Upper Bounds of Rates of Complex Orthogonal Space-Time Block Codes*

Haiquan Wang   Xiang-Gen Xia

Abstract

In this paper, we derive some upper bounds of the rates of (generalized) complex orthogonal space-time block codes. We first present some new properties of complex orthogonal designs and then show that the rates of complex orthogonal space-time block codes for more than 2 transmit antennas are upper bounded by 3/4. We show that the rates of generalized complex orthogonal space-time block codes for more than 2 transmit antennas are upper bounded by 4/5, where the norms of column vectors may not be necessarily the same. We also present another upper bound under a certain condition.

For a (generalized) complex orthogonal design, its variables are not restricted to any alphabet sets but are on the whole complex plane. In this paper, a (generalized) complex orthogonal design with variables over some alphabet sets on the complex plane is also considered. We obtain a condition on the alphabet sets such that a (generalized) complex orthogonal design with variables over these alphabet sets is also a conventional (generalized) complex orthogonal design and therefore the above upper bounds on its rate also hold. We show that commonly used QAM constellations of sizes above 4 satisfy this condition.

Keywords: complex orthogonal designs, complex orthogonal space-time block codes, Hurwitz-Radon theory, Hurwitz family, Hermitian compositions of quadratic forms.

1 Introduction

The first real/complex orthogonal space-time block code was proposed by Alamouti [1] for two transmit antennas. It was then generalized to real/complex orthogonal space-time block codes for more than two transmit antennas by Tarokh, Jafarkhani and Calderbank [3]. There

*The authors are with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716. Email: {hwang, xxia}@ee.udel.edu; Phone: (302)831-8038. Their work was supported in part by the Air Force Office of Scientific Research (AFOSR) under Grant No. F49620-02-1-0157 and the National Science Foundation (NSF) under Grants MIP-9703377 and CCR-0097240. The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Lausanne, Switzerland, July 2002.

are two important properties of real/complex orthogonal space-time block codes: (i) they have fast maximum-likelihood (ML) decoding, namely symbol-by-symbol decoding; (ii) they have the full diversity. These two properties make real/complex orthogonal space-time block codes attractive in space-time code designs. By utilizing the Hurwitz-Radon theory [16, 17, 18, 22, 25], Tarokh, Jafarkhani and Calderbank [3] provided a systematic method to construct *real* orthogonal space-time block codes of size $p \times n$ and rate 1 for k PAM symbols, where n is the number of transmit antennas, p is the time delay (or block size), and $R = k/p = 1$ is the code rate. They also provided a construction of rate 1/2 *complex* orthogonal space-time block codes for PSK and QAM symbols using real orthogonal space-time block codes of rate 1. In order to maintain the fast ML decoding and the full diversity of a space-time block code, the orthonormality in the sense that the norms of all column vectors are the same can be relaxed to a general orthogonality where the norms of column vectors may not be necessarily the same [3]. A complex orthogonal space-time block code with the generalized orthonormality is called a generalized complex orthogonal space-time block code. In [2, 3], it has been shown that the rate $R \leq 1$ for both real and complex orthogonal space-time block codes for any number of transmit antennas. While the maximal rate 1, i.e., $R = 1$, is reachable for real orthogonal space-time block codes as we mentioned above from the Hurwitz-Radon's constructive theory, it has been recently shown in [8] that $k \leq p - 1$ when $n > 2$, i.e., $R < 1$ and $R = 1$ is not reachable for (generalized) complex orthogonal space-time block codes no matter what a time delay p is unless the number of transmit antennas is 2, i.e., the Alamouti's scheme. Notice that, if condition $p = n$ is required, i.e., *square* codes or *square* complex orthogonal designs, then $R < 1$ when $n > 2$ directly follows from the results on amicable designs [17, 20, 21, 22, 3, 5, 6, 7] that have small rates when $n \geq 8$. While both square and non-square *real* orthogonal designs (or compositions of quadratic forms) are well understood, not much is known for non-square *complex* orthogonal designs (or Hermitian compositions of quadratic forms [25]), [3, 25, 26].

In this paper, we derive some upper bounds on the rates R of (generalized) complex orthogonal space-time block codes (or complex orthogonal designs). We emphasize that the sizes of (generalized) complex orthogonal space-time block codes (or complex orthogonal designs) here are general and they may not be square, i.e., p may not be equal to n. We show that, when the number of transmit antennas is more than 2, i.e., $n > 2$, the rates of complex orthogonal space-time block codes are upper bounded by 3/4, i.e.,

$$R \leq \frac{3}{4},$$

and the rates of generalized complex orthogonal space-time block codes are upper bounded by 4/5, i.e.,

$$R \leq \frac{4}{5}.$$

Note that rate 3/4 complex orthogonal space-time block codes for 3 and 4 transmit antennas have appeared in [3, 4, 5, 6]. Therefore, the above upper bound tells us that these complex orthogonal space-time block codes have already reached the optimal rate. Also note that the above upper bound 3/4 on the rates is not new for *square* complex orthogonal designs. In fact, it has been shown and reviewed from amicable designs in [17, 20, 21, 22, 3, 5, 6, 7]. However, this upper bound is *new* for non-square complex orthogonal designs. In the meantime, it is known that to generate orthogonal space-time codes, a square orthogonal design is not necessary [3].

In a conventional (generalized) complex orthogonal design, its variables may take any values in the complex plane. However, as we shall see later, to generate a space-time code, the variables only take values in some finite subsets, called alphabet sets, on the complex plane. The question then becomes whether it is helpful to produce more (generalized) complex orthogonal designs of high rates when their variables are restricted to some alphabet sets. This question has been partially studied lately in [12, 13, 15, 14]. For square real orthogonal designs, when their variables are restricted to finite or infinite subsets of the real line (or field), they are called restricted orthogonal designs in [12] and pseudo orthogonal designs in [13]. It is shown in [12, 13] that there does not exist new square real orthogonal designs even when their variables are restricted to subsets of the real line, if the number of the elements of the alphabet set is greater than 2. For square complex orthogonal designs, it is known that (also as mentioned previously) the maximal rate of 4×4 complex orthogonal designs is

3/4 when all the variables can take any values on the complex plane. However, examples of rate 1 complex orthogonal designs of size 4×4 have been shown in [15] when their variables take some alphabet sets on the complex plane, where in the examples all the alphabet sets are PAM and a rotation of PAM, i.e., all points in an alphabet set are collinear. In this paper, we also consider this problem. We obtain a condition on the alphabet sets such that a (generalized) complex orthogonal design with variables over these alphabet sets is also a conventional (generalized) complex orthogonal design and therefore the above upper bounds on its rate also hold. We show that commonly used QAM signal constellations of size above 4 do satisfy this condition and therefore, a (generalized) complex orthogonal design with their variables over QAM constellations of size above 4 is also a conventional (generalized) complex orthogonal design. For convenience, in what follows we adopt the name "*restricted* (generalized) complex orthogonal design" as used in [12] for real orthogonal designs, when their variables are restricted to some alphabet sets.

This paper is organized as follows. In Section 2, we provide some preparations and new properties on (generalized) complex orthogonal designs. In Section 3, we prove several upper bounds. In Section 4, we study restricted (generalized) complex orthogonal designs. Some necessary lemmas (including a new form shown in Lemma 6 of singular value decompositions of a family of special matrices) and their proofs are put in Appendix.

2 Some Preliminaries and New Properties on Complex Orthogonal Designs

In this section, we present some properties of a (generalized) complex orthogonal design used in a (generalized) complex orthogonal space-time block code. In what follows, $\mathbb{C}$ denotes the field of all complex numbers and $\mathbb{R}$ denotes the field of all real numbers. For convenience, symbol 0 means scalar 0 or all 0 matrices of possibly different sizes and I means the identity matrices of possibly different sizes unless specified otherwise. For two matrices A and B of same number of rows, $(A\ B)$ denotes the concatenation matrix of A and B, i.e., $(A\ B)$ is a new matrix with the columns of A as its first part columns and the columns of B as its second part columns.

A *complex orthogonal design* $\mathcal{G}(x_1, x_2, \cdots, x_k)$ of size $p \times n$ is a $p \times n$ matrix satisfying the following conditions:

- the entries of $\mathcal{G}(x_1, x_2, \cdots, x_k)$ are complex linear combinations of $x_1, x_2, \cdots, x_k$ and their complex conjugates $x_1^*, x_2^*, \cdots, x_k^*$;
- the orthonormality $(\mathcal{G}(x_1, x_2, \cdots, x_k))^H \mathcal{G}(x_1, x_2, \cdots, x_k) = (|x_1|^2 + |x_2|^2 + \cdots + |x_k|^2)I$ holds for *any* complex values x_i, $i = 1, 2, \cdots, k$, where ^{H} stands for the complex conjugate transpose and I is the $n \times n$ identity matrix.

The orthonormality in the above definition can be generalized to the orthogonality as follows for preserving the full-diversity and the fast ML decoding [3].

A *generalized complex orthogonal design* $\mathcal{G}(x_1, x_2, \cdots, x_k)$ of size $p \times n$ is a $p \times n$ matrix satisfying the following conditions:

- the entries of $\mathcal{G}(x_1, x_2, \cdots, x_k)$ are complex linear combinations of $x_1, x_2, \cdots, x_k$ and their complex conjugates $x_1^*, x_2^*, \cdots, x_k^*$;
- the orthogonality $(\mathcal{G}(x_1, x_2, \cdots, x_k))^H \mathcal{G}(x_1, x_2, \cdots, x_k) = (|x_1|^2 D_1 + |x_2|^2 D_2 + \cdots + |x_k|^2 D_k)$ holds for *any* complex values x_i, $i = 1, 2, \cdots, k$, where D_i, $i = 1, 2, \cdots, k$, are $n \times n$ diagonal positive definite constant matrices, i.e., their diagonal elements are all positive constants.

Let $\mathcal{A}$ denote a signal constellation alphabet set and $\mathcal{C} = \{\mathcal{G}(x_1, x_2, \cdots, x_k) : \ x_i \in \mathcal{A}\}$. Then, $\mathcal{C}$ is called a complex (or generalized) orthogonal space-time block code. For this block code, every p time slots carries k information symbols, $x_1, x_2, \cdots, x_k$. The *rate* of this complex orthogonal space-time (or generalize complex orthogonal space-time) block code is defined as k/p and denoted by R, i.e., $R = k/p$. Without any confusion in understanding, in what follows we use complex orthogonal space-time (or generalize orthogonal complex space-time) block code $\mathcal{C}$ and (generalized) complex orthogonal design $\mathcal{G}(x_1, x_2, \cdots, x_k)$ interchangeably.

For a real orthogonal design, x_i are real-valued in the above definition and the coefficients in the linear combinations of x_i of components of $\mathcal{G}(x_1, x_2, \cdots, x_k)$ are all real. It is known that there exist real orthogonal designs with $R = 1$ for any number, n, of transmit antennas, see [18, 22, 25, 3]. We refer the reader to [1, 3] for the properties of the fast ML decoding and the full diversity of a complex orthogonal space-time (or generalized complex orthogonal space-time) block code, where the full diversity means that any difference matrix of two different complex orthogonal space-time (or generalized complex orthogonal space-time) block codewords (or code matrices) has full rank. The main goal of this paper is to show: (i) if $\mathcal{G} = \mathcal{G}(x_1, x_2, \cdots, x_k)$ of size $p \times n$ is a complex orthogonal design and $n \geq 3$, then its rate $R = k/p \leq 3/4$; (ii) if $\mathcal{G} = \mathcal{G}(x_1, x_2, \cdots, x_k)$ of size $p \times n$ is a generalized complex orthogonal design and $n \geq 3$, then its rate $R = k/p \leq 4/5$. To do so, let us have some preparations.

Let $\mathcal{G} = \mathcal{G}(x_1, x_2, \cdots, x_k)$ be a matrix of size $p \times n$, where its entries are the complex linear combinations of $x_1, x_2, \cdots, x_k$ and their complex conjugates $x_1^*, x_2^*, \cdots, x_k^*$. Then, $\mathcal{G}$ can be expressed in terms of its column vectors as follows:

$$\mathcal{G} = (A_1\mathbf{x} + B_1\mathbf{x}^* \;\; A_2\mathbf{x} + B_2\mathbf{x}^* \;\; \cdots \;\; A_n\mathbf{x} + B_n\mathbf{x}^*), \tag{1}$$

where A_i, B_i, $i = 1, \cdots, n$, are $p \times k$ constant complex matrices, $\mathbf{x} = (x_1, \cdots, x_k)^t$, and ^{t} stands for the transpose, and * stands for the complex conjugate.

For the $n \times n$ diagonal matrices D_i given in the above definition of a generalized complex orthogonal design, we denote $D_i = \text{diag}(d_1^i, d_2^i, \cdots, d_n^i)$. For each j, $j = 1, \cdots, n$, all the (j, j)-entries d_j^i of matrices D_i, $i = 1, \cdots, k$, form a new $k \times k$ diagonal matrix E_j as follows:

$$E_j \triangleq \text{diag}(d_j^1, d_j^2, \cdots, d_j^k). \tag{2}$$

Clearly, when all D_i are positive definite, all E_j are positive definite. Using these matrices, we can transfer the orthogonal condition on $\mathcal{G}$ into the conditions on the matrices A_i, B_j, $1 \leq i, j \leq n$.

