New Results on Hermitian Matrix Rank-One Decomposition

Wenbao Ai * Yongwei Huang † Shuzhong Zhang ‡

June 18, 2009

Abstract

In this paper, we present several new rank-one decomposition theorems for Hermitian positive semidefinite matrices, which generalize our previous results in [18, 2]. The new matrix rank-one decomposition theorems appear to have wide applications in theory as well as in practice. On the theoretical side, for example, we show how to further extend some of the classical results including a lemma due to Yuan [27], the classical results on the convexity of the joint numerical ranges [23, 4], and the so-called Finsler's lemma [9, 4]. On the practical side, we show that the new results can be applied to solve two typical problems in signal processing and communication: one for radar code optimization and the other for robust beamforming. The new matrix decomposition theorems are proven by construction in this paper, and we demonstrate that the constructive procedures can be implemented efficiently, stably, and accurately. The URL of our Matlab programs is given in this paper. We strongly believe that the new decomposition procedures, as a means to solve non-convex quadratic optimization with a few quadratic constraints, are useful for many other potential engineering applications.

Keywords: positive semidefinite Hermitian matrix, matrix rank-one decomposition, joint numerical range, quadratic optimization.

Mathematics Subject Classification: 90C33, 90C51, 90C05.

*School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, People's Republic of China. Email: wenbaoai@gmail.com. Project10701016 supported by Chinese NSFC.

†Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: ywhuang@se.cuhk.edu.hk

‡Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong. Email: zhang@se.cuhk.edu.hk. Research supported by Hong Kong RGC Earmarked Grants CUHK418505 and CUHK418406.

1 Introduction

In a series of recently papers ([25, 18, 2]) we have developed the technique of decomposing a positive semidefinite matrix into the sum of rank-one matrices with some desirable properties. This matrix rank-one decomposition approach proves to be useful in a variety of ways. Most noticeably, it can be used to derive the rank-one optimal solutions for an Semidefinite Program when the number of constraints is small; see [26]. In that way, the method has found wide applications in signal processing and communication; see [12, 13].

The first such type matrix rank-one decomposition method was introduced by Sturm and Zhang in [25] as a means to characterize the matrix cone whose quadratic form is co-positive over a given domain. This naturally connects to the S-lemma of Yakubovich, since the S-lemma is concerned with the positivity of a quadratic form over the domain defined by the level set of another quadratic form. As a matter of fact, these results can be viewed as a duality pair. The matrix rank-one decomposition procedure proposed in [25] is easy to implement, and can be considered as a constructive proof for the S-lemma. For a survey on the S-lemma, we refer to Polik and Terlaky [24]. Applications of such matrix decomposition technique can be quite versatile. For instance, Sturm and Zhang in [25] showed how one can solve the quadratic optimization problem with the constraint set being the intersection of an ellipsoid and a half-plane, via semidefinite programing (SDP) relaxation followed by a rank-one decomposition procedure. Following the approaches adopted in Sturm and Zhang [25], Huang and Zhang [18] generalized the constructive technique to the domain of Hermitian PSD matrices, proving the complex version of Yakubovich's S-lemma, and also deriving an upper bound on the lowest rank among all the optimal solutions for a standard complex SDP problem. Pang and Zhang [23] and Huang and Zhang [18] gave alternative proofs for some convexity properties of the joint numerical ranges ([10, 4]) using the matrix rank-one decomposition techniques. Up to that point, the matrix rank-one decomposition was meant to be a complete decomposition. Ai and Zhang [2] obtained a *partial* rank-one decomposition result for the real symmetric positive semidefinite matrices, and used this result to fully characterize the condition under which the strong duality holds for the (nonconvex) quadratic optimization problem over the intersection of two ellipsoids: the problem is known as the CDT subproblem in the literature needed by the trust region method for nonlinear programming.

In the current paper, we present some new rank-one decomposition results for the Hermitian positive semidefinite matrices, generalizing the matrix decomposition theorems in [18, 2]. We illustrate the potentials of the new decomposition results from both practical and theoretical aspects. On the practical side, we shall demonstrate two applications of the new decomposition theorems in signal processing, viz. the optimal radar code selection problem and the robust beamforming problem. On the theoretical side, we present a generalization of Yuan's lemma [27], and show the connections to the famous convexity results with regard to the joint numerical ranges (cf. [23, 4]) and a generalized

Finsler's lemma (cf. [9, 4]). Finally, we discuss the numerical performance of the matrix rank-one decomposition algorithms. The URL of our Matlab programs can be found at Section 6. We believe that the algorithms are useful in many other applications.

This paper is organized as follows. In Section 2, we present three new Hermitian p.s.d. matrix rank-one decomposition theorems and some initial analysis of the results. Section 3 is devoted to the proofs of these three new theorems. To showcase the potential applications of the new results, we illustrate in Section 4 applications of the decomposition theorems arising from signal processing. To put our new results in perspective, we present in Subsection 5.1 an extension of Yuan's lemma [27], and in Subsection 5.2 the equivalence between one of the new decomposition theorems and some other well-established results in the literature. All of our proofs are constructive, and they can be efficiently implemented in the form of algorithms. In order to evaluate these algorithms, in Section 6 we present the performance of the constructive procedures; we pay special attention to the numerical stability and accuracy of those solution procedures.

Notation. Throughout, we denote $\bar{a}$ to be the conjugate of a complex number a, $\mathbf{C}^n$ to be the space of n-dimension complex vectors, $\boldsymbol{i}$ to be the imaginary unit, i.e., $\boldsymbol{i}^2 = -1$. For a given vector $z \in \mathbf{C}^n$, z^{H} denotes the conjugate transpose of z. The space of $n \times n$ real symmetric and complex Hermitian matrices are denoted by $\mathcal{S}^n$ and $\mathcal{H}^n$, respectively. For a matrix $Z \in \mathcal{H}^n$, we denote Re Z and Im Z as the real and imaginary part of Z respectively. Matrix Z being Hermitian implies that Re Z is symmetric and Im Z is skew-symmetric. We denote by $\mathcal{S}^n_+$ ($\mathcal{S}^n_{++}$) and $\mathcal{H}^n_+$ ($\mathcal{H}^n_{++}$) the cones of real symmetric positive semidefinite (positive definite) and complex Hermitian positive semidefinite (positive definite) matrices respectively. The notation $Z \succeq (\succ 0)$ means that Z is positive semidefinite (positive definite). For two complex matrices Y and Z, their inner product $Y \bullet Z = \mathrm{Re}\ (\mathrm{tr}\ Y^{\mathrm{H}}Z) = \mathrm{tr}\ \left[(\mathrm{Re}\ Y)^{\mathrm{T}}(\mathrm{Re}\ Z) + (\mathrm{Im}\ Y)^{\mathrm{T}}(\mathrm{Im}\ Z)\right]$, where 'tr' denotes the trace of a matrix and $^{\mathrm{T}}$ denotes the transpose of a matrix, and $^{\mathrm{H}}$ denotes the conjugate transpose of a matrix. For a square matrix M, the notation 'rank(M)', 'Range(M)', 'Null(M)' stand for, respectively, the rank, the range space, and the null space of M.

2 The Main Results

Let us start by presenting the main results of this paper, in the form of three theorems. The first theorem is about decomposing a Hermitian positive semidefinite matrix into the sum of rank-one matrices, with a specific condition, as we see below.

Theorem 2.1. *Let $A_1, A_2, A_3 \in \mathcal{H}^n$, and $X \in \mathcal{H}^n_+$ be a nonzero Hermitian positive semidefinite matrix. If $r = \mathrm{rank}\,(X) \geq 3$, then one can find in polynomial-time a rank-one decomposition of X,*

$X = \sum_{i=1}^{r} x_i x_i^{\mathrm{H}}$, *such that*

$$\begin{cases} A_1 \bullet x_i x_i^{\mathrm{H}} = A_1 \bullet X/r, \; i = 1, ..., r; \\ A_2 \bullet x_i x_i^{\mathrm{H}} = A_2 \bullet X/r, \; i = 1, ..., r; \\ A_3 \bullet x_i x_i^{\mathrm{H}} = A_3 \bullet X/r, \; i = 1, ..., r-2. \end{cases}$$

The real-case counterpart of the above theorem is in Ai and Zhang [2]. In the real case, the result actually appears to be weaker. The next theorem mainly deals with the situation where Theorem 2.1 does not apply; in particular, when $\operatorname{rank}(X) = 2$. In Theorem 2.2, we no longer seek for a complete rank-one decomposition of X. Rather, we look for a rank-one matrix solution to a system of linear matrix equations, within a slightly expanded range space (of X).

Theorem 2.2. *Suppose $n \geq 3$. Let $A_1, A_2, A_3 \in \mathcal{H}^n$, and $X \in \mathcal{H}_+^n$ be a nonzero Hermitian positive semidefinite matrix of rank r. If $r \geq 3$, then one can find in polynomial-time a nonzero vector $y \in \operatorname{Range}(X)$ such that*

$$\begin{cases} A_1 \bullet yy^{\mathrm{H}} = A_1 \bullet X, \\ A_2 \bullet yy^{\mathrm{H}} = A_2 \bullet X, \\ A_3 \bullet yy^{\mathrm{H}} = A_3 \bullet X, \end{cases}$$

with $X - \frac{1}{r} yy^{\mathrm{H}} \succeq 0$ and $\operatorname{rank}(X - \frac{1}{r} yy^{\mathrm{H}}) \leq r - 1$. If $r = 2$, then for any $z \notin \operatorname{Range}(X)$ there exists $y \in \mathbf{C}^n$ in the linear subspace spanned by z and $\operatorname{Range}(X)$, such that

$$\begin{cases} A_1 \bullet yy^{\mathrm{H}} = A_1 \bullet X, \\ A_2 \bullet yy^{\mathrm{H}} = A_2 \bullet X, \\ A_3 \bullet yy^{\mathrm{H}} = A_3 \bullet X, \end{cases}$$

with $X + zz^{\mathrm{H}} - \frac{1}{r} yy^{\mathrm{H}} \succeq 0$ and $\operatorname{rank}(X + zz^{\mathrm{H}} - \frac{1}{r} yy^{\mathrm{H}}) \leq 2$.

Although Theorem 2.2 is best possible in general, with one more regularity condition we will be able to extend the result to satisfy four matrix equations.

Theorem 2.3. *Suppose $n \geq 3$. Let $A_1, A_2, A_3, A_4 \in \mathcal{H}^n$, and $X \in \mathcal{H}_+^n$ be a nonzero Hermitian positive semidefinite matrix of rank r. Furthermore, suppose that*

$$(A_1 \bullet Y, A_2 \bullet Y, A_3 \bullet Y, A_4 \bullet Y) \neq (0, 0, 0, 0),$$

for any nonzero matrix $Y \in \mathcal{H}_+^n$. If $r \geq 3$, then one can find in polynomial-time a nonzero vector $y \in \operatorname{Range}(X)$ such that

$$\begin{cases} A_1 \bullet yy^{\mathrm{H}} = A_1 \bullet X, \\ A_2 \bullet yy^{\mathrm{H}} = A_2 \bullet X, \\ A_3 \bullet yy^{\mathrm{H}} = A_3 \bullet X, \\ A_4 \bullet yy^{\mathrm{H}} = A_4 \bullet X. \end{cases}$$

If $r = 2$, then for any $z \notin \text{Range}(X)$ there exists $y \in \mathbf{C}^n$ in the linear subspace spanned by z and $\text{Range}(X)$, such that

$$\begin{cases} A_1 \bullet yy^{\mathrm{H}} = A_1 \bullet X, \\ A_2 \bullet yy^{\mathrm{H}} = A_2 \bullet X, \\ A_3 \bullet yy^{\mathrm{H}} = A_3 \bullet X, \\ A_4 \bullet yy^{\mathrm{H}} = A_4 \bullet X. \end{cases}$$

Before discussing the proofs, let us start by presenting the following three examples, demonstrating that the conditions in these theorems are all necessary. Examples 1 and 2 show that the assumption $n \geq 3$ is necessary for Theorems 2.1 and 2.3, and Example 3 shows that the assumption $(A_1 \bullet Y, A_2 \bullet Y, A_3 \bullet Y, A_4 \bullet Y) \neq (0, 0, 0, 0)$ for all nonzero $Y \succeq 0$ is necessary for Theorem 2.3.

Example 1. Consider

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \ A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ A_3 = \begin{bmatrix} 0 & \boldsymbol{i} \\ -\boldsymbol{i} & 0 \end{bmatrix}, \ X = I_2,$$

where I_2 represents the 2×2 identity matrix. We have $A_1 \bullet X = A_2 \bullet X = A_3 \bullet X = 0$. Notice that the system $A_1 \bullet yy^{\mathrm{H}} = A_2 \bullet yy^{\mathrm{H}} = A_3 \bullet yy^{\mathrm{H}} = 0$ is equivalent to

$$|y_1|^2 - |y_2|^2 = 0, \ \text{Re}\ (\bar{y}_1 y_2) = 0, \ \text{Im}\ (\bar{y}_1 y_2) = 0,$$

which imply that $y_1 = y_2 = 0$. Thus there can be no rank-one matrix X to the system of equations: $A_1 \bullet X = A_2 \bullet X = A_3 \bullet X = 0$, $X \succeq 0$. This shows that $n \geq 3$ is necessary for Theorems 2.1 and 2.2.

