begin{pmatrix} 1 & & \sqrt{\epsilon} \\ & -1 & \sqrt{\epsilon} \\ \sqrt{\epsilon} & \sqrt{\epsilon} & 0 \end{pmatrix}.$$

A straightforward computation shows that $\tilde{A}^{-1}\tilde{B}$ has simple eigenvalues $\{-1, 0, 1\}$ whence $\left\{\tilde{A}, \tilde{B}\right\}$ is SDC.

The fact that $\left\{\bar{A}, \bar{B}\right\}$ is not SDC is equivalent to the statement: there does *not* exist a $\mathbb{C}$-independent set $\{p_1, p_2\} \in \mathbb{C}^2$ such that the quadratic forms $x^*\bar{A}x$ and $x^*\bar{B}x$ can be expressed as

$$\begin{aligned} x^*\bar{A}x &= \alpha_1 \left|p_1^* x\right|^2 + \alpha_2 \left|p_2^* x\right|^2, \quad \text{and} \\ x^*\bar{B}x &= \beta_1 \left|p_1^* x\right|^2 + \beta_2 \left|p_2^* x\right|^2, \end{aligned}$$

for some $\alpha_i, \beta_i \in \mathbb{R}$. On the other hand, the fact that $\left\{\tilde{A}, \tilde{B}\right\}$ is SDC shows that there exists a $\mathbb{C}$-spanning set $\{p_1, p_2, p_3\} \subseteq \mathbb{C}^2$ and $\alpha_i, \beta_i \in \mathbb{R}$ such that

$$\begin{aligned} x^*\bar{A}x &= \alpha_1 \left|p_1^* x\right|^2 + \alpha_2 \left|p_2^* x\right|^2 + \alpha_3 \left|p_3^* x\right|^2, \quad \text{and} \\ x^*\bar{B}x &= \beta_1 \left|p_1^* x\right|^2 + \beta_2 \left|p_2^* x\right|^2 + \beta_3 \left|p_3^* x\right|^2. \end{aligned}$$

Intuitively, the ASDC property asks whether a set of quadratic forms can be (almost) diagonalized using n (the ambient dimension)-many linear forms whereas the SDC property may be forced to use a smaller number of linear forms. ☐

Theorem 2. *Let $\{A, B\} \subseteq \mathbb{H}^n$. If $\{A, B\}$ is singular, then it is ASDC.*

Proof. Without loss of generality, suppose A is a max-rank element of $\operatorname{span}(\{A, B\})$. We will apply Proposition 3 to $\{A, B\}$. Note that as A is singular, we have $m_3 + m_4 \geq 1$. We will break our proof into three cases: where $m \geq 2$ and $m_4 = 1$, where $m \geq 2$, $m_3 \geq 1$ and $m_4 = 0$, and where $m = 1$.

For cases 1 and 2, we will make the following simplifying assumptions: We will assume without loss of generality that

$$A = \begin{pmatrix} \bar{A} & \\ & A_m \end{pmatrix}, \qquad B = \begin{pmatrix} \bar{B} & \\ & B_m \end{pmatrix}.$$

In case 1, $A_m = B_m = 0_{1\times 1}$. In case 2,

$$A_m = \begin{pmatrix} & & F_{n_m} \\ & 0 & \\ F_{n_m} & & \end{pmatrix}, \qquad B_m = G_{2n_m+1}.$$

Furthermore, we may assume without loss of generality that $\bar{A}$ is invertible (else perturb $\bar{A}$), $\bar{A}^{-1}\bar{B}$ has simple eigenvalues (else apply Lemma 3), and that $\bar{A}^{-1}\bar{B}$ has only non-real eigenvalues (else consider only the submatrices of $\bar{A}$ and $\bar{B}$ corresponding to the complex eigenvalues of $\bar{A}^{-1}\bar{B}$). In particular, $\bar{A}, \bar{B} \in \mathbb{H}^{2k}$ for some $k \in \mathbb{N}$. Finally, we will work in the basis furnished by Proposition 3 for $\mathbb{C}^{2k}$ so that

$$\bar{A} = \left(\begin{array}{cc|c|cc} & 1 & & & \\ 1 & & & & \\ \hline & & \ddots & & \\ \hline & & & & 1 \\ & & & 1 & \end{array}\right), \qquad \bar{B} = \left(\begin{array}{cc|c|cc} & \lambda_1 & & & \\ \lambda_1^* & & & & \\ \hline & & \ddots & & \\ \hline & & & & \lambda_k \\ & & & \lambda_k^* & \end{array}\right). \tag{4}$$

Case 1. Set

$$\tilde{A}_\epsilon = \left(\begin{array}{cc|c|cc|c} & 1 & & & & \\ 1 & & & & & \\ \hline & & \ddots & & & \\ \hline & & & & 1 & \\ & & & 1 & & \\ \hline & & & & & \epsilon \end{array}\right), \qquad \tilde{B}_\epsilon = \left(\begin{array}{cc|c|cc|c} & \lambda_1 & & & & \sqrt{\epsilon}\alpha_1 \\ \lambda_1^* & & & & & \sqrt{\epsilon} \\ \hline & & \ddots & & & \vdots \\ \hline & & & & \lambda_k & \sqrt{\epsilon}\alpha_k \\ & & & \lambda_k^* & & \sqrt{\epsilon} \\ \hline \sqrt{\epsilon}\alpha_1^* & \sqrt{\epsilon} & \cdots & \sqrt{\epsilon}\alpha_k^* & \sqrt{\epsilon} & \epsilon z \end{array}\right)$$

for some $\alpha \in \mathbb{C}^k$, $z \in \mathbb{R}$, and $\epsilon > 0$. The eigenvalues of

$$\tilde{A}_\epsilon^{-1}\tilde{B}_\epsilon = \left(\begin{array}{cc|c|cc|c} \lambda_1^* & & & & & \sqrt{\epsilon} \\ & \lambda_1 & & & & \sqrt{\epsilon}\alpha_1 \\ \hline & & \ddots & & & \vdots \\ \hline & & & \lambda_k^* & & \sqrt{\epsilon} \\ & & & & \lambda_k & \sqrt{\epsilon}\alpha_k \\ \hline \frac{\alpha_1^*}{\sqrt{\epsilon}} & \frac{1}{\sqrt{\epsilon}} & \cdots & \frac{\alpha_k^*}{\sqrt{\epsilon}} & \frac{1}{\sqrt{\epsilon}} & z \end{array}\right)$$

are the roots (in the variable ξ) of

$$(z-\xi)\prod_{i=1}^{k}(\lambda_i-\xi)(\lambda_i^*-\xi) - \sum_{i=1}^{k}\left(2\operatorname{Re}(\alpha_i\lambda_i^*) - 2\operatorname{Re}(\alpha_i)\xi\right)\prod_{j\neq i}(\lambda_j-\xi)(\lambda_j^*-\xi) \tag{5}$$

and are independent of ϵ. By parameterizing $\alpha_i = \frac{y_i}{2} - \mathrm{i}\frac{x_i+\operatorname{Re}(\lambda_i)y_i}{2\operatorname{Im}(\lambda_i)}$ for $x_i, y_i \in \mathbb{R}$, the characteristic polynomial becomes

$$(z-\xi)\prod_{i=1}^{k}(\lambda_i-\xi)(\lambda_i^*-\xi) + \sum_{i=1}^{k}(x_i+y_i\xi)\prod_{j\neq i}(\lambda_j-\xi)(\lambda_j^*-\xi). \tag{6}$$

It suffices to show that there exist $x, y \in \mathbb{R}^n$ and $z \in \mathbb{R}$ such that the roots of (6) are all real, as we may take $\epsilon > 0$ to zero independently of our choice of x, y, z.

Define the following polynomials.

$$f_i(\xi) \coloneqq \prod_{j\neq i}(\lambda_j-\xi)(\lambda_j^*-\xi), \quad g_i(\xi) \coloneqq \xi f_i(\xi),\ \forall i \in [k], \quad \text{and}$$

$$h(\xi) \coloneqq \prod_{i=1}^{k}(\lambda_i-\xi)(\lambda_i^*-\xi).$$

As $\{\lambda_1, \lambda_1^*, \ldots, \lambda_k, \lambda_k^*\}$ are distinct values in $\mathbb{C}\setminus\mathbb{R}$, we have that $\{f_1, g_1, \ldots, f_k, g_k, h\}$ are a basis for the degree-$2k$ polynomials in ξ.

Now pick $2k+1$ distinct values $\xi_1, \ldots, \xi_{2k+1} \in \mathbb{R}$. Note that $\{\xi_1, \ldots, \xi_{2k+1}\}$ are the roots to (6) if and only if $x, y \in \mathbb{R}^n$ and $z \in \mathbb{R}$ satisfy

$$\begin{pmatrix} f_1(\xi_1) & g_1(\xi_1) & \cdots & f_k(\xi_1) & g_k(\xi_1) & h(\xi_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ f_1(\xi_{2k+1}) & g_1(\xi_{2k+1}) & \cdots & f_k(\xi_{2k+1}) & g_k(\xi_{2k+1}) & h(\xi_{2k+1}) \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \\ \vdots \\ x_k \\ y_k \\ z \end{pmatrix} = \begin{pmatrix} \xi_1 h(\xi_1) \\ \vdots \\ \xi_{2k+1}h(\xi_{2k+1}) \end{pmatrix}. \tag{7}$$

Note that the matrix on the left is invertible (as $\{f_1, g_1, \ldots, f_k, g_k, h\}$ is independent and the ξ_i are distinct) and real (as the ξ_i are real). Consequently, the matrix on the left has a real inverse. Note also that the vector on the right is real. We deduce that there exist $x, y \in \mathbb{R}^n$ and $z \in \mathbb{R}$ such that the eigenvalues of $\tilde{A}_\epsilon^{-1}\tilde{B}$ are real and simple.

Case 2. Set

$$\tilde{A}_\epsilon = \left(\begin{array}{c|c|c|c|c|c}
\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} & & & & & \\ \hline
 & \ddots & & & & \\ \hline
 & & \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} & & & \\ \hline
 & & & & & F_{n_m} \\ \hline
 & & & & \epsilon & \\ \hline
 & & & F_{n_m} & &
\end{array}\right), \quad
\tilde{B}_\epsilon = \left(\begin{array}{c|c|c|c|c|c}
\begin{smallmatrix} & \lambda_1 \\ \lambda_1^* & \end{smallmatrix} & & & & \begin{smallmatrix} \sqrt{\epsilon}\alpha_1 \\ \sqrt{\epsilon} \end{smallmatrix} & \\ \hline
 & \ddots & & & \vdots & \\ \hline
 & & \begin{smallmatrix} & \lambda_k \\ \lambda_k^* & \end{smallmatrix} & & \begin{smallmatrix} \sqrt{\epsilon}\alpha_k \\ \sqrt{\epsilon} \end{smallmatrix} & \\ \hline
 & & & & & G_{n_m} \\ \hline
\sqrt{\epsilon}\alpha_1^* \ \sqrt{\epsilon} & \cdots & \sqrt{\epsilon}\alpha_k^* \ \sqrt{\epsilon} & & \epsilon z & e_1^* \\ \hline
 & & & G_{n_m} & e_1 &
\end{array}\right)$$

for some $\alpha \in \mathbb{C}^k$, $z \in \mathbb{R}$, and $\epsilon > 0$. The eigenvalues of

$$\tilde{A}_\epsilon^{-1}\tilde{B}_\epsilon = \left(\begin{array}{c|c|c|c|c|c}
\begin{smallmatrix} \lambda_1^* & \\ & \lambda_1 \end{smallmatrix} & & & & \begin{smallmatrix} \sqrt{\epsilon} \\ \sqrt{\epsilon}\alpha_1 \end{smallmatrix} & \\ \hline
 & \ddots & & & \vdots & \\ \hline
 & & \begin{smallmatrix} \lambda_k^* & \\ & \lambda_k \end{smallmatrix} & & \begin{smallmatrix} \sqrt{\epsilon} \\ \sqrt{\epsilon}\alpha_k \end{smallmatrix} & \\ \hline
 & & & F_{n_m}G_{n_m} & e_{n_m} & \\ \hline
\frac{\alpha_1^*}{\sqrt{\epsilon}} \ \frac{1}{\sqrt{\epsilon}} & \cdots & \frac{\alpha_k^*}{\sqrt{\epsilon}} \ \frac{1}{\sqrt{\epsilon}} & & z & \frac{e_1^*}{\epsilon} \\ \hline
 & & & & & F_{n_m}G_{n_m}
\end{array}\right)$$

are the roots (in the variable ξ) of

$$\xi^{2n_m}\Bigg((z-\xi)\prod_{i=1}^{k}(\lambda_i-\xi)(\lambda_i^*-\xi) - \sum_{i=1}^{k}\left(2\operatorname{Re}(\alpha_i\lambda_i^*) - 2\operatorname{Re}(\alpha_i)\xi\right)\prod_{j\neq i}(\lambda_j-\xi)(\lambda_j^*-\xi)\Bigg)$$

and are independent of ϵ. As in Case 1 (cf. (5)), we may pick $\alpha \in \mathbb{C}^k$ and $z \in \mathbb{R}$ such that $\tilde{A}_\epsilon^{-1}\tilde{B}_\epsilon$ has real (but no longer necessarily simple) eigenvalues. Finally, applying Theorem 1, we deduce that for all $\epsilon > 0$, $\left\{\tilde{A}_\epsilon, \tilde{B}_\epsilon\right\}$ is ASDC. We conclude that $\{A, B\}$ is ASDC.