Proposition 1 *Matrix $\mathcal{G}$ in (1) is a generalized complex orthogonal design, i.e.,*

$$\mathcal{G}^H\mathcal{G} = |x_1|^2 D_1 + |x_2|^2 D_2 + \cdots + |x_k|^2 D_k,$$

for some $n \times n$ diagonal positive definite constant matrices D_i, $1 \leq i \leq k$, if and only if there exist diagonal positive definite matrices E_i, $i = 1, 2, \cdots, n$, such that its associated matrices

A_i and B_i, $i = 1, \cdots, n$, in (1) satisfy the following conditions:

$$A_i^H A_j + B_j^t B_i^* = \delta_{ij} E_i, \quad A_i^H B_j + B_j^t A_i^* = 0, \quad B_i^H A_j + A_j^t B_i^* = 0, \tag{3}$$

or equivalently,

$$\begin{pmatrix} A_i & B_i \\ B_j^* & A_j^* \end{pmatrix}^H \begin{pmatrix} A_j & B_j \\ B_i^* & A_i^* \end{pmatrix} = \delta_{ij} \begin{pmatrix} E_i & 0 \\ 0 & E_j \end{pmatrix}, \tag{4}$$

for all $i, j = 1, \cdots, n$, where $\delta_{ij} = 1$ when $i = j$ and $\delta_{ij} = 0$ when $i \neq j$.

In particular, $\mathcal{G}$ is a complex orthogonal design if and only if (3) or (4) holds for $E_i = I$ for $1 \leq i \leq n$.

Proof: By the orthogonality of a generalized complex orthogonal design in terms of its column vectors, we have

$$(A_i \mathbf{x} + B_i \mathbf{x}^*)^H (A_j \mathbf{x} + B_j \mathbf{x}^*) = \mathbf{x}^H \delta_{ij} E_i \mathbf{x},$$

i.e.,

$$\mathbf{x}^H A_i^H A_j \mathbf{x} + \mathbf{x}^H A_i^H B_j \mathbf{x}^* + \mathbf{x}^t B_i^H A_j \mathbf{x} + \mathbf{x}^t B_i^H B_j \mathbf{x}^* = \mathbf{x}^H \delta_{ij} E_i \mathbf{x},$$

where E_j are from D_i as in (2) and therefore they are positive definite. Note that

$$\mathbf{x}^t B_i^H B_j \mathbf{x}^* = (\mathbf{x}^t B_i^H B_j \mathbf{x}^*)^t = \mathbf{x}^H B_j^t B_i^* \mathbf{x},$$

the above equation can be rewritten as

$$\mathbf{x}^H (A_i^H A_j + B_j^t B_i^* - \delta_{ij} E_i) \mathbf{x} + \mathbf{x}^H A_i^H B_j \mathbf{x}^* + \mathbf{x}^t B_i^H A_j \mathbf{x} = 0, \quad \text{for any } \mathbf{x} \in \mathbb{C}^k.$$

By Lemma 1 in Appendix, we obtain $A_i^H A_j + B_j^t B_i^* = \delta_{ij} E_i$, $A_i^H B_j + (A_i^H B_j)^t = 0$ and $B_i^H A_j + (B_i^H A_j)^t = 0$.

The sufficiency part is easy to verify. **q.e.d.**

As a remark, Lemma 1 in Appendix was first obtained in [8]. For the completeness, its proof is given in Appendix. Another remark is that equation $A_i^H B_j + (A_i^H B_j)^t = 0$ holds is equivalent to matrix $A_i^H B_j$ is skew-symmetry[1], which are used interchangeably in what follows.

[1] A matrix $S = (s_{ij})$ is called skew-symmetric if $s_{ij} = -s_{ji}$. For a $k \times k$ skew-symmetric matrix $S = (s_{ij})$, we always have $\mathbf{x}^t S \mathbf{x} = 0$ for any $k \times 1$ vector $\mathbf{x} \in \mathbb{C}^k$.

We next investigate some properties of a generalized complex orthogonal design $\mathcal{G}$ under a unitary transformation. Let U be a unitary matrix and $\mathcal{G}(\mathbf{x})$ be a generalized complex orthogonal design, then $\mathcal{G}(U\mathbf{x})$ may not be a generalized complex orthogonal design due to the fact that $U^H E_i U$ may not be diagonal, i.e., unitary transform on variables x_i does not preserve a generalized complex orthogonal design. On the other hand, if $\mathcal{G}(\mathbf{x})$ is a complex orthogonal design, then $\mathcal{G}(U\mathbf{x})$ is also a complex orthogonal design due to $E_i = I$ and $U^H E_i U = I$, i.e., unitary transform on variables x_i preserves a complex orthogonal design.

In order to implement unitary transformations on variables of a generalized complex orthogonal design to simplify its corresponding matrices, we introduce the following concept of Hurwitz families, which is preserved by a unitary transformation as we can see below.

Definition 1 *A set of $p \times 2k$ matrices $\{(A_1\ B_1), (A_2\ B_2), \cdots, (A_n\ B_n)\}$ is called a Hurwitz family if there exist n positive definite matrices E_i, $i = 1, 2, \cdots, n$, such that*

$$A_i^H A_j + B_j^t B_i^* = \delta_{ij} E_i, \ \ 1 \leq i, j \leq n, \tag{5}$$

and

$$A_i^H B_j + B_j^t A_i^* = 0, \ \ B_i^H A_j + A_j^t B_i^* = 0, \ \ 1 \leq i \neq j \leq n. \tag{6}$$

In the above definition of a Hurwitz family, the diagonality of the matrices E_i is *not* required. Clearly, by Proposition 1, the matrices $\{(A_1\ B_1), (A_2\ B_2), \cdots, (A_n\ B_n)\}$ of a generalized complex orthogonal design $\mathcal{G}(\mathbf{x})$ form a Hurwitz family, and

$$\{(A_1 U\ B_1 U^*), (A_2 U\ B_2 U^*), \cdots, (A_n U\ B_n U^*)\}$$

of $\mathcal{G}(U\mathbf{x})$ for a unitary transform U also form a Hurwitz family.

Note that in (6), we have the restriction $i \neq j$ due to the fact that it can not be deduced for $i = j$ when E_i is not the identity matrix when a unitary transform is applied to a generalized complex orthogonal design as we shall see after the proof of Lemma 7 in Appendix. Thus, the condition for a Hurwitz family is weaker than the one for a generalized complex orthogonal design. Also note that the above definition coincides with the one in [22] when $B_i = 0$, A_i are real and $E_i = I$, i.e., the real case.

For a Hurwitz family $\{(A_1\ B_1), (A_2\ B_2), \cdots, (A_n\ B_n)\}$, by using some proper unitary transformations, we can diagonalize the first matrix $(A_1\ B_1)$ as follows, which plays a key role in the proof of our main theorem in next section.

Proposition 2 *Let*

$$\mathcal{G} = (A_1\mathbf{x} + B_1\mathbf{x}^*\ A_2\mathbf{x} + B_2\mathbf{x}^* \cdots A_n\mathbf{x} + B_n\mathbf{x}^*)$$

be a generalized complex orthogonal design. Then, there exists a Hurwitz family

$$\{(\tilde{A}_1\ \tilde{B}_1), (\tilde{A}_2\ \tilde{B}_2), \cdots, (\tilde{A}_n\ \tilde{B}_n)\} \tag{7}$$

with the same parameters, p, n, k, as $\mathcal{G}$ and

$$\tilde{A}_1^H\tilde{A}_1 + \tilde{B}_1^t\tilde{B}_1^* = I,\quad \tilde{A}_1^H\tilde{B}_1 + \tilde{B}_1^t\tilde{A}_1^* = 0,\quad \tilde{B}_1^H\tilde{A}_1 + \tilde{A}_1^t\tilde{B}_1^* = 0, \tag{8}$$

and furthermore, $\tilde{A}_1$ and $\tilde{B}_1$ have the following forms

$$\tilde{A}_1 = \begin{pmatrix} D_\lambda & 0 \\ 0 & I_{k-s} \\ 0 & 0 \\ 0 & 0 \end{pmatrix},\quad \tilde{B}_1 = \begin{pmatrix} 0_{s\times s} & 0 \\ 0_{(k-s)\times s} & 0 \\ D_\mu & 0 \\ 0 & 0 \end{pmatrix}, \tag{9}$$

where $k - s \geq 2k - p$, $D_\lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_s)$, $D_\mu = \mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_s)$ and $\lambda_i^2 + \mu_i^2 = 1$, $1 > \lambda_i \geq \sqrt{1/2} \geq \mu_j > 0$, $i, j = 1, 2, \cdots, s$, $k + s = \kappa$, and $\kappa \triangleq \mathrm{rank}((A\ B)) \geq k$. In particular, if $\mathcal{G}$ is a complex orthogonal design, then there exists a complex orthogonal design $\tilde{\mathcal{G}}$ with the same parameters, p, n, k, as $\mathcal{G}$ such that its corresponding matrices $\tilde{A}_1$ and $\tilde{B}_1$ have the forms in (9).

The proof of this proposition is put in Appendix.

In the proof of the main theorem in next section, we need the following rank inequalities.

1. (*Sylvester's inequality*) Let A be a $k \times p$ matrix and B be a $p \times n$ matrix. Then

$$\mathrm{rank}(A) + \mathrm{rank}(B) - p \leq \mathrm{rank}(AB).$$

2. Let A be a $k \times p$ matrix and B be a $p \times n$ matrix. Then

$$\mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}.$$

3. Let A, B be two $n \times m$ matrices and E be an $m \times m$ positive definite matrix. If $A^H A + B^H B = E$, then

$$\operatorname{rank}(A) + \operatorname{rank}(B) \geq m.$$

The above rank inequalities 1 and 2 are fundamental and can be found in linear algebra books, e.g., see [24]. Rank inequality 3 can be obtained from

$$\operatorname{rank}(A) + \operatorname{rank}(B) = \operatorname{rank}(A^H A) + \operatorname{rank}(B^H B) \geq \operatorname{rank}(A^H A + B^H B) = m.$$

3 Upper Bounds of Rates for Three or More Antennas

In this section, we present several upper bounds of the rates for both complex orthogonal designs and generalized complex orthogonal designs.

Theorem 1 *Let $\mathcal{G} = \mathcal{G}(x_1, x_2, \cdots, x_k)$ be a generalized complex orthogonal design of size $p \times n$. If $n \geq 3$, then, its rate is upper bounded by $4/5$, i.e.,*

$$R = \frac{k}{p} \leq \frac{4}{5}. \tag{10}$$

If $\mathcal{G}$ is a complex orthogonal design and $n \geq 3$, then its rate is upper bounded by $3/4$, i.e.,

$$R = \frac{k}{p} \leq \frac{3}{4}. \tag{11}$$

Proof: We first want to prove the first part of this theorem. Let $\mathcal{G} = (A_1\mathbf{x} + B_1\mathbf{x}^* \ A_2\mathbf{x} + B_2\mathbf{x}^* \ \cdots \ A_n\mathbf{x} + B_n\mathbf{x}^*)$. By Proposition 2, we can assume that A_1 and B_1 have the following forms:

$$A_1 = \begin{pmatrix} D_\lambda & 0 \\ 0 & I_{k-s} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \ B_1 = \begin{pmatrix} 0_{s\times s} & 0 \\ 0_{(k-s)\times s} & 0 \\ D_\mu & 0 \\ 0 & 0 \end{pmatrix},$$

where $D_\lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_s)$, $D_\mu = \operatorname{diag}(\mu_1, \cdots, \mu_s)$, I_{k-s} is the identity matrix of size $k-s$, and $1 > \lambda_i \geq \mu_j > 0$, $i, j = 1, \cdots, s$, and $\{(A_1 \ B_1), (A_2 \ B_2), \cdots, (A_n \ B_n)\}$ is a Hurwitz family with $A_1^H A_1 + B_1^t B_1^* = I$.

Divide $p \times k$ matrices A_i and B_i into block matrices as follows:

$$A_i = \begin{pmatrix} A_{i1} & A_{i2} \\ A_{i3} & A_{i4} \\ A_{i5} & A_{i6} \\ A_{i7} & A_{i8} \end{pmatrix}, \quad B_i = \begin{pmatrix} B_{i1} & B_{i2} \\ B_{i3} & B_{i4} \\ B_{i5} & B_{i6} \\ B_{i7} & B_{i8} \end{pmatrix},$$

where A_{i1} and B_{i1} are $s \times s$ matrices, A_{i3} and B_{i3} are $(k-s) \times s$ matrices, A_{i5} and B_{i5} are $s \times s$ matrices, A_{i7} and B_{i7} are $(p-k-s) \times s$ matrices, A_{i2} and B_{i2} are $s \times (k-s)$ matrices, A_{i4} and B_{i4} are $(k-s) \times (k-s)$ matrices, A_{i6} and B_{i6} are $s \times (k-s)$ matrices, and A_{i8} and B_{i8} are $(p-k-s) \times (k-s)$ matrices.