Example 2. Consider

$$A_1 = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \ A_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \ A_3 = \begin{bmatrix} 1 & \boldsymbol{i} \\ -\boldsymbol{i} & 1 \end{bmatrix}, \ A_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

It is easily verified that all the assumptions in Theorem 2.3 are satisfied, except for '$n \geq 3$'. Notice that

$$A_1 \bullet I_2 = A_2 \bullet I_2 = A_3 \bullet I_2 = A_4 \bullet I_2 = 2.$$

Assume that there is a two-dimensional complex vector $y = [y_1, y_2]^{\mathrm{T}}$ such that $A_1 \bullet yy^{\mathrm{H}} = A_2 \bullet yy^{\mathrm{H}} = A_3 \bullet yy^{\mathrm{H}} = A_4 \bullet yy^{\mathrm{H}} = 2$. These equalities amount to

$$|y_1|^2 = 1, \ |y_1|^2 + |y_2|^2 + 2\,\text{Re}\ (\bar{y}_1 y_2) = 2, \ |y_1|^2 + |y_2|^2 - 2\,\text{Im}\ (\bar{y}_1 y_2) = 2, \ |y_1|^2 + |y_2|^2 = 2,$$

which lead to

$$|y_1| = |y_2| = 1, \ \bar{y}_1 y_2 = 0,$$

a clear contradiction. Hence there can be no $y \in \mathbf{C}^2 \setminus \{0\}$ satisfying $A_i \bullet yy^{\mathrm{H}} = A_i \bullet I_2$, $i = 1, 2, 3, 4$.

Example 3. Consider

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; A_2 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; A_3 = \begin{bmatrix} 1 & -\boldsymbol{i} & 0 \\ \boldsymbol{i} & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; A_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -1/2 \end{bmatrix},$$

and

$$X = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}.$$

Notice that $A_i \bullet I_3 = 0$, $i = 1, 2, 3, 4$, and $A_i \bullet X = 1$, $i = 1, 2, 3$, and $A_4 \bullet X = 5/4$. Suppose that there is a three-dimension complex vector $y = [y_1, y_2, y_3]^{\mathrm{T}}$ satisfying $A_i \bullet yy^{\mathrm{H}} = 1$, $i = 1, 2, 3$, and $A_4 \bullet yy^{\mathrm{H}} = 5/4$. Then, these equalities would lead to

$$|y_1|^2 - |y_2|^2 = 1,\; |y_1|^2 - |y_2|^2 + 2\,\mathrm{Re}\,(\bar{y}_1 y_2) = 1,\; |y_1|^2 - |y_2|^2 + 2\,\mathrm{Im}\,(\bar{y}_1 y_2) = 1,\; |y_1|^2 - \frac{1}{2}|y_2|^2 - \frac{1}{2}|y_3|^2 = 5/4,$$

which is clearly impossible.

3 Proofs of the Main Results

In this section we shall present proofs for the three theorems that we have just discussed above. To accomplish this task, we need the following two technical results, Lemma 3.1 and Proposition 3.2, as the main tools.

Lemma 3.1. *For any positive numbers $c_{-1} > 0, c_0 > 0$, any complex numbers a_i, b_i, c_i, $i = 1, 2, 3$, and any real numbers a_4, b_4, c_4, the following system of equations*

$$\begin{aligned} \mathrm{Re}\,(a_1 \bar{x} y) + \mathrm{Re}\,(a_2 \bar{x} z) + \mathrm{Re}\,(a_3 \bar{y} z) + a_4 |z|^2 &= 0, \\ \mathrm{Re}\,(b_1 \bar{x} y) + \mathrm{Re}\,(b_2 \bar{x} z) + \mathrm{Re}\,(b_3 \bar{y} z) + b_4 |z|^2 &= 0, \\ c_{-1}|x|^2 - c_0 |y|^2 + \mathrm{Re}\,(c_1 \bar{x} y) + \mathrm{Re}\,(c_2 \bar{x} z) + \mathrm{Re}\,(c_3 \bar{y} z) + c_4 |z|^2 &= 0, \end{aligned}$$

always admits a non-zero complex-valued solution.

Our proof for the above lemma is constructive, but it is rather tedious. In order not to distract the flow of our discussion, we delegate the detailed proof to Appendix A.

The following proposition is due to Huang and Zhang [18] (Theorem 2.1), where the proof is based on a simple (polynomial-time) construction.

Proposition 3.2. *Suppose that $A_1, A_2 \in \mathcal{H}^n$ and $X \in \mathcal{H}_+^n$. Then there exists a rank-one decomposition of X, $X = x_1x_1^{\rm H} + \cdots + x_rx_r^{\rm H}$, such that*

$$A_1 \bullet x_ix_i^{\rm H} = A_1 \bullet X/r, \quad A_2 \bullet x_ix_i^{\rm H} = A_2 \bullet X/r, \quad i = 1,...,r,$$

where $r = \operatorname{rank}(X)$.

3.1 Proof of Theorem 2.1

To simply the notation, in this and subsequent proofs we shall denote $\delta_i = A_i \bullet X$, for all index i.

It follows from Proposition 3.2 that there is a rank-one decomposition $X = \sum_{i=1}^r x_ix_i^{\rm H}$ such that

$$A_1 \bullet x_ix_i^{\rm H} = \delta_1/r, \quad A_2 \bullet x_ix_i^{\rm H} = \delta_2/r, \quad i = 1, \ldots, r.$$

If $A_3 \bullet x_ix_i^{\rm H} = \delta_3/r$ for $i = 1, \ldots, r-2$, then the theorem would follow. Otherwise, let us suppose that $A_3 \bullet x_1x_1^{\rm H} - \delta_3/r > 0$ and $A_3 \bullet x_2x_2^{\rm H} - \delta_3/r < 0$. Let

$$y = \frac{\alpha_1x_1 + \alpha_2x_2 + \alpha_3x_3}{\sqrt{|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2}},$$

where $\alpha_i, i = 1, 2, 3$, are three complex-valued parameters to be specified. Then the equation (with respect to $\alpha_1, \alpha_2, \alpha_3$)

$$\begin{cases} A_1 \bullet yy^{\rm H} = \delta_1/r, \\ A_2 \bullet yy^{\rm H} = \delta_2/r, \\ A_3 \bullet yy^{\rm H} = \delta_3/r, \end{cases}$$

is equivalent to

$$\begin{cases} 0 = \operatorname{Re}(2x_1^{\rm H}A_1x_2\bar{\alpha}_1\alpha_2) + \operatorname{Re}(2x_2^{\rm H}A_1x_3\bar{\alpha}_2\alpha_3) + \operatorname{Re}(2x_3^{\rm H}A_1x_1\bar{\alpha}_3\alpha_1), \\ 0 = \operatorname{Re}(2x_1^{\rm H}A_2x_2\bar{\alpha}_1\alpha_2) + \operatorname{Re}(2x_2^{\rm H}A_2x_3\bar{\alpha}_2\alpha_3) + \operatorname{Re}(2x_3^{\rm H}A_2x_1\bar{\alpha}_3\alpha_1), \\ 0 = (x_1^{\rm H}A_3x_1 - \delta_3/r)|\alpha_1|^2 + (x_2^{\rm H}A_3x_2 - \delta_3/r)|\alpha_2|^2 + \operatorname{Re}(2x_1^{\rm H}A_3x_2\bar{\alpha}_1\alpha_2) \\ \quad + \operatorname{Re}(2x_2^{\rm H}A_3x_3\bar{\alpha}_2\alpha_3) + \operatorname{Re}(2x_3^{\rm H}A_3x_1\bar{\alpha}_3\alpha_1) + (x_3^{\rm H}A_3x_3 - \delta_3/r)|\alpha_3|^2. \end{cases} \tag{3.1}$$

By Lemma 3.1 we know that the above system of equations has a nonzero solution $(\alpha_1^0, \alpha_2^0, \alpha_3^0)$. Let us normalize the solution to be

$$\alpha := (\alpha_1^0, \alpha_2^0, \alpha_3^0)^{\rm T}/\sqrt{|\alpha_1^0|^2 + |\alpha_2^0|^2 + |\alpha_3^0|^2}.$$

One easily verifies that $I_3 - \alpha\alpha^{\rm H}$ is a positive semidefinite matrix of rank 2, and

$$x_1x_1^{\rm H} + x_2x_2^{\rm H} + x_3x_3^{\rm H} - yy^{\rm H} = [x_1, x_2, x_3](I_3 - \alpha\alpha^{\rm H})[x_1, x_2, x_3]^{\rm H}.$$

Therefore $X - yy^{\rm H}$ remains positive semidefinite and $\operatorname{rank}(X - yy^{\rm H}) = r - 1$.

Let us update x_1 and X by $x_1 := y$ and $X := X - yy^{\mathrm{H}}$, and then repeat the above procedure until rank $(X) = 2$. Then we have $A_i \bullet x_1 x_1^{\mathrm{H}} = \cdots = A_i \bullet x_{r-2} x_{r-2}^{\mathrm{H}} = \delta_i/r, i = 1, 2, 3$. Finally, applying Proposition 3.2 to the resulting X leads to $A_i \bullet x_{r-1} x_{r-1}^{\mathrm{H}} = A_i \bullet x_r x_r^{\mathrm{H}} = \delta_i/r, i = 1, 2$. The proof is thus complete. □

We remark that Theorem 2.1 is a generalization of Theorem 3.4 of [2], Theorem 2.1 of [18] and Corollary 4 of [25].

3.2 Proof of Theorem 2.2

If rank $(X) \geq 3$, then applying Theorem 2.1, we have that there is a rank-one decomposition $X = \sum_{i=1}^{r} x_i x_i^{\mathrm{H}}$ such that

$$A_1 \bullet x_i x_i^{\mathrm{H}} = \delta_1/r,\ A_2 \bullet x_i x_i^{\mathrm{H}} = \delta_2/r,\ i = 1, \ldots, r;\ A_3 \bullet x_i x_i^{\mathrm{H}} = \delta_3/r,\ i = 1, \ldots, r-2,$$

where $r = \text{rank}\,(X)$. Observing that $x_i^{\mathrm{H}} z = 0$, for all $z \in \text{Null}(X)$, hence $x_i \in \text{Range}(X)$, $i = 1, \ldots, r$. Take $y = \sqrt{r} x_1$. We have $A_1 \bullet yy^{\mathrm{H}} = \delta_1, A_2 \bullet yy^{\mathrm{H}} = \delta_2, A_3 \bullet yy^{\mathrm{H}} = \delta_3$. Also, clearly in this case $X - \frac{1}{r} yy^{\mathrm{H}} \succeq 0$ and rank $(X - \frac{1}{r} yy^{\mathrm{H}}) \leq r - 1$, and the proof is complete.

If rank $(X) = 2$, then by Proposition 3.2 there exists a rank-one decomposition of X, $X = x_1 x_1^{\mathrm{H}} + x_2 x_2^{\mathrm{H}}$, such that

$$A_1 \bullet x_1 x_1^{\mathrm{H}} = A_1 \bullet x_2 x_2^{\mathrm{H}} = \delta_1/2; \quad A_2 \bullet x_1 x_1^{\mathrm{H}} = A_2 \bullet x_2 x_2^{\mathrm{H}} = \delta_2/2.$$

If $A_3 \bullet x_1 x_1^{\mathrm{H}} = \delta_3/2$, then choosing $y = \sqrt{2} x_1$ would complete the proof. Otherwise, in the remainder let us consider the case that $(A_3 \bullet x_1 x_1^{\mathrm{H}} - \delta_3/2)(A_3 \bullet x_2 x_2^{\mathrm{H}} - \delta_3/2) < 0$.

Since $n \geq 3$, there must be a vector x_3 which is linearly independent of x_1 and x_2. We claim that there exists a nonzero $(\alpha_1, \alpha_2, \alpha_3)$, such that

$$y = \frac{\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3}{\sqrt{|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2}}; \quad A_1 \bullet yy^{\mathrm{H}} = \delta_1/2, \quad A_2 \bullet yy^{\mathrm{H}} = \delta_2/2, \quad A_3 \bullet yy^{\mathrm{H}} = \delta_3/2.$$

Indeed, the above can be equivalently expanded to the following system of equations:

$$\begin{cases} \text{Re}\,(2x_1^{\mathrm{H}} A_1 x_2 \bar{\alpha}_1 \alpha_2) + \text{Re}\,(2x_2^{\mathrm{H}} A_1 x_3 \bar{\alpha}_2 \alpha_3) + \text{Re}\,(2x_3^{\mathrm{H}} A_1 x_1 \bar{\alpha}_3 \alpha_1) + (x_3^{\mathrm{H}} A_1 x_3 - \delta_1/2)|\alpha_3|^2 & = & 0, \\ \text{Re}\,(2x_1^{\mathrm{H}} A_2 x_2 \bar{\alpha}_1 \alpha_2) + \text{Re}\,(2x_2^{\mathrm{H}} A_2 x_3 \bar{\alpha}_2 \alpha_3) + \text{Re}\,(2x_3^{\mathrm{H}} A_2 x_1 \bar{\alpha}_3 \alpha_1) + (x_3^{\mathrm{H}} A_2 x_3 - \delta_2/2)|\alpha_3|^2 & = & 0, \\ (x_1^{\mathrm{H}} A_3 x_1 - \delta_3/2)|\alpha_1|^2 + (x_2^{\mathrm{H}} A_3 x_2 - \delta_3/2)|\alpha_2|^2 + (x_3^{\mathrm{H}} A_3 x_3 - \delta_3/2)|\alpha_3|^2 & & \\ \quad + \text{Re}\,(2x_1^{\mathrm{H}} A_3 x_2 \bar{\alpha}_1 \alpha_2) + \text{Re}\,(2x_2^{\mathrm{H}} A_3 x_3 \bar{\alpha}_2 \alpha_3) + \text{Re}\,(2x_3^{\mathrm{H}} A_3 x_1 \bar{\alpha}_3 \alpha_1) & = & 0. \end{cases} \tag{3.2}$$