Case 3. In the final case, we have that $m = m_3 + m_4 = 1$. If $m_4 = 1$ (so that $A = B = 0$), it is clear that $\{A, B\}$ is actually SDC. Finally, suppose $m_3 = 1$ so that

$$A = \begin{pmatrix} & & F_{n_m} \\ & 0 & \\ F_{n_m} & & \end{pmatrix}, \qquad B = G_{2n_m+1}.$$

Then for $\epsilon \neq 0$, set

$$\tilde{A}_\epsilon = \begin{pmatrix} & & F_{n_m} \\ & \epsilon & \\ F_{n_m} & & \end{pmatrix}.$$

Note that $\tilde{A}^{-1}B$ is upper triangular with all diagonal entries equal to zero. Then applying Theorem 1, we deduce that for all $\epsilon \neq 0$, $\left\{\tilde{A}_\epsilon, B\right\}$ is ASDC. We conclude that $\{A, B\}$ is ASDC. ■

4 The ASDC property of nonsingular Hermitian triples

In this section, we will prove the following characterization of the ASDC property for nonsingular Hermitian triples.

Theorem 3. *Let $\{A, B, C\} \subseteq \mathbb{H}^n$ and suppose A is invertible. Then, $\{A, B, C\}$ is ASDC if and only if $\{A^{-1}B, A^{-1}C\}$ are a pair of commuting matrices with real eigenvalues.*

As always, the forward direction follows trivially from Proposition 1 and continuity. For the reverse direction, we will extend an inductive argument due to Motzkin and Taussky [20] to show that we may repeatedly perturb either $A^{-1}B$ or $A^{-1}C$ to increase the number of simple eigenvalues. In contrast to the original argument in [20], which establishes that any commuting pair $\{S, T\} \subseteq \mathbb{C}^{n\times n}$ is almost simultaneously diagonalizable *via similarity* (and thus only needs to inductively maintain commutativity of S and T), for our proof we will further need to maintain that A, B, C are Hermitian matrices and that $A^{-1}B$ and $A^{-1}C$ have real eigenvalues.

Our proof will require two technical facts about block matrices consisting of upper triangular Toeplitz blocks. We present these facts below and defer their proofs to Appendix B.

Definition 5. $T \in \mathbb{C}^{n_i\times n_j}$ is an *upper triangular Toeplitz matrix* if T is of the form

$$T = \left(\begin{array}{c|cccc} 0_{n_i\times(n_j-n_i)} & t^{(1)} & t^{(2)} & \cdots & t^{(n_i)} \\ & & t^{(1)} & \ddots & \vdots \\ & & & \ddots & t^{(2)} \\ & & & & t^{(1)} \end{array}\right) \quad \text{or} \quad T = \left(\begin{array}{cccc} t^{(1)} & t^{(2)} & \cdots & t^{(n_j)} \\ & t^{(1)} & \ddots & \vdots \\ & & \ddots & t^{(2)} \\ & & & t^{(1)} \\ \hline \multicolumn{4}{c}{0_{(n_i-n_j)\times n_j}} \end{array}\right)$$

if $n_i \leq n_j$ and $n_j \leq n_i$ respectively. $\square$

Definition 6. Let $(n_1, \ldots, n_k)$ such that $\sum_i n_i = n$. Let $\mathbb{T}(n_1, \ldots, n_k) \subseteq \mathbb{C}^{n\times n}$ denote the linear subspace of matrices T such that each block $T_{i,j}$ (when the rows and columns of T are partitioned according to $(n_1, \ldots, n_k)$) is an upper triangular Toeplitz matrix. When the partition $(n_1, \ldots, n_k)$ is clear from context, we will simply write $\mathbb{T}$. $\square$

The following well-known fact characterizes the set of matrices which commute with a nilpotent Jordan chain (see for example [24, Theorem 6]).

Lemma 4. *Let $(n_1, \ldots, n_k)$ with $\sum_i n_i = n$. Let $J \in \mathbb{C}^{n\times n}$ be a block diagonal matrix with diagonal block $J_{i,i} = F_{n_i}G_{n_i}$, i.e., a Jordan block of size n_i and eigenvalue zero. Then, $T \in \mathbb{C}^{n\times n}$ commutes with J if and only if $T \in \mathbb{T}$.*

Definition 7. Let $(n_1, \ldots, n_k)$ such that $\sum_i n_i = n$. Define the linear map $\Pi_{(n_1,\ldots,n_k)} : \mathbb{T}(n_1, \ldots, n_k) \to \mathbb{C}^{k\times k}$ by

$$(\Pi_{(n_1,\ldots,n_k)}(T))_{i,j} = \begin{cases} T^{(1)}_{i,j} & \text{if } n_i = n_j, \\ 0 & \text{else.} \end{cases}$$

When the partition $(n_1, \ldots, n_k)$ is clear from context, we will simply write Π. $\square$

The following fact follows from the observation that the characteristic polynomial of a matrix $T \in \mathbb{T}$ depends on only a few of its entries (see Lemma 8).

Lemma 5. *Let $(n_1, \ldots, n_k)$ such that $\sum_i n_i = n$. Then, for any $T \in \mathbb{T}$, the matrices T and $\Pi(T)$ have the same eigenvalues.*

We are now ready to prove Theorem 3.

Proof of Theorem 3. Suppose $n \geq 3$ and inductively assume that the statement is true for all smaller n (it is possible to check via elementary arguments that the statement is true for $n \leq 2$). We will break our proof into four cases: First, we will consider when either $A^{-1}B$ or $A^{-1}C$ (without loss of generality $A^{-1}B$) has multiple eigenvalues. Failing this case, we may then assume by considering the basis $\{A, B+\lambda_B A, C+\lambda_C A\}$ (for an appropriate choice of λ_B and λ_C) of $\operatorname{span}(\{A,B,C\})$, that $A^{-1}B$ and $A^{-1}C$ are nilpotent (note that this reduction requires $A^{-1}B$ and $A^{-1}C$ to have real eigenvalues). The remaining three cases will consider when the Jordan block structure of $A^{-1}B$ has: multiple block sizes, multiple blocks of the same size, and a single block.

Without loss of generality, we will work in the basis furnished by Proposition 3 so that $A^{-1}B$ is in Jordan canonical form. We may further assume that the blocks of $A^{-1}B$ are ordered first according to increasing eigenvalues then increasing block sizes.

Case 1. Suppose $A^{-1}B$ has ℓ-many distinct eigenvalues. Write C as an $\ell \times \ell$ block matrix according to the partition induced by the eigenvalues of $A^{-1}B$. Then, as $A^{-1}C$ and $A^{-1}B$ commute (whence $A^{-1}C$ respects the generalized eigenspaces of $A^{-1}B$), we have that $A^{-1}C$ (perforce C) is block diagonal. Thus, according to the block structure induced by the eigenvalues of $A^{-1}B$, the matrices A, B, C are jointly block diagonal, with each diagonal block satisfying the conditions of the inductive hypothesis. We conclude that $\{A, B, C\}$ is ASDC.

Case 2. Suppose $A^{-1}B$ and $A^{-1}C$ are nilpotent and that $A^{-1}B$ has distinct block sizes. For concreteness, suppose $A^{-1}B$ has k blocks of size $\eta = n_1 = \cdots = n_k < n_{k+1} \leq \cdots \leq n_m$. By Proposition 3,

$$A = \operatorname{Diag}(\sigma_1 F_\eta, \ldots, \sigma_k F_\eta, \sigma_{k+1} F_{n_{k+1}}, \ldots, \sigma_m F_{n_m})$$

for some $\sigma_i \in \{\pm 1\}$. Set

$$\tilde{C} = C + \epsilon \operatorname{Diag}(\sigma_1 F_\eta, \ldots, \sigma_k F_\eta, 0_{n_{k+1}}, \ldots, 0_{n_m}).$$

Applying Lemma 4, we have that $A^{-1}\tilde{C}$ commutes with $A^{-1}B$ and that $\tilde{C} \in \mathbb{H}^n$. Let Π denote the linear map furnished by Lemma 5. As $n_i \neq n_j$ for all $i \leq k$ and $j \geq k+1$, we have that $\Pi(A^{-1}C)$ can be written as a block diagonal matrix

$$\Pi(A^{-1}C) = \begin{pmatrix} \Pi(A^{-1}C)_{1,1} & \\ & \Pi(A^{-1}C)_{2,2} \end{pmatrix}$$

with blocks of size $\eta k \times \eta k$ and $(n - \eta k) \times (n - \eta k)$ respectively. As Π preserves eigenvalues for inputs in $\mathbb{T}$, we have that $\Pi(A^{-1}C)_{1,1}$ and $\Pi(A^{-1}C)_{2,2}$ are both nilpotent. Then, as $A^{-1}\tilde{C}$ has the same eigenvalues as

$$\Pi(A^{-1}\tilde{C}) = \begin{pmatrix} \Pi(A^{-1}C)_{1,1} + \epsilon I_{\eta k} & \\ & \Pi(A^{-1}C)_{2,2} \end{pmatrix},$$

we deduce that $A^{-1}\tilde{C}$ has eigenvalues $\{0, \epsilon\}$. We have reduced to case (1) whence $\left\{A, B, \tilde{C}\right\}$ is ASDC. We conclude that $\{A, B, C\}$ is ASDC.

Case 3. Suppose $A^{-1}B$ and $A^{-1}C$ are nilpotent and that $A^{-1}B$ has Jordan blocks all of the same dimension. For concreteness, suppose $A^{-1}B$ has $k \geq 2$ Jordan blocks of dimension η. In this case Proposition 3 states that

$$A = \operatorname{Diag}(\sigma_1, \ldots, \sigma_k) \otimes F_\eta \quad \text{and} \quad B = \operatorname{Diag}(\sigma_1, \ldots, \sigma_k) \otimes G_\eta$$

where $\sigma_i \in \{\pm 1\}$. Write C as a $k \times k$ block matrix with blocks $C_{i,j} \in \mathbb{C}^{\eta\times\eta}$. By Lemma 4, $A^{-1}C \in \mathbb{T}$ and we may write

$$C_{i,j} = F_\eta \left(\gamma_{i,j}^{(1)} I_\eta + \sum_{\ell=2}^{\eta} \gamma_{i,j}^{(\ell)} (F_\eta G_\eta)^{\ell-1} \right).$$

Let Π denote the linear map furnished by Lemma 5. Let $\bar{A} = \mathrm{Diag}(\sigma_1, \dots, \sigma_k)$ and

$$\bar{C} = \left(\gamma_{i,j}^{(1)} \right).$$

Note that as $C \in \mathbb{H}^n$, we have $\gamma_{i,j}^{(1)} = \left(\gamma_{j,i}^{(1)} \right)^*$, whence $\bar{A}, \bar{C} \in \mathbb{H}^k$. As Π preserves the eigenvalues for inputs in $\mathbb{T}$ and $\bar{A}^{-1}\bar{C} = \Pi(A^{-1}C)$, we deduce that $\bar{A}^{-1}\bar{C}$ has real eigenvalues (in fact, the single eigenvalue 0). Then applying Lemma 3, there exists $\bar{C}' \in \mathbb{H}^k$ such that $\left\| \bar{C} - \bar{C}' \right\| \le \epsilon$ and $\bar{A}^{-1}\bar{C}'$ has k-many distinct real eigenvalues. Finally, set

$$\tilde{C} = C + (\bar{C}' - \bar{C}) \otimes F_\eta.$$

Then Lemma 4 implies that $A^{-1}B$ and $A^{-1}\tilde{C}$ commute. Furthermore, by construction, $A^{-1}\tilde{C}$ has upper triangular Toeplitz blocks so that its eigenvalues are the same as the eigenvalues of $\Pi(A^{-1}\tilde{C}) = \bar{A}^{-1}\bar{C}'$. We have reduced to case (1) and $\left\{A, B, \tilde{C}\right\}$ is ASDC. We conclude that $\{A, B, C\}$ is also ASDC.

Case 4. Suppose $A^{-1}B$ and $A^{-1}C$ are nilpotent and that $A^{-1}B$ is a single Jordan block. Then, by Proposition 3,

$$A = \sigma F_n \quad \text{and} \quad B = \sigma G_n$$

for some $\sigma \in \{\pm 1\}$. Furthermore, by Lemma 4 and the assumption that $A^{-1}C$ is nilpotent, we may write

$$C = \sigma F_n \left(\sum_{i=2}^{n} c_i (F_n G_n)^{i-1} \right)$$

for some $c_2, \dots, c_n \in \mathbb{R}$. Now, for $\epsilon > 0$, set

$$\tilde{B} = B + \epsilon\sigma(e_1 e_n^* + e_n e_1^*)$$
$$\tilde{C} = C + \sigma(e_n \gamma^* + \gamma e_n^*)$$

where $\gamma \in \mathbb{R}^n$ is defined recursively as $\gamma_n = \gamma_{n-1} = 0$ and $\gamma_i = \epsilon(c_{i+1} + \gamma_{i+1})$ for $i \in [n-2]$. A straightforward calculation shows that $A^{-1}\tilde{B}$ and $A^{-1}\tilde{C}$ commute and both have real eigenvalues. Finally, as $A^{-1}\tilde{B}$ has distinct eigenvalues $\{0, \epsilon\}$, we have reduced to case (1) and $\left\{A, \tilde{B}, \tilde{C}\right\}$ is ASDC. We conclude that $\{A, B, C\}$ is also ASDC. ■

5 Restricted SDC

In this section, we investigate a variant of the ASDC property that will be useful for applications in Section 7. We will see soon that we have in fact already seen this property before in Section 3.

Definition 8. Let $\mathcal{A} \subseteq \mathbb{H}^n$ and $d \in \mathbb{N}$. We will say that $\mathcal{A}$ is *d-restricted SDC* (d-RSDC) if there exists a mapping $f : \mathcal{A} \to \mathbb{H}^{n+d}$ such that

- for all $A \in \mathcal{A}$, the top-left $n \times n$ principal submatrix of $f(A)$ is A, and
- $f(\mathcal{A})$ is SDC. □

We record some simple consequences of the d-RSDC property that follow from Observation 1 and Lemma 2.