Since $A_1^H A_i + B_i^t B_1^* = 0$ for $i > 1$, we have

$$\begin{pmatrix} D_\lambda & 0 & 0 & 0 \\ 0 & I_{k-s} & 0 & 0 \end{pmatrix} \begin{pmatrix} A_{i1} & A_{i2} \\ A_{i3} & A_{i4} \\ A_{i5} & A_{i6} \\ A_{i7} & A_{i8} \end{pmatrix} + \begin{pmatrix} B_{i1}^t & B_{i3}^t & B_{i5}^t & B_{i7}^t \\ B_{i2}^t & B_{i4}^t & B_{i6}^t & B_{i8}^t \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ D_\mu & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

This matrix equation implies

$$A_{i2} = 0, A_{i4} = 0, \quad i = 2, \cdots, n.$$

From the skew-symmetry of $B_1^H A_i$, we obtain $A_{i6} = 0$ for $i = 2, \cdots, n$.

Define $\{(\hat{A}_2 \ \hat{B}_2), (\hat{A}_3 \ \hat{B}_3), \cdots, (\hat{A}_n \ \hat{B}_n)\}$ as follows:

$$\hat{A}_i = \begin{pmatrix} 0 \\ 0 \\ 0 \\ A_{i8} \end{pmatrix}, \quad \hat{B}_i = \begin{pmatrix} B_{i2} \\ B_{i4} \\ B_{i6} \\ B_{i8} \end{pmatrix},$$

i.e., $\hat{A}_i$ and $\hat{B}_i$ are the second columns of the block matrices in A_i and B_i, respectively.

Since $A_i^H A_j + B_j^t B_i^* = \delta_{ij} E_i$ for $i, j \geq 2$, we have

$$\begin{pmatrix} A_{i1}^H & A_{i3}^H & A_{i5}^H & A_{i7}^H \\ 0 & 0 & 0 & A_{i8}^H \end{pmatrix} \begin{pmatrix} A_{j1} & 0 \\ A_{j3} & 0 \\ A_{j5} & 0 \\ A_{j7} & A_{j8} \end{pmatrix} + \begin{pmatrix} B_{j1}^t & B_{j3}^t & B_{j5}^t & B_{j7}^t \\ B_{j2}^t & B_{j4}^t & B_{j6}^t & B_{j8}^t \end{pmatrix} \begin{pmatrix} B_{i1}^* & B_{i2}^* \\ B_{i3}^* & B_{i4}^* \\ B_{i5}^* & B_{i6}^* \\ B_{i7}^* & B_{i8}^* \end{pmatrix} = \delta_{ij} E_i,$$

where E_i are positive definite. By noting that the second row and the second column in the above products, we obtain

$$\hat{A}_i^H \hat{A}_j + \hat{B}_j^t \hat{B}_i^* = \delta_{ij} \hat{E}_i, \quad i, j \geq 2, \tag{12}$$

where $\hat{E}_i$ are the $(k-s)\times(k-s)$ matrix taken from the last $k-s$ rows and the last $k-s$ columns of E_i, and therefore, $\hat{E}_i$ are also positive definite. By similarly showing other conditions, $\{(\hat{A}_2\ \hat{B}_2),\cdots,(\hat{A}_n\ \hat{B}_n)\}$ is also a Hurwitz family of size $(k-s)\times p$ matrices.

By the rank inequality 3 in the end of Section 2, (12) implies

$$\mathrm{rank}(A_{i8})+\mathrm{rank}(\hat{B}_i)\geq k-s,\quad i=2,\cdots,n. \tag{13}$$

Since $n\geq 3$, there exist a pair i and j with $i\neq j\geq 2$. When $i\neq j\geq 2$, we have $\hat{A}_i^H\hat{A}_j+\hat{B}_j^t\hat{B}_i^*=0$, that is, $A_{i8}^HA_{j8}+\hat{B}_j^t\hat{B}_i^*=0$, which implies

$$\mathrm{rank}(\hat{B}_i)+\mathrm{rank}(\hat{B}_j)-p\leq \mathrm{rank}\ (\hat{B}_j^t\hat{B}_i^*)=\mathrm{rank}\ (A_{i8}^HA_{j8})\leq p-k-s, \tag{14}$$

where the first inequality is due to Sylvester's inequality and the row size of $\hat{B}_i$ and $\hat{B}_j$ is p, and the last inequality is because A_{i8} and A_{j8} are all of size $(p-k-s)\times(k-s)$ and the rank inequality 2 in the end of Section 2. Hence, from (14) and (13), $\mathrm{rank}(A_{j8})\leq p-k-s$ and $\mathrm{rank}(A_{i8})\leq p-k-s$, we have

$$\begin{aligned}
p-k-s &\geq \mathrm{rank}(\hat{B}_i)+\mathrm{rank}(\hat{B}_j)-p\\
&\geq k-s-\mathrm{rank}(A_{i8})+k-s-\mathrm{rank}(A_{j8})-p\\
&\geq k-s-(p-k-s)+k-s-(p-k-s)-p=4k-3p,
\end{aligned}$$

which implies

$$4p-5k\geq s\geq 0.$$

Therefore, the first half of the theorem is proved.

We next want to show the second half of the theorem and assume that $\mathcal{G}$ is a complex orthogonal design. All the above derivations still hold for $\mathcal{G}$ and are adopted in the following proof. Since $\mathcal{G}$ is a complex orthogonal design, by Proposition 2,

$$\{(A_2\ B_2),(A_3\ B_3),\cdots,(A_n\ B_n)\}$$

satisfies (3) in Proposition 1 with $E_i=I$. Therefore, it is not hard to see that

$$\{(\hat{A}_2\ \hat{B}_2),(\hat{A}_3\ \hat{B}_3),\cdots,(\hat{A}_n\ \hat{B}_n)\}$$

also satisfies (3) in Proposition 1 with $\hat{E}_i = I$.

Consider matrix A_{28} in $\hat{A}_2$, which has size $(p-k-s)\times(k-s)$. There exist a $(p-k-s)\times(p-k-s)$ unitary matrix U and a $(k-s)\times(k-s)$ unitary matrix R such that

$$UA_{28}R = \begin{pmatrix} 0 & 0 \\ D_\beta & 0 \end{pmatrix},$$

where $D_\beta = \mathrm{diag}(\beta_1, \beta_2, \cdots, \beta_r)$, r is the rank of the matrix A_{28} and $\beta_i > 0$, $i = 1, 2, \cdots, r$, are the positive square roots of the eigenvalues of $A_{28}A_{28}^H$. Clearly,

$$r \le p-k-s. \tag{15}$$

Using these unitary matrices U and R, we rewrite the matrix pairs $\{(\hat{A}_2\ \hat{B}_2), \cdots, (\hat{A}_n\ \hat{B}_n)\}$ as follows.

Let $P = \begin{pmatrix} I & 0 \\ 0 & U \end{pmatrix}$, where I is the identity matrix of size $k+s$. Then P is a $p\times p$ unitary matrix and $\{(P\hat{A}_2R\ \ P\hat{B}_2R^*), \cdots, (P\hat{A}_nR\ \ P\hat{B}_nR^*)\}$ also satisfies (3) in Proposition 1 with $E_i = I$. Furthermore, $\{P\hat{A}_2R, P\hat{B}_2R^*, P\hat{A}_3R, P\hat{B}_3R^*\}$ can be written as follows

$$\left\{ \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ D_\beta & 0 \end{pmatrix}, \begin{pmatrix} \tilde{B}_{22} & \bar{B}_{22} \\ \tilde{B}_{24} & \bar{B}_{24} \\ \tilde{B}_{26} & \bar{B}_{26} \\ \tilde{B}_{281} & \bar{B}_{282} \\ \tilde{B}_{283} & \bar{B}_{284} \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \tilde{A}_{381} & \bar{A}_{382} \\ \tilde{A}_{383} & \bar{A}_{384} \end{pmatrix}, \begin{pmatrix} \tilde{B}_{32} & \bar{B}_{32} \\ \tilde{B}_{34} & \bar{B}_{34} \\ \tilde{B}_{36} & \bar{B}_{36} \\ \tilde{B}_{381} & \bar{B}_{382} \\ \tilde{B}_{383} & \bar{B}_{384} \end{pmatrix} \right\}, \tag{16}$$

where the sizes of $\tilde{B}_{i2}, \tilde{B}_{i4}, \tilde{B}_{i6}$ are $s\times r$, $(k-s)\times r$, $s\times r$, respectively, the sizes of $\bar{B}_{i2}, \bar{B}_{i4}, \bar{B}_{i6}$ are $s\times(k-s-r)$, $(k-s)\times(k-s-r)$, $s\times(k-s-r)$, respectively, the sizes of $\tilde{B}_{281}, \tilde{A}_{381}, \tilde{B}_{381}$ are $(p-k-s-r)\times r$, the sizes of $\tilde{B}_{283}, \tilde{A}_{383}, \tilde{B}_{383}$ are $r\times r$, the sizes of $\bar{B}_{282}, \bar{A}_{382}, \bar{B}_{382}$ are $(p-k-s-r)\times(k-s-r)$, and the sizes of $\bar{B}_{284}, \bar{A}_{384}, \bar{B}_{384}$ are $r\times(k-s-r)$.

We next want to show that $\bar{B}_{284} = 0$ and $\bar{B}_{384} = 0$. From (3) in Proposition 1, the matrix $(P\hat{A}_2R)^H(P\hat{B}_2R^*)$ is skew-symmetry, i.e.,

$$\begin{pmatrix} 0 & 0 & 0 & 0 & D_\beta \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{B}_{22} & \bar{B}_{22} \\ \tilde{B}_{24} & \bar{B}_{24} \\ \tilde{B}_{26} & \bar{B}_{26} \\ \tilde{B}_{281} & \bar{B}_{282} \\ \tilde{B}_{283} & \bar{B}_{284} \end{pmatrix}$$

is skew-symmetry, which implies $D_\beta\bar{B}_{284} = 0$, therefore, $\bar{B}_{284} = 0$ because D_β is invertible. Similarly, the matrix $(P\hat{A}_2R)^H(P\hat{B}_3R^*)$ is also skew-symmetry, which implies $\bar{B}_{384} = 0$.

Again by (3) in Proposition 1, we have

$$(P\hat{A}_2R)^H(P\hat{A}_3R)+(P\hat{B}_3R^*)^t(P\hat{B}_2R^*)^*=0,$$

i.e.,

$$\begin{pmatrix} 0 & 0 & 0 & 0 & D_\beta \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \tilde{A}_{381} & \bar{A}_{382} \\ \tilde{A}_{383} & \bar{A}_{384} \end{pmatrix}+\begin{pmatrix} \tilde{B}^t_{32} & \tilde{B}^t_{34} & \tilde{B}^t_{36} & \tilde{B}^t_{381} & \tilde{B}^t_{383} \\ \bar{B}^t_{32} & \bar{B}^t_{34} & \bar{B}^t_{36} & \bar{B}^t_{382} & 0 \end{pmatrix}\begin{pmatrix} \tilde{B}^*_{22} & \bar{B}^*_{22} \\ \tilde{B}^*_{24} & \bar{B}^*_{24} \\ \tilde{B}^*_{26} & \bar{B}^*_{26} \\ \tilde{B}^*_{281} & \bar{B}^*_{282} \\ \tilde{B}^*_{283} & 0 \end{pmatrix}=0.$$

From the second row and the second column, the above equation implies

$$\bar{B}^t_3\bar{B}^*_2=0, \tag{17}$$

where $\bar{B}_i=(\bar{B}^t_{i2}\ \bar{B}^t_{i4}\ \bar{B}^t_{i6}\ \bar{B}^t_{i82})^t$ for $i=2,3$. By Sylvester's inequality, and noting that the size of matrices $\bar{B}_2$ and $\bar{B}_3$ is $(p-r)\times(k-s-r)$, (17) implies

$$\operatorname{rank}(\bar{B}_2)+\operatorname{rank}(\bar{B}_3)\le p-r. \tag{18}$$

We next want to determine the ranks of $\bar{B}_2$ and $\bar{B}_3$.