Following the constructive proof of Lemma 3.1, we can find such $(\alpha_1, \alpha_2, \alpha_3)$. Thus, letting $z := x_3$ and $y := \sqrt{2} y$ produces the required solution. □

Remark. The mere existence of y in Theorem 2.2 follows from a result of Bohnenblust [9]; however, Bohnenblust's proof is nonconstructive. In the case that $(A_1 \bullet X, A_2 \bullet X, A_3 \bullet X) \neq (0,0,0)$, one may also prove Theorem 2.2 constructively by using some purification techniques (see e.g. Theorem 5.1 of [18], or Appendix A of [5], or [21]). However, if $A_i \bullet X = 0$ for all $i = 1,2,3$, then these purification techniques do not work. Even if $(A_1 \bullet X, A_2 \bullet X, A_3 \bullet X) \neq (0,0,0)$, such y obtained by the purification technique may not satisfy the property that $\operatorname{rank}\left(X - \frac{1}{r}yy^{\mathrm{H}}\right) \leq r-1$. Here is such an example. Suppose that

$$A_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \; A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \; A_3 = \begin{bmatrix} 1 & \boldsymbol{i} & 0 \\ -\boldsymbol{i} & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The system of matrix equations

$$\begin{cases} A_1 \bullet X = \delta_1 = 2, \\ A_2 \bullet X = \delta_2 = 0, \\ A_3 \bullet X = \delta_3 = 2, \end{cases} \tag{3.3}$$

has a matrix solution $X = I_3$, the 3×3 identity matrix. Following a standard rank reduction procedure (e.g. [18]), we obtain from I_3 a rank-one solution

$$x_1 x_1^{\mathrm{H}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

However, one verifies that

$$\operatorname{rank}\left(X - \frac{1}{3}x_1 x_1^{\mathrm{H}}\right) = \operatorname{rank}\left(\begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 3.$$

If we apply the procedure as in the proof of Theorem 2.2 to the above example, then the output of our Matlab program (to be introduced later) is

$$y = \begin{pmatrix} -0.0737 - 1.6514\boldsymbol{i} \\ 0.2745 - 0.0789\boldsymbol{i} \\ -0.3168 + 0.2922\boldsymbol{i} \end{pmatrix},$$

and the rank of $X - \frac{1}{3}yy^{\mathrm{H}}$ is indeed 2.

3.3 Proof of Theorem 2.3

Since $(\delta_1, \delta_2, \delta_3, \delta_4) \neq 0$, without loss of generality, let us assume $\delta_4 \neq 0$. Then we have

$$\left(A_i - \frac{\delta_i}{\delta_4}A_4\right) \bullet X = 0, \; i = 1,2,3.$$

It follows from Theorem 2.2 that there exists a nonzero n-dimensional complex vector $y \in \mathbf{C}^n$ such that

$$\left(A_i - \frac{\delta_i}{\delta_4} A_4\right) \bullet yy^{\mathrm{H}} = 0, \;\; i = 1, 2, 3.$$

Denote $t = A_4 \bullet yy^{\mathrm{H}}/\delta_4$. Then we have

$$A_i \bullet yy^{\mathrm{H}} = t\delta_i, \;\; i = 1, 2, 3, 4.$$

It is clear that $t \neq 0$, since $(A_1 \bullet Y, A_2 \bullet Y, A_3 \bullet Y, A_4 \bullet Y) \neq (0, 0, 0, 0)$ for all nonzero $Y \succeq 0$. Furthermore, $t > 0$, because if $t < 0$ then by setting $Y = X - \frac{1}{t} yy^H \succeq 0$ we will have $A_i \bullet Y = 0$ for all $i = 1, 2, 3, 4$, which violates the condition of the theorem. Letting $y := \frac{1}{\sqrt{t}} y$, we get

$$A_i \bullet yy^{\mathrm{H}} = \delta_i, \;\; i = 1, 2, 3, 4.$$

□

4 Applications in Signal Processing and Communication

In this section, we will illustrate applications of the matrix decomposition results presented in Section 2 to solve some specific problems in signal processing and communication. In particular, we shall study one application from the radar code optimization and another from robust beamforming. These two case-studies will be presented in the next two subsections, to show how non-convex (complex) quadratic optimization models arise in engineering applications, and how our new results help to solve such non-convex quadratic programs.

4.1 Application in Radar Space-Time Adaptive Processing (STAP)

Various algorithms for radar signal design have been proposed in the recent years. One approach in the radar signal design, known as the radar coding, relies on the modulation of a pulse train parameters (amplitude, phase, and frequency), so as to synthesize waveforms to satisfy some specified properties. A substantial amount of work has been done and recorded in the literature on this topic; the interested reader is referred to our recent paper and the references therein [14]. In this subsection, we introduce the code optimization problem for radar STAP in the presence of colored Gaussian disturbances. Our main goal, however, is to demonstrate how the decomposition theorems developed in Section 2 of the paper help solve the resulting optimization problems.

The problem is to design an optimal radar code, which is used to modulate transmitted signals belonging to the class of coded pulse trains, in such a way that the detection performance of radar is maximized under a control on the achievable values of the temporal and spatial Doppler estimation accuracy, as well as on the degree of similarity with a pre-fixed radar code. The resulting optimization problem is formulated as following:

$$
(\mathrm{QP}) \begin{cases} \max_{c\in\mathbf{C}^n} & c^{\mathrm{H}} R_0 c \\ \text{s.t.} & c^{\mathrm{H}} c = 1, \\ & c^{\mathrm{H}} R_t c \ge \delta_t, \\ & c^{\mathrm{H}} R_s c \ge \delta_s, \\ & \|c - \hat{c}\|^2 \le \epsilon. \end{cases}
$$

The objective function is the output Signal-to-Interference-plus-Noise Ratio (SINR), with respect to which the detection probability (P_d) of radar is increasing, and $R_0 \succ 0$. The first and fourth constraints are normalized code constraint and similarity constraint with a pre-fixed code $\hat{c}$ ($\hat{c}^{\mathrm{H}}\hat{c} = 1$), respectively, and the second and third constraints rule the region of achievable temporal and spatial Doppler estimation accuracies, respectively. For the detailed derivation, one is referred to [14]. It is easy to verify that (QP) is tantamount to the following homogenous quadratic programming:

$$
(\mathrm{HQP}) \begin{cases} \max_{c\in\mathbf{C}^n} & c^{\mathrm{H}} R_0 c \\ \text{s.t.} & c^{\mathrm{H}} c = 1, \\ & c^{\mathrm{H}} R_t c \ge \delta_t, \\ & c^{\mathrm{H}} R_s c \ge \delta_s, \\ & c^{\mathrm{H}} C_0 c \ge \delta_\epsilon, \end{cases}
$$

where $C_0 = \hat{c}\hat{c}^{\mathrm{H}}$ and $\delta_\epsilon = (1 - \epsilon/2)^2$, in the sense that they have the same optimal value and $c^* = \tilde{c}e^{\boldsymbol{i}\arg(\tilde{c}^H\hat{c})}$ is an optimal solution of (QP) if $\tilde{c}$ is optimal to (HQP). For this reason, from now on we shall focus on (HQP). Note that the SDP relaxation for (HQP) and its dual are:

$$
(\mathrm{SDR}) \begin{cases} \max & R_0 \bullet C \\ \text{s.t.} & I \bullet C = 1, \\ & R_t \bullet C \ge \delta_t, \\ & R_s \bullet C \ge \delta_s, \\ & C_0 \bullet C \ge \delta_\epsilon, \\ & C \succeq 0, \end{cases} \quad \text{and } (\mathrm{DSDR}) \begin{cases} \min & y_1 - y_2\delta_t - y_3\delta_s - y_4\delta_\epsilon \\ \text{s.t.} & y_1 I - y_2 R_t - y_3 R_s - y_4 C_0 \succeq R_0, \\ & y_1 \in \Re,\ y_2 \ge 0,\ y_3 \ge 0,\ y_4 \ge 0, \end{cases}
$$

respectively.

Suppose that both the primal and dual SDPs are solvable. (To this end, it suffices to assume that the initial code $\hat{c}$ is strictly feasible to (QP), which ensures (SDR) to be strictly feasible, while (DSDR) is always strictly feasible). Suppose that the primal-dual feasible pair $(C^*, y_1^*, y_2^*, y_3^*, y_4^*)$ is optimal to (SDR) and (DSDR) respectively. Consequently, it satisfies the complementary conditions:

$$
\begin{aligned}
(y_1^* I - y_2^* R_t - y_3^* R_s - y_4^* C_0 - R_0) \bullet C^* &= 0, && (4.1)\\
(R_t \bullet C^* - \delta_t) y_2^* &= 0, && (4.2)\\
(R_s \bullet C^* - \delta_s) y_3^* &= 0, && (4.3)\\
(C_0 \bullet C^* - \delta_\epsilon) y_4^* &= 0. && (4.4)
\end{aligned}
$$

If $\operatorname{rank}(C^*) \geq 3$, then apply Theorem 2.3 to (C^*, I, R_t, R_s, C_0) we find a vector $c^* \in \operatorname{Range}(C^*)$ such that

$$I \bullet c^* c^{*\mathrm{H}} = I \bullet C^*,\ R_t \bullet c^* c^{*\mathrm{H}} = R_t \bullet C^*,\ R_s \bullet c^* c^{*\mathrm{H}} = R_s \bullet C^*,\ C_0 \bullet c^* c^{*\mathrm{H}} = C_0 \bullet C^*.$$

Since the primal-dual pair $(c^* c^{*\mathrm{H}}, y_1^*, y_2^*, y_3^*, y_4^*)$ is feasible and satisfies the complementary conditions, we conclude that $c^* c^{*\mathrm{H}}$ is optimal to (SDR) and c^* is optimal to (HQP).

If $\operatorname{rank}(C^*) = 2$ and one of the inequality constraints of (SDR) is non-binding (non-active) at optimality, then a rank-one solution of (SDR) can be obtained by resorting to Proposition 3.2. We remarked in [14, Section IV-C] that when we pick the parameters R_0, R_t, and R_s from the standard Knowledge Aided Sensor Signal Processing and Expert Reasoning (KASSPER) datacube, and set $\hat{c}$ to be a generalized Barker sequence, and randomly choose δ_t, δ_s and δ_ϵ from proper intervals respectively, our simulation shows that only 10 out of 10,000 instances of (QP) can be classified as in the 'hard case', meaning that the output solution of the corresponding (SDR) is exactly of rank 2 and all inequality constraints are active at the optimality. For these instances classified as the 'hard case' above, we then apply Theorem 2.3 to get a high quality feasible solution for (HQP); specifically, we may choose any z vector as required in Theorem 2.3 to produce a feasible solution based on that vector z. The z vector can be picked randomly and we can choose the best one among the random samples. Our simulation results show that this heuristic for the 'hard case' instances works extremely well in practice.

To summarize, in this approach the main steps to solve the radar code optimization problem involve: (1) using an SDP solver (e.g. SeDuMi) to solve (SDR); (2) implementing the appropriate decomposition theorem to further process the solution of (SDR) as we discussed above; (3) assembling an optimal solution (or high quality feasible solution) for (HQP). In the last section of the paper, we will test the numerical performance of the decomposition procedures stipulated in Theorems 2.1 - 2.3.

4.2 Application in Triply Constrained Robust Capon Beamformer Design

Beamforming is an important task in array signal processing. The so-called Capon beamformer selects the beamforming weight vector adaptively so as to minimize the array output power subject to the linear constraint that the signal of interest does not suffer from any distortion ([19, 20]). As long as the array steering vector (i.e., the response of the array to an arriving plane wave, termed also the *array manifold* in the literature) corresponding to the signal of interest is known, the Capon beamformer has a better performance than the standard data-independent beamformer; see [20]. However, like other adaptive beamformers, the Capon beamformer is sensitive to the steering vector to the signal of interest, and the performance of the Capon beamformer will be substantially degraded if the mismatch between the presumed steering vector and the true steering vector occurs. In practical situations, this type of mismatch can easily occur due, for instance, to look directions and signal pointing errors, among other possible causes. To cope with the performance degradation caused by the mismatch and improve the robustness of the adaptive beamformers, engineers have proposed various remedies in the past three decades (cf. [19] and the references therein). In this subsection, we shall focus on the robust Capon beamforming problem as introduced in [20], and we shall demonstrate how to resolve this class of problems using the decomposition theorems presented in this paper.

Consider an array of n sensors, and denote by R the theoretical covariance matrix of the array output vector. Assume that R has the following form:

$$R = \sigma^2 a a^{\mathrm{H}} + \sum_{k=1}^{K} \sigma_k^2 b_k b_k^{\mathrm{H}} + Q, \tag{4.5}$$

where $(\sigma^2, \sigma_1^2, \ldots, \sigma_K^2)$ are the powers of the $(K+1)$ uncorrelated signals impinging on the array, a is the steering vector to the signal of interest, $(b_1, \ldots, b_K)$ are the steering vectors to the K signals treated as interference, and $Q \succ 0$ is the noise covariance. In practical applications, R is replaced by the sample covariance of recently received samples of the array output, e.g.,

$$\hat{R} = \frac{1}{N} \sum_{n=1}^{N} y_n y_n^{\mathrm{H}}.$$

The robust Capon beamforming problem that we deal with is to determine the power of the signal of interest (i.e., σ^2 in (4.5)), where the only knowledge about its steering vector a is that it belongs to an uncertainty set. More specifically, we assume the uncertainty set is given by

$$A = \{a \in \mathbf{C}^n : \|a\|^2 = n,\ (a - a_1)^{\mathrm{H}} D_1 (a - a_1) \le \epsilon_1,\ (a - a_2)^{\mathrm{H}} D_2 (a - a_2) \ge \epsilon_2\}. \tag{4.6}$$

In the above set, the norm constraint is posed to avoid ambiguities (see [20]); in the second constraint (i.e., the similarity constraint), $D_1 \succeq 0$ and $\epsilon_1 > 0$ define the similarity scale, and a_1 is a presumed nominal steering vector; by posing the third constraint, the steering vector is forced to be away from the prefixed steering vector a_2 corresponding to a known co-existing source emitting "contaminating" (or interfering) signals, and the parameters $D_2 \succeq 0$ and $\epsilon_2 \ge 0$. In particular, if $D_1 = I$ and $\epsilon_2 = 0$, then the uncertainty region is exactly the same as the one considered in [20].