Observation 4. *Let $\mathcal{A} \subseteq \mathbb{H}^n$ and $d \in \mathbb{N}$. Then*

- *$\mathcal{A}$ is d-RSDC if and only if $\{A_1, \dots, A_m\}$ is d-RSDC for some basis $\{A_1, \dots, A_m\}$ of* $\operatorname{span}(\mathcal{A})$.
- *if $\mathcal{A}$ is d-RSDC, then $\mathcal{A}$ is d'-RSDC for all $d' \geq d$.*

The following lemma explains the connection between the d-RSDC property and the ASDC property.

Lemma 6. *Let $A_1, \dots, A_m \in \mathbb{H}^n$ and let $d \in \mathbb{N}$. If $\mathcal{A} = \{A_1, \dots, A_m\}$ is d-RSDC, then*

$$\mathcal{A} \oplus 0_d \coloneqq \left\{ \begin{pmatrix} A_i & \\ & 0_d \end{pmatrix} : i \in [m] \right\}$$

is ASDC. On the other hand, if $\mathcal{A} \oplus 0_d$ is ASDC, then for all $\epsilon > 0$, there exist $\tilde{A}_1, \dots, \tilde{A}_m \in \mathbb{H}^n$ such that

- *for all $i \in [m]$, the spectral norm $\left\| A_i - \tilde{A}_i \right\| \leq \epsilon$, and*
- *$\left\{ \tilde{A}_1, \dots, \tilde{A}_m \right\}$ is d-RSDC.*

Proof. First, suppose $\{A_1, \dots, A_m\}$ is d-RSDC. Let f denote the map furnished by d-RSDC and define $\tilde{A}_i = f(A_i)$. Next, let $\epsilon > 0$ and set

$$P = \begin{pmatrix} I_n & \\ & \sqrt{\epsilon} I_d \end{pmatrix}.$$

Clearly, P is invertible so that $\left\{ P^* \tilde{A}_i P : i \in [m] \right\}$ is also SDC. Then, note that

$$P^* \tilde{A}_i P = P^* \begin{pmatrix} A_i & (\tilde{A}_i)_{1,2} \\ (\tilde{A}_i)_{1,2}^* & (\tilde{A}_i)_{2,2} \end{pmatrix} P = \begin{pmatrix} A_i & \sqrt{\epsilon}(\tilde{A}_i)_{1,2} \\ \sqrt{\epsilon}(\tilde{A}_i)_{1,2}^* & \epsilon(\tilde{A}_i)_{2,2} \end{pmatrix}$$

so that $\mathcal{A} \oplus 0_d$ is ASDC.

Next, suppose $\mathcal{A} \oplus 0_d$ is ASDC and let $\epsilon > 0$. Then, there exist $\bar{A}_1, \dots, \bar{A}_m \in \mathbb{H}^{n+d}$ such that $\left\| \bar{A}_i - A_i \oplus 0_d \right\| \leq \epsilon$ and $\left\{ \bar{A}_1, \dots, \bar{A}_m \right\}$ is SDC. Finally, note that $\left\| A_1 - (\bar{A}_1)_{1,1} \right\| \leq \epsilon$. ■

Remark 3. While the restriction of an SDC set does not necessarily result in an SDC set, there is a setting arising naturally when analyzing QCQPs in which the restriction of an SDC set is again SDC. Specifically, let $Q_1, \dots, Q_m \in \mathbb{H}^{n+1}$ where Q_i has A_i as its top-left $n \times n$ principal submatrix. Furthermore suppose that there exists a positive definite matrix in $\operatorname{span}(\{A_1, \dots, A_m\})$. Then, if $\{Q_1, \dots, Q_m, e_{n+1}e_{n+1}^*\}$ is SDC, so is $\{A_1, \dots, A_m\}$. In words, if the homogenized quadratic forms in a QCQP, along with $e_{n+1}e_{n+1}^*$, are SDC, then so are the original quadratic forms (under a standard "definiteness" assumption). See Appendix D for details. □

Finally, we record a recasting of Theorem 2 in terms of these new definitions.

Theorem 4. *Let $A, B \in \mathbb{H}^n$. Then for every $\epsilon > 0$, there exist $\tilde{A}, \tilde{B} \in \mathbb{H}^n$ such that $\left\|A - \tilde{A}\right\|, \left\|B - \tilde{B}\right\| \leq \epsilon$ and $\left\{\tilde{A}, \tilde{B}\right\}$ is 1-RSDC. Furthermore, if A is invertible and $A^{-1}B$ has simple eigenvalues, then $\{A, B\}$ is itself 1-RSDC.*

Corollary 2. *Let $n \in \mathbb{N}$ and let $(A, B) \in \mathbb{H}^n \times \mathbb{H}^n$ be a pair of matrices jointly sampled according to an absolutely continuous probability measure on $\mathbb{H}^n \times \mathbb{H}^n$. Then, $\{A, B\}$ is 1-RSDC almost surely.*

6 Obstructions to further generalization

In this section, we record explicit counterexamples to *a priori* plausible extensions to Theorems 1 to 3.

6.1 Singular Hermitian triples

In Theorem 2, we showed that any singular Hermitian pair is ASDC. A natural question to ask is whether any singular set of Hermitian matrices (regardless of the dimension of its span) is ASDC. The following theorem presents an obstruction to generalizations in this direction. Specifically, in contrast to Theorem 2 (where it was shown that singularity implies ASDC in the context of Hermitian pairs), Theorem 5 below shows that even Hermitian *triples* with *"large amounts"* of singularity can fail to be ASDC.

Theorem 5. *Let $\{A = I_n, B, C\} \subseteq \mathbb{H}^n$. Then, if $d < \operatorname{rank}([B, C])/2$, the set*

$$\left\{\begin{pmatrix} A & \\ & 0_d \end{pmatrix}, \begin{pmatrix} B & \\ & 0_d \end{pmatrix}, \begin{pmatrix} C & \\ & 0_d \end{pmatrix}\right\}$$

is not ASDC.

Proof. Suppose for the sake of contradiction that this set is ASDC. Let $\epsilon \in (0, 1/2)$ and let $\left\{\tilde{A}, \tilde{B}, \tilde{C}\right\} \subseteq \mathbb{H}^{n+d}$ denote an SDC set furnished by the ASDC assumption. Without loss of generality, $\tilde{A}$ has rank $n + d$. Write

$$\tilde{A} = \begin{pmatrix} \tilde{A}_{1,1} & \tilde{A}_{1,2} \\ \tilde{A}_{1,2}^* & \tilde{A}_{2,2} \end{pmatrix}.$$

Similarly decompose $\tilde{B}$ and $\tilde{C}$. As $\epsilon \in (0, 1/2)$, we have that $\tilde{A}_{1,1}$ is invertible. Let

$$P = \begin{pmatrix} \tilde{A}_{1,1}^{-1/2} & -\tilde{A}_{1,1}^{-1}\tilde{A}_{1,2} \\ 0 & I_d \end{pmatrix}.$$

Then as P is invertible, $\left\{P^*\tilde{A}P, P^*\tilde{B}P, P^*\tilde{C}P\right\}$ is again SDC. Note that $P^*\tilde{A}P$ has the form

$$P^*\tilde{A}P = \begin{pmatrix} I_n & \\ & \tilde{A}_{2,2} - \tilde{A}_{1,2}^*\tilde{A}_{1,1}^{-1}\tilde{A}_{1,2} \end{pmatrix}.$$

Furthermore,

$$\begin{aligned}
&\left\|P^*\tilde{B}P - B\right\| \\
&\quad = \left\|(P - I_{n+d})^*\tilde{B}(P - I_{n+d}) + \tilde{B}(P - I_{n+d}) + (P - I_{n+d})^*\tilde{B} + (\tilde{B} - B)\right\| \\
&\quad \leq \left\|\tilde{B}\right\| \|P - I_{n+d}\|^2 + 2\left\|\tilde{B}\right\| \|P - I_{n+d}\| + \epsilon.
\end{aligned}$$

We claim that $\|P - I_{n+d}\|$ can be bounded in terms of ϵ:

$$\begin{aligned}
\|P - I_{n+d}\| &\leq \left\|\tilde{A}_{1,1}^{-1/2} - I\right\| + \left\|\tilde{A}_{1,1}^{-1}\right\| \left\|\tilde{A}_{1,2}\right\| \\
&\leq \max\left\{\frac{1}{\sqrt{1-\epsilon}} - 1,\, 1 - \frac{1}{\sqrt{1+\epsilon}}\right\} + \frac{\epsilon}{1-\epsilon} \\
&\leq \frac{2\epsilon}{1-\epsilon}.
\end{aligned}$$

Here, we have used the fact that $\left\|\tilde{A} - A \oplus 0_d\right\| \leq \epsilon$, so that $\left\|\tilde{A}_{1,1} - I_n\right\| \leq \epsilon$ and $\left\|\tilde{A}_{1,2}\right\| \leq \epsilon$. Consequently, as we may also bound $\left\|\tilde{B}\right\| \leq \|B\| + \epsilon$, we deduce that for any $\delta > 0$, we can pick $\epsilon \in (0, 1/2)$ small enough such that $\left\|P^*\tilde{B}P - B\right\| \leq \delta$. An identical calculation holds for $\left\|P^*\tilde{C}P - C\right\|$. We conclude that for all $\delta > 0$, there exist $\bar{A}, \bar{B}, \bar{C}$ of the form

$$\bar{A} = \begin{pmatrix} I_n & \\ & \bar{A}_{2,2} \end{pmatrix}, \quad \bar{B} = \begin{pmatrix} \bar{B}_{1,1} & \bar{B}_{1,2} \\ \bar{B}_{1,2}^* & \bar{B}_{2,2} \end{pmatrix}, \quad \bar{C} = \begin{pmatrix} \bar{C}_{1,1} & \bar{C}_{1,2} \\ \bar{C}_{1,2}^* & \bar{C}_{2,2} \end{pmatrix}$$

such that $\left\{\bar{A}, \bar{B}, \bar{C}\right\}$ is SDC, $\left\|A - \bar{A}\right\|, \left\|B - \bar{B}\right\|, \left\|C - \bar{C}\right\| \leq \delta$, and $\bar{A}_{2,2}$ is invertible. Then by Proposition 1, the top-left block of the commutator $[\bar{A}^{-1}\bar{B}, \bar{A}^{-1}\bar{C}]$ is equal to 0_n. Expanding this top-left block, we deduce

$$[\bar{B}_{1,1}, \bar{C}_{1,1}] = \bar{C}_{1,2}\bar{A}_{2,2}^{-1}\bar{B}_{1,2}^* - \bar{B}_{1,2}\bar{A}_{2,2}^{-1}\bar{C}_{1,2}^*. \tag{8}$$

Finally, by lower semi-continuity of rank, we have $\operatorname{rank}([\bar{B}_{1,1}, \bar{C}_{1,1}]) \geq \operatorname{rank}([B, C])$ for all $\delta > 0$ small enough. This is a contradiction as the expression on the right of (8) has rank at most $2d < \operatorname{rank}([B, C])$. ■

This same construction can be viewed as an obstruction to generalizations of Theorem 4 to Hermitian triples with constant d.

Corollary 3. *Let $\{A = I_n, B, C\} \subseteq \mathbb{H}^n$. Then $A^{-1}B$ and $A^{-1}C$ are both diagonalizable with real eigenvalues and $\{A, B, C\}$ is not d-RSDC for any $d < \operatorname{rank}([B, C])/2$.*

Remark 4. Note that for all $n \in \mathbb{N}$, there exist B, $C \in \mathbb{H}^{2n}$ such that $\operatorname{rank}([B, C]) = 2n$. For example, set

$$B = \begin{pmatrix} I_n & \\ & -I_n \end{pmatrix}, \qquad C = \begin{pmatrix} & I_n \\ I_n & \end{pmatrix}.$$

Then, $\{A = I_{2n}, B, C\} \subseteq \mathbb{H}^{2n}$ is a nonsingular Hermitian triple such that $A^{-1}B$ and $A^{-1}C$ are both diagonalizable. On the other hand, Theorem 6 and Corollary 3 imply that

$$\left\{\begin{pmatrix} A & \\ & 0_{n-1} \end{pmatrix}, \begin{pmatrix} B & \\ & 0_{n-1} \end{pmatrix}, \begin{pmatrix} C & \\ & 0_{n-1} \end{pmatrix}\right\}$$

is not ASDC and $\{A, B, C\}$ is not $(n-1)$-RSDC. □

6.2 Nonsingular five-tuples

We may reinterpret Theorems 1 and 3 as saying that if $\mathcal{A}$ satisfies $\dim(\operatorname{span}(\mathcal{A})) \leq 3$ and contains an invertible matrix S, then $\mathcal{A}$ is ASDC if and only if $S^{-1}\mathcal{A}$ consists of a set of commuting matrices with real eigenvalues. A natural question to ask is whether the same statement holds without any assumption on the dimension of the span of $\mathcal{A}$. Theorem 6 below presents an obstruction to generalizations in this direction. Specifically, Theorem 6 constructs a non-ASDC set $\mathcal{A} = \{A_1, \dots, A_5\} \subseteq \mathbb{H}^4$ where A_1 is invertible and $A_1^{-1}\mathcal{A}$ consists of a set of commuting matrices with real eigenvalues.

The following lemma adapts a technique introduced by O'meara and Vinsonhaler [22] for studying the almost simultaneously diagonalizable via similarity property of subsets of $\mathbb{C}^{n\times n}$.

Lemma 7. *Let $\mathcal{A} = \{A_1, \dots, A_m\} \subseteq \mathbb{H}^n$ where $A_1 \in \mathcal{A}$ is invertible. If $\mathcal{A}$ is SDC, then $\dim(\mathbb{R}[A_1^{-1}\mathcal{A}]) \leq n$. Here, $\mathbb{R}[A_1^{-1}\mathcal{A}]$ is the real algebra generated by $A_1^{-1}\mathcal{A}$.*

Proof. Let P denote the invertible matrix furnished by SDC and suppose $P^* A_i P = D_i$. Then,

$$\dim\left(\mathbb{R}\left[A_1^{-1}\mathcal{A}\right]\right) = \dim\left(\mathbb{R}\left[\left\{D_1^{-1}D_i : i \in [m]\right\}\right]\right) \leq n. \qquad \blacksquare$$

The following corollary then follows by lower semi-continuity.