Because $(P\hat{A}_2R)^H(P\hat{A}_2R)+(P\hat{B}_2R^*)^t(P\hat{B}_2R^*)^*=I$, we have

$$\begin{pmatrix} 0 & 0 & 0 & 0 & D_\beta \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ D_\beta & 0 \end{pmatrix}+\begin{pmatrix} \tilde{B}^t_{22} & \tilde{B}^t_{24} & \tilde{B}^t_{26} & \tilde{B}^t_{281} & \tilde{B}^t_{283} \\ \bar{B}^t_{22} & \bar{B}^t_{24} & \bar{B}^t_{26} & \bar{B}^t_{282} & 0 \end{pmatrix}\begin{pmatrix} \tilde{B}^*_{22} & \bar{B}^*_{22} \\ \tilde{B}^*_{24} & \bar{B}^*_{24} \\ \tilde{B}^*_{26} & \bar{B}^*_{26} \\ \tilde{B}^*_{281} & \bar{B}^*_{282} \\ \tilde{B}^*_{283} & 0 \end{pmatrix}=I.$$

Therefore, by noting that the second row and the second column, we have $\bar{B}^t_2\bar{B}^*_2=I$, hence,

$$\operatorname{rank}(\bar{B}_2)=k-s-r. \tag{19}$$

For the rank of $\bar{B}_3$, we first use the fact that $(P\hat{A}_3R)^H(P\hat{A}_3R)+(P\hat{B}_3R^*)^t(P\hat{B}_3R^*)^*=I$ and we then use the forms of $P\hat{A}_3R$ and $P\hat{B}_3R^*$ in (16) and expand the summation. We then conclude that $\bar{A}^H_{382}\bar{A}_{382}+\bar{A}^H_{384}\bar{A}_{384}+\bar{B}^t_3\bar{B}^*_3=I$. Finally, from the rank inequality 3 in the end of Section 2, we have

$$\operatorname{rank}(\bar{B}_3)\ge k-s-r-r_1, \tag{20}$$

where r_1 is the rank of matrix $(\bar{A}_{382}^t\ \bar{A}_{384}^t)^t$ that has $p-k-s$ rows from (16). Thus, we also have

$$r_1 \leq p-k-s. \tag{21}$$

Combining formulas (15) and (18)-(21), we have

$$2k-p \leq 2s+r+r_1 \leq 2s+(p-k-s)+(p-k-s)=2p-2k,$$

i.e.,

$$\frac{k}{p} \leq \frac{3}{4}.$$

This proves Theorem 1. **q.e.d.**

From the above proof, one can see that the difference between the above upper bounds we obtained on the rates of complex orthogonal designs and generalized orthogonal designs comes from whether the property

$$\hat{A}_i^H \hat{B}_i + \hat{B}_i^t \hat{A}_i^* = 0 \tag{22}$$

holds. It holds for the complex orthogonal designs due to the orthonormality but may not hold for generalized complex orthogonal designs.

As another application of Proposition 2, we have another upper bound for the rates of a complex orthogonal space-time block code, which sharpens the result in Theorem 1 if an additional condition is satisfied. Let $\mathcal{G} = (A_1\mathbf{x}+B_1\mathbf{x}^*\ \cdots\ A_n\mathbf{x}+B_n\mathbf{x}^*)$ be a generalized complex orthogonal design. Define

$$\rho \triangleq \max_{i=1,2,\cdots,n} \operatorname{rank}((A_i\ B_i)). \tag{23}$$

Theorem 2 *Let* $\mathcal{G} = (A_1\mathbf{x}+B_1\mathbf{x}^*\ A_2\mathbf{x}+B_2\mathbf{x}^* \cdots A_n\mathbf{x}+B_n\mathbf{x}^*)$ *be a generalized complex orthogonal design. If* $\rho = p$ *and* $n \geq 2$*, then the rate of* $\mathcal{G}$ *is upper bounded by* $n/(2n-2)$*, i.e.,*

$$R = \frac{k}{p} \leq \frac{n}{2n-2}.$$

Proof: If $2k \leq p$, then, the above theorem is proved. So, in what follows, we assume $2k > p$. Without loss of generality, we assume $\mathrm{rank}((A_1\ B_1)) = p$. By Proposition 2, we may assume A_1 and B_1 have the following forms:

$$A_1 = \begin{pmatrix} D_\lambda & 0 \\ 0 & I_{k-s} \\ 0 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0_{s\times s} & 0 \\ 0_{(k-s)\times s} & 0 \\ D_\mu & 0 \end{pmatrix},$$

where $s = p - k$, $D_\lambda = \mathrm{diag}\,(\lambda_1, \cdots, \lambda_s)$, $D_\mu = \mathrm{diag}\,(\mu_1, \cdots, \mu_s)$, I_{k-s} is the identity matrix of size $k - s$, and $1 > \lambda_i > \mu_j > 0$, $i, j = 1, \cdots, s$, and $\{(A_1\ B_1), (A_2\ B_2), \cdots, (A_n\ B_n)\}$ is a Hurwitz family.

Divide matrices A_i and B_i into block matrices as follows:

$$A_i = \begin{pmatrix} A_{i1} & A_{i2} \\ A_{i3} & A_{i4} \\ A_{i5} & A_{i6} \end{pmatrix}, \quad B_i = \begin{pmatrix} B_{i1} & B_{i2} \\ B_{i3} & B_{i4} \\ B_{i5} & B_{i6} \end{pmatrix},$$

where the sizes of A_{i1} and B_{i1} are $s \times s$, the sizes of A_{i3} and B_{i3} are $(k-s) \times s$, and the sizes of the rest submatrices can be determined accordingly. By the properties of a Hurwitz family in Definition 1, we have $A_1^H A_i + B_i^t B_1^* = 0$ for $i > 1$, that is,

$$\begin{pmatrix} D_\lambda & 0 & 0 \\ 0 & I & 0 \end{pmatrix} \begin{pmatrix} A_{i1} & A_{i2} \\ A_{i3} & A_{i4} \\ A_{i5} & A_{i6} \end{pmatrix} + \begin{pmatrix} B_{i1}^t & B_{i3}^t & B_{i5}^t \\ B_{i2}^t & B_{i4}^t & B_{i6}^t \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ D_\mu & 0 \end{pmatrix} = 0,$$

which implies $A_{i2} = A_{i4} = 0$ for $i > 1$. Similarly, $A_{i6} = 0$ can been obtained from the skew-symmetry of $A_i^H B_1$.

Define

$$\hat{A}_i = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \hat{B}_i = \begin{pmatrix} B_{i2} \\ B_{i4} \\ B_{i6} \end{pmatrix}.$$

Then, by the same method used in the proof of Theorem 1, $\{(\hat{A}_2\ \hat{B}_2), \cdots, (\hat{A}_n\ \hat{B}_n)\}$ is also a Hurwitz family.

By applying the properties of a Hurwitz family, it is not hard to verify that the $(n-1)(k-s)$ column vectors of the following matrix

$$\begin{pmatrix} B_{22} & B_{32} & \cdots & B_{n2} \\ B_{24} & B_{34} & \cdots & B_{n4} \\ B_{26} & B_{36} & \cdots & B_{n6} \end{pmatrix}$$

are linearly independent in $\mathbb{C}^p$. Therefore, by noticing $s = p - k$, we have

$$(n-1)(k-s) = (n-1)(k-(p-k)) \le p,$$

which establishes the theorem. **q.e.d.**

It is not hard to check that the rate 3/4 complex orthogonal space-time block codes in [3, 4, 5, 6] for 3 and 4 transmit antennas do not satisfy the condition in Theorem 2 and therefore, the upper bound in Theorem 2 does not apply to them.

4 Restricted (Generalized) Complex Orthogonal Designs

In the previous sections, we consider the conventional (generalized) complex orthogonal designs in the sense that all the variables in the designs may take any values on the complex plane. In this section, we consider restricted (generalized) complex orthogonal designs where the variables only take values from subsets of the complex plane. To do so, we first introduce some necessary notations and concepts.

Let $\mathcal{A}$ be a subset (finite or infinite) of the complex plane, which is called alphabet set. The *difference* set of $\mathcal{A}$, denoted $\Delta\mathcal{A}$, is defined by

$$\Delta\mathcal{A} \triangleq \{z_1 - z_2 \mid \text{for any } z_1, z_2 \in \mathcal{A}\}. \tag{24}$$

Note that for any alphabet set $\mathcal{A}$, we have $0 \in \Delta\mathcal{A}$.

For an alphabet set $\mathcal{A}$, it is called *admissible* if it contains at least three distinct points such that they are *not* collinear, i.e., not lie on a straight line on the complex plane, or more precisely, there exist $z_j = p_j + \mathbf{j}q_j \in \mathcal{A}$, $j = 1, 2, 3$, such that

$$\det\begin{pmatrix} p_1 - p_2 & q_1 - q_2 \\ p_1 - p_3 & q_1 - q_3 \end{pmatrix} \ne 0, \tag{25}$$

where $\mathbf{j} = \sqrt{-1}$. We next want to see the admissibility condition (25) on the difference set $\Delta\mathcal{A}$. It is clear that condition (25) is equivalent to anyone of the following :

$$\det\begin{pmatrix} p_2 - p_1 & q_2 - q_1 \\ p_2 - p_3 & q_2 - q_3 \end{pmatrix} \ne 0, \quad \det\begin{pmatrix} p_3 - p_1 & q_3 - q_1 \\ p_3 - p_2 & q_3 - q_3 \end{pmatrix} \ne 0. \tag{26}$$

Let $x_1 = z_1 - z_2, x_2 = z_2 - z_3, x_3 = z_3 - z_1$. Then $x_j \in \Delta\mathcal{A}$, furthermore, condition (25) can be rewritten as

$$\det \begin{pmatrix} \mathrm{Re}(x_1) & \mathrm{Im}(x_1) \\ -\mathrm{Re}(x_3) & -\mathrm{Im}(x_3) \end{pmatrix} \neq 0, \tag{27}$$

where $\mathrm{Re}(x)$ and $\mathrm{Im}(x)$ are the real and image parts of x, respectively. By some simple calculations, we may find that condition (27) or (25) is equivalent to

$$\det \begin{pmatrix} x_1 & x_1^* \\ x_3 & x_3^* \end{pmatrix} \neq 0. \tag{28}$$

Similarly, (26) is equivalent to, respectively,

$$\det \begin{pmatrix} x_1 & x_1^* \\ x_2 & x_2^* \end{pmatrix} \neq 0, \quad \det \begin{pmatrix} x_3 & x_3^* \\ x_2 & x_2^* \end{pmatrix} \neq 0. \tag{29}$$

In summary, an alphabet set $\mathcal{A}$ is admissible if and only if there exist at least three points $\{x_1, x_2, x_3\}$ in $\Delta\mathcal{A}$ such that condition (28) or anyone of the two in (29) holds. As a remark, a constellation set M-PSK ($M > 2$) or M-QAM ($M > 2$) is admissible.

We next give the definition of a *restricted* generalized complex orthogonal design. Let

$$\mathcal{G} = (A_1\mathbf{z} + B_1\mathbf{z}^* \ \cdots \ A_n\mathbf{z} + B_n\mathbf{z}^*)$$

be a $p \times n$ matrix, where $\mathbf{z} = (z_1, z_2, \cdots, z_k)^t \in \mathbb{C}^k$, and A_i, B_i, $i = 1, \cdots, n$, are $p \times k$ complex constant matrices.

Definition 2 *Let $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_k$ be k complex point sets. $\{\mathcal{G}; \mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_k\}$ is called a restricted orthogonal space-time design if for any $\mathbf{z} = (z_1, z_2, \cdots, z_k)^t \in (\mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_k)^t$, the following orthogonality holds:*

$$\begin{aligned} \mathcal{G}^H\mathcal{G} &= (A_1\mathbf{z} + B_1\mathbf{z}^* \cdots A_n\mathbf{z} + B_n\mathbf{z}^*)^H (A_1\mathbf{z} + B_1\mathbf{z}^* \cdots A_n\mathbf{z} + B_n\mathbf{z}^*) \\ &= |\mathbf{z}_1|^2 D_1 + |\mathbf{z}_2|^2 D_2 + \cdots + |\mathbf{z}_k|^2 D_k, \end{aligned} \tag{30}$$

where D_i, $i = 1, 2, \cdots, k$, are some $n \times n$ diagonal positive definite constant matrices.

For a restricted generalized complex orthogonal design, we have the following theorem.

Theorem 3 *Let $\{\mathcal{G}; \mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_k\}$ be a restricted generalized complex orthogonal design. If for each i with $1 \leq i \leq k$, the following conditions hold for alphabet set $\mathcal{A}_i$:*

(i) $\mathcal{A}_i$ is admissible, i.e., $\mathcal{A}_i$ does not contain only collinear points;

(ii) There exists $z = p + q\mathbf{j} \in \mathcal{A}_i$ with $pq \neq 0$ such that $z \neq z^*$ and $z^* \in \mathcal{A}_i$;

(iii) There exist $z_j = p_j + q_j\mathbf{j} \in \mathcal{A}_i$, $j = 1, 2, 3$, such that

$$det \begin{pmatrix} p_1^2 - p_2^2 & q_1^2 - q_2^2 \\ p_1^2 - p_3^2 & q_1^2 - q_3^2 \end{pmatrix} \neq 0, \tag{31}$$

then $\mathcal{G}$ is also a generalized complex orthogonal design.