Using the reformulation of the robust Capon beamforming problem in [20], to which the uncertainty set (4.6) is appended, we obtain a robust estimate of σ^2, without calculating the beamforming weight vector:

$$\text{(RCB)} \begin{cases} \max_{\sigma^2, a} & \sigma^2 \\ \text{s.t.} & R - \sigma^2 a a^{\mathrm{H}} \succeq 0, \ \forall a \in A. \end{cases}$$

It is easily checked that the optimal value is attained $\hat{\sigma}^2 = 1/(a^{*\mathrm{H}} R^{-1} a^*)$, where a^* is an optimal solution of the following equivalent problem

$$\text{(RCB)} \begin{cases} \min & a^{\mathrm{H}} R^{-1} a \\ \text{s.t.} & \|a\|^2 = n, \\ & (a - a_1)^{\mathrm{H}} D_1 (a - a_1) \le \epsilon_1, \\ & (a - a_2)^{\mathrm{H}} D_2 (a - a_2) \ge \epsilon_2. \end{cases}$$

Therefore, all we need to do now is to solve (RCB), which can be further homogenized as:

$$\text{(HRCB)} \begin{cases} \min & x^{\mathrm{H}} R_0 x \\ \text{s.t.} & x^{\mathrm{H}} R_1 x = n, \\ & x^{\mathrm{H}} R_2 x \le \epsilon_1, \\ & x^{\mathrm{H}} R_3 x \ge \epsilon_2, \\ & x^{\mathrm{H}} R_4 x = 1, \end{cases}$$

where

$$x = \begin{bmatrix} a \\ t \end{bmatrix}, \ R_0 = \begin{bmatrix} R^{-1} & 0_{n\times 1} \\ 0_{1\times n} & 0 \end{bmatrix}, \ R_1 = \begin{bmatrix} I & 0_{n\times 1} \\ 0_{1\times n} & 0 \end{bmatrix},$$

$$R_2 = \begin{bmatrix} D_1 & -D_1 a_1 \\ -a_1^{\mathrm{H}} D_1 & a_1^{\mathrm{H}} D_1 a_1 \end{bmatrix}, \ R_3 = \begin{bmatrix} D_2 & -D_2 a_2 \\ -a_2^{\mathrm{H}} D_2 & a_2^{\mathrm{H}} D_2 a_2 \end{bmatrix}, \ R_4 = \begin{bmatrix} 0_{n\times n} & 0_{n\times 1} \\ 0_{1\times n} & 1 \end{bmatrix}.$$

In other words, (RCB) and (HRCB) have the same optimal value, and a^*/t^* is an optimal solution if $x^* = [a^{*\mathrm{T}}, t^{*\mathrm{T}}]^{\mathrm{T}}$ is an optimal solution of (HRCB). The SDP relaxation of (HRCB) and its dual are:

$$\text{(SDR1)} \begin{cases} \min & R_0 \bullet X \\ \text{s.t.} & R_1 \bullet X = n, \\ & R_2 \bullet X \le \epsilon_1, \\ & R_3 \bullet X \ge \epsilon_2, \\ & R_4 \bullet X = 1, \\ & X \succeq 0, \end{cases} \quad \text{and (DSDR1)} \begin{cases} \max & n y_1 + \epsilon_1 y_2 + \epsilon_2 y_3 + y_4 \\ \text{s.t.} & R_0 - y_1 R_1 - y_2 R_2 - y_3 R_3 - y_4 R_4 \succeq 0, \\ & y_1 \in \Re, \ y_2 \le 0, \ y_3 \ge 0, \ y_4 \in \Re. \end{cases}$$

Suppose that both (SDR1) and (DSDR1) are solvable, and let $(X^*, y_1^*, y_2^*, y_3^*, y_4^*)$ be an optimal primal-dual pair. Similarly as in the previous application case, if the rank of X^* is no less than three, then an optimal solution of (HRCB) can be found using Theorem 2.3; if the rank of X^* is equal to two and one of the two inequality constraints of (SDR1) is non-binding at the point X^*, then an optimal solution of (HRCB) can be found by Proposition 3.2; if the rank of X^* is equal to two and the two inequality constraints of (SDR1) are binding at the point X^*, then a good feasible solution of (HRCB) can be found by resorting to Theorem 2.3 where one may randomly choose the 'dimension extending direction' z.

Here we shall briefly summarize our findings in this section. As far as optimization formulations are concerned, our new matrix decomposition results can be applied, in combination of SDP relaxation, to solve any non-convex, complex-valued, homogeneous quadratically constrained quadratic programs, where the number of constraints is at most four and the number of variables is at least three. Our approach solves a vast majority of the instances to optimality, leaving only one possible unsettled case, whose frequency of occurrence is exceedingly small. Even when the unsettled case does occur, our decomposition theorem (Theorem 2.3) can still be applied to yield a good feasible solution (though not necessarily optimal anymore in this case). Such non-convex quadratic optimization models appear to be useful for a variety of practical applications, and according to our numerical experiences the approach proposed in this paper works extremely well, for instance, in solving problems arising from signal processing.

5 Theoretical Implications of the New Results

5.1 Extension of Yuan's lemma and the $\mathcal{S}$-lemma

Yuan [27] proved the following result (Lemma 2.3 in Yuan [27]) which turns out to be useful in several applications.

Theorem 5.1. *Let A_1 and A_2 be in $\mathcal{S}^n$. Then the following are equivalent:*

(i) $\max\{x^{\mathrm{T}}A_1x, x^{\mathrm{T}}A_2x\} \geq 0$ *for all* $x \in \Re^n$ *(resp.* > 0 *for all* $x \neq 0$*).*

(ii) There exist $\mu_1 \geq 0, \mu_2 \geq 0, \mu_1 + \mu_2 = 1$ *such that* $\mu_1 A_1 + \mu_2 A_2 \succeq 0$ *(resp.* $\succ 0$*).*

Our new results lead to the following extension for Yuan's theorem.

Theorem 5.2. *Suppose that* $n \geq 3$, $A_i \in \mathcal{H}^n, i = 1,2,3,4$, *and there are* $\lambda_i \in \Re, i = 1,2,3,4$, *such that* $\lambda_1 A_1 + \lambda_2 A_2 + \lambda_3 A_3 + \lambda_4 A_4 \succ 0$. *If*

$$\max\{z^{\mathrm{H}}A_1z, z^{\mathrm{H}}A_2z, z^{\mathrm{H}}A_3z, z^{\mathrm{H}}A_4z\} \geq 0, \forall z \in \mathbf{C}^n \,(\textit{resp.} \; > 0, \forall z \neq 0),$$

then there are $\mu_i \geq 0, i = 1,2,3,4$, *such that* $\mu_1 + \mu_2 + \mu_3 + \mu_4 = 1$ *and* $\mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 + \mu_4 A_4 \succeq 0 \,(\textit{resp.} \; \succ 0)$.

Proof. If the conditions of the theorem hold, then $\max\{A_1 \bullet Z, A_2 \bullet Z, A_3 \bullet Z, A_4 \bullet Z\} \geq 0$ for all $Z \succeq 0$. To see this, suppose by contradiction that there is $Z \succeq 0$ such that $A_i \bullet Z < 0$, for $i = 1, 2, 3, 4$. It follows from Theorem 2.3 that there is nonzero $y \in \mathbf{C}^n$, with $A_i \bullet yy^{\rm H} < 0, i = 1, 2, 3, 4$, which contradicts to the condition that $\max\{z^{\rm H}A_1z, z^{\rm H}A_2z, z^{\rm H}A_3z, z^{\rm H}A_4z\} \geq 0, \forall z \in \mathbf{C}^n$. Therefore, the optimal value of the following SDP problem

$$\begin{cases} \min & t \\ \text{s.t.} & t - A_i \bullet Z \geq 0,\ i = 1, 2, 3, 4 \\ & I \bullet Z = 1, \\ & Z \succeq 0 \end{cases}$$

is nonnegative. At the same time, the dual of the above problem is

$$\begin{cases} \max & \mu_5 \\ \text{s.t.} & \mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 + \mu_4 A_4 \succeq \mu_5 I, \\ & 1 - \mu_1 - \mu_2 - \mu_3 - \mu_4 = 0, \\ & \mu_1 \geq 0, \mu_2 \geq 0, \mu_3 \geq 0, \mu_4 \geq 0. \end{cases}$$

It is easily seen that both the primal and dual problems have a strictly feasible solution. Hence the strong duality holds, with both primal and dual problems having attainable optimal solutions, to be denoted by (t^*, Z^*) and $(\mu_1^*, \mu_2^*, \mu_3^*, \mu_4^*, \mu_5^*)$ respectively. Note that $\mu_i^* \geq 0, i = 1, 2, 3, 4$, $\mu_1^* + \mu_2^* + \mu_3^* + {} + \mu_4^* = 1$, and $\mu_1^* A_1 + \mu_2^* A_2 + \mu_3^* A_3 + \mu_4^* A_4 \succeq \mu_5^* I = t^* I \succeq 0$, completing the proof. □

We remark that in the same vein, Yuan's result (Theorem 5.1) can be proven analogously (but then in the real domain), using Proposition 3 of Sturm and Zhang [25]. In fact, there are several different ways to prove Theorem 5.1, including of course the original one presented in Yuan [27]. Remark here that Hiriart-Urruty and Torki [16] used a separation argument to prove Theorem 5.1, based on a convexity result of the joint numerical range. The situation is similar for the proofs of the $\mathcal{S}$-lemma (see Pólik and Terlaky [24]). Particularly, in Section 2.4 of [24], the authors used Theorem 5.1 to prove the $\mathcal{S}$-lemma.

5.2 Joint numerical ranges

To put the results in perspective, we present in this subsection some other connections between the convexity of the joint numerical ranges, generalized Finsler's lemma, and our new rank-one decomposition theorem.

Theorem 5.3. *The following statements are all equivalent and correct:*

1) Theorem 2.2, where $A_i \bullet X = 0$, $i = 1, 2, 3$.

2) Suppose $n \geq 3$, and $A_1, A_2, A_3 \in \mathcal{H}^n$, then the set

$$W = \left\{(z^{\mathrm{H}}A_1z, z^{\mathrm{H}}A_2z, z^{\mathrm{H}}A_3z) \mid z^{\mathrm{H}}z = 1, z \in \mathbf{C}^n\right\}$$

is convex.

3) Suppose $n \geq 3$, and $A_1, A_2, A_3 \in \mathcal{H}^n$ satisfy $\rho(z) = \sqrt{(z^{\mathrm{H}}A_1z)^2 + (z^{\mathrm{H}}A_2z)^2 + (z^{\mathrm{H}}A_3z)^2} > 0$, for all z with $\|z\| = 1$. Then the set

$$S := \left\{\left(z^{\mathrm{H}}A_1z/\rho(z), z^{\mathrm{H}}A_2z/\rho(z), z^{\mathrm{H}}A_3z/\rho(z)\right) \mid z^{\mathrm{H}}z = 1\right\}$$

is a closed region contained in a half-sphere of the 3-dimensional unit ball, with the property that $x \in S \implies -x \notin S$.

4) Suppose $n \geq 3$, and $A_1, A_2, A_3 \in \mathcal{H}^n$ satisfy $(z^{\mathrm{H}}A_1z, z^{\mathrm{H}}A_2z, z^{\mathrm{H}}A_3z) \neq (0,0,0), \forall z \in \mathbf{C}^n \setminus \{0\}$. Then there exist $\lambda_1, \lambda_2, \lambda_3 \in \Re$ such that

$$\lambda_1 A_1 + \lambda_2 A_2 + \lambda_3 A_3 \succ 0.$$

Proof. 1) $\implies$ 2). Let $W' = \{(A_1 \bullet Z, A_2 \bullet Z, A_3 \bullet Z) \mid Z \bullet I = 1, Z \succeq 0\}$. Clearly W' is a convex set, and $W \subseteq W'$. We shall show $W' \subseteq W$. Let $(v_1, v_2, v_3) = (A_1 \bullet Z, A_2 \bullet Z, A_3 \bullet Z) \in W'$, and $A_i' := A_i - v_i I$, $i = 1,2,3$. Then $A_i' \bullet Z = 0$, $i = 1,2,3$. It follows by Theorem 2.2 that there is a non-zero vector $y \in \mathbf{C}^n$ such that $A_i' \bullet yy^{\mathrm{H}} = 0$, $i = 1,2,3$; hence

$$A_i \bullet \left(\frac{y}{\sqrt{y^{\mathrm{H}}y}} \frac{y^{\mathrm{H}}}{\sqrt{y^{\mathrm{T}}y}}\right) = v_i,\ i = 1,2,3,$$

implying that (v_1, v_2, v_3) belongs to W as well, and consequently $W' = W$.