Corollary 4. *Let $\mathcal{A} = \{A_1, \dots, A_m\} \subseteq \mathbb{H}^n$ where $A_1 \in \mathcal{A}$ is invertible. If $\mathcal{A}$ is ASDC, then $\dim(\mathbb{R}[A_1^{-1}\mathcal{A}]) \leq n$. Here, $\mathbb{R}[A_1^{-1}\mathcal{A}]$ is the real algebra generated by $A_1^{-1}\mathcal{A}$.*

Theorem 6. *There exists a set $\mathcal{A} = \{A_1, \dots, A_5\} \subseteq \mathbb{H}^4$ such that A_1 is invertible, $A_1^{-1}\mathcal{A}$ is a set of commuting matrices with real eigenvalues, and $\mathcal{A}$ is* not *ASDC.*

Proof. Set

$$A_1 = \begin{pmatrix} & & & 1\\ & & 1 & \\ & 1 & & \\ 1 & & & \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & 1\\ & & 1 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & -\mathrm{i}\\ & & \mathrm{i} & 0 \end{pmatrix},$$
$$A_4 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 1 \end{pmatrix}.$$

Note that A_1 is invertible. It is not hard to verify that $A_1^{-1}\mathcal{A}$ forms a set of commuting matrices with real eigenvalues. On the other hand, note that

$$\mathbb{R}[A_1^{-1}\mathcal{A}] = \left\{ \begin{pmatrix} a & & b+c\mathrm{i} & e\\ & a & d & b-c\mathrm{i}\\ & & a & \\ & & & a \end{pmatrix} : a, b, c, d, e \in \mathbb{R} \right\}$$

so that $\dim(\mathbb{R}[A_1^{-1}\mathcal{A}]) = 5 > 4 = n$. We deduce from Corollary 4 that $\mathcal{A}$ is *not* ASDC. $\blacksquare$

7 Applications to quadratically constrained quadratic programming

In this section, we suggest several applications of the ASDC and d-RSDC properties to optimizing QCQPs. We break from the convention thus far and state the ideas in this section in terms of real symmetric matrices and QCQPs over $\mathbb{R}^n$. Nevertheless, the same ideas can be applied to the Hermitian setting and QCQPs over $\mathbb{C}^n$—a setting which arises frequently in signal processing and communication applications [1, 11, 12].

Consider a QCQP over $\mathbb{R}^n$ of the form

$$\mathrm{Opt} \coloneqq \inf_{x\in\mathbb{R}^n} \left\{ q_1(x) : \begin{array}{l} q_i(x) \le 0,\ \forall i \in [2,m] \\ x \in \mathcal{L} \end{array} \right\},$$

where for each $i \in [m]$, $q_i : \mathbb{R}^n \to \mathbb{R}$ is a quadratic function of the form $q_i(x) = x^\intercal A_i x + 2b_i^\intercal x + c_i$ with $A_i \in \mathbb{S}^n$, $b_i \in \mathbb{R}^n$, $c_i \in \mathbb{R}$, and $\mathcal{L} \subseteq \mathbb{R}^n$ is a polytope.

Utilizing the ASDC property. We make the following trivial observation.

Observation 5. *Suppose $A_1, \dots, A_m \in \mathbb{S}^n$ and $\tilde{A}_1, \dots, \tilde{A}_m \in \mathbb{S}^n$ satisfy $\left\| A_i - \tilde{A}_i \right\| \le \epsilon$ for all $i \in [m]$. Furthermore, suppose $\mathcal{L} \subseteq B(0,R)$, the ball of radius R centered at the origin, and set $\delta = \epsilon R^2$. Then, define*

$$\tilde{\mathrm{Opt}} \coloneqq \inf_{x\in\mathbb{R}^n} \left\{ \tilde{q}_1(x) : \begin{array}{l} \tilde{q}_i(x) \le 0,\ \forall i \in [2,m] \\ x \in \mathcal{L} \end{array} \right\}, \text{ and}$$

$$\mathrm{Opt}_\pm \coloneqq \inf_{x\in\mathbb{R}^n} \left\{ q_1(x) \pm \delta : \begin{array}{l} q_i(x) \pm \delta \le 0,\ \forall i \in [2,m] \\ x \in \mathcal{L} \end{array} \right\},$$

where $\tilde{q}_i(x) = x^\intercal \tilde{A}_i x + 2b_i^\intercal x + c_i$. Then, $\mathrm{Opt}_+ \ge \tilde{\mathrm{Opt}} \ge \mathrm{Opt}_-$.

In particular, when $\{A_1, \dots, A_m\}$ is ASDC, we may set $\delta > 0$ to be arbitrarily small and further have that $\left\{ \tilde{A}_1, \dots, \tilde{A}_m \right\}$ is SDC. In other words, if we are willing to lose arbitrarily small additive errors in both the objective function and the constraints of a QCQP with ASDC quadratic forms, then we may approximate the QCQP by a diagonalizable QCQP.

Utilizing the d-RSDC property. Next, suppose $\{A_1, \dots, A_m\}$ is d-RSDC and let $\tilde{A}_1, \dots, \tilde{A}_m \in \mathbb{S}^{n+d}$ denote the furnished matrices. Then defining the quadratic functions $\tilde{q}_i : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ by

$$\tilde{q}_i(x,y) \coloneqq \begin{pmatrix} x \\ y \end{pmatrix}^\intercal \tilde{A}_i \begin{pmatrix} x \\ y \end{pmatrix} + 2b_i^\intercal x + c_i$$

for all $i \in [m]$, we have that

$$\mathrm{Opt} = \inf_{(x,y)\in\mathbb{R}^n\times\mathbb{R}^d} \left\{ \tilde{q}_1(x,y) : \begin{array}{l} \tilde{q}_i(x,y) \le 0,\ \forall i \in [2,m] \\ x \in \mathcal{L},\ y = 0 \end{array} \right\}.$$

In particular, the augmented QCQP (which by construction is diagonalizable) is an exact reformulation of the original QCQP.

7.1 Numerical results

In this subsection, we present some preliminary numerical results on the d-RSDC property and its applicability in solving QCQP problems with only two quadratic forms. While the experiments are performed on synthetic instances and are perhaps not representative of "real-world" instances, *we believe they shed light on unexplored questions in the area and highlight interesting future directions.*

We will consider random instances of the following problem

$$\min_{x\in\mathbb{R}^n} \left\{ x^\intercal A_1 x + 2b_1^\intercal x : \begin{array}{l} x^\intercal A_2 x + 2b_2^\intercal x \le 1 \\ Lx \le 1 \end{array} \right\} \tag{9}$$

where A_1, $A_2 \in \mathbb{S}^n$, $b_2 \in \mathbb{R}^n$, and $L \in \mathbb{R}^{m\times n}$. We will additionally ensure that $\{x : Lx \le 1\}$ is a polytope.

Algorithm 1 Construction for 1-RSDC pair

Given $A_1,\, A_2 \in \mathbb{S}^n$ such that $A_1^{-1}A_2$ has simple eigenvalues

1. Let $P \in \mathbb{R}^{n\times n}$ denote the invertible matrix guaranteed by [26]; this can be computed using an eigenvalue decomposition of $A_1^{-1}A_2$. Then $P^\intercal A_1 P = \operatorname{Diag}(\sigma_1, \dots, \sigma_r, F_2, \dots, F_2)$ and $P^\intercal A_2 P = \operatorname{Diag}(\sigma_1\mu_1, \dots, \sigma_r\mu_r, T_1, \dots, T_k)$. Here, $\sigma_1, \dots, \sigma_r \in \{\pm 1\}$, $\mu_1, \dots, \mu_r \in \mathbb{R}$ and for $i \in [k]$, the matrix T_i has the form
$$T_i = \begin{pmatrix} \operatorname{Im}(\lambda_i) & \operatorname{Re}(\lambda_i) \\ \operatorname{Re}(\lambda_i) & -\operatorname{Im}(\lambda_i) \end{pmatrix}$$
for some $\lambda_i \in \mathbb{C} \setminus \mathbb{R}$.
2. Choose an arbitrary set of $2k+1$ distinct points $\xi_1, \dots, \xi_{2k+1} \in \mathbb{R}$.
3. Solve for $x,\, y \in \mathbb{R}^k$ and $z \in \mathbb{R}$ in the linear system (7).
4. Let $\alpha,\, \beta \in \mathbb{R}^k$ so that
$$x_i = \operatorname{Im}(\lambda_i)(\beta_i^2 - \alpha_i^2) - 2\operatorname{Re}(\alpha_i\beta_i), \quad \text{and} \quad y_i = 2\alpha_i\beta_i,\, \forall i \in [k]$$
and define $\gamma \in \mathbb{R}^{r+2k}$ as
$$\gamma = \begin{pmatrix} 0_{1\times k} & \alpha_1 & \beta_1 & \dots & \alpha_k & \beta_k \end{pmatrix}^\intercal.$$
5. Let $Q = P^{-1} \oplus I_1$ and return
$$\tilde{A}_1 = Q^\intercal \begin{pmatrix} P^\intercal A_1 P & \\ & 1 \end{pmatrix} Q, \qquad \tilde{A}_2 = Q^\intercal \begin{pmatrix} P^\intercal A_2 P & \gamma \\ \gamma^\intercal & z \end{pmatrix} Q.$$

Random model. We will consider a family of distributions over instances of (9) parameterized by $n \in \mathbb{N}$ and $k \in \mathbb{N}_0$. Here, k will parameterize the number of (pairs of) complex eigenvalues of $A^{-1}B$. Specifically, given (n, k) such that $2k \le n$:

1. Let $r = n - 2k$
2. Generate a random orthogonal matrix V by taking M to be a random $n \times n$ matrix with entries i.i.d. $N(0,1)$ and then taking V to be a matrix of left singular vectors of M. Let $\sigma_1, \dots, \sigma_r$ be i.i.d. Rademacher random variables. Let $\mu_1, \dots, \mu_r$ be i.i.d. $N(0,1)$ random variables. Let $x_1, \dots, x_k, y_1, \dots, y_k$ be i.i.d. $N(0,1)$ random variables. Then, set
$$\begin{aligned} A_1 &= V^\intercal \operatorname{Diag}(\sigma_1, \dots, \sigma_r, F_2, \dots, F_2)V \\ A_2 &= V^\intercal \operatorname{Diag}(\sigma_1\mu_1, \dots, \sigma_r\mu_r, T_1, \dots, T_k)V. \end{aligned}$$
Here, $T_i \in \mathbb{S}^2$ is the random matrix $\left(\begin{smallmatrix} x_i & y_i \\ y_i & -x_i \end{smallmatrix}\right)$.
3. Let the entries of b_1, b_2 and L be i.i.d. $N(0,1)$ random variables.
4. If $\{x :\, Lx \le 1\}$ is unbounded, then reject this instance and resample.

Note that Corollary 2 implies that $\{A_1,\, A_2\}$ is almost surely 1-RSDC (whence also 2-RSDC) in this random model.

Algorithm 2 Construction for 2-RSDC pair

Given $A_1,\, A_2 \in \mathbb{S}^n$ such that $A_1^{-1}A_2$ has simple eigenvalues

1. Let $P \in \mathbb{R}^{n\times n}$ denote the invertible matrix guaranteed by [26]; this can be computed using an eigenvalue decomposition of $A_1^{-1}A_2$. Then $P^\intercal A_1 P = \operatorname{Diag}(\sigma_1,\dots,\sigma_r,F_2,\dots,F_2)$ and $P^\intercal A_2 P = \operatorname{Diag}(\sigma_1\mu_1,\dots,\sigma_r\mu_r,T_1,\dots,T_k)$. Here, $\sigma_1,\dots,\sigma_r \in \{\pm 1\}$, $\mu_1,\dots,\mu_r \in \mathbb{R}$ and for $i \in [k]$, the matrix T_i has the form
$$T_i = \begin{pmatrix} \operatorname{Im}(\lambda_i) & \operatorname{Re}(\lambda_i) \\ \operatorname{Re}(\lambda_i) & -\operatorname{Im}(\lambda_i) \end{pmatrix}$$
for some $\lambda_i \in \mathbb{C}\setminus\mathbb{R}$.
2. Choose an arbitrary set of $k+1$ distinct points $\xi_1,\dots,\xi_{k+1} \in \mathbb{R}$.
3. Solve for $z \in \mathbb{C}^{k+1}$ in the linear system (16). (Take a square root for each z_i^2, $i = 1,\dots,k$.)
4. Let $a = \operatorname{Re}(z)$ and $b = \operatorname{Im}(z)$ and define $\gamma \in \mathbb{R}^{2\times(r+2k)}$ as
$$\gamma = \begin{pmatrix} 0_{1\times k} & b_1 & a_1 & \cdots & b_k & a_k \\ 0_{1\times k} & a_1 & -b_1 & & a_k & -b_k \end{pmatrix}^\intercal.$$
5. Let $Q = P^{-1} \oplus I_2$ and return
$$\tilde{A}_1 = Q^\intercal \left(\begin{array}{c|cc} P^\intercal A_1 P & & \\ \hline & & 1 \\ & 1 & \end{array}\right) Q, \qquad \tilde{A}_2 = Q^\intercal \left(\begin{array}{c|cc} P^\intercal A_2 P & \multicolumn{2}{c}{\gamma} \\ \hline \gamma^\intercal & b_{k+1} & a_{k+1} \\ & a_{k+1} & -b_{k+1} \end{array}\right) Q.$$

Solution methods. We implemented the following five methods for solving instances of (9) in Gurobi 9.03 [6] through its Matlab interface:

- `oriQCQP` solves (9) directly using Gurobi's built-in nonconvex quadratic optimization solver.
- `SDCQCQP` is a solution method which can only be applied when A_1 and A_2 are both already SDC. In this case (letting P denote the corresponding invertible matrix), `SDCQCQP` reformulates (9) as
$$\min_{x\in\mathbb{R}^n} \left\{ x^\intercal(P^\intercal A_1 P)x + 2(P^\intercal b_1)^\intercal x : \begin{array}{l} x^\intercal(P^\intercal A_2 P)x + 2(P^\intercal b_2)^\intercal x \le 1 \\ Lx \le 1, \end{array} \right\} \tag{10}$$
and solves this reformulation using Gurobi's built-in nonconvex quadratic optimization solver.
- `1-RSDCQCQP` applies the construction in the proof of Theorem 2 (consolidated as Algorithm 1) to construct an SDC pair $\left\{\tilde{A}_1,\, \tilde{A}_2\right\} \in \mathbb{S}^{n+1}$ whose top-left $n\times n$ principal submatrices are A_1 and A_2, respectively; see also Appendix C.2. Let $P \in \mathbb{R}^{(n+1)\times(n+1)}$ denote the invertible matrix furnished by the SDC property of $\left\{\tilde{A}_1,\, \tilde{A}_2\right\}$. Also, set $\tilde{b}_i = (b_i^\intercal, 0)^\intercal$ and $\tilde{L} = (L, 0_{m,1})$. Then, `1-RSDCQCQP` reformulates (9) as
$$\inf_{w\in\mathbb{R}^{n+1}} \left\{ w^\intercal(P^\intercal \tilde{A}_1 P)w + 2(P^\intercal \tilde{b}_1)^\intercal w : \begin{array}{l} w^\intercal(P^\intercal \tilde{A}_2 P)w + 2(P^\intercal \tilde{b}_2)^\intercal w \le 1 \\ (\tilde{L}P)w \le 1 \\ (Pw)_{n+1} = 0 \end{array} \right\} \tag{11}$$
and solves this reformulation using Gurobi's built-in nonconvex quadratic optimization solver.

- `2-RSDCQCQP` applies the construction in Appendix C.3 (consolidated as Algorithm 2) to construct an SDC pair $\left\{\tilde{A}_1,\, \tilde{A}_2\right\} \in \mathbb{S}^{n+2}$ whose top-left $n \times n$ principal submatrices are A_1 and A_2, respectively. Let $P \in \mathbb{R}^{(n+2)\times(n+2)}$ denote the invertible matrix furnished by the SDC property of $\left\{\tilde{A}_1,\, \tilde{A}_2\right\}$. Also, set $\tilde{b}_i = (b_i^\intercal, 0, 0)^\intercal$ and $\tilde{L} = (L, 0_{m,2})$. Then, `2-RSDCQCQP` reformulates (9) as

$$\inf_{w\in\mathbb{R}^{n+1}} \left\{ w^\intercal(P^\intercal \tilde{A}_1 P)w + 2(P^\intercal \tilde{b}_1)^\intercal w : \begin{array}{l} w^\intercal(P^\intercal \tilde{A}_2 P)w + 2(P^\intercal \tilde{b}_2)^\intercal w \leq 1 \\ (\tilde{L}P)w \leq 1 \\ (Pw)_{n+1} = (Pw)_{n+2} = 0 \end{array} \right\} \tag{12}$$

  and solves this reformulation using Gurobi's built-in nonconvex quadratic optimization solver.

- `eigQCQP` first performs an eigenvalue decomposition on A_1 to write $D_1 = P_1^\intercal A_1 P_1$, where D_1 is a diagonal matrix. Then, it performs a second eigenvalue decomposition to write $D_2 = P_2^\intercal (P_1^\intercal A_2 P_1) P_2$, where D_2 is a diagonal matrix. Finally, `eigQCQP` reformulates (9) as

$$\inf_{y,z\in\mathbb{R}^{n}} \left\{ y^\intercal D_1 y + 2(P_1^\intercal b_1)^\intercal y : \begin{array}{l} z^\intercal D_2 z + 2(P_1^\intercal b_2)^\intercal y + c_2 \leq 1 \\ (LP_1)y \leq 1 \\ y = P_2 z \end{array} \right\} \tag{13}$$

  and solves this reformulation using Gurobi's built-in nonconvex quadratic optimization solver.

For each method, we also compute lower and upper bounds for each decision variable using only the polytope constraints and pass the corresponding bounds to Gurobi.

Remark 5. `SDCQCQP`, `1-RSDCQCQP`, `2-RSDCQCQP`, and `eigQCQP` can be thought of as different reformulations within a parameterized family of reformulations of (9). Specifically, these four algorithms reformulate (9) (where possible) as diagonal QCQPs with n, $n+1$, $n+2$, and $2n$ variables respectively. □

Experiment setup. We tested the solution methods on random instances of size $n \in \{10, 15, 20\}$, $m = 100$, and $k \in \{0, 1, 2, 3\}$. For $k = 0$, i.e., the case where A_1 and A_2 are guaranteed to be SDC, we compared `oriQCQP`, `SDCQCQP`, and `eigQCQP` (note that in this case `SDCQCQP`, `1-RSDCQCQP`, and `2-RSDCQCQP` reformulate (9) identically). For $k \in \{1, 2, 3\}$, we compared `oriQCQP`, `1-RSDCQCQP`, `2-RSDCQCQP`, and `eigQCQP`.

Each procedure was terminated when the CPU time reached 300 seconds or when the relative gap (between the objective value of the current solution and the best lower bound) fell below the default tolerance threshold, 10^{-4}. Detailed numerical results for 5 random instances in the settings $k = 0$ and $k \in \{1, 2, 3\}$ are reported in Tables 1 and 2, respectively.

Summary and discussion. Table 1 indicates that `SDCQCQP` generally outperforms both `eigQCQP` and `oriQCQP` when A_1 and A_2 are already SDC.

Table 2 indicates that `1-RSDCQCQP`, `2-RSDCQCQP` and `eigQCQP` generally outperform `oriQCQP`. `eigQCQP` performs well across all settings of the parameters n and k that we tested, while the performance of `1-RSDCQCQP` and `2-RSDCQCQP` seem to depend on the parameters n and k:

- For $n = 10$, `1-RSDCQCQP` and `2-RSDCQCQP` both solved almost all of the instances relatively quickly (except two outliers for `1-RSDCQCQP`), but were generally outperformed by `eigQCQP`.

n	time			objective value		
	`ori`	`SDC`	`eig`	`ori`	`SDC`	`eig`
10	3.11	**0.54**	0.67	-3.953	-3.953	-3.953
10	16.37	**0.61**	0.78	-3.495	-3.495	-3.495
10	2.32	0.62	**0.55**	-2.715	-2.715	-2.715
10	3.62	**0.75**	0.89	-4.322	-4.322	-4.322
10	5.96	**0.73**	1.15	-3.552	-3.552	-3.552
15	*	12.86	**6.38**	-3.021	-3.055	-3.055
15	*	**2.94**	4.47	-3.874	-3.875	-3.875
15	*	**8.17**	17.24	-5.65	-5.665	-5.665
15	*	*	*	-5.854	**-5.946**	-5.847
15	*	241.93	**239.22**	-5.638	-5.67	-5.67
20	*	**22.53**	*	-11.24	-11.26	-11.26
20	*	2.59	**2.32**	-7.01	-7.103	-7.103
20	*	**211.26**	*	-7.662	-7.891	-7.887
20	*	**4.91**	5.09	-10.31	-10.35	-10.35
20	*	50.97	**34.76**	-6.467	-6.796	-6.796

Table 1: Comparison of different methods for solving (9) on 5 instances of size $n \in \{10, 15, 20\}$ and $m = 100$, where A_1 and A_2 are SDC. In each row, the solution method with the lowest solution time is highlighted. For instances where all three methods time out (300 seconds, denoted by "*") before reaching optimality, the solution method with the lowest objective value is highlighted.

- For $n = 15$, `1-RSDCQCQP` and `2-RSDCQCQP` outperformed `eigQCQP` for $k = 1$, were comparable with `eigQCQP` for $k = 2$, and were outperformed by `eigQCQP` for $k = 3$.
- For $n = 20$, `2-RSDCQCQP` slightly outperformed `eigQCQP` which in turn slightly outperformed `1-RSDCQCQP`.

We comment on three interesting trends in Table 2: First, for a fixed k, `1-RSDCQCQP` and `2-RSDCQCQP` seem to perform better (compared to `eigQCQP`) as n increases. We believe this can be explained by the fact that the numbers of variables in the reformulations (11), (12) and (13) are $n+1$, $n+2$ and $2n$ respectively. In particular, the relative "computational savings" we might expect from reformulations (11) and (12) over (13) should grow with n. Second, for a fixed n, `1-RSDCQCQP` and `2-RSDCQCQP` seem to perform worse (compared to `eigQCQP`) as k increases. We believe that this trend can be explained by observing that the condition numbers of the P matrices (i.e., $\|P\| \, \|P^{-1}\|$) that we construct for (11) and (12) are likely to "blow up" as k increases (see the two rightmost columns of Table 2). Specifically, we observed that the lower and upper bounds that we precomputed for the decision variables in `oriQCQP` and `eigQCQP` were relatively small intervals, while the corresponding bounds for those in `1-RSDCQCQP` were often much larger (e.g., on the order of 1000 times larger for $k = 3$). Finally, comparing the rightmost two columns of Table 2, we see that the condition numbers of the invertible matrices P that we construct are often much smaller for `2-RSDCQCQP` than for `1-RSDCQCQP`, especially as n and k get larger. We believe that this explains why `2-RSDCQCQP` generally outperforms `1-RSDCQCQP` for larger values of the parameters n and k.

Future directions. Inspired by our numerical results, we raise two important questions which we believe are out of reach of our current understanding of the ASDC and d-RSDC properties.

(n,k)	time				objective value				condition number	
	`ori`	`1-RSDC`	`2-RSDC`	`eig`	`ori`	`1-RSDC`	`2-RSDC`	`eig`	`1-RSDC`	`2-RSDC`
(10,1)	2.03	0.72	0.81	**0.55**	-4.765	-4.765	-4.765	-4.764	1.73e+01	6.34
(10,1)	4.8	0.99	1.16	**0.97**	-2.882	-2.882	-2.882	-2.882	1.27e+01	6.68
(10,1)	1.53	0.81	0.7	**0.64**	-3.374	-3.374	-3.374	-3.374	3.49e+01	6.20
(10,1)	2.72	**0.79**	0.83	0.93	-4.153	-4.153	-4.153	-4.153	5.99	2.85
(10,1)	6.2	1.76	**1.62**	2.04	-4.195	-4.195	-4.195	-4.195	2.25e+01	1.26e+01
(10,2)	5.71	2.44	1.37	**1.09**	-3.735	-3.735	-3.735	-3.735	5.67e+01	1.24e+01
(10,2)	2.44	**0.66**	0.68	0.89	-3.775	-3.774	-3.775	-3.775	3.18e+01	8.73
(10,2)	2.45	1.52	0.96	**0.64**	-3.785	-3.785	-3.785	-3.785	5.52e+02	3.27e+01
(10,2)	1.5	4.73	2.95	**0.74**	-5.274	-5.274	-5.274	-5.274	4.81e+02	5.08e+01
(10,2)	0.93	2.98	3.55	**0.70**	-6.202	-6.202	-6.202	-6.202	1.78e+02	4.44e+01
(10,3)	2.11	7.84	4.17	**0.69**	-3.143	-3.143	-3.143	-3.143	2.59e+02	4.80e+01
(10,3)	6.97	3.78	3.32	**1.41**	-2.803	-2.803	-2.803	-2.803	2.18e+02	2.95e+01
(10,3)	1.19	0.89	0.73	**0.49**	-4.166	-4.166	-4.166	-4.166	5.82e+01	1.69e+01
(10,3)	3.71	300.2	27.49	**0.84**	-4.398	-4.397	-4.398	-4.398	1.93e+03	1.77e+02
(10,3)	2.52	300.22	8.61	**0.73**	-4.590	-4.590	-4.590	-4.590	2.78e+03	6.80e+01
(15,1)	*	4.97	**3.79**	6.61	-3.097	-3.122	-3.122	-3.122	3.00	4.92
(15,1)	*	**20.15**	25.64	44.75	-3.963	-3.964	-3.964	-3.964	2.20e+01	2.28e+01
(15,1)	244.68	**18.78**	33.74	219.66	-5.073	-5.073	-5.073	-5.073	1.53e+01	5.13
(15,1)	221.48	**13.22**	16.4	*	-5.117	-5.117	-5.117	-5.117	5.03	2.58
(15,1)	*	**192.00**	222.99	*	-6.226	-6.276	-6.276	-6.216	9.57	3.07
(15,2)	*	212.37	**166.87**	*	-2.843	-2.877	-2.877	-2.877	3.97e+01	1.03e+01
(15,2)	267.23	2.49	2.42	**1.37**	-4.090	-4.090	-4.090	-4.090	1.04e+01	3.42
(15,2)	295.23	33.19	56.13	**23.38**	-6.491	-6.491	-6.491	-6.491	2.49e+01	1.20e+01
(15,2)	*	124.68	110.08	**5.94**	-4.181	-4.195	-4.195	-4.195	1.91e+02	3.39e+02
(15,2)	*	*	*	**29.87**	-7.575	-7.594	-7.595	-7.595	4.53e+02	5.52e+01
(15,3)	289.17	*	*	**1.61**	-2.891	-2.87	-2.884	-2.891	2.10e+04	3.02e+02
(15,3)	143.4	*	11.97	**1.26**	-6.878	-6.878	-6.878	-6.878	4.33e+02	5.38e+01
(15,3)	56.79	*	256.24	**6.03**	-7.110	-7.107	-7.110	-7.110	1.42e+03	9.30e+01
(15,3)	*	*	154.77	**33.27**	-5.254	-5.255	-5.264	-5.264	3.79e+02	3.07e+01
(15,3)	*	*	21.54	**2.33**	-6.837	-6.838	-6.838	-6.838	2.99e+02	2.89e+01
(20,1)	*	*	*	*	-5.273	-5.553	**-5.555**	-5.529	2.77e+01	6.39
(20,1)	*	**120.66**	*	*	-6.269	-6.403	-6.403	-6.403	2.78	2.02
(20,1)	*	*	*	*	-8.154	-8.166	**-8.188**	-8.125	1.32e+01	1.28e+01
(20,1)	*	40.3	74.39	**39.95**	-5.498	-5.708	-5.708	-5.708	2.72	2.99
(20,1)	*	*	*	*	-6.242	-6.279	**-6.295**	-6.289	1.25e+01	6.02
(20,2)	*	*	*	**7.89**	-9.499	-9.633	-9.642	-9.643	5.39e+02	5.29e+01
(20,2)	*	*	*	**234.35**	-4.733	-5.049	-5.053	-5.054	1.13e+02	2.31e+01
(20,2)	*	*	*	*	-9.734	-9.933	-9.946	**-9.960**	6.20e+02	8.78e+01
(20,2)	*	26.34	89.63	**3.81**	-8.548	-8.559	-8.559	-8.559	2.70e+01	5.07e+01
(20,2)	*	*	*	*	-8.384	-8.550	**-8.558**	-8.263	9.12e+02	2.41e+01
(20,3)	*	*	*	*	-10.02	-10.09	**-10.12**	**-10.12**	4.72e+05	7.41e+02
(20,3)	*	*	*	*	-5.546	-5.644	**-5.678**	-5.667	1.22e+04	1.62e+02
(20,3)	*	*	*	*	-6.052	-6.286	**-6.296**	**-6.296**	2.62e+02	4.06e+01
(20,3)	*	*	*	**7.88**	-6.098	-6.163	-6.177	-6.188	1.29e+03	5.44e+01
(20,3)	*	*	*	*	-6.260	-6.347	**-6.419**	-6.383	8.01e+02	3.86e+01