Proof. By Proposition 1, it is enough to prove that under the conditions of this theorem, the following matrix equations hold:

$$A_i^H A_j + B_j^t B_i^* = \delta_{ij} E_i, \quad i, j = 1, 2, \cdots, n, \tag{32}$$

$$A_i^H B_j + B_j^t A_i^* = 0, \quad i, j = 1, 2, \cdots, n, \tag{33}$$

where E_i are positive definite diagonal matrices, which can be proved by using Lemma 8 similar to the proof of Proposition 1 by using Lemma 1. We omit the details. **q.e.d.**

This result can be thought of a generalization of the results in [12, 13] in the sense that the size of an orthogonal design is not necessarily square, i.e., p does not have to be equal to n, and it is in the complex field instead of the real field, and the orthonormality is generalized to the orthogonality. One can see that PSK constellations do not satisfy the condition (iii) in the above theorem. However, it is not difficult to check that, the commonly used QAM signal constellations of sizes above 4 located on a square lattice satisfy the conditions (i)-(iii) in the above theorem. Therefore, we have the following corollary.

Corollary 1 *A restricted generalized complex orthogonal design with its variables restricted to QAM constellations of sizes above 4 on square lattices is also a generalized complex orthogonal design and therefore, the upper bounds on its rate in Section 3 hold.*

The fact that PSK constellations do not satisfy the condition (iii) in Theorem 3 shows that that the admissibility (25) does not imply the condition (iii) in Theorem 3. On the other hand, by considering points on a straight line (it is neither the x-axis nor the y-axis), the condition (iii) in Theorem 3 may hold. This shows that the condition (iii) in Theorem 3 does not imply the admissibility (25) in general.

5 Conclusion

In this paper, we have shown that the rates of complex orthogonal space-time block codes for three or more transmit antennas are upper bounded by 3/4 and the rates of generalized complex orthogonal space-time codes for three or more transmit antennas are upper bounded by 4/5. We have presented another sharper upper bound for the rates under a certain condition. Notice that the maximal rate of real orthogonal space-time codes is 1 for any number transmit antennas, which is achievable using the Hurwitz-Radon constructive proof. For complex orthogonal space-time block codes or generalized complex orthogonal space-time block codes, the maximal rate 1 is reached only for two transmit antennas. For generalized complex space-time block codes, rate 7/11 and 3/5 generalized complex orthogonal designs for $n = 5$ and $n = 6$ have been constructed in [9], which are 9/55 and 1/5 away from the upper bound 4/5 we derived for generalized complex orthogonal space-time block codes in this paper, respectively. For complex orthogonal space-time codes, rate 2/3 complex orthogonal design for $n = 5$ has been constructed in [11], which is 1/12 away from the upper bound 3/4. For a general n, we conjecture that the upper bound 3/4 of the rate of complex orthogonal designs can be sharpened as

$$R \leq \frac{\lceil \frac{n}{2} \rceil + 1}{2\lceil \frac{n}{2} \rceil},$$

which can be achieved for $n = 1, 2, 3, 4, 5$.

As a remark, the upper bound of the rates, $R \leq 3/4$ when $n > 2$, was proved in [10] for a special family of complex orthogonal space-time block codes from the complex orthogonal designs $\mathcal{G}$ where the entries of $\mathcal{G}$ do not consist of any linear processing of $\mathbf{x}_i$ and $\mathbf{x}_i^*$, $i = 1, 2, \cdots, k$, and can only be 0 or single variables $\pm\mathbf{x}_i$ or $\pm\mathbf{x}_i^*$, $i = 1, 2, \cdots, k$, and these variables do not repeat in any column of $\mathcal{G}$. The method used in [10] was based more on a combinatorial argument that is different from what was used in this paper.

In the last part of this paper, we have considered the restricted generalized complex orthogonal designs by restricting the variables to subsets of the complex plane. We have obtained a condition on the alphabet sets such that a restricted generalized complex orthogonal design is a generalized complex orthogonal design. The commonly used QAM constellations of size above 4 on square lattices do satisfy the condition. Thus, the upper bounds on the rates presented in this paper also apply to restricted generalized complex orthogonal designs for commonly used QAM signal constellations of sizes above 4. This result can be thought of as a generalization of the results in [12, 13] from square real orthogonal designs to (not necessary square) generalized complex orthogonal designs.

Acknowledgement

The authors would like to thank Xue-Bin Liang and Weifeng Su for their careful reading of this manuscript and for their useful comments. They also would like to thank the anonymous referees for their careful reading and detailed comments that have improved the clarity of this paper. In particular, one of the referees pointed the references [14] and [13] to the authors, which has led them to study the restricted (pseudo) orthogonal designs.

Appendix

The following Lemma 1 is from [8]. For the completeness, its proof in [8] is given below.

Lemma 1 *Let A, B and C be three $m \times m$ complex constant matrices. If for any $\mathbf{x} \in \mathbb{C}^m$,*

$$\mathbf{x}^H A\mathbf{x} + \mathbf{x}^H B\mathbf{x}^* + \mathbf{x}^t C\mathbf{x} = 0,$$

then

$$A = B + B^t = C + C^t = 0.$$

Proof: Let $A = A_1 + \mathbf{j}A_2$, $(B + B^t)/2 = B_1 + \mathbf{j}B_2$ and $(C + C^t)/2 = C_1 + \mathbf{j}C_2$, where A_i, B_i, C_i are all $m \times m$ real matrices, $i = 1, 2$. It is obvious that B_1, B_2, C_1, C_2 are symmetric matrices. Rewrite the condition as follows

$$\begin{aligned}
&\mathbf{x}^H A\mathbf{x} + \mathbf{x}^H B\mathbf{x}^* + \mathbf{x}^t C\mathbf{x} = \mathbf{x}^H A\mathbf{x} + \frac{1}{2}(\mathbf{x}^H(B + B^t)\mathbf{x}^*) + \frac{1}{2}(\mathbf{x}^t(C + C^t)\mathbf{x}) \\
&= \mathbf{x}^H(A_1 + \mathbf{j}A_2)\mathbf{x} + \mathbf{x}^H(B_1 + \mathbf{j}B_2)\mathbf{x}^* + \mathbf{x}^t(C_1 + \mathbf{j}C_2)\mathbf{x} = 0
\end{aligned} \tag{34}$$

for all $\mathbf{x} \in \mathbb{C}^m$.

First consider the case when $\mathbf{x} = \mathbf{a} \in \mathbb{R}^m$. From (34), we have

$$\mathbf{a}^t(A_1 + \mathbf{j}A_2)\mathbf{a} + \mathbf{a}^t(B_1 + \mathbf{j}B_2)\mathbf{a} + \mathbf{a}^t(C_1 + \mathbf{j}C_2)\mathbf{a} = 0,$$

for all $\mathbf{a} \in \mathbb{R}^m$. Hence, by equating the real and imaginary parts on the two sides,

$$\mathbf{a}^t(A_1 + B_1 + C_1)\mathbf{a} = 0,$$

$$\mathbf{a}^t(A_2 + B_2 + C_2)\mathbf{a} = 0.$$

Then, consider the case when $\mathbf{x} = \mathbf{ja}$ with $\mathbf{a} \in \mathbb{R}^m$ in (34):

$$\mathbf{a}^t(-A_1 + B_1 + C_1)\mathbf{a} = 0,$$

$$\mathbf{a}^t(-A_2 + B_2 + C_2)\mathbf{a} = 0.$$

From the above four equalities, we have

$$\mathbf{a}^t A_1\mathbf{a} = \mathbf{a}^t A_2\mathbf{a} = \mathbf{a}^t(B_1 + C_1)\mathbf{a} = \mathbf{a}^t(B_2 + C_2)\mathbf{a} = 0$$

for all $\mathbf{a} \in \mathbb{R}^m$. Therefore the matrices $A_1, A_2, B_1 + C_1, B_2 + C_2$ are skew-symmetric. But B_1, B_2 and C_1, C_2 are symmetric, we obtain

$$A_1 = -A_1^t, A_2 = -A_2^t, B_1 = -C_1, B_2 = -C_2.$$

Finally, using the relationships obtained above with $\mathbf{x} = \mathbf{a} + \mathbf{jb}$, where $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$, (34) becomes

$$\mathbf{x}^H A\mathbf{x} + \mathbf{x}^H B\mathbf{x}^* + \mathbf{x}^t C\mathbf{x} = 2\mathbf{a}^t(2B_2 - A_2)\mathbf{b} + 2\mathbf{ja}^t(A_1 - 2B_1)\mathbf{b} = 0,$$

which gives

$$\mathbf{a}^t(2B_2 - A_2)\mathbf{b} = 0, \quad \mathbf{a}^t(A_1 - 2B_1)\mathbf{b} = 0.$$

Because $\mathbf{a}$ and $\mathbf{b}$ are arbitrary, we have $A_1 = 2B_1$ and $A_2 = 2B_2$. Since A_1, A_2 are skew-symmetric and B_1, B_2 are symmetric, we conclude $A_1 = A_2 = B_1 = B_2 = 0$. Therefore, $C_1 = C_2 = 0$. This proves the lemma. **q.e.d.**

In the following, we want to prove Proposition 2. To do so, we need several lemmas. First we have two direct consequences from Proposition 1.

Lemma 2 *Let* $\mathcal{G} = (A_1\mathbf{x} + B_1\mathbf{x}^* \; A_2\mathbf{x} + B_2\mathbf{x}^* \; \cdots \; A_n\mathbf{x} + B_n\mathbf{x}^*)$ *be a generalized complex orthogonal design. Then, for any* $\mathbf{x}, \mathbf{y} \in \mathbb{C}^k$,

$$\|A_i\mathbf{x} + B_i\mathbf{y}\|^2 + \|B_i^*\mathbf{x} + A_i^*\mathbf{y}\|^2 = \mathbf{x}^H E_i\mathbf{x} + \mathbf{y}^H E_i\mathbf{y},$$

where $\|\cdot\|^2 = \langle\cdot,\cdot\rangle$, *and* $\langle,\rangle$ *denotes the inner product* $\langle\mathbf{x},\mathbf{y}\rangle = \mathbf{x}^H\mathbf{y}$ *in* $\mathbb{C}^k$, *and* E_i *are from (3).*

Proof: It is directly from Proposition 1 when $j = i$. **q.e.d.**

Lemma 3 *Let*

$$\mathcal{G} = (A_1\mathbf{x} + B_1\mathbf{x}^* \; A_2\mathbf{x} + B_2\mathbf{x}^* \; \cdots \; A_n\mathbf{x} + B_n\mathbf{x}^*)$$

be a generalized complex orthogonal design. Then, $\mathcal{G}$ *can be reduced to a new generalized complex orthogonal design* $\tilde{\mathcal{G}}$ *with the same parameters* p, k, n *as in* $\mathcal{G}$ *as follows:*

$$\tilde{\mathcal{G}} = (\tilde{A}_1\mathbf{y} + \tilde{B}_1\mathbf{y}^* \; \tilde{A}_2\mathbf{y} + \tilde{B}_2\mathbf{y}^* \; \cdots \; \tilde{A}_n\mathbf{y} + \tilde{B}_n\mathbf{y}^*)$$

with $\tilde{A}_1^H\tilde{A}_1 + \tilde{B}_1^t\tilde{B}_1^* = I$, *that is,* $E_1 = I$ *in (3) for* $\tilde{A}_1$ *and* $\tilde{B}_1$, *where* $\mathbf{y} = (y_1 \; y_2 \cdots y_k)^t$.

Proof: By Proposition 1, E_1 is diagonal positive definite. Let $U = \sqrt{E_1^{-1}}$ and then U is also diagonal positive definite. Make the transformation $\mathbf{x} = U\mathbf{y}$ and let $\tilde{A}_i = A_iU$, $\tilde{B}_i = B_iU$ and $\tilde{E}_i = UE_iU$, then $\tilde{A}_i$ and $\tilde{B}_i$ satisfy (3) and $\tilde{E}_i$ are all diagonal positive definite. Furthermore, $\tilde{A}_1^H\tilde{A}_1 + \tilde{B}_1^t\tilde{B}_1^* = UE_1U = I$. **q.e.d.**

Going back to Proposition 2, one can see that it has a similar form as a classical singular value decomposition (SVD) for the matrix $(A_1\ B_1)$ but the Hurwitz family properties need to be maintained during the unitary transformations. Thus, the basic idea in the following to prove Proposition 2 is similar to the SVD by carefully selecting the unitary transforms. To do so, we have other two lemmas.

Lemma 4 *Let A and B be two $p\times k$ matrices and satisfy conditions: $A^HA+B^tB^*=I$, and A^HB and B^HA are skew-symmetric. Let $M=\max\left\{\|A\mathbf{x}+B\mathbf{y}\|^2:\ \|\mathbf{x}\|^2+\|\mathbf{y}\|^2=1\right\}$. Then,*

(i) $1\geq M\geq 1/2$ and there is a vector $\mathbf{u}=(\mathbf{u}_1^t\ \mathbf{u}_2^t)\in\mathbb{C}^{2k}$ with $\|\mathbf{u}\|=1$ such that $M=\|A\mathbf{u}_1+B\mathbf{u}_2\|^2$;

(ii) If $M>1/2$, then, vectors $\mathbf{u}_1$ and $\mathbf{u}_2^*$ in (i) are orthogonal, i.e., $\langle\mathbf{u}_1,\mathbf{u}_2^*\rangle=0$;

(iii) If $M=1/2$, then, $A^HA=B^HB=\frac{1}{2}I$, and $A^HB=0$.