2) $\implies$ 3). We begin by showing that S does not contain opposite points. Suppose that this is not the case. Let $(a,b,c) \in S$, while $-(a,b,c) \in S$; that is, there are $x, y \in \mathbf{C}^n$ with $\|x\| = \|y\| = 1$ and

$$\rho(x) \times (a,b,c) = (x^{\mathrm{H}}A_1x, x^{\mathrm{H}}A_2x, x^{\mathrm{H}}A_3x), -\rho(y) \times (a,b,c) = (y^{\mathrm{H}}A_1y, y^{\mathrm{H}}A_2y, y^{\mathrm{H}}A_3y).$$

Since W is convex, we have

$$\frac{\rho(y)}{\rho(x)+\rho(y)}(x^{\mathrm{H}}A_1x, x^{\mathrm{H}}A_2x, x^{\mathrm{H}}A_3x) + \frac{\rho(x)}{\rho(x)+\rho(y)}(y^{\mathrm{H}}A_1y, y^{\mathrm{H}}A_2y, y^{\mathrm{H}}A_3y) \in W,$$

and so

$$(0,0,0) = \frac{\rho(y)}{\rho(x)+\rho(y)}\rho(x)(a,b,c) + \frac{\rho(x)}{\rho(x)+\rho(y)}\rho(y)(-a,-b,-c) \in W.$$

Hence there is z with $\|z\| = 1$ such that $(z^{\mathrm{H}}A_1z, z^{\mathrm{H}}A_2z, z^{\mathrm{H}}A_3z) = (0,0,0)$, i.e., $\rho(z) = 0$, which is a contradiction.

Observe that S is the image of W on the sphere of the unit ball, projected from the origin. Since W is a compact convex set which does not contain the origin, its image S must be contained strictly in one half of the unit ball.

3) $\Longrightarrow$ 4). Since S is a closed spherical region contained strictly in one half of the 3-dimensional unit ball, and does not contain two opposite points with respect to the origin, its convex hull $\bar{S}$ must not contain the origin either. By the separation theorem, there is a plane going through 0, with $\bar{S}$ lying properly on one half of the plane. That is, if we let (a_0, b_0, c_0) the the normal direction of the separating plane, then the angle between any vector in $\bar{S}$ and the direction (a_0, b_0, c_0) should be strictly less than $\theta < \pi/2$. In other words,

$$\cos\theta = a_0 z^{\mathrm{H}} A_1 z/\rho(z) + b_0 z^{\mathrm{H}} A_2 z/\rho(z) + c_0 z^{\mathrm{H}} A_3 z/\rho(z) > 0, \forall z : \|z\| = 1.$$

This implies that $a_0 A_1 + b_0 A_2 + c_0 A_3 \succ 0$.

4) $\Longrightarrow$ 1). A key step in the proof of Theorem 2.2 is that the system of equations (3.2) has a nontrivial solution. Recall that (3.2) can be explicitly written as

$$(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3) \begin{bmatrix} 0 & x_1^{\mathrm{H}} A_1 x_2 & x_3^{\mathrm{H}} A_1 x_1 \\ & 0 & x_2^{\mathrm{H}} A_1 x_3 \\ & & x_3^{\mathrm{H}} A_1 x_3 - \frac{\delta_1}{2} \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = 0$$
$$(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3) \begin{bmatrix} 0 & x_1^{\mathrm{H}} A_2 x_2 & x_3^{\mathrm{H}} A_2 x_1 \\ & 0 & x_2^{\mathrm{H}} A_2 x_3 \\ & & x_3^{\mathrm{H}} A_2 x_3 - \frac{\delta_2}{2} \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = 0$$
$$(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3) \begin{bmatrix} x_1^{\mathrm{H}} A_3 x_1 - \frac{\delta_3}{2} & x_1^{\mathrm{H}} A_3 x_2 & x_3^{\mathrm{H}} A_3 x_1 \\ & x_2^{\mathrm{H}} A_3 x_2 - \frac{\delta_3}{2} & x_2^{\mathrm{H}} A_3 x_3 \\ & & x_3^{\mathrm{H}} A_3 x_3 - \frac{\delta_3}{2} \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = 0,$$

where for clarity we only displayed the upper part of the matrices – the lower part is determined by (Hermitian) symmetry. The first two diagonal elements of these three matrices are $(0, 0)$, $(0, 0)$, and $(x_1^{\mathrm{H}} A_3 x_1 - \frac{\delta_3}{2}, x_2^{\mathrm{H}} A_3 x_2 - \frac{\delta_3}{2})$. The last vector has one positive element and one negative element, and so it is impossible to combine these vectors into a positive vector, let alone the three entire matrices. By 4), we know that there is a non-zero solution of (3.2). The rest of the proof is identical to that in the proof of Theorem 2.2. The essential difference between this argument, and the proof presented in Section 2 is that the latter is constructive; i.e., following that proof we can actually produce a rank-one solution of a system of linear matrix equations, which makes the result more useful in practical engineering applications. □

Another immediate consequence of Theorem 2.3 is the following result, which generalizes the famous theorems of Hausdorff [15] and Brickman [10].

Theorem 5.4. *Suppose $n \geq 3$. Let $A_1, A_2, A_3, A_4 \in \mathcal{H}^n$ satisfy*

$$(A_1 \bullet Y, A_2 \bullet Y, A_3 \bullet Y, A_4 \bullet Y) \neq (0,0,0,0),$$

for any nonzero matrix $Y \in \mathcal{H}^n_+$. Then,

$$\{(z^{\rm H}A_1z, z^{\rm H}A_2z, z^{\rm H}A_3z, z^{\rm H}A_4z) \mid z \in \mathbf{C}^n\}$$

is a convex cone.

Proof. Let us denote

$$\mathcal{K}_1 := \{(z^{\rm H}A_1z, z^{\rm H}A_2z, z^{\rm H}A_3z, z^{\rm H}A_4z) \mid z \in \mathbf{C}^n\},$$

and

$$\mathcal{K}_2 := \{(A_1 \bullet Z, A_2 \bullet Z, A_3 \bullet Z, A_4 \bullet Z) \mid Z \in \mathcal{H}^n_+\}.$$

Evidently, $\mathcal{K}_1 \subseteq \mathcal{K}_2$. Next we shall show the other containing relation, thus establishing the equality. Take any $0 \neq Z \in \mathcal{H}^n_+$. If $\text{rank}\,(Z) = 1$, then

$$(A_1 \bullet Z, A_2 \bullet Z, A_3 \bullet Z, A_4 \bullet Z) \in \mathcal{K}_1;$$

if $\text{rank}\,(Z) \geq 3$, then Theorem 2.3 asserts that there is $z \in \text{Range}(Z)$ such that

$$(A_1 \bullet Z, A_2 \bullet Z, A_3 \bullet Z, A_4 \bullet Z) = (z^{\rm H}A_1z, z^{\rm H}A_2z, z^{\rm H}A_3z, z^{\rm H}A_4z) \in \mathcal{K}_1;$$

if $\text{rank}\,(Z) = 2$, then, since $n \geq 3$ we have $\mathbf{C}^n \backslash \text{Range}(Z) \neq \emptyset$, and so by taking any $x \in \mathbf{C}^n \backslash \text{Range}(Z)$, Theorem 2.3 guarantees the existence of z which is in the span of x and $\text{Range}(Z)$, such that

$$(A_1 \bullet Z, A_2 \bullet Z, A_3 \bullet Z, A_4 \bullet Z) = (z^{\rm H}A_1z, z^{\rm H}A_2z, z^{\rm H}A_3z, z^{\rm H}A_4z) \in \mathcal{K}_1.$$

This shows that $\mathcal{K}_2 \subseteq \mathcal{K}_1$, and consequently $\mathcal{K}_1 = \mathcal{K}_2$ as a result. □

In case $A_4 = I_n$, then the condition of the theorem (regarding the A_i matrices) holds. It thus follows that

$$\{(z^{\rm H}A_1z, z^{\rm H}A_2z, z^{\rm H}A_3z) \mid z^{\rm H}z = 1\}$$

is a convex set whenever $n \geq 3$, which is an enhancement of Hausdorff's theorem [15], and also a generalization of Brickman's theorem [10] (see also Part 2 of Theorem 5.3). The real number counterpart of Theorem 5.4 includes a result of Polyak [22] as a special case; in Polyak's paper, a seemingly different and stronger condition is used, which is that the A_i matrices can be linearly combined into a positive definite one. Interestingly, the statements as in Theorem 5.4 can be slightly strengthened along this line as well.

Theorem 5.5. *Suppose $n \geq 3$, and $A_1, A_2, A_3, A_4 \in \mathcal{H}^n$. If*

$$(A_1 \bullet Y, A_2 \bullet Y, A_3 \bullet Y, A_4 \bullet Y) \neq (0,0,0,0),$$

for any nonzero matrix $Y \in \mathcal{H}^n_+$, then there exist $\alpha_i \in \Re$, $i = 1,2,3,4$ such that $\sum_{i=1}^4 \alpha_i A_i \succ 0$, and

$$\{(z^{\mathrm{H}} A_1 z, z^{\mathrm{H}} A_2 z, z^{\mathrm{H}} A_3 z, z^{\mathrm{H}} A_4 z) \mid z \in \mathbf{C}^n\}$$

is a pointed, closed, convex cone.

Proof. Let us denote

$$\mathcal{K} := \{(z^{\mathrm{H}} A_1 z, z^{\mathrm{H}} A_2 z, z^{\mathrm{H}} A_3 z, z^{\mathrm{H}} A_4 z) \mid z \in \mathbf{C}^n\}$$

which is a convex cone by Theorem 5.4. Moreover, $\mathcal{K}$ is also pointed, because, if there is $0 \neq d \in \mathcal{K}$, with $-d \in \mathcal{K}$, then by denoting

$$d = (z_1^{\mathrm{H}} A_1 z_1, z_1^{\mathrm{H}} A_2 z_1, z_1^{\mathrm{H}} A_3 z_1, z_1^{\mathrm{H}} A_4 z_1) \text{ and } -d = (z_2^{\mathrm{H}} A_1 z_2, z_2^{\mathrm{H}} A_2 z_2, z_2^{\mathrm{H}} A_3 z_2, z_2^{\mathrm{H}} A_4 z_2)$$

we would have

$$\big(A_1 \bullet (z_1 z_1^{\mathrm{H}} + z_2 z_2^{\mathrm{H}}), A_2 \bullet (z_1 z_1^{\mathrm{H}} + z_2 z_2^{\mathrm{H}}), A_3 \bullet (z_1 z_1^{\mathrm{H}} + z_2 z_2^{\mathrm{H}}), A_4 \bullet (z_1 z_1^{\mathrm{H}} + z_2 z_2^{\mathrm{H}})\big) = (0,0,0,0)$$

which contradicts the condition of the theorem. We proceed to show that $\mathcal{K}$ is a closed set. For this purpose, let us take any sequence $v^k \in \mathcal{K}$ with $\lim_{k\to\infty} v^k = \hat{v}$. Clearly, if $\hat{v} = 0$ then $\hat{v} \in \mathcal{K}$. Let us consider the nontrivial case where $\hat{v} \neq 0$. Suppose that

$$\begin{aligned} v^k &= (z_k^{\mathrm{H}} A_1 z_k, z_k^{\mathrm{H}} A_2 z_k, z_k^{\mathrm{H}} A_3 z_k, z_k^{\mathrm{H}} A_4 z_k) \\ &= \|z_k\|^2 \left(\left(\frac{z_k}{\|z_k\|}\right)^{\mathrm{H}} A_1 \frac{z_k}{\|z_k\|}, \left(\frac{z_k}{\|z_k\|}\right)^{\mathrm{H}} A_2 \frac{z_k}{\|z_k\|}, \left(\frac{z_k}{\|z_k\|}\right)^{\mathrm{H}} A_3 \frac{z_k}{\|z_k\|}, \left(\frac{z_k}{\|z_k\|}\right)^{\mathrm{H}} A_4 \frac{z_k}{\|z_k\|} \right). \end{aligned}$$

Without losing generality we may assume $\lim_{k\to\infty} z_k/\|z_k\| = \hat{z}$. Therefore,

$$\hat{v} = \lim_{k\to\infty} v^k = \lim_{k\to\infty} \|z_k\|^2 (\hat{z}^{\mathrm{H}} A_1 \hat{z}, \hat{z}^{\mathrm{H}} A_2 \hat{z}, \hat{z}^{\mathrm{H}} A_3 \hat{z}, \hat{z}^{\mathrm{H}} A_4 \hat{z}).$$

By the condition of the theorem, we know that $(\hat{z}^{\mathrm{H}} A_1 \hat{z}, \hat{z}^{\mathrm{H}} A_2 \hat{z}, \hat{z}^{\mathrm{H}} A_3 \hat{z}, \hat{z}^{\mathrm{H}} A_4 \hat{z}) \neq 0$, and therefore $\lim_{k\to\infty} \|z_k\|^2$ exists and is finite; hence $\hat{v} \in \mathcal{K}$. This shows that $\mathcal{K}$ is closed. To summarize, we have shown that $\mathcal{K}$ is a pointed, closed, and convex cone. Using the separation theorem, we conclude that there is $\alpha_i \in \Re$, $i = 1,2,3,4$, such that $\sum_{i=1}^4 \alpha_i z^{\mathrm{H}} A_i z > 0$ for any $0 \neq z \in \mathbf{C}^n$. Stated differently, $\sum_{i=1}^4 \alpha_i A_i \succ 0$. □

6 Algorithms and Numerical Results

Our recent investigations indicate that there are amply applications of the rank-one decomposition theorems in engineering, arising from signal processing, radar, wireless communication and so forth (see for example [12, 13, 14]). It is therefore helpful to provide workable Matlab codes implementing these rank-one decomposition theorems, for the benefits of the users. Based on our constructive proofs of the main results in the previous sections, we implemented the algorithms to get the rank-one solutions. Our Matlab programs can be found at the following website:

`http://www.se.cuhk.edu.hk/~ywhuang/dcmp/paper.html`

[username: `dcmp` ; password: `dcmpdcmp`]

The theme of this section is to test the numerical stability of the implemented programs, and consequently to make the algorithms accessible for practical problems. Lastly, the codes also serve to cross checking the validity of our proofs, at least numerically, since the programs are coded quite literally following the proofs presented in this paper. To start with, let us outline the procedure to implement the decomposition as stipulated in Theorem 2.1 (refer to the proof of Theorem 2.1 in [18]):

Algorithm 1: Computing the decomposition as in Theorem 2.1

Input: $A_1, A_2, A_3 \in \mathcal{H}^n$, and $X \in \mathcal{H}^n_+$ with $r = \mathrm{rank}\,(X) \geq 3$.
Output: $X = \sum_{i=1}^r \hat{x}_i \hat{x}_i^{\mathrm{H}}$, a rank-one decomposition of X, such that $A_1 \bullet \hat{x}_i \hat{x}_i^{\mathrm{H}} = A_1 \bullet X/r$, $A_2 \bullet \hat{x}_i \hat{x}_i^{\mathrm{H}} = A_2 \bullet X/r$, $i = 1, ..., r$; and $A_3 \bullet \hat{x}_i \hat{x}_i^{\mathrm{H}} = A_3 \bullet X/r$, $i = 1, ..., r-2$.