Table 2: Comparison of different methods for solving (9) on 5 instances with $n \in \{10, 15, 20\}$, $k \in \{1, 2, 3\}$ and $m = 100$. In each row, the solution method with the lowest solution time is highlighted. For instances where all three methods time out (300 seconds, denoted by "*") before reaching optimality, the solution method with the lowest objective value is highlighted.

1. The 1-RSDC construction used in the proof of Theorem 2 and implemented in `1-RSDCQCQP` has been observed to lead to ill-conditioned P matrices even for only moderately large k. Are there other 1-RSDC constructions with better-conditioned P matrices? What if we are allowed to first perturb A_1 and A_2 by small *constant* amounts?
2. We can think of (13) as the n-RSDC version of (11) and (12). Is there an interesting parameterized construction of the d-RSDC property for $1 \leq d \leq n$? If so, how do the condition numbers of the corresponding P matrices trade off with the extra dimensions d?

Acknowledgments

The authors would like to thank Kevin Pratt for offering an idea which helped to simplify the proof of Lemma 5. The second author is supported in part by NSFC 11801087.

References

[1] W. Ai, Y. Huang, and S. Zhang. New results on Hermitian matrix rank-one decomposition. *Math. Program.*, 128:253–283, 2011.

[2] A. Ben-Tal and D. den Hertog. Hidden conic quadratic representation of some nonconvex quadratic optimization problems. *Math. Program.*, 143:1–29, 2014.

[3] A. Ben-Tal and M. Teboulle. Hidden convexity in some nonconvex quadratically constrained quadratic programming. *Math. Program.*, 72:51–63, 1996.

[4] S. Burer and Y. Ye. Exact semidefinite formulations for a class of (random and non-random) nonconvex quadratic programs. *Math. Program.*, pages 1–17, 2019.

[5] M. D. Bustamante, P. Mellon, and M. V. Velasco. Solving the problem of simultaneous diagonalization of complex symmetric matrices via congruence. *SIAM Journal on Matrix Analysis and Applications*, 41(4):1616–1629, 2020.

[6] LLC Gurobi Optimization. Gurobi optimizer reference manual, 9.0, 2020.

[7] J. Hiriart-Urruty. Potpourri of conjectures and open questions in nonlinear analysis and optimization. *SIAM review*, 49(2):255–273, 2007.

[8] N. Ho-Nguyen and F. Kılınç-Karzan. A second-order cone based approach for solving the trust-region subproblem and its variants. *SIAM J. Optim.*, 27(3):1485–1512, 2017.

[9] R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press, 2012.

[10] Y. Hsia and R. Sheu. Trust region subproblem with a fixed number of additional linear inequality constraints has polynomial complexity. *arXiv preprint*, (arXiv:1312.1398), 2013.

[11] K. Huang and N. D. Sidiropoulos. Consensus-ADMM for general quadratically constrained quadratic programming. *IEEE Transactions on Signal Processing*, 64(20):5297–5310, 2016.

[12] Y. Huang and S. Zhang. Complex matrix decomposition and quadratic programming. *Math. Oper. Res.*, 32(3):758–768, 2007.

[13] J. Jeyakumar and G. Li. Trust-region problems with linear inequality constraints: exact SDP relaxation, global optimality and robust optimization. *Math. Program.*, 147:171–206, 2014.

[14] R. Jiang and D. Li. Simultaneous diagonalization of matrices and its applications in quadratically constrained quadratic programming. *SIAM J. Optim.*, 26(3):1649–1668, 2016.

[15] L. Kronecker. *Collected works.* American Mathematical Society, 1968.

[16] P. Lancaster and L. Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Review*, 47(3):407–443, 2005.

[17] T. H. Le and T. N. Nguyen. Simultaneous diagonalization via congruence of hermitian matrices: some equivalent conditions and a numerical solution. *arXiv preprint*, (arXiv:2007.14034), 2020.

[18] M. Locatelli. Exactness conditions for an SDP relaxation of the extended trust region problem. *Oper. Res. Lett.*, 10(6):1141–1151, 2016.

[19] H. Luo, Y. Chen, X. Zhang, and D. Li. Effective algorithms for optimal portfolio deleveraging problem with cross impact. *arXiv preprint*, (arXiv:2012.07368), 2020.

[20] T. S. Motzkin and O. Taussky. Pairs of matrices with property L. II. *Trans. Amer. Math. Soc.*, 80(2):387–401, 1955.

[21] T. Nguyen, V. Nguyen, T. Le, and R. Sheu. On simultaneous diagonalization via congruence of real symmetric matrices. *arXiv preprint*, (arXiv:2004.06360), 2020.

[22] K. O'meara and C. Vinsonhaler. On approximately simultaneously diagonalizable matrices. *Linear Algebra Appl.*, 412:39–74, 2006.

[23] B. T. Polyak. Convexity of quadratic transformations and its use in control and optimization. *J. Optim. Theory Appl.*, 99(3):553–583, 1998.

[24] D. A. Suprunenko and R. I. Tyshkevich. *Commutative matrices*. Academic Press, 1968.

[25] H. W. Turnbull and A. C. Aitken. *An introduction to the theory of canonical matrices*. Dover, 1961.

[26] F. Uhlig. A canonical form for a pair of real symmetric matrices that generate a nonsingular pencil. *Linear Algebra Appl.*, 14(3):189–209, 1976.

[27] F. Uhlig. A recurring theorem about pairs of quadratic forms and extensions: A survey. *Linear Algebra Appl.*, 25:219–237, 1979.

[28] R. Vollgraf and Klaus K. Obermayer. Quadratic optimization for simultaneous matrix diagonalization. *IEEE Trans. Signal Process.*, 54(9):3270–3278, 2006.

[29] A. L. Wang and F. Kılınç-Karzan. The generalized trust region subproblem: solution complexity and convex hull results. *Math. Program.*, 2020. doi: 10.1007/s10107-020-01560-8. Forthcoming.

[30] A. L. Wang and F. Kılınç-Karzan. A geometric view of SDP exactness in QCQPs and its applications. *arXiv preprint*, (arXiv:2011.07155), 2020.

[31] A. L. Wang and F. Kılınç-Karzan. On the tightness of SDP relaxations of QCQPs. *Math. Program.*, 2021. doi: 10.1007/s10107-020-01589-9. Forthcoming.

[32] K. Weierstrass. Zur Theorie der quadratischen und bilinearen Formen. *Monatsber. Akad. Wiss., Berlin*, pages 310–338, 1868.

[33] J. Zhou and Z. Xu. A simultaneous diagonalization based SOCP relaxation for convex quadratic programs with linear complementarity constraints. *Optim. Lett.*, 13(7):1615–1630, 2019.

[34] J. Zhou, S. Chen, S. Yu, and Y. Tian. A simultaneous diagonalization-based quadratic convex reformulation for nonconvex quadratically constrained quadratic program. *Optimization*, pages 1–17, 2020.

A Proof of Propositions 1 and 2

Proposition 1. *Let $\mathcal{A} \subseteq \mathbb{H}^n$ and suppose $S \in \operatorname{span}(\mathcal{A})$ is nonsingular. Then, $\mathcal{A}$ is SDC if and only if $S^{-1}\mathcal{A}$ is a commuting set of diagonalizable matrices with real eigenvalues.*

Proof. ($\Rightarrow$) Let $P \in \mathbb{C}^{n\times n}$ furnished by SDC. For $A \in \mathcal{A}$, note that

$$P^{-1}S^{-1}AP = (P^*SP)^{-1}(P^*AP).$$

Then, as P^*SP and P^*AP are both diagonal matrices with real entries, we deduce that $S^{-1}A$ is diagonalizable with real eigenvalues. The fact that $S^{-1}\mathcal{A}$ is a set of commuting matrices follows similarly.

($\Leftarrow$) Recall that a commuting set of diagonalizable matrices can be simultaneously diagonalized via a similarity transformation, i.e., there exists an invertible $P \in \mathbb{C}^{n\times n}$ such that $P^{-1}S^{-1}AP$ is diagonal for each $A \in \mathcal{A}$ [9]. The diagonal entries of $P^{-1}S^{-1}AP$ are furthermore real by the assumption that $S^{-1}A$ has a real spectrum. For each $A \in \mathcal{A}$, define

$$\bar{A} \coloneqq P^*AP, \qquad D_A \coloneqq P^{-1}S^{-1}AP.$$

Next, note that the identity $P^{-1}S^{-1}AP = (P^*SP)^{-1}(P^*AP)$ can be expressed as $D_A = \bar{S}^{-1}\bar{A}$. Or, equivalently, $\bar{S}D_A = \bar{A}$ for all $A \in \mathcal{A}$. For $i, j \in [n]$, we have the identity

$$\bar{S}_{i,j}(D_A)_{j,j} = \bar{A}_{i,j} = \left(\bar{A}_{j,i}\right)^* = \left(\bar{S}_{j,i}(D_A)_{i,i}\right)^* = \bar{S}_{i,j}(D_A)_{i,i}.$$

Here, we have used that $\bar{S}$ and $\bar{A}$ are Hermitian and D_A is real diagonal. In particular, if there exists some $A \in \mathcal{A}$ such that $(D_A)_{i,i} \neq (D_A)_{j,j}$, then $\bar{S}_{i,j} = \bar{A}_{i,j} = 0$. Furthermore, by the relation $\bar{S}D_B = \bar{B}$, we also have that $\bar{B}_{i,j} = 0$ for all other $B \in \mathcal{A}$.

We conclude that by permuting the columns of P if necessary (so that $[n]$ is grouped according to the equivalence relation: $i \sim j$ if and only if $(D_A)_{i,i} = (D_A)_{j,j}$ for all $A \in \mathcal{A}$), we can write $\bar{S}$ as a block diagonal matrix $\bar{S} = \operatorname{Diag}(S^{(1)}, \dots, S^{(k)})$. Furthermore, for every $A \in \mathcal{A}$, there exists $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that $\bar{A} = \operatorname{Diag}(\lambda_1 S^{(1)}, \dots, \lambda_k S^{(k)})$. It remains to note that each block $S^{(i)}$ can be diagonalized separately. ∎

Proposition 2. *Let $\mathcal{A} \subseteq \mathbb{H}^n$ and suppose $S \in \operatorname{span}(\mathcal{A})$ is a max-rank element of $\operatorname{span}(\mathcal{A})$. Then, $\mathcal{A}$ is SDC if and only if $\operatorname{range}(A) \subseteq \operatorname{range}(S)$ for every $A \in \mathcal{A}$ and $\left\{A|_{\operatorname{range}(S)} : A \in \mathcal{A}\right\}$ is SDC.*

Proof. It suffices to show that if $\mathcal{A}$ is SDC then $\operatorname{range}(A) \subseteq \operatorname{range}(S)$ for every $A \in \mathcal{A}$ as then applying Lemma 2 completes the proof.