Proof: Under the condition of this lemma and using Lemma 2, it is clear that if $\|\mathbf{x}\|^2+\|\mathbf{y}\|^2=1$, then

$$\|A\mathbf{x}+B\mathbf{y}\|^2+\|B^*\mathbf{x}+A^*\mathbf{y}\|^2=1. \tag{35}$$

By noting that $\|B^*\mathbf{x}+A^*\mathbf{y}\|^2=\|B\mathbf{x}^*+A\mathbf{y}^*\|^2$ and that the unit ball $\|\mathbf{x}\|^2+\|\mathbf{y}\|^2=1$ in a finite dimensional space is compact, (i) is not hard to see.

We next want to prove (ii). Let $C=(A\ B)$ be the $p\times 2k$ matrix and consider the Hermitian matrix C^HC. Obviously, for any a vector $\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix}\in\mathbb{C}^{2k}$,

$$(\mathbf{x}^H\ \mathbf{y}^H)C^HC\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix}=\|A\mathbf{x}+B\mathbf{y}\|^2.$$

Because C^HC is a Hermitian matrix, by Rayleigh-Ritz theorem ([24], p.176), M is equal to the maximum eigenvalue of C^HC. It is easy to check that for any $\mathbf{u}$ as in (i), $\mathbf{u}$ is an eigenvector of M. Therefore, we have $C^HC\mathbf{u}=M\mathbf{u}$, that is,

$$\begin{aligned}(\mathbf{u}_1^HA^H+\mathbf{u}_2^HB^H)A &= M\mathbf{u}_1^H, && (36)\\ (\mathbf{u}_1^HA^H+\mathbf{u}_2^HB^H)B &= M\mathbf{u}_2^H. && (37)\end{aligned}$$

We take the inner products of (36) and (37) with $\mathbf{u}_2^*$ and $\mathbf{u}_1^*$, respectively, and sum these inner products up at both sides of equations (36) and (37). By doing so and noting the previous footnote 1 on a skew-symmetric matrix, we then obtain

$$(2M-1)\mathbf{u}_1^H\mathbf{u}_2^* = 0.$$

Thus, $\mathbf{u}_1^H\mathbf{u}_2^* = 0$. Thus, we have proved (ii).

If $M = 1/2$, by (35) we know that the minimum of $\|A\mathbf{x}+B\mathbf{y}\|^2$ on the unit ball is 1/2. Hence, $\|A\mathbf{x}+B\mathbf{y}\|^2 = \frac{1}{2}$ for any unit vector $(\mathbf{x}^t\ \mathbf{y}^t)$, which is equivalent to

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}^H (C^HC - \frac{1}{2}I) \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = 0,$$

for any unit vector $(\mathbf{x}^t\ \mathbf{y}^t)$, i.e., $C^HC = \frac{1}{2}I$. Therefore, (iii) is proved. **q.e.d.**

Lemma 5 *Let* $\mathbb{V}$ *be a subspace of* $\mathbb{C}^{2k}$ *with* $\dim(\mathbb{V}) = 2s$, *and* $\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}$ *and* $\begin{pmatrix} \mathbf{u}_2^* \\ \mathbf{u}_1^* \end{pmatrix}$ *be two orthonormal vectors in* $\mathbb{V}$, *where* $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{C}^k$. *Then, there exist* $s-1$ *unit vectors* $\begin{pmatrix} \mathbf{a}_2 \\ \mathbf{b}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{a}_3 \\ \mathbf{b}_3 \end{pmatrix}, \cdots, \begin{pmatrix} \mathbf{a}_s \\ \mathbf{b}_s \end{pmatrix}$ *in* $\mathbb{V}$ *such that*

$$\left\{ \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2^* \\ \mathbf{u}_1^* \end{pmatrix}, \begin{pmatrix} \mathbf{a}_2 \\ \mathbf{b}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{b}_2^* \\ \mathbf{a}_2^* \end{pmatrix}, \begin{pmatrix} \mathbf{a}_3 \\ \mathbf{b}_3 \end{pmatrix}, \begin{pmatrix} \mathbf{b}_3^* \\ \mathbf{a}_3^* \end{pmatrix}, \cdots, \begin{pmatrix} \mathbf{a}_s \\ \mathbf{b}_s \end{pmatrix}, \begin{pmatrix} \mathbf{b}_s^* \\ \mathbf{a}_s^* \end{pmatrix} \right\}$$

forms an orthonormal basis of $\mathbb{V}$.

Proof: First of all, it is easy to check that vectors $\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}$ and $\begin{pmatrix} \mathbf{v}_2^* \\ \mathbf{v}_1^* \end{pmatrix}$ are orthogonal if and only if $\mathbf{v}_1^t\mathbf{v}_2 = \mathbf{v}_2^t\mathbf{v}_1 = 0$.

We then use induction on s to prove the proposition. If $s = 1$, it is trivial. Assume that $s = s_1 < k$, the proposition is true. Let $s = s_1 + 1$. Let

$$\mathbb{V}_1 = \mathrm{Span}\{ \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2^* \\ \mathbf{u}_1^* \end{pmatrix}, \begin{pmatrix} \mathbf{a}_2 \\ \mathbf{b}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{b}_2^* \\ \mathbf{a}_2^* \end{pmatrix}, \cdots, \begin{pmatrix} \mathbf{a}_{s_1} \\ \mathbf{b}_{s_1} \end{pmatrix}, \begin{pmatrix} \mathbf{b}_{s_1}^* \\ \mathbf{a}_{s_1}^* \end{pmatrix} \}$$

and $\mathbb{U}_1 = \mathbb{V}_1^\perp$, the orthogonal complementary space of $\mathbb{V}_1$ in $\mathbb{V}$. Then, $\dim(\mathbb{U}_1) = 2s - 2s_1 = 2$.

Take two linear independent vectors $\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{y}_1 \end{pmatrix}$ and $\begin{pmatrix} \mathbf{x}_2 \\ \mathbf{y}_2 \end{pmatrix}$ in $\mathbb{U}_1$. We want to construct vectors $\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix}$ and $\begin{pmatrix} \mathbf{y}_0^* \\ \mathbf{x}_0^* \end{pmatrix}$ in $\mathbb{U}_1$ such that they are orthogonal.

If $\mathbf{x}_1^t\mathbf{y}_1 = 0$, let $\mathbf{x}_0 = \mathbf{x}_1, \mathbf{y}_0 = \mathbf{y}_1$; if $\mathbf{x}_2^t\mathbf{y}_2 = 0$, let $\mathbf{x}_0 = \mathbf{x}_2, \mathbf{y}_0 = \mathbf{y}_2$. If $\mathbf{x}_1^t\mathbf{y}_1 \neq 0$ and $\mathbf{x}_2^t\mathbf{y}_2 \neq 0$, we take a constant c as one of the solutions of the following quadratic equation

$$c^2\mathbf{x}_1^t\mathbf{y}_1 + (\mathbf{x}_2^t\mathbf{y}_1 + \mathbf{x}_1^t\mathbf{y}_2)c + \mathbf{x}_2^t\mathbf{y}_2 = 0$$

and let

$$\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix} = c\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{y}_1 \end{pmatrix} + \begin{pmatrix} \mathbf{x}_2 \\ \mathbf{y}_2 \end{pmatrix}.$$

Then, $0 \neq \begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix} \in \mathbb{U}_1$ and $\mathbf{x}_0^t\mathbf{y}_0 = 0$. By the argument in the beginning of the proof, vectors $\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix}$, and $\begin{pmatrix} \mathbf{y}_0^* \\ \mathbf{x}_0^* \end{pmatrix}$ are orthogonal. Since they are in the orthogonal complementary space of $\mathbb{V}_1$, they are orthogonal to all vectors

$$\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{u}_2^* \\ \mathbf{u}_1^* \end{pmatrix}, \begin{pmatrix} \mathbf{a}_2 \\ \mathbf{b}_2 \end{pmatrix}, \begin{pmatrix} \mathbf{b}_2^* \\ \mathbf{a}_2^* \end{pmatrix}, \begin{pmatrix} \mathbf{a}_3 \\ \mathbf{b}_3 \end{pmatrix}, \begin{pmatrix} \mathbf{b}_3^* \\ \mathbf{a}_3^* \end{pmatrix}, \cdots, \begin{pmatrix} \mathbf{a}_{s_1} \\ \mathbf{b}_{s_1} \end{pmatrix}, \begin{pmatrix} \mathbf{b}_{s_1}^* \\ \mathbf{a}_{s_1}^* \end{pmatrix}.$$

By normalizing vectors $\begin{pmatrix} \mathbf{x}_0 \\ \mathbf{y}_0 \end{pmatrix}$ and $\begin{pmatrix} \mathbf{y}_0^* \\ \mathbf{x}_0^* \end{pmatrix}$, and use the induction, we have proved the lemma. **q.e.d.**

The following lemma plays the key role in the proof of Proposition 2 and can be also thought of as an independent result in linear algebra on special SVD forms of special matrices.

Lemma 6 *Let A and B be two $p \times k$ matrices and satisfy conditions: $A^H A + B^t B^* = I$, and $A^H B$ and $B^H A$ are skew-symmetric. Then, there exist a unitary matrix V of size $p \times p$ and a unitary matrix U of size $2k \times 2k$ such that the $p \times 2k$ matrix $(A\ B)$ can be diagonalized as follows*

$$V(A\ B)U = \Sigma \triangleq \begin{pmatrix} D_\lambda & 0 & 0 & 0 \\ 0 & I_{k-s} & 0 & 0 \\ 0 & 0 & D_\mu & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_{p \times 2k}, \tag{38}$$

where $k - s \geq 2k - p$, $D_\lambda = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_s)$, $D_\mu = \mathrm{diag}(\mu_1, \mu_2, \cdots, \mu_s)$ and $\lambda_i^2 + \mu_i^2 = 1$, $1 > \lambda_i \geq \sqrt{1/2} \geq \mu_j > 0$, $i, j = 1, 2, \cdots, s$, $k + s = \kappa$, and $\kappa = \mathrm{rank}(A\ B) \geq k$, and furthermore the $2k \times 2k$ unitary matrix U has the following form

$$U = \begin{pmatrix} W_1 & W_2 \\ W_2^* & W_1^* \end{pmatrix}, \tag{39}$$

where W_i, $i = 1, 2$, are $k \times k$ matrices.

Note that the speciality of the above SVD of matrix $(A\ B)$ comes from the special form of U in (39) that may not hold for a SVD of a general matrix.

Proof. First of all, let us prove that $\kappa = \text{rank}((A\ B)) \geq k$. The conditions $A^H A + B^t B^* = I$, and $A^H B$ and $B^H A$ are skew-symmetric are equivalent to

$$\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}^H \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} = I_{2k}, \tag{40}$$

where I_{2k} is the $2k \times 2k$ identity matrix. Thus,

$$\text{rank}\left(\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}\right) = 2k,$$

which implies $\text{rank}((A\ B))+\text{rank}((B\ A)^*) \geq 2k$. Since the rank of $(A\ B)$ is the same as that of $(B\ A)^*$, we have $\kappa = \text{rank}((A\ B)) \geq k$. As a consequence, we have $p \geq \kappa \geq k$, which coincides with the result that the rate $R \leq 1$ obtained in [2, 3].

We next want to prove the diagonalization form of $(A\ B)$ in (38) by using the induction method on k.

At first, we consider the case when $k = 1$. Since $A^H B$ is skew-symmetric, $A^H B = -(A^H B)^t$ but $A^H B$ is a scalar number. Thus, $A^H B = 0$, i.e., $p \times 1$ vectors A and B are orthogonal each other.

If none of A and B is zero, then, $\|A\| > 0$ and $\|B\| > 0$. Moreover, by the condition, $\|A\|^2 + \|B\|^2 = 1$. Without loss of generality, we assume $\|A\| \geq \|B\|$. Thus, $1 > \|A\| \geq \sqrt{1/2} \geq \|B\| > 0$. Let $\mathbf{v}_i$, $i = 3, \cdots, p$, be $p \times 1$ unit vectors such that $V = (A/\|A\|\ B/\|B\|\ \mathbf{v}_3 \cdots \mathbf{v}_p)^H$ is a $p \times p$ unitary matrix. In this case,

$$V(A\ B) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \mu_1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix},$$

where $\lambda_1 = \|A\|$ and $\mu_1 = \|B\|$, $1 > \lambda_1 \geq \sqrt{1/2} \geq \mu_1 > 0$ and $\lambda_1^2 + \mu_1^2 = 1$. In other words, (38) with $U = I$ is proved.