0) Let $r := \mathrm{rank}\,(X)$, $\delta_i := A_i \bullet X, i = 1, 2, 3$, and $s := 1$.

1) Repeat the following steps if $r \geq 3$:

1.1) Use Proposition 3.2, obtaining $X = \sum_{i=1}^r x_i x_i^{\mathrm{H}}$ such that $A_1 \bullet x_i x_i^{\mathrm{H}} = \delta_1/r, A_2 \bullet x_i x_i^{\mathrm{H}} = \delta_2/r, i = 1, \ldots, r$.

1.2) If all $A_3 \bullet x_i x_i^{\mathrm{H}} = \delta_3/r$ for $i = 1, \ldots, r-2$, then set $\hat{x}_i = x_i$, $i = 1, \ldots, r-2$, break and terminate; else go to Step 1.3.

1.3) Find $i_1, i_2 \in \{1, \ldots, r\}$ such that $A_3 \bullet x_{i_1} x_{i_1}^{\mathrm{H}} - \delta_3/r > 0$ and $A_3 \bullet x_{i_2} x_{i_2}^{\mathrm{H}} - \delta_3/r < 0$.

1.4) Pick any $i_3 \in \{1, \ldots, r\} \setminus \{i_1, i_2\}$. W.l.o.g., suppose $i_1 = 1, i_2 = 2$ and $i_3 = 3$.

1.5) Determine $\alpha_1, \alpha_2, \alpha_3 \in \mathbf{C}$ such that

$$\left(\frac{\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3}{\sqrt{|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2}}\right)^{\mathrm{H}} A_i \frac{\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3}{\sqrt{|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2}} = \frac{\delta_i}{r}, i = 1, 2, 3,$$

by solving the system of equations (3.1) (see Lemma 3.1).

1.6) Return $\hat{x}_s = \frac{\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3}{\sqrt{|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2}}$ and update $X = X - \hat{x}_s \hat{x}_s^{\mathrm{H}}$, $r = r - 1$ and $s = s + 1$.

2) Use Proposition 3.2 obtaining $X = x_1 x_1^{\mathrm{H}} + x_2 x_2^{\mathrm{H}}$ such that $A_1 \bullet x_i x_i^{\mathrm{H}} = A_1 \bullet X/r, A_2 \bullet x_i x_i^{\mathrm{H}} = A_2 \bullet X/r, i = 1, 2$. Return $\hat{x}_{r-1} = x_1$ and $\hat{x}_r = x_2$.

We test the numerical performance of **Algorithm 1** as follows. We generate 300 trials. At each trial, data A_1, A_2, A_3 and X are randomly generated, with matrix size `n=6+floor(30*rand)` and X's rank `r=3+floor((n-1)*rand)` (where '`rand`' is uniformly drawn from $[0, 1]$). The performance of each trial is measured by the error mean; that is, the average of the following $3r - 2$ terms:

$$\left\{\left|x_i^{\mathrm{H}} A_1 x_i - \frac{A_1 \bullet X}{r}\right|, \left|x_i^{\mathrm{H}} A_2 x_i - \frac{A_2 \bullet X}{r}\right|, 1 \leq i \leq r;\ \left|x_i^{\mathrm{H}} A_3 x_i - \frac{A_3 \bullet X}{r}\right|, 1 \leq i \leq r-2\right\}.$$

In Figure 1, we plot the performance of the 300 trials, and the red line in the figure is the mean of the performance measure. As we see, the algorithm performed remarkably stable and accurate. Note that the accuracy was set to be 10^{-8} in our numerical tests.

Likewise, we implement the matrix decomposition procedures as described by Theorem 2.2.

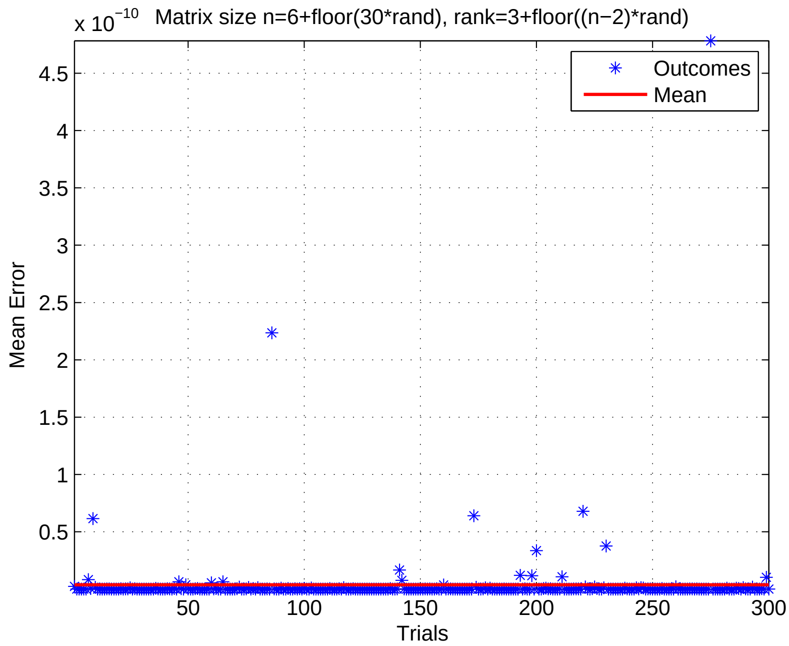

Figure 1: The performance of Algorithm 1 for 300 trial runs; the given accuracy is $\epsilon = 10^{-8}$.

Algorithm 2.1: Computing the decomposition as in Theorem 2.2 ($r \geq 3$)

Input: $A_1, A_2, A_3 \in \mathcal{H}^n$, and $X \in \mathcal{H}^n_+$ with $r = \mathrm{rank}\,(X) \geq 3$.
Output: Vector y such that $A_i \bullet yy^{\mathrm{H}} = A_i \bullet X, i = 1, 2, 3$, $X - \frac{1}{r} yy^{\mathrm{H}} \succeq 0$, and $\mathrm{rank}\,(X - \frac{1}{r} yy^{\mathrm{H}}) \leq r - 1$.

0) Let $\delta_i := A_i \bullet X, i = 1, 2, 3$.

1) Call Algorithm 1 to obtain $X = \sum_{i=1}^{r} \hat{x}_i \hat{x}_i^{\mathrm{H}}$ such that $A_1 \bullet \hat{x}_i \hat{x}_i^{\mathrm{H}} = \delta_1/r$, $A_2 \bullet \hat{x}_i \hat{x}_i^{\mathrm{H}} = \delta_2/r$, $i = 1, ..., r$; and $A_3 \bullet \hat{x}_i \hat{x}_i^{\mathrm{H}} = \delta_3/r$, $i = 1, ..., r-2$.

2) Return $y = \sqrt{r}\hat{x}_1$.

Algorithm 2.2: Computing the decomposition as in Theorem 2.2 ($r = 2$)

Input: $A_1, A_2, A_3 \in \mathcal{H}^n$ with $n \geq 3$, $X \in \mathcal{H}^n_+$ with $r = \text{rank}\,(X) = 2$, and a nonzero vector $z \notin \text{Range}\,(X)$.
Output: Vector y in the linear space spanned by z and Range (X), such that $A_i \bullet yy^{\text{H}} = A_i \bullet X, i = 1, 2, 3$, $X + zz^{\text{H}} - \frac{1}{r}yy^{\text{H}} \succeq 0$, and rank $(X + zz^{\text{H}} - \frac{1}{r}yy^{\text{H}}) \leq 2$.

0) Let $\delta_i := A_i \bullet X, i = 1, 2, 3$.

1) Use Proposition 3.2 to obtain $X = x_1x_1^{\text{H}} + x_2x_2^{\text{H}}$, such that $A_1 \bullet x_ix_i^{\text{H}} = \delta_1/2; A_2 \bullet x_ix_i^{\text{H}} = \delta_2/2, i = 1, 2$.

2) If $A_3 \bullet x_1x_1^{\text{H}} = \delta_3/2$, then set $y = \sqrt{r}x_1$ and terminate; otherwise, go to 3).

3) Set $x_3 = z$.

4) Determine $\alpha_1, \alpha_2, \alpha_3 \in \mathbf{C}$ such that

$$\left(\frac{\alpha_1x_1 + \alpha_2x_2 + \alpha_3x_3}{\sqrt{|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2}}\right)^{\text{H}} A_i \frac{\alpha_1x_1 + \alpha_2x_2 + \alpha_3x_3}{\sqrt{|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2}} = \frac{\delta_i}{2}, i = 1, 2, 3,$$

by solving the system of equations (3.2).

5) Return $y = \sqrt{r}\frac{\alpha_1x_1+\alpha_2x_2+\alpha_3x_3}{\sqrt{|\alpha_1|^2+|\alpha_2|^2+|\alpha_3|^2}}$.

We report numerical implementation of **Algorithm 2.2** since for rank $(X) \geq 3$, the algorithm can be considered as a special case of **Algorithm 1**. We execute **Algorithm 2.2** for 300 trial runs. At each run, data matrices A_1, A_2, A_3 and X are randomly generated, with the size `n=6+floor(30*rand)` and X's rank `r=2`. The performance of each trial run is measured by the error mean; that is, the average of the following three terms:

$$\left\{\left|y^{\text{H}}A_1y - A_1 \bullet X\right|, \left|y^{\text{H}}A_2y - A_2 \bullet X\right|, \left|y^{\text{H}}A_3y - A_3 \bullet X\right|\right\}.$$

Figure 2 summarizes the performance of these 300 trial runs, where the red line in the figure is the mean of the performance measure.

Finally, we implement the matrix decomposition procedure as described in Theorem 2.3.

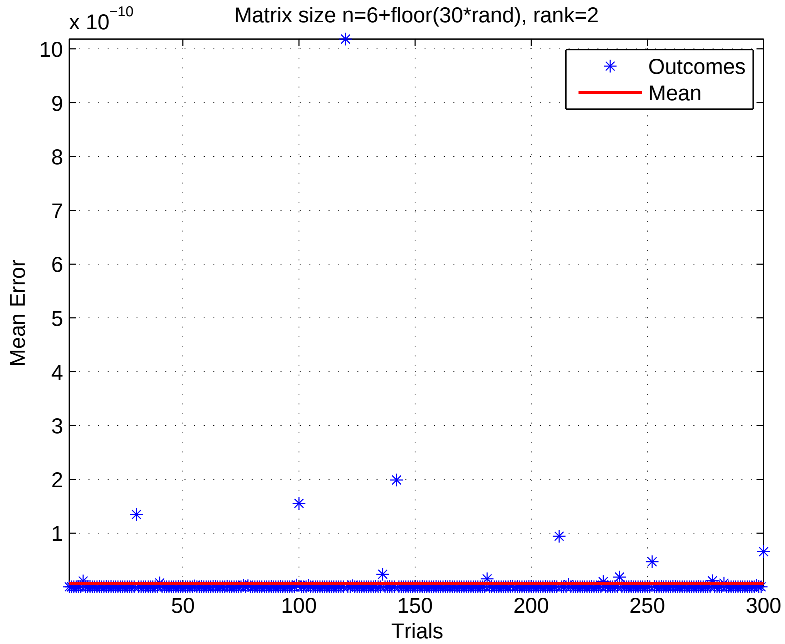

Figure 2: Performance of Algorithm 2.2 for 300 trial runs; the given accuracy is $\epsilon = 10^{-8}$.

Algorithm 3: Computing the decomposition as in Theorem 2.3

Input: $X \in \mathcal{H}_+^n$ with $\operatorname{rank}(X) \geq 2$ and $n \geq 3$, and $A_1, A_2, A_3, A_4 \in \mathcal{H}^n$ such that $(A_1 \bullet Y, A_2 \bullet Y, A_3 \bullet Y, A_4 \bullet Y) \neq (0,0,0,0)$ for all nonzero $Y \succeq 0$.
Output: Vector y such that $A_i \bullet yy^{\mathrm{H}} = A_i \bullet X, i = 1,2,3,4$.