Let $r = \operatorname{rank}(S)$. Let $P \in \mathbb{C}^{n\times n}$ furnished by SDC. Note that by permuting the columns of P if necessary, we may assume that P^*SP is a diagonal matrix with support contained in its first r-many diagonal entries. As S is a max-rank element of $\operatorname{span}(\mathcal{A})$, we similarly have that for every $A \in \mathcal{A}$, the matrix P^*AP is a diagonal matrix with support contained in its first r-many diagonal entries. For $A \in \mathcal{A}$, write $P^*AP = \operatorname{Diag}(\bar{A}, 0_{(n-r)\times(n-r)})$ where $\bar{A}$ is a diagonal $r \times r$ matrix. Then,

$$\operatorname{range}(A) = \operatorname{range}(P^{-*}P^*APP^{-1}) \subseteq \operatorname{span}\{q_1, \dots, q_r\}.$$

Here, $q_i \in \mathbb{C}^n$ is the ith column of P^{-*}. On the other hand, as $\bar{S}$ has full rank, $\operatorname{range}(S) = \operatorname{span}\{q_1, \dots, q_r\}$. ∎

B Facts about matrices with upper triangular Toeplitz blocks

Lemma 8. *Let $(n_1, \dots, n_k)$ with $\sum_i n_i = n$. Suppose $T \in \mathbb{T}$. Then, the characteristic polynomial of T depends only on the entries $\left\{t_{i,j}^{(1)} : n_i = n_j\right\}$.*

Proof. In this proof, we will use $a, b \in [n]$ to index entries in T (specifically, $T_{a,b} \in \mathbb{C}$ is a number, not a matrix block). For each $a \in [n]$, let $i_a \in [k]$ denote the block containing a, and let $\ell_a \in [n_k]$ denote the position of a within block i_a. By the assumption that $T \in \mathbb{T}$, we have

$$T_{a,b} \neq 0 \implies \min\{n_{i_a}, n_{i_b}\} - n_{i_a} + (\ell_a - \ell_b) \geq 0.$$

Now, for each $a \in [n]$, assign the weight $w_a := \ell_a - \frac{n_{i_a}}{2}$. Note that by construction, if $T_{a,b} \neq 0$, then

$$w_a - w_b = \frac{n_{i_b}}{2} - \frac{n_{i_a}}{2} + (\ell_a - \ell_b) \geq 0.$$

Furthermore, note that if $T_{a,b} \neq 0$ and $w_a - w_b = 0$, then $n_{i_a} = n_{i_b}$ and $\ell_a = \ell_b$.

Next, consider a permutation $\sigma \in S_n$ such that $\prod_{a=1}^n T_{a,\sigma(a)} \neq 0$. Note that

$$\sum_{a=1}^n w_a - w_{\sigma(a)} = \sum_{a=1}^n w_a - \sum_{a=1}^n w_{\sigma(a)} = 0.$$

Then, by the above paragraph, we conclude that σ satisfies $n_{i_a} = n_{i_{\sigma(a)}}$ and $\ell_a = \ell_\sigma(a)$ for all $a \in [n]$.

Returning to the previous notation, the characteristic polynomial of T depends only on the entries $\left\{t_{i,j}^{(1)} : n_i = n_j\right\}$. ■

Lemma 5. *Let $(n_1, \dots, n_k)$ such that $\sum_i n_i = n$. Then, for any $T \in \mathbb{T}$, the matrices T and $\Pi(T)$ have the same eigenvalues.*

Proof. Without loss of generality, suppose $n_1 \leq \cdots \leq n_k$ and let $T \in \mathbb{T}$. By Lemma 8, T has the same eigenvalues as the matrix $\hat{T} \in \mathbb{T}$ with entries

$$\hat{T}_{i,j}^{(\ell)} = \begin{cases} T_{i,j}^{(\ell)} & \text{if } n_i = n_j,\ \ell = 1, \\ 0 & \text{else.} \end{cases}$$

Now, suppose that there are m distinct block sizes $s_1, \dots, s_m$. Partitioning both $\Pi(T)$ and $\hat{T}$according to $s_1, \dots, s_m$, we have that

$$\Pi(T) = \mathrm{Diag}(\tilde{T}_1, \dots, \tilde{T}_m) \quad \text{and} \quad \bar{T} = \mathrm{Diag}(\tilde{T}_1 \otimes I_{s_1}, \dots, \tilde{T}_m \otimes I_{s_m}).$$

We conclude that $\Pi(T)$ and $\bar{T}$ have the same eigenvalues. ■

C Details for the real symmetric case

Let $\mathbb{S}^n$ denote the set of $n \times n$ real symmetric matrices. For a vector $v \in \mathbb{R}^n$ and a matrix $A \in \mathbb{R}^{n \times n}$, let $v^\intercal$ and $A^\intercal$ denote the transpose of v and A respectively.

C.1 Definitions and theorem statements

Almost all of our results extend verbatim to the real symmetric setting. For brevity, we only state our more interesting definitions and results as adapted to this setting.

Definition 9. A set of real symmetric matrices $\mathcal{A} \subseteq \mathbb{S}^n$ is *simultaneously diagonalizable via congruence* (SDC) if there exists an invertible $P \in \mathbb{R}^{n\times n}$ such that $P^\intercal AP$ is diagonal for all $A \in \mathcal{A}$. $\square$

Definition 10. A set of real symmetric matrices $\mathcal{A} \subseteq \mathbb{S}^n$ is *almost simultaneously diagonalizable via congruence* (ASDC) if there exists a mapping $f : \mathcal{A} \times \mathbb{N} \to \mathbb{S}^n$ such that

- for all $A \in \mathcal{A}$, the limit $\lim_{j\to\infty} f(A, j)$ exists and is equal to A, and
- for all $j \in \mathbb{N}$, the set $\{f(A, j) : A \in \mathcal{A}\}$ is SDC. $\square$

Definition 11. A set of real symmetric matrices $\mathcal{A} \subseteq \mathbb{S}^n$ is *nonsingular* if there exists a nonsingular $A \in \operatorname{span}(\mathcal{A})$. Else, it is *singular*. $\square$

Definition 12. Given a set of real symmetric matrices $\mathcal{A} \subseteq \mathbb{S}^n$, we will say that $S \in \mathcal{A}$ is a *max-rank element* of $\operatorname{span}(\mathcal{A})$ if $\operatorname{rank}(S) = \max_{A\in\mathcal{A}} \operatorname{rank}(A)$. $\square$

Theorem 7. *Let $A, B \in \mathbb{S}^n$ and suppose A is invertible. Then, $\{A, B\}$ is ASDC if and only if $A^{-1}B$ has real eigenvalues.*

Theorem 8. *Let $\{A, B\} \subseteq \mathbb{S}^n$. If $\{A, B\}$ is singular, then it is ASDC.*

Theorem 9. *Let $\{A, B, C\} \subseteq \mathbb{S}^n$ and suppose A is invertible. Then, $\{A, B, C\}$ is ASDC if and only if $\{A^{-1}B, A^{-1}C\}$ are a pair of commuting matrices with real eigenvalues.*

Theorem 10. *Let $\{A = I_n, B, C\} \subseteq \mathbb{S}^n$. Then, if $d < \operatorname{rank}([B, C])/2$, the set*

$$\left\{\begin{pmatrix} A & \\ & 0_d \end{pmatrix}, \begin{pmatrix} B & \\ & 0_d \end{pmatrix}, \begin{pmatrix} C & \\ & 0_d \end{pmatrix}\right\}$$

is not ASDC.

Theorem 11. *There exists a set $\mathcal{A} = \{A_1, \dots, A_7\} \subseteq \mathbb{S}^6$ such that A_1 is invertible, $A_1^{-1}\mathcal{A}$ is a set of commuting matrices with real eigenvalues, and $\mathcal{A}$ is not ASDC.*

Definition 13. Let $\mathcal{A} \subseteq \mathbb{S}^n$ and $d \in \mathbb{N}$. We will say that $\mathcal{A}$ is *d-restricted SDC* (d-RSDC) if there exists a mapping $f : \mathcal{A} \to \mathbb{S}^{n+d}$ such that

- for all $A \in \mathcal{A}$, the top-left $n \times n$ principal submatrix of $f(A)$ is A, and
- $f(\mathcal{A})$ is SDC. $\square$

Theorem 12. *Let $A, B \in \mathbb{S}^n$. Then for every $\epsilon > 0$, there exist $\tilde{A}, \tilde{B} \in \mathbb{S}^n$ such that $\left\|A - \tilde{A}\right\|, \left\|B - \tilde{B}\right\| \leq \epsilon$ and $\left\{\tilde{A}, \tilde{B}\right\}$ is 1-RSDC. Furthermore, if A is invertible and $A^{-1}B$ has simple eigenvalues, then $\{A, B\}$ is itself 1-RSDC.*

Corollary 5. *Let $n \in \mathbb{N}$ and let $(A, B) \in \mathbb{S}^n \times \mathbb{S}^n$ be a pair of matrices jointly sampled according to an absolutely continuous probability measure on $\mathbb{S}^n \times \mathbb{S}^n$. Then, $\{A, B\}$ is 1-RSDC almost surely.*

C.2 Necessary modifications

Next, we discuss technical changes that need to be made to adapt our proofs from the Hermitian setting to the real symmetric setting. For brevity, we only list changes beyond the trivial changes, e.g., replacing $\mathbb{H}^n$ by $\mathbb{S}^n$, $\mathbb{C}^{n\times n}$ by $\mathbb{R}^{n\times n}$, and * by $^\intercal$.

- In the real symmetric version of Proposition 3, the m_2-many blocks corresponding to non-real eigenvalues (previously (1)) will have the form

$$S_i = F_{2n_i}, \qquad T_i = F_{n_i} \otimes \begin{pmatrix} \operatorname{Im}(\lambda_i) & \operatorname{Re}(\lambda_i) \\ \operatorname{Re}(\lambda_i) & -\operatorname{Im}(\lambda_i) \end{pmatrix} + G_{n_i} \otimes F_2$$

  where $n_i \in \mathbb{N}$ and $\lambda_i \in \mathbb{C} \setminus \mathbb{R}$. See [16, Theorem 9.2] for further details.

- In the proof of Lemma 3, we will set the perturbed blocks $\tilde{T}_i$ corresponding to non-real eigenvalues (previously (2)) to be

$$\tilde{T}_i = T_i + \eta F_{2n_i} + \epsilon H_{n_i} \otimes F_2,\ \forall i \in [r+1, m].$$

  Then, note that for all $i \in [r+1, m]$, the block

$$S_i^{-1}\tilde{T}_i = I_{n_i} \otimes \begin{pmatrix} \operatorname{Re}(\lambda_i) & -\operatorname{Im}(\lambda_i) \\ \operatorname{Im}(\lambda_i) & \operatorname{Re}(\lambda_i) \end{pmatrix} + F_{n_i}G_{n_i} \otimes I_2 + \eta_i I_{2n_i} + \epsilon F_{n_i} H_{n_i} \otimes I_2$$

  is similar (via $P = I_{n_i} \otimes \left(\begin{smallmatrix} i & -i \\ 1 & 1 \end{smallmatrix}\right)$) to

$$I_{n_i} \otimes \begin{pmatrix} \lambda_i + \eta_i & \\ & \lambda_i^* + \eta_i \end{pmatrix} + (F_{n_i}G_{n_i} + \epsilon F_{n_i} H_{n_i}) \otimes I_2.$$

  This is, up to a permutation of rows and columns, a direct sum of Toeplitz tridiagonal matrices. The remainder of the proof is unchanged.

- In the proof of Theorem 2, we will work in the basis furnished by the real symmetric version of Proposition 3 for $\mathbb{R}^{2k}$. That is, we may assume that $\bar{A}$ and $\bar{B}$ (previously (4)) have the form

$$\bar{A} = \left(\begin{array}{c|c|c} \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} & & \\ \hline & \ddots & \\ \hline & & \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} \end{array}\right), \qquad \bar{B} = \left(\begin{array}{c|c|c} \begin{smallmatrix} \operatorname{Im}(\lambda_1) & \operatorname{Re}(\lambda_1) \\ \operatorname{Re}(\lambda_1) & -\operatorname{Im}(\lambda_1) \end{smallmatrix} & & \\ \hline & \ddots & \\ \hline & & \begin{smallmatrix} \operatorname{Im}(\lambda_k) & \operatorname{Re}(\lambda_k) \\ \operatorname{Re}(\lambda_k) & -\operatorname{Im}(\lambda_k) \end{smallmatrix} \end{array}\right).$$

  We will set $\tilde{A}_\epsilon$ as in the Hermitian case for both Cases 1 and 2. We will set $\tilde{B}_\epsilon$ to be

$$\tilde{B}_\epsilon = \left(\begin{array}{c|c|c|c} \begin{smallmatrix} \operatorname{Im}(\lambda_1) & \operatorname{Re}(\lambda_1) \\ \operatorname{Re}(\lambda_1) & -\operatorname{Im}(\lambda_1) \end{smallmatrix} & & & \begin{smallmatrix} \sqrt{\epsilon}\alpha_1 \\ \sqrt{\epsilon}\beta_1 \end{smallmatrix} \\ \hline & \ddots & & \vdots \\ \hline & & \begin{smallmatrix} \operatorname{Im}(\lambda_k) & \operatorname{Re}(\lambda_k) \\ \operatorname{Re}(\lambda_k) & -\operatorname{Im}(\lambda_k) \end{smallmatrix} & \begin{smallmatrix} \sqrt{\epsilon}\alpha_k \\ \sqrt{\epsilon}\beta_k \end{smallmatrix} \\ \hline \sqrt{\epsilon}\alpha_1\ \sqrt{\epsilon}\beta_k & \cdots & \sqrt{\epsilon}\alpha_k\ \sqrt{\epsilon}\beta_k & \epsilon z \end{array}\right)$$

and

$$
\tilde{B}_\epsilon = \left(\begin{array}{cc|c|cc|c|c|c}
\mathrm{Im}(\lambda_1) & \mathrm{Re}(\lambda_1) & & & & & \sqrt{\epsilon}\alpha_1 & \\
\mathrm{Re}(\lambda_1) & -\mathrm{Im}(\lambda_1) & & & & & \sqrt{\epsilon}\beta_1 & \\ \hline
& & \ddots & & & & \vdots & \\ \hline
& & & \mathrm{Im}(\lambda_k) & \mathrm{Re}(\lambda_k) & & \sqrt{\epsilon}\alpha_k & \\
& & & \mathrm{Re}(\lambda_k) & -\mathrm{Im}(\lambda_k) & & \sqrt{\epsilon}\beta_k & \\ \hline
& & & & & & & G_{n_m} \\ \hline
\sqrt{\epsilon}\alpha_1 & \sqrt{\epsilon}\beta_1 & \cdots & \sqrt{\epsilon}\alpha_k & \sqrt{\epsilon}\beta_k & & \epsilon z & e_1^\intercal \\ \hline
& & & & & G_{n_m} & e_1 &
\end{array}\right)
$$