If one of A and B is zero, without loss of generality, say $B = 0$, then, $\|A\| = 1$ because $A^H A + B^t B^* = I$. We let $\mathbf{v}_1 = A$ and $\mathbf{v}_i$, $i = 2, \cdots, p$, be $p \times 1$ unit vectors such that

$V = (\mathbf{v}_1\ \mathbf{v}_2\ \cdots\ \mathbf{v}_p)^H$ is a $p \times p$ unitary matrix. Then,

$$V(A\ B) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{pmatrix},$$

which has the form (38) with $U = I$.

Since $A^H A + B^t B^* = I$, A and B can not be both 0. Therefore, (38) is proved for the case when $k = 1$.

Assume that, for all matrices A and B of column sizes less than k and satisfying the conditions in the lemma, the decomposition (38) holds. We next want to show that, for matrices A and B of column size k and satisfying the conditions in the lemma, (38) still holds.

To do so, let $\mathcal{B}$ be the unit ball in $\mathbb{C}^{2k}$. Let $M = \max\{\|A\mathbf{x}+B\mathbf{y}\|^2 :\ (\mathbf{x}^t\ \mathbf{y}^t) \in \mathcal{B}\}$. Then, by Lemma 4 (i), we have $\frac{1}{2} \leq M \leq 1$.

Case 1: $M = \frac{1}{2}$.

In this case, by Lemma 4 (iii), $A^H A = B^H B = \frac{1}{2} I$ and $A^H B = 0$. We claim $p \geq 2k$. In fact, if $p < 2k$, then the equation $B^*\mathbf{x} + A^*\mathbf{y} = 0$ has at least a non-trivial solution for $\mathbf{x}$ and $\mathbf{y}$. Let $B^*\mathbf{x}_0 + A^*\mathbf{y}_0 = 0$ and $\|\mathbf{x}_0\|^2 + \|\mathbf{y}_0\|^2 = 1$. Then, by the relation (35), $\|A\mathbf{x}_0 + B\mathbf{y}_0\|^2 = 1$. This implies $M = 1$, which contradicts with the assumption $M = 1/2$.

Since $A^H A = B^H B = \frac{1}{2} I$ and $A^H B = 0$, it is possible to select a $p \times (p - 2k)$ matrix C such that $V = (\sqrt{2}A\ \sqrt{2}B, C)^H$ is a unitary matrix. Then,

$$V(A\ B) = \begin{pmatrix} \sqrt{2}A^H \\ \sqrt{2}B^H \\ C^H \end{pmatrix} (A\ B) = \begin{pmatrix} \sqrt{2}A^H A & \sqrt{2}A^H B \\ \sqrt{2}B^H A & \sqrt{2}B^H B \\ C^H A & C^H B \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} I & 0 \\ 0 & \frac{1}{\sqrt{2}} I \\ 0 & 0 \end{pmatrix},$$

which corresponds to (38) with $s = k$, $D_\lambda = D_\mu = \frac{1}{\sqrt{2}} I$, and $U = I_{2k}$. Therefore, in this case, the lemma is proved.

Case 2: $M > \frac{1}{2}$.

Let $(\mathbf{u}_1^t\ \mathbf{u}_2^t) \in \mathcal{B}$ such that $\|A\mathbf{u}_1 + B\mathbf{u}_2\|^2 = M$. Since $\mathbf{u}_1$ and $\mathbf{u}_2^*$ are orthogonal from Lemma 4 (ii), vectors $\begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix}$ and $\begin{pmatrix} \mathbf{u}_2^* \\ \mathbf{u}_1^* \end{pmatrix}$ are orthogonal too. Thus, by Lemma 5, there exist vectors

$$\begin{pmatrix} \mathbf{a}_2 \\ \mathbf{b}_2 \end{pmatrix}, \cdots, \begin{pmatrix} \mathbf{a}_k \\ \mathbf{b}_k \end{pmatrix} \in \mathbb{C}^{2k}$$

such that

$$\hat{U}_1 = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2^* & \mathbf{a}_2 & \mathbf{b}_2^* & \cdots & \mathbf{a}_k & \mathbf{b}_k^* \\ \mathbf{u}_2 & \mathbf{u}_1^* & \mathbf{b}_2 & \mathbf{a}_2^* & \cdots & \mathbf{b}_k & \mathbf{a}_k^* \end{pmatrix}$$

is a $2k \times 2k$ unitary matrix.

We now divide this case into two subcases to select the unitary matrix V.

Subcase 1: $\frac{1}{2} < M < 1$. In this case, let $M_1 = M$ and

$$\mathbf{v}_1 = \frac{1}{\sqrt{M}}(A\mathbf{u}_1 + B\mathbf{u}_2), \quad \mathbf{v}_2 = \frac{1}{\sqrt{1-M}}(A\mathbf{u}_2^* + B\mathbf{u}_1^*).$$

It is not hard to see that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal to each other because $\mathbf{u}_1^t\mathbf{u}_2 = 0$. Then, there exist $p-2$ vectors $\mathbf{v}_3, \cdots, \mathbf{v}_p$ such that $V_1 \triangleq (\mathbf{v}_1\ \mathbf{v}_2\ \mathbf{v}_3 \cdots \mathbf{v}_p)^H$ is a $p \times p$ unitary matrix. Therefore,

$$\begin{aligned}
\left(\hat{A}_1\ \hat{B}_1\right) &\triangleq V_1(A\ B)\hat{U}_1 \\
&= \begin{pmatrix} \mathbf{v}_1^H \\ \mathbf{v}_2^H \\ \vdots \\ \mathbf{v}_p^H \end{pmatrix} \begin{pmatrix} A\mathbf{u}_1 + B\mathbf{u}_2 & A\mathbf{u}_2^* + B\mathbf{u}_1^* & \cdots & A\mathbf{a}_k + B\mathbf{b}_k & A\mathbf{b}_k^* + B\mathbf{a}_k^* \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{v}_1^H \\ \mathbf{v}_2^H \\ \vdots \\ \mathbf{v}_p^H \end{pmatrix} \begin{pmatrix} \sqrt{M}\mathbf{v}_1 & \sqrt{1-M}\mathbf{v}_2 & A\mathbf{a}_2 + B\mathbf{b}_2 & \cdots & A\mathbf{b}_k^* + B\mathbf{a}_k^* \end{pmatrix} \\
&= \begin{pmatrix} \sqrt{M} & 0 & \mathbf{v}_1^H(A\mathbf{a}_2 + B\mathbf{b}_2) & \cdots & \mathbf{v}_1^H(A\mathbf{b}_k^* + B\mathbf{a}_k^*) \\ 0 & \sqrt{1-M} & \mathbf{v}_2^H(A\mathbf{a}_2 + B\mathbf{b}_2) & \cdots & \mathbf{v}_2^H(A\mathbf{b}_k^* + B\mathbf{a}_k^*) \\ 0 & 0 & & T & \end{pmatrix}
\end{aligned}$$

where T is a $(p-2) \times (2k-2)$ matrix. Let

$$\mathbf{z}_1 = \begin{pmatrix} \mathbf{v}_1^H(A\mathbf{a}_2 + B\mathbf{b}_2) \\ \vdots \\ \mathbf{v}_1^H(A\mathbf{b}_k^* + B\mathbf{a}_k^*) \end{pmatrix}_{(2k-2)\times 1}, \quad \mathbf{z}_2 = \begin{pmatrix} \mathbf{v}_2^H(A\mathbf{a}_2 + B\mathbf{b}_2) \\ \vdots \\ \mathbf{v}_2^H(A\mathbf{b}_k^* + B\mathbf{a}_k^*) \end{pmatrix}_{(2k-2)\times 1}$$

and consider the unit vector

$$\xi_1 = \left(\frac{1}{M + \mathbf{z}_1^H\mathbf{z}_1}\right)^{1/2} \begin{pmatrix} \sqrt{M} \\ 0 \\ \mathbf{z}_1 \end{pmatrix},$$

then

$$\left\|(A\ B)\hat{U}_1\xi_1\right\|^2 = \left\|V_1^H(A\ B)\hat{U}_1\xi_1\right\|^2 = \left\|(\hat{A}_1\ \hat{B}_1)\xi_1\right\|^2$$
$$= (\frac{1}{M+\mathbf{z}_1^H\mathbf{z}_1})\left[(M+\mathbf{z}_1^H\mathbf{z}_1)^2 + \|T\mathbf{z}_1\|^2\right] \geq M+\mathbf{z}_1^H\mathbf{z}_1,$$

where $\|\hat{U}_1\xi_1\| = \|\xi_1\| = 1$, i.e., $\hat{U}_1\xi_1 \in \mathcal{B}$. Since M is the maximal value of $\|(A\ B)\mathbf{x}\|$ for $\mathbf{x} \in \mathcal{B}$, we conclude that $\mathbf{z}_1 = 0$, that is,

$$\mathbf{v}_1^H(A\mathbf{a}_i + B\mathbf{b}_i) = 0, \quad \mathbf{v}_1^H(A\mathbf{b}_i^* + B\mathbf{a}_i^*) = 0,\ 2 \leq i \leq k. \tag{41}$$

Since

$$\begin{aligned}
\sqrt{1-M}\mathbf{v}_2^H(A\mathbf{a}_i + B\mathbf{b}_i) &= (A\mathbf{u}_2^* + B\mathbf{u}_1^*)^H(A\mathbf{a}_i + B\mathbf{b}_i) \\
&= \mathbf{u}_2^t A^H A\mathbf{a}_i + \mathbf{u}_1^t B^H A\mathbf{a}_i + \mathbf{u}_2^t A^H B\mathbf{b}_i + \mathbf{u}_1^t B^H B\mathbf{b}_i \\
&= \mathbf{u}_2^t\mathbf{a}_i + \mathbf{u}_1^t\mathbf{b}_i - \mathbf{u}_2^t B^t B^*\mathbf{a}_i - \mathbf{u}_1^t A^t A^*\mathbf{b}_i - \mathbf{u}_1^t A^t B^*\mathbf{a}_i - \mathbf{u}_2^t B^t A^*\mathbf{b}_i \\
&= -(A\mathbf{u}_1 + B\mathbf{u}_2)^t(A\mathbf{b}_i^* + B\mathbf{a}_i^*)^* = \sqrt{M}(\mathbf{v}_1^H(A\mathbf{b}_i^* + B\mathbf{a}_i^*))^*,
\end{aligned}$$

and $\mathbf{v}_1^H(A\mathbf{b}_i^* + B\mathbf{a}_i^*) = 0$, we have

$$\mathbf{v}_2^H(A\mathbf{a}_i + B\mathbf{b}_i) = 0, \quad 2 \leq i \leq k, \tag{42}$$

and similarly,

$$\mathbf{v}_2^H(A\mathbf{b}_i^* + B\mathbf{a}_i^*) = 0, \quad 2 \leq i \leq k. \tag{43}$$

In other words, we have $\mathbf{z}_2 = 0$. Therefore,

$$(\hat{A}_1\ \hat{B}_1) = \begin{pmatrix} \sqrt{M} & 0 & 0 \\ 0 & \sqrt{1-M} & 0 \\ 0 & 0 & T \end{pmatrix}.$$

Due to the particular form of $\hat{U}_1$, there is a permutation matrix P such that

$$\hat{U}_1 P = \begin{pmatrix} \mathbf{u}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_k & \mathbf{u}_2^* & \mathbf{b}_2^* & \cdots & \mathbf{b}_k^* \\ \mathbf{u}_2 & \mathbf{b}_2 & \cdots & \mathbf{b}_k & \mathbf{u}_1^* & \mathbf{a}_2^* & \cdots & \mathbf{a}_k^* \end{pmatrix}.$$

Let $U_1 = \hat{U}_1 P$. Then, it has the form in (39). Furthermore,

$$V_1(A\ B)U_1 = \begin{pmatrix} \sqrt{M} & 0 & 0 \\ 0 & \sqrt{1-M} & 0 \\ 0 & 0 & T \end{pmatrix} P = \begin{pmatrix} \sqrt{M} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{1-M} & 0 \\ 0 & A_1 & 0 & B_1 \end{pmatrix}, \tag{44}$$

where

$$A_1 = \begin{pmatrix} \mathbf{v}_3^H \\ \vdots \\ \mathbf{v}_p^H \end{pmatrix} (A\ B) \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix}, \quad B_1 = \begin{pmatrix} \mathbf{v}_3^H \\ \vdots \\ \mathbf{v}_p^H \end{pmatrix} (A\ B) \begin{pmatrix} \mathbf{b}_2^* & \cdots & \mathbf{b}_k^* \\ \mathbf{a}_2^* & \cdots & \mathbf{a}_k^* \end{pmatrix}$$

are $(p-2) \times (k-1)$ matrices.