0) Let $\delta_i := A_i \bullet X, i = 1,2,3,4$.

1) Pick an $i_0 \in \{1,2,3,4\}$ such that $\delta_{i_0} \neq 0$, say $i_0 = 4$.

2) Apply Theorem 2.2 to $A_i - \frac{\delta_i}{\delta_4} A_4, i = 1,2,3,$ and X, obtaining a vector y such that $(A_i - \frac{\delta_i}{\delta_4} A_4) \bullet yy^{\mathrm{H}} = 0,\ i = 1,2,3.$

3) Set $t = A_4 \bullet yy^{\mathrm{H}} / \delta_4$.

4) Update and return $y = \frac{1}{\sqrt{t}} y$.

We report the numerical performance of **Algorithm 3**. Like before, we apply **Algorithm 3** for 300 trial runs, and at each trial run the data matrices A_1, A_2, A_3 and X are randomly generated, with the matrix size `n=6+floor(30*rand)` and X's rank `r=2`. Specially, data A_4 is randomly generated such that it is positive or negative definite (by this, A_1, A_2, A_3, A_4 are legal inputs for the theorem).

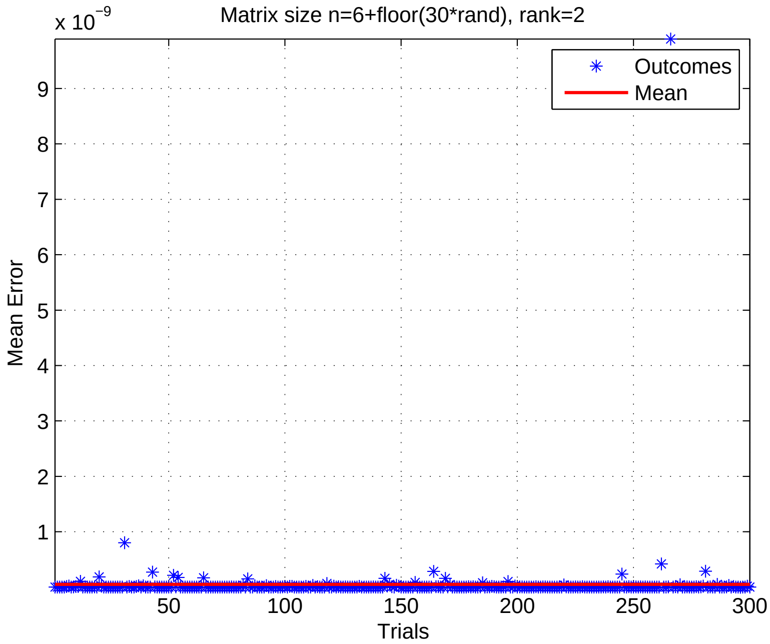

Figure 3: Performance of Algorithm 3 for 300 trial runs; the given accuracy is $\epsilon = 10^{-8}$.

The performance of each trial run is measured by the error mean; that is, the average of the following four terms:

$$\left\{ \left| y^{\mathrm{H}} A_1 y - A_1 \bullet X \right|, \left| y^{\mathrm{H}} A_2 y - A_2 \bullet X \right|, \left| y^{\mathrm{H}} A_3 y - A_3 \bullet X \right|, \left| y^{\mathrm{H}} A_4 y - A_4 \bullet X \right| \right\}.$$

Figure 3 plots the performance of these 300 trial runs, and the red line in the figure is the mean of the performance measure.

References

[1] W. Ai, Y.W. Huang and S. Zhang. On the low rank solutions for linear matrix inequalities. *Mathematics of Operations Research*, 33(4): 965 - 975, 2008.

[2] W. Ai and S. Zhang. Strong duality for the CDT subproblem: A necessary and sufficient condition. *SIAM Journal on Optimization*, 19(4): 1735 - 1756, 2009.

[3] Y.H. Au-Yeung. A theorem on a mapping from a sphere to the circle and the simultaneous diagonalization of two Hermitian matrices, *Proceedings of the American Mathematical Society*, 20(2): 545 - 548, 1969.

[4] Y.H. Au-Yeung and Y.T. Poon. A remark on the convexity and positive definiteness concerning Hermitian matrices. *Southeast Asian Bulletin of Mathematics*, 3: 85 - 92, 1979.

[5] A. Beck. Quadratic matrix programming. *SIAM Journal on Optimization*, 17(4): 1224 - 1238, 2007.

[6] A. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete and Computational Geometry*, 12: 189 - 202, 1995.

[7] A. Barvinok. A remark on the rank of positive semidefinite matrices subject to affine constraints. *Discrete and Computational Geometry*, 25: 23 - 31, 2001.

[8] D.P. Bertsekas, A. Nedić and A.E. Ozdaglar. *Convex analysis and optimization.* Athena Scientific, 2003

[9] F. Bohnenblust. Joint positiveness of matrices. *Unpublished Manuscript.*

[10] L. Brickman. On the field of values of a matrix. *Proceedings of the American Mathematical Society*, 12: 61 - 66, 1961.

[11] X. Chen and Y.X. Yuan. A note on quadratic forms. *Mathematical Programming*, 86:187 - 197, 1997.

[12] A. De Maio, S. De Nicola, Y.W. Huang, S. Zhang and A. Farina. Code design to optimize radar detection performance under accuracy and similarity constraints. *IEEE Transactions on Signal Processing*, 56(11): 5618 - 5629, 2008

[13] A. De Maio, S. De Nicola, Y.W. Huang, S. Zhang and A. Farina. Adaptive detection and steering estimation in the presence of useful signal and interference mismatches. *IEEE Transactions on Signal Processing*, 57(2): 436 - 450, 2009.

[14] A. De Maio, S. De Nicola, Y.W. Huang, D. P. Palomar, S. Zhang and A. Farina. Code Design for Radar STAP via Optimization Theory. *Technical Report SEEM2009-04*, Department of Systems Engineering & Engineering Management, The Chinese University of Hong Kong, 2009. Accepted for publication in *IEEE Transaction on Signal Processing.*

[15] F. Hausdorff. Der Wertvorrat einer Bilinearform, *Mathematische Zeitschrift*, 3: 314 - 316, 1919.

[16] J.-B. Hiriart-Urruty and M. Torki. Permanently going back and forth between the "quadratic world" and the "convexity world" in optimization. *Applied Mathematics Optimization*, 45: 169 - 184, 2002.

[17] R.A. Horn and C.R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, 1985.

[18] Y.W. Huang and S. Zhang. Complex matrix decomposition and quadratic programming. *Mathematics of Operations Research*, 32(3): 758 - 768, 2007.

[19] J. Li and P. Stoica. *Robust adaptive beamforming.* Wiley-Interscience, Hoboken, N.J., 2006.

[20] J. Li, P. Stoica and Z. Wang. Doubly constrained robust Capon beamformer. *IEEE Transactions on Signal Processing*, 52(9): 2407 - 2423, 2004.

[21] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of Operations Research*, 23(2): 339 - 358, 1998.

[22] B.T. Polyak. Convexity of quadratic transformations and its use in control and optimization. *Journal of Optimization Theory and Applications*, 99: 53 - 583, 1998.

[23] J.S. Pang and S. Zhang. The joint numerical range and quadratic optimization. *Unpublished uanuscript.* January 2004.

[24] I. Pólik and T. Terlaky. S-lemma: a survey. *SIAM Review*, 49(3): 371 - 418, 2007.

[25] J.F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, 28: 246 - 267, 2003.

[26] Y. Ye and S. Zhang. New results on quadratic minimization. *SIAM Journal on Optimization*, 14: 245 - 267, 2003.

[27] Y.X. Yuan. On a subproblem of trust region algorithms for constrained optimization. *Mathematical Programming*, 47:53 - 63, 1990.

A Proof of Lemma 3.1

Lemma A.1. *Let a_2, a_3 be given complex numbers, a_4 be a given real number. Then the following equation has a complex-valued solution:*

$$\text{Im }(\bar{x}y) + \text{Re }(a_2\bar{y}) + \text{Re }(a_3x) + a_4 = 0.$$

PROOF. Let $a_2 = u_2e^{\boldsymbol{i}\xi_2}, a_3 = u_3e^{\boldsymbol{i}\xi_3}$, and $x = r_1e^{\boldsymbol{i}\theta_1}, y = r_2e^{\boldsymbol{i}\theta_2}$. Then the equation can be rewritten equivalently as

$$r_1r_2\sin(\theta_2-\theta_1) + r_2u_2\cos(\xi_2-\theta_2) + r_1u_3\cos(\xi_3+\theta_1) + a_4 = 0.$$

If $a_4 = 0$, then $x = y = 0$ is a solution. Assume $a_4 < 0$ (the case $a_4 > 0$ is symmetric). If $u_2 = u_3 = 0$, then a solution is found by setting $r_1 = 1, \theta_1 = 0, r_2 = |a_4|, \theta_2 = \frac{\pi}{2}$. If $u_2 \neq 0$ or $u_3 \neq 0$, say $u_2 \neq 0$, then a solution is $r_1 = 0, \theta_1 = 0, r_2 = \frac{|a_4|}{u_2}, \theta_2 = \xi_2$. □

Lemma A.2. *Let a_2, a_3, b_2, b_3 be given complex numbers, a_4, b_4 be given real numbers. Then the following system of equations has a complex-valued solution:*

$$\text{Im }(\bar{x}y) + \text{Re }(a_2\bar{y}) + \text{Re }(a_3x) + a_4 = 0, \quad \text{(A.1)}$$

$$\text{Re }(\bar{x}y) + \text{Re }(b_2\bar{y}) + \text{Re }(b_3x) + b_4 = 0. \quad \text{(A.2)}$$

PROOF. If $a_4 \neq 0$, it follows from Lemma A.1 that there is a solution (x_0, y_0) solving Equation (A.1), then using variable transformation: $x = x' + x_0, y = y' + y_0$, we may turn the equations (A.1) and (A.2) to

$$\begin{aligned} \text{Im }(\bar{x'}y') + \text{Re }(a_2'\bar{y'}) + \text{Re }(a_3'x') &= 0, \\ \text{Re }(\bar{x'}y') + \text{Re }(b_2'\bar{y'}) + \text{Re }(b_3'x') + b_4' &= 0, \end{aligned}$$

where $a_2' = a_2 + \boldsymbol{i}x_0, a_3' = a_3 + \boldsymbol{i}\bar{y}_0, b_2' = b_2 + x_0, b_3' = b_3 + \bar{y}_0, b_4' = \text{Re }(\bar{x}_0y_0) + \text{Re }(b_2\bar{y}_0) + \text{Re }(b_3x_0) + b_4$. Hence, without losing generality we need only consider $a_4 = 0$. Let $b_4 \neq 0$ (otherwise $x = y = 0$ is a solution). Let $x = r_1e^{\boldsymbol{i}\theta_1}, y = r_2e^{\boldsymbol{i}\theta_2}, a_2 = u_2e^{\boldsymbol{i}\xi_2}, a_3 = u_3e^{\boldsymbol{i}\xi_3}, b_2 = v_2e^{\boldsymbol{i}\eta_2}, b_3 = v_3e^{\boldsymbol{i}\eta_3}$. Then the equations (A.1) and (A.2) become

$$r_1r_2\sin(\theta_2-\theta_1) + r_2u_2\cos(\xi_2-\theta_2) + r_1u_3\cos(\xi_3+\theta_1) = 0, \quad \text{(A.3)}$$

$$r_1r_2\cos(\theta_2-\theta_1) + r_2v_2\cos(\eta_2-\theta_2) + r_1v_3\cos(\eta_3+\theta_1) + b_4 = 0. \quad \text{(A.4)}$$

Assume $b_4 < 0$. We choose $\theta_2 = \theta_1$. (If $b_4 > 0$, then we choose $\theta_2 = \theta_1 + \pi$; the remaining discussion is parallel). Three cases are considered here:

Case 1. $u_2 = 0$ or $u_3 = 0$. Suppose $u_2 = 0$. Then it is easy to check that $r_1 = r_2$ and $\theta_2 = \theta_1 = \pi/2 - \xi_3$ solve equation (A.3), and equation (A.4) becomes

$$r_1^2 + r_1(v_2 \cos(\eta_2 - \theta_2) + v_3 \cos(\eta_3 + \theta_1)) + b_4 = 0,$$

which clearly has a positive solution since $b_4 < 0$. Denote this positive solution to be $\hat{r}$. Therefore a solution of (A.3) and (A.4) can be set as $\theta_2 = \theta_1 = \pi/2 - \xi_3$, and $r_1 = r_2 = \hat{r}$.

Case 2. $u_2 \neq 0$, $u_3 \neq 0$ and $\xi_2 + \xi_3 = k\pi$ for some integer k. Similar to *Case 1*, one verifies that $\theta_2 = \theta_1 = \pi/2 - \xi_3$ and $r_1 = r_2 = \hat{r}$ for some $\hat{r} > 0$ is a solution.

Case 3. $u_2 \neq 0$, $u_3 \neq 0$ and $\xi_2 + \xi_3 \neq k\pi$ for every integer k. Take $\theta_2 = \theta_1 = \frac{\pi}{2} + \frac{\xi_2 - \xi_3}{2}$. Since $\cos(\xi_2 - \theta_2) = -\cos(\xi_3 + \theta_1) = \sin\frac{\xi_2+\xi_3}{2} \neq 0$, equation (A.3) leads to $\frac{r_2}{r_1} = \frac{u_3}{u_2} =: \lambda > 0$. By this relation and (A.4), it follows that

$$\lambda r_1^2 + r_1(\lambda v_2 \cos(\eta_2 - \theta_2) + v_3 \cos(\eta_3 + \theta_1)) + b_4 = 0,$$

which has a positive solution since $b_4 < 0$. Denote this positive root to be $\hat{r}$. Clearly, $r_1 = \hat{r}, r_2 = \lambda\hat{r}$ and $\theta_2 = \theta_1 = \frac{\pi}{2} - \frac{\xi_3 - \xi_2}{2}$ is a solution in this case. □

Now let us recall a particular case of Lemma 1 from [3], which we will need to use.