for Cases 1 and 2, respectively. Here, $\alpha, \beta \in \mathbb{R}^k$, $z \in \mathbb{R}$, and $\epsilon > 0$. The eigenvalues of $\tilde{A}_\epsilon^{-1}\tilde{B}_\epsilon$ are the roots (in the variable ξ) of

$$
\begin{aligned}
&(z - \xi)\prod_{i=1}^k (\lambda_i - \xi)(\lambda_i^* - \xi) \\
&\quad + \sum_{i=1}^k \Big(\mathrm{Im}(\lambda_i)(\beta_i^2 - \alpha_i^2) - 2\alpha_i\beta_i(\mathrm{Re}(\lambda_i) - \xi)\Big)\prod_{j \neq i}(\lambda_j - \xi)(\lambda_j^* - \xi)
\end{aligned}
$$

and

$$
\begin{aligned}
\xi^{2n_m}\bigg(&(z - \xi)\prod_{i=1}^k (\lambda_i - \xi)(\lambda_i^* - \xi) \\
&\quad + \sum_{i=1}^k \Big(\mathrm{Im}(\lambda_i)(\beta_i^2 - \alpha_i^2) - 2\alpha_i\beta_i(\mathrm{Re}(\lambda_i) - \xi)\Big)\prod_{j \neq i}(\lambda_j - \xi)(\lambda_j^* - \xi)\bigg)
\end{aligned}
$$

in Cases 1 and 2, respectively. Finally, note that for any $x, y \in \mathbb{R}^k$, there exist $\alpha, \beta \in \mathbb{R}^k$ such that $x_i = \mathrm{Im}(\lambda_i)(\beta_i^2 - \alpha_i^2) - 2\,\mathrm{Re}(\lambda_i)\alpha_i\beta_i$ and $y_i = 2\alpha_i\beta_i$ for all $i \in [k]$. The remainder of the proof follows unchanged.

- In the real symmetric setting, the statement in Theorem 6 should be changed to: "There exists a set $\mathcal{A} = \{A_1, \dots, A_7\} \subseteq \mathbb{S}^6$ such that A_1 is invertible, $A_1^{-1}\mathcal{A}$ is a set of commuting matrices with real eigenvalues and $\mathcal{A}$ is not ASDC." The proof is unchanged after setting

$$
A_1 = \begin{pmatrix} &&&&&1\\ &&&&1&\\ &&&1&&\\ &&1&&&\\ &1&&&&\\ 1&&&&& \end{pmatrix},\quad
A_2 = \begin{pmatrix} 0&&&&&\\ &0&&&&\\ &&0&&&\\ &&&1&&\\ &&&&0&\\ &&&&&0 \end{pmatrix},\quad
A_3 = \begin{pmatrix} 0&&&&&\\ &0&&&&\\ &&0&&&\\ &&&&1&\\ &&&1&&\\ &&&&&0 \end{pmatrix},
$$

$$
A_4 = \begin{pmatrix} 0&&&&&\\ &0&&&&\\ &&0&&&\\ &&&&&1\\ &&&&0&\\ &&&1&& \end{pmatrix},\quad
A_5 = \begin{pmatrix} 0&&&&&\\ &0&&&&\\ &&0&&&\\ &&&0&&\\ &&&&1&\\ &&&&&0 \end{pmatrix},
$$

$$
A_6 = \begin{pmatrix} 0&&&&&\\ &0&&&&\\ &&0&&&\\ &&&0&&\\ &&&&&1\\ &&&&1& \end{pmatrix},\quad
A_7 = \begin{pmatrix} 0&&&&&\\ &0&&&&\\ &&0&&&\\ &&&0&&\\ &&&&0&\\ &&&&&1 \end{pmatrix}.
$$

C.3 A 2-RSDC construction

In this subsection, we will present an explicit construction which shows that any real symmetric[8] pair $\{A, B\} \in \mathbb{S}^n$ where A is invertible and $A^{-1}B$ has simple eigenvalues is 2-RSDC. While this is

[8] As in Section 7, we will present the results of this section only for the real symmetric setting. An analogous construction works for the Hermitian setting.

of no theoretical consequence (we have already show in Theorem 4 that such real symmetric pairs are in fact 1-RSDC), this construction has been observed to produce change-of-basis matrices which are better conditioned than those arising from our 1-RSDC construction.

The construction is very similar to our construction for the 1-RSDC property. We will assume without loss of generality, that A, B are of the form

$$A = \left(\begin{array}{c|c|c} \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} & & \\ \hline & \ddots & \\ \hline & & \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} \end{array}\right), \qquad B = \left(\begin{array}{c|c|c} \begin{smallmatrix} \mathrm{Im}(\lambda_1) & \mathrm{Re}(\lambda_1) \\ \mathrm{Re}(\lambda_1) & -\mathrm{Im}(\lambda_1) \end{smallmatrix} & & \\ \hline & \ddots & \\ \hline & & \begin{smallmatrix} \mathrm{Im}(\lambda_k) & \mathrm{Re}(\lambda_k) \\ \mathrm{Re}(\lambda_k) & -\mathrm{Im}(\lambda_k) \end{smallmatrix} \end{array}\right).$$

We will set

$$\tilde{A} = \left(\begin{array}{c|c} A & \\ \hline & \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} \end{array}\right) = \left(\begin{array}{c|c|c|c} \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} & & & \\ \hline & \ddots & & \\ \hline & & \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} & \\ \hline & & & \begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix} \end{array}\right)$$

and

$$\tilde{B} = \left(\begin{array}{c|c|c|c} \begin{smallmatrix} \mathrm{Im}(\lambda_1) & \mathrm{Re}(\lambda_1) \\ \mathrm{Re}(\lambda_1) & -\mathrm{Im}(\lambda_1) \end{smallmatrix} & & & \begin{smallmatrix} b_1 & a_1 \\ a_1 & -b_1 \end{smallmatrix} \\ \hline & \ddots & & \vdots \\ \hline & & \begin{smallmatrix} \mathrm{Im}(\lambda_k) & \mathrm{Re}(\lambda_k) \\ \mathrm{Re}(\lambda_k) & -\mathrm{Im}(\lambda_k) \end{smallmatrix} & \begin{smallmatrix} b_k & a_k \\ a_k & -b_k \end{smallmatrix} \\ \hline \begin{smallmatrix} b_1 & a_1 \\ a_1 & -b_1 \end{smallmatrix} & \cdots & \begin{smallmatrix} b_k & a_k \\ a_k & -b_k \end{smallmatrix} & \begin{smallmatrix} b_{k+1} & a_{k+1} \\ a_{k+1} & -b_{k+1} \end{smallmatrix} \end{array}\right)$$

for some $a, b \in \mathbb{R}^{k+1}$, whence

$$\tilde{A}^{-1}\tilde{B} = \left(\begin{array}{c|c|c|c} \begin{smallmatrix} \mathrm{Re}(\lambda_1) & -\mathrm{Im}(\lambda_1) \\ \mathrm{Im}(\lambda_1) & \mathrm{Re}(\lambda_1) \end{smallmatrix} & & & \begin{smallmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{smallmatrix} \\ \hline & \ddots & & \vdots \\ \hline & & \begin{smallmatrix} \mathrm{Re}(\lambda_k) & -\mathrm{Im}(\lambda_k) \\ \mathrm{Im}(\lambda_k) & \mathrm{Re}(\lambda_k) \end{smallmatrix} & \begin{smallmatrix} a_k & -b_k \\ b_k & a_k \end{smallmatrix} \\ \hline \begin{smallmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{smallmatrix} & \cdots & \begin{smallmatrix} a_k & -b_k \\ b_k & a_k \end{smallmatrix} & \begin{smallmatrix} a_{k+1} & -b_{k+1} \\ b_{k+1} & a_{k+1} \end{smallmatrix} \end{array}\right).$$

Note that $\tilde{A}^{-1}\tilde{B}$ is similar to

$$\left(\begin{array}{c|c} \begin{smallmatrix} \lambda_1 & & & a_1+b_1\mathrm{i} \\ & \ddots & & \vdots \\ & & \lambda_k & a_k+b_k\mathrm{i} \\ a_1+b_1\mathrm{i} & \ldots & a_k+b_k\mathrm{i} & a_{k+1}+b_{k+1}\mathrm{i} \end{smallmatrix} & \\ \hline & \begin{smallmatrix} \lambda_1^* & & & a_1-b_1\mathrm{i} \\ & \ddots & & \vdots \\ & & \lambda_k^* & a_k-b_k\mathrm{i} \\ a_1-b_1\mathrm{i} & \ldots & a_k-b_k\mathrm{i} & a_{k+1}-b_{k+1}\mathrm{i} \end{smallmatrix} \end{array}\right). \tag{14}$$

The top-left and bottom-right blocks of this matrix are complex conjugates so that the eigenvalues of the bottom-right block are the complex conjugates of the eigenvalues of the top-left block. Thus, it suffices to choose $a, b \in \mathbb{R}^{k+1}$ such that the top-left block has real and simple eigenvalues. For notational convenience, let $z_i = a_i + b_i\mathrm{i}$.

The eigenvalues of the top-left block are the roots (in the variable ξ) of

$$(z_{k+1} - \xi)\prod_{i=1}^{k}(\lambda_i - \xi) - \sum_{i=1}^{k} z_i^2 \prod_{j\neq i}(\lambda_j - \xi). \tag{15}$$

Define the following polynomials.

$$f_i(\xi) := \prod_{j\neq i}(\lambda_j - \xi), \quad \forall i \in [k], \quad \text{and} \quad h(\xi) := \prod_{i=1}^{k}(\lambda_i - \xi).$$

As $\{\lambda_1, \ldots, \lambda_k\}$ are distinct, we have that $\{f_1, \ldots, f_k, h\}$ are a basis for the degree-k polynomials in ξ.

Now pick $k+1$ distinct values $\xi_1, \ldots, \xi_{k+1} \in \mathbb{R}$. Note that $\{\xi_1, \ldots, \xi_{k+1}\}$ are the roots of (15) if and only if $z \in \mathbb{C}^{k+1}$ satisfies

$$\begin{pmatrix} f_1(\xi_1) & \cdots & f_k(\xi_1) & h(\xi_1) \\ \vdots & \ddots & \vdots & \vdots \\ f_1(\xi_{k+1}) & \cdots & f_k(\xi_{k+1}) & h(\xi_{k+1}) \end{pmatrix} \begin{pmatrix} z_1^2 \\ \vdots \\ z_k^2 \\ z_{k+1} \end{pmatrix} = \begin{pmatrix} \xi_1 h(\xi_1) \\ \vdots \\ \xi_{k+1} h(\xi_{k+1}) \end{pmatrix}. \tag{16}$$

Note that the matrix on the left is invertible (as $\{f_1, \ldots, f_k, h\}$ is independent and the ξ_i are distinct). We deduce that there exists $z \in \mathbb{C}^{k+1}$ such that the eigenvalues of the top-left block of (14) are real and simple. In turn, there exist $a, b \in \mathbb{R}^{k+1}$ such that $\tilde{A}^{-1}\tilde{B}$ has real eigenvalues and is diagonalizable.

D An example where the SDC property is preserved under restriction

In this section, we give an example of a setting in which the restriction of an SDC set to one of its principal submatrices results in another SDC set. This setting arises for example in QCQPs [14].

Proposition 4. *Let $A_1, \ldots, A_m \in \mathbb{H}^n$ such that* $\operatorname{span}(\{A_1, \ldots, A_m\})$ *contains a positive definite matrix. Let $b_1, \ldots, b_m \in \mathbb{C}^n$ and $c_1, \ldots, c_m \in \mathbb{R}$, and define*

$$Q_i = \begin{pmatrix} A_i & b_i \\ b_i^* & c_i \end{pmatrix} \in \mathbb{H}^{n+1}.$$

If $\{Q_1, \ldots, Q_m, e_{n+1}e_{n+1}^*\}$ is SDC, then so is $\{A_1, \ldots, A_m\}$.

Proof. Without loss of generality, let $A_1 \succ 0$. Note that for all $\lambda \in \mathbb{R}$ large enough, the matrix $S_\lambda := Q_1 + \lambda e_{n+1}e_{n+1}^* \succ 0$. By the inverse formula for a block matrix [9], we have that for all λ large enough,

$$S_\lambda^{-1} = \begin{pmatrix} A_1^{-1} + \frac{A_1^{-1}b_1b_1^*A_1^{-1}}{\lambda+(c_1-b_1^*A_1b_1)} & \frac{-A_1^{-1}b_1}{\lambda+(c_1-b_1A_1^{-1}b_1)} \\ \frac{-b_1^*A_1^{-1}}{\lambda+(c_1-b_1A_1^{-1}b_1)} & \frac{1}{\lambda+(c_1-b_1A_1^{-1}b_1)} \end{pmatrix}.$$

In particular,

$$\lim_{\lambda\to\infty} S_\lambda^{-1} = \begin{pmatrix} A_1^{-1} & \\ & 0 \end{pmatrix}.$$

On the other hand, by Lemma 1, we have that for all $i, j \in [m]$,

$$0 = \left[S_\lambda^{-1}Q_i,\, S_\lambda^{-1}Q_j\right].$$

Finally, by continuity we have that

$$0 = \lim_{\lambda \to \infty} \left[S_\lambda^{-1} Q_i,\, S_\lambda^{-1} Q_j \right] = \begin{pmatrix} \left[A_1^{-1} A_i,\, A_1^{-1} A_j \right] & \\ & 0 \end{pmatrix}.$$

We conclude that $A_1^{-1}\{A_1, \ldots, A_m\}$ commute, whence by Lemma 1 this set is SDC. ■