Lemma A.3. *For any real numbers $a_2, a_3, a_4, b_2, b_3, b_4$, the following system of equations with real variables:*

$$\begin{aligned} xy + a_2x + a_3y + a_4 &= 0, \\ x^2 - y^2 + b_2x + b_3y + b_4 &= 0, \end{aligned}$$

has a solution (zero solution is allowed).

Notice that the proof of Lemma 1 of [3] is constructive, and indeed we can find a solution of the above equations easily.

Lemma A.4. *Let $a_2, a_3, b_2, b_3, c_2, c_3$ be any given complex numbers, and a_4, b_4, c_4 be any given real numbers. Then the system of equations has a complex-valued solution:*

$$\text{Im }(\bar{x}y) + \text{Re }(a_2\bar{y}) + \text{Re }(a_3x) + a_4 = 0, \tag{A.5}$$

$$\text{Re }(\bar{x}y) + \text{Re }(b_2\bar{y}) + \text{Re }(b_3x) + b_4 = 0, \tag{A.6}$$

$$|x|^2 - |y|^2 + \text{Re }(c_2\bar{y}) + \text{Re }(c_3x) + c_4 = 0. \tag{A.7}$$

Proof. We may assume $a_4 = b_4 = 0$, because, if $(a_4, b_4) \neq 0$, then by Lemma A.2 there exists a solution (x_0, y_0) to (A.5) and (A.6). Using a translation $x = x' + x_0$ and $y = y' + y_0$, the new zero-degree terms of equations (A.5) and (A.6), say a_4' and b_4', will be zero.

Let $a_i = u_i e^{\boldsymbol{i}\xi_i}, b_i = v_i e^{\boldsymbol{i}\eta_i}, c_i = w_i e^{\boldsymbol{i}\zeta_i}$, $i = 2, 3$, and $x = r_1 e^{\boldsymbol{i}\theta_1}, y = r_2 e^{\boldsymbol{i}\theta_2}$. Then the system of equations can be written as

$$\begin{aligned} r_1 r_2 \sin(\theta_2 - \theta_1) + r_2 u_2 \cos(\xi_2 - \theta_2) + r_1 u_3 \cos(\xi_3 + \theta_1) &= 0, & \text{(A.8)} \\ r_1 r_2 \cos(\theta_2 - \theta_1) + r_2 v_2 \cos(\eta_2 - \theta_2) + r_1 v_3 \cos(\eta_3 + \theta_1) &= 0, & \text{(A.9)} \\ (r_1^2 - r_2^2) + r_2 w_2 \cos(\zeta_2 - \theta_2) + r_1 w_3 \cos(\zeta_3 + \theta_1) + c_4 &= 0. & \text{(A.10)} \end{aligned}$$

We first wish to show that there is (θ_1, θ_2) such that

$$t := \frac{\sin(\theta_2 - \theta_1)}{\cos(\theta_2 - \theta_1)} = \frac{u_2 \cos(\xi_2 - \theta_2)}{v_2 \cos(\eta_2 - \theta_2)} = \frac{u_3 \cos(\xi_3 + \theta_1)}{v_3 \cos(\eta_3 + \theta_1)}, \tag{A.11}$$

where $t \in (-\infty, +\infty)$. If $t = \infty$, we proceed the proof by considering

$$s := 1/t = \frac{\cos(\theta_2 - \theta_1)}{\sin(\theta_2 - \theta_1)} = \frac{v_2 \cos(\eta_2 - \theta_2)}{u_2 \cos(\xi_2 - \theta_2)} = \frac{v_3 \cos(\eta_3 + \theta_1)}{u_3 \cos(\xi_3 + \theta_1)}. \tag{A.12}$$

Solving (A.11) yields the following equations:

$$\begin{aligned} \tan(\theta_2 - \theta_1) &= t, & \text{(A.13)} \\ \tan\theta_2 &= \frac{t v_2 \cos\eta_2 - u_2 \cos\xi_2}{u_2 \sin\xi_2 - t v_2 \sin\eta_2}, & \text{(A.14)} \\ \tan\theta_1 &= \frac{t v_3 \cos\eta_3 - u_3 \cos\xi_3}{t v_3 \sin\eta_3 - u_3 \sin\xi_3}. & \text{(A.15)} \end{aligned}$$

Since $\tan(\theta_2 - \theta_1) = (\tan\theta_2 - \tan\theta_1)/(1 + \tan\theta_2 \tan\theta_1)$, we substitute (A.14) and (A.15) into (A.13), and get the following after some arrangements:

$$t^3 v_2 v_3 \cos(\eta_2 + \eta_3) + t^2 p_2 + t p_1 - u_2 u_3 \sin(\xi_2 + \xi_3) = 0, \tag{A.16}$$

where $p_2 = -u_2 v_3 \cos(\xi_2 + \eta_3) - v_2 u_3 \cos(\eta_2 + \xi_3) - v_2 v_3 \sin(\eta_2 + \eta_3)$ and $p_1 = u_2 u_3 \cos(\xi_2 + \xi_3) + u_2 v_3 \sin(\xi_2 + \eta_3) + v_2 u_3 \sin(\eta_2 + \xi_3)$. In order to investigate the solution of equation (A.16), we proceed by the following three cases:

Case 1. $v_2 v_3 \cos(\eta_2 + \eta_3) \neq 0$. Then equation (A.16) has a real solution $t_0 \in (-\infty, +\infty)$ (including zero), thus (θ_2^0, θ_1^0) are obtained from (A.14) and (A.15). It follows that the equations (A.8), (A.9) and (A.10) reduce to

$$\begin{cases} r_1 r_2 \cos(\theta_2^0 - \theta_1^0) + r_2 v_2 \cos(\eta_2 - \theta_2^0) + r_1 v_3 \cos(\eta_3 + \theta_1^0) = 0, \\ (r_1^2 - r_2^2) + r_2 w_2 \cos(\zeta_2 - \theta_2^0) + r_1 w_3 \cos(\zeta_3 + \theta_1^0) + c_4 = 0. \end{cases} \tag{A.17}$$

Since $t_0 \neq \infty$, then $\cos(\theta_2^0 - \theta_1^0) \neq 0$. From Lemma A.3, we conclude that the system (A.17) has always a solution, say (r_1^0, r_2^0).

Observe that $(\theta_1^0 + k_1\pi, \theta_2^0 + k_2\pi, t_0)$ is also a solution of (A.13), (A.14) and (A.15) for any integers k_1, k_2. Then we can adjust the signs of r_1^0 and r_2^0 by selecting appropriate k_1, k_2 such that

$r_1^0 \geq 0, r_2^0 \geq 0$ (for example, if $r_1^0 < 0$, we may replace (r_1^0, θ_1^0) with $(-r_1^0, \theta_1^0 + \pi)$ in the system (A.17)). Hence a solution $(r_1 e^{\boldsymbol{i}\theta_1^0}, r_2 e^{\boldsymbol{i}\theta_2^0})$ for the system of equations (A.5), (A.6) and (A.7) is found.

Case 2. $v_2 v_3 \cos(\eta_2 + \eta_3) = 0, u_2 u_3 \sin(\xi_2 + \xi_3) \neq 0$. Instead of (A.11), we consider (A.12), which yield

$$-s^3 u_2 u_3 \sin(\xi_2 + \xi_3) + v_2 v_3 \cos(\eta_2 + \eta_3) + s^2 p_1 + s p_2 = 0, \tag{A.18}$$

where p_1, p_2 are the same as those in (A.16). Obviously $s = 0$ is a solution to (A.18), and the remaining argument follows *Case 1.*

Case 3. $v_2 v_3 \cos(\eta_2 + \eta_3) = 0, u_2 u_3 \sin(\xi_2 + \xi_3) = 0$. In the current case, note that $t = 0$ is a solution of (A.16), and then the remaining discussion is similar to *Case 1.* □

Now we are ready to prove Lemma 3.1.

Proof of Lemma 3.1. Redefine

$$\begin{array}{lll} x := \sqrt{c_{-1}}\, x, & y := \sqrt{c_0}\, y, & \\ a_1 := \frac{a_1}{\sqrt{c_{-1} c_0}}, & b_1 := \frac{b_1}{\sqrt{c_{-1} c_0}}, & c_1 := \frac{c_1}{\sqrt{c_{-1} c_0}}, \\ a_2 := \frac{a_2}{\sqrt{c_{-1}}}, & b_2 := \frac{b_2}{\sqrt{c_{-1}}}, & b_2 := \frac{b_2}{\sqrt{c_{-1}}}, \\ a_3 := \frac{a_3}{\sqrt{c_0}}, & b_3 := \frac{b_3}{\sqrt{c_0}}, & c_3 := \frac{c_3}{\sqrt{c_0}}. \end{array}$$

Then the above system becomes

$$\begin{array}{rcl} \mathrm{Re}\,(a_1 \bar{x} y) + \mathrm{Re}\,(a_2 \bar{x} z) + \mathrm{Re}\,(a_3 \bar{y} z) + a_4 |z|^2 & = & 0, \\ \mathrm{Re}\,(b_1 \bar{x} y) + \mathrm{Re}\,(b_2 \bar{x} z) + \mathrm{Re}\,(b_3 \bar{y} z) + b_4 |z|^2 & = & 0, \\ |x|^2 - |y|^2 + \mathrm{Re}\,(c_1 \bar{x} y) + \mathrm{Re}\,(c_2 \bar{x} z) + \mathrm{Re}\,(c_3 \bar{y} z) + c_4 |z|^2 & = & 0. \end{array} \tag{A.19}$$

Since $\mathrm{Re}\,(a_1 \bar{x} y) = \mathrm{Re}\,(a_1) \mathrm{Re}\,(\bar{x} y) - \mathrm{Im}\,(a_1) \mathrm{Im}\,(\bar{x} y)$ (similarly for $\mathrm{Re}\,(b_1 \bar{x} y)$ and $\mathrm{Re}\,(c_1 \bar{x} y)$), the system (A.19) can be rewritten in matrix form as:

$$A \begin{pmatrix} \mathrm{Im}\,(\bar{x} y) \\ \mathrm{Re}\,(\bar{x} y) \\ |x|^2 - |y|^2 \end{pmatrix} + \begin{pmatrix} \mathrm{Re}\,(a_2 \bar{x} z) + \mathrm{Re}\,(a_3 \bar{y} z) + a_4 |z|^2 \\ \mathrm{Re}\,(b_2 \bar{x} z) + \mathrm{Re}\,(b_3 \bar{y} z) + b_4 |z|^2 \\ \mathrm{Re}\,(c_2 \bar{x} z) + \mathrm{Re}\,(c_3 \bar{y} z) + c_4 |z|^2 \end{pmatrix} = 0, \tag{A.20}$$

where

$$A = \begin{bmatrix} -\mathrm{Im}\, a_1 & \mathrm{Re}\, a_1 & 0 \\ -\mathrm{Im}\, b_1 & \mathrm{Re}\, b_1 & 0 \\ -\mathrm{Im}\, c_1 & \mathrm{Re}\, c_1 & 1 \end{bmatrix}.$$

To complete the analysis, we consider the following two cases, exclusively.

Case 1. $\det A \neq 0$.

Setting $z = 1$ and left-multiplying A^{-1} on both sides of the equations (A.20) yields

$$\begin{cases} \text{Im } (\bar{x}y) + \text{Re } (a'_2\bar{y}) + \text{Re } (a'_3 y) + a'_4 = 0, \\ \text{Re } (\bar{x}y) + \text{Re } (b'_2\bar{y}) + \text{Re } (b'_3 y) + b'_4 = 0, \\ |x|^2 - |y|^2 + \text{Re } (c'_2\bar{y}) + \text{Re } (c'_3 y) + c'_4 = 0, \end{cases} \tag{A.21}$$

where

$$\begin{bmatrix} a'_2 & a'_3 & a'_4 \\ b'_2 & b'_3 & b'_4 \\ c'_2 & c'_3 & c'_4 \end{bmatrix} = A^{-1} \begin{bmatrix} a_2 & a_3 & a_4 \\ b_2 & b_3 & b_4 \\ c_2 & c_3 & c_4 \end{bmatrix}.$$

Now, (A.21) has a solution due to Lemma A.4, which yields a nonzero solution for (A.20), and consequently a nonzero solution as well for (A.19).

Case 2. $\det A = 0$.

We set $z = 0$. Then, (A.20) reduces to

$$A \begin{bmatrix} \text{Im } (\bar{x}y) \\ \text{Re } (\bar{x}y) \\ |x|^2 - |y|^2 \end{bmatrix} = 0,$$

which has a nonzero real solution (s_1, s_2, s_3); that is

$$\text{Im } (\bar{x}y) = s_1, \quad \text{Re } (\bar{x}y) = s_2, \quad |x|^2 - |y|^2 = s_3. \tag{A.22}$$

The relation (A.22) gives $|x|^2|y|^2 = s_1^2 + s_2^2$, $|x|^2 - |y|^2 = s_3$, from which one can always assemble a solution $(x, y) \neq (0, 0)$ to satisfy (A.19). □