We next want to show $A_1^H A_1 + B_1^t B_1^* = I$, and $A_1^H B_1$ and $B_1^H A_1$ are skew-symmetric, i.e., A_1 and B_1 have the properties as A and B do, while their column sizes are $k-1$. Note that

$$(\mathbf{v}_3 \cdots \mathbf{v}_p)(\mathbf{v}_3 \cdots \mathbf{v}_p)^H = I - \mathbf{v}_1 \mathbf{v}_1^H - \mathbf{v}_2 \mathbf{v}_2^H,$$

and $\mathbf{z}_1 = 0$, and $\mathbf{z}_2 = 0$, we have

$$\begin{aligned}
& A_1^H A_1 + B_1^t B_1^* \\
= & \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix}^H (A\ B)^H (\mathbf{v}_3 \cdots \mathbf{v}_p) \begin{pmatrix} \mathbf{v}_3^H \\ \vdots \\ \mathbf{v}_p^H \end{pmatrix} (A\ B) \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix} \\
& + \begin{pmatrix} \mathbf{b}_2 & \cdots & \mathbf{b}_k \\ \mathbf{a}_2 & \cdots & \mathbf{a}_k \end{pmatrix}^H (A\ B)^t (\mathbf{v}_3^* \cdots \mathbf{v}_p^*) \begin{pmatrix} \mathbf{v}_3^t \\ \vdots \\ \mathbf{v}_p^t \end{pmatrix} (A^*\ B^*) \begin{pmatrix} \mathbf{b}_2 & \cdots & \mathbf{b}_k \\ \mathbf{a}_2 & \cdots & \mathbf{a}_k \end{pmatrix} \\
= & \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix}^H (A\ B)^H (I - \mathbf{v}_1 \mathbf{v}_1^H - \mathbf{v}_2 \mathbf{v}_2^H)(A\ B) \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix} \\
& + \begin{pmatrix} \mathbf{b}_2 & \cdots & \mathbf{b}_k \\ \mathbf{a}_2 & \cdots & \mathbf{a}_k \end{pmatrix}^H (A\ B)^t (I - \mathbf{v}_1 \mathbf{v}_1^H - \mathbf{v}_2 \mathbf{v}_2^H)^* (A^*\ B^*) \begin{pmatrix} \mathbf{b}_2 & \cdots & \mathbf{b}_k \\ \mathbf{a}_2 & \cdots & \mathbf{a}_k \end{pmatrix} \\
= & \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix}^H (A\ B)^H (A\ B) \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix} \\
& + \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix}^H \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} (A\ B)^t (A^*\ B^*) \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix} \\
= & I,
\end{aligned}$$

where the third equation is from (41), (42), and (43). Similarly, we have $A_1^H B_1 + B_1^t A_1^* = 0$ and $B_1^H A_1 + A_1^t B_1^* = 0$, that is, $A_1^H B_1$ and $B_1^H A_1$ are skew-symmetric.

By using the induction assumption to A_1 and B_1 of size $(p-2) \times (k-1)$, they have the decomposition (38) with k replaced by $k-1$, i.e., there exist a $(p-2) \times (p-2)$ unitary matrix V_2 and a $2(k-1)\times 2(k-1)$ unitary matrix U_2 of the form in (39) such that

$$V_2(A_1\ B_1)U_2 = \begin{pmatrix} D_\lambda & 0 & 0 & 0 \\ 0 & I_{k-1-s_1} & 0 & 0 \\ 0 & 0 & D_\mu & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{45}$$

where $D_\lambda = \mathrm{diag}(\lambda_2, \cdots, \lambda_{s_1+1})$, $D_\mu = \mathrm{diag}(\mu_2, \cdots, \mu_{s_1+1})$, $1 > \lambda_i \geq \sqrt{1/2} \geq \mu_j > 0$, and $\lambda_i^2 + \mu_i^2 = 1$, $2 \leq i, j \leq s_1 + 1$. Let

$$U_2 = \begin{pmatrix} U_{21} & U_{22} \\ U_{22}^* & U_{21}^* \end{pmatrix}, \quad V_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & V_2 \end{pmatrix}, \quad \text{and} \quad U_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & U_{21} & 0 & U_{22} \\ 0 & 0 & 1 & 0 \\ 0 & U_{22}^* & 0 & U_{21}^* \end{pmatrix},$$

where V_3 and U_3 are clearly $p \times p$ and $2k \times 2k$ unitary matrices, respectively, and U_3 also has the form in (39). Furthermore, by using (44) and some calculations, it is not hard to see

$$V_3V_1(A\ B)U_1U_3 = \Sigma_1 \triangleq \begin{pmatrix} \sqrt{M} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{1-M} & 0 & 0 \\ 0 & D_\lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{k-1-s_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & D_\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{46}$$

Let P_1 be a row permutation that moves the second row of Σ_1 in (46) to the $(k+1)$th row, then $P_1V_3V_1(A\ B)U_1U_3$ has the form in (38) with $s = s_1 + 1$ and $1 > \lambda_1 = \sqrt{M} \geq \sqrt{1/2}$ and $\sqrt{1/2} \geq \mu_1 = \sqrt{1-M} > 0$ and the rest are the same as in (45). It is easy to check that the form of any product of matrices of the form in (39) remains. If we let $V = P_1V_3V_1$ and $U = U_1U_3$, then (38) is achieved for A and B of column size k.

Subcase 2: $M = 1$. In this case, $\mathbf{v}_1$ is chosen as above, i.e.,

$$\mathbf{v}_1 = \frac{1}{\sqrt{M}}(A\mathbf{u}_1 + B\mathbf{u}_2).$$

Take vectors $\mathbf{v}_2, \cdots, \mathbf{v}_p$ such that $V_1 = (\mathbf{v}_1\ \mathbf{v}_2\ \cdots\ \mathbf{v}_p)^H$ is a $p \times p$ unitary matrix.

As in Subcase 1 and use the same $\hat{U}_1$ as in Subcase 1, we obtain

$$V_1(A\ B)\hat{U}_1 = \begin{pmatrix} \sqrt{M} & 0 & 0 \\ 0 & 0 & T \end{pmatrix},$$

where T has size $(p-1)\times(2k-2)$. Similar to (44) and use the same U_1 as in Subcase 1, we have

$$V_1(A\ B)U_1 = \begin{pmatrix} \sqrt{M} & 0 & 0 & 0 \\ 0 & A_1 & 0 & B_1 \end{pmatrix}, \tag{47}$$

where

$$A_1 = \begin{pmatrix} \mathbf{v}_2^H \\ \mathbf{v}_3^H \\ \vdots \\ \mathbf{v}_p^H \end{pmatrix} (A\ B) \begin{pmatrix} \mathbf{a}_2 & \cdots & \mathbf{a}_k \\ \mathbf{b}_2 & \cdots & \mathbf{b}_k \end{pmatrix}, \quad B_1 = \begin{pmatrix} \mathbf{v}_2^H \\ \mathbf{v}_3^H \\ \vdots \\ \mathbf{v}_p^H \end{pmatrix} (A\ B) \begin{pmatrix} \mathbf{b}_2^* & \cdots & \mathbf{b}_k^* \\ \mathbf{a}_2^* & \cdots & \mathbf{a}_k^* \end{pmatrix}$$

are $(p-1)\times(k-1)$ matrices. Similar to the proof in Subcase 1, it can also be shown that $A_1^H A_1 + B_1^t B_1^* = I$, and $A_1^H B_1$ and $B_1^H A_1$ are skew-symmetric, i.e., matrices A_1 and B_1 have the same properties as A and B do but of column sizes $k-1$.

By using the induction assumption to A_1 and B_1 of size $(p-1)\times(k-1)$, they have the decomposition (38) with k replaced by $k-1$, i.e., there exist $(p-1)\times(p-1)$ unitary matrices V_2 and $2(k-1)\times 2(k-1)$ unitary matrices U_2 of the form (39) such that

$$V_2(A_1\ B_1)U_2 = \begin{pmatrix} D_\lambda & 0 & 0 & 0 \\ 0 & I_{k-1-s} & 0 & 0 \\ 0 & 0 & D_\mu & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{48}$$

where $D_\lambda = \operatorname{diag}(\lambda_1, \cdots, \lambda_s)$, $D_\mu = \operatorname{diag}(\mu_1, \cdots, \mu_s)$, $1 > \lambda_i \geq \sqrt{1/2} \geq \mu_j > 0$, and $\lambda_i^2 + \mu_i^2 = 1$, $1 \leq i, j \leq s$. Let

$$U_2 = \begin{pmatrix} U_{21} & U_{22} \\ U_{22}^* & U_{21}^* \end{pmatrix}, \quad V_3 = \begin{pmatrix} 1 & 0 \\ 0 & V_2 \end{pmatrix}, \quad \text{and} \quad U_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & U_{21} & 0 & U_{22} \\ 0 & 0 & 1 & 0 \\ 0 & U_{22}^* & 0 & U_{21}^* \end{pmatrix},$$

where V_3 and U_3 are clearly $p\times p$ and $2k\times 2k$ unitary matrices, respectively, and U_3 also has the form in (39). Furthermore, by using (47) and some calculations, it is not hard to see

$$V_3V_1(A\ B)U_1U_3 = \Sigma_1 \triangleq \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & D_\lambda & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{k-1-s} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & D_\mu & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{49}$$

We now want to move the element 1 from the position at the first row and the first column to the position at the kth row and the kth column, and the element 0 from the position at the first row and the $(k+1)$th column to the position at the kth row and $2k$th column. Let P_1 be a row permutation that moves the first row of Σ_1 in (49) to the kth row. Let P_2 be the column permutation that moves the first column of Σ_1 to the kth column and the

$(k+1)$th column to the $2k$th column. Clearly, P_2 has the form in (39). By doing so, it is not hard to check that $P_1V_3V_1(A\ B)U_1U_3P_2$ has the form in (38). If we let $V = P_1V_3V_1$ and $U = U_1U_3P_2$. Similar to Subcase 1, U has the form (39) and therefore (38) is achieved for A and B of column size k.

By summarizing the above cases and using the induction, Lemma 6 is proved. **q.e.d.**

As a consequence of Lemma 6, if the rank of $(A\ B)$ in Lemma 6 is k, then $s = 0$ in (38) and therefore, all the diagonal elements are 1, i.e., all singular values of $(A\ B)$ are 1. Another remark is that, when $p = k$, i.e., A and B are square, then the above proof can be simplified as follows. When $p = k$, the matrix $\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}$ is square. Then, the two matrices in the product in the left hand side of (40) can be exchanged, i.e.,

$$\begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix} \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}^H = I_{2k}.$$

In this case, if we take $U = \begin{pmatrix} A & B \\ B^* & A^* \end{pmatrix}^H$ that certainly has the form of (39), then (38) is proved.

We next want to make a transformation to the variables of a generalized complex orthogonal design.

Lemma 7 *Let*

$$\mathcal{G} = (A_1\mathbf{x} + B_1\mathbf{x}^*\ A_2\mathbf{x} + B_2\mathbf{x}^* \cdots A_n\mathbf{x} + B_n\mathbf{x}^*)$$

be a generalized complex orthogonal design and matrix $\begin{pmatrix} W_1 & W_2 \\ W_2^* & W_1^* \end{pmatrix}$ *and matrix* V *be unitary. Make the transformation* $\mathbf{x} = W_1\mathbf{y} + W_2\mathbf{y}^*$ *and let* $\tilde{A}_i = VA_iW_1 + VB_iW_2^*$ *and* $\tilde{B}_i = VA_iW_2 + VB_iW_1^*$*, then*

$$\tilde{A}_i^H\tilde{A}_j + \tilde{B}_j^t\tilde{B}_i^* = \delta_{ij}\tilde{E}_i,\ \ 1 \le i, j \le n,$$

and

$$\tilde{A}_i^H\tilde{B}_j + \tilde{B}_j^t\tilde{A}_i^* = 0,\ \ \tilde{B}_i^H\tilde{A}_j + \tilde{A}_j^t\tilde{B}_i^* = 0 \quad 1 \le i \ne j \le n, \tag{50}$$

where $\tilde{E}_i = W_1^HE_iW_1 + W_2^tE_iW_2^*$ *are positive definite. In other words,* $\{(\tilde{A}_1\ \tilde{B}_1), \cdots,$ $(\tilde{A}_n\ \tilde{B}_n)\}$ *form a Hurwitz family. In particular, if* $\mathcal{G}$ *is a complex orthogonal design, then its* *transformation* $\tilde{\mathcal{G}} = (\tilde{A}_1\mathbf{x} + \tilde{B}_1\mathbf{x}^* \ \tilde{A}_2\mathbf{x} + \tilde{B}_2\mathbf{x}^* \ \cdots \ \tilde{A}_n\mathbf{x} + \tilde{B}_n\mathbf{x}^*)$ *is also a complex orthogonal design.*

Proof: It is enough to notice that

$$\begin{pmatrix} \tilde{A}_i & \tilde{B}_i \\ \tilde{B}_j^* & \tilde{A}_j^* \end{pmatrix} = \begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix} \begin{pmatrix} A_i & B_i \\ B_j^* & A_j^* \end{pmatrix} \begin{pmatrix} W_1 & W_2 \\ W_2^* & W_1^* \end{pmatrix}. \quad \textbf{q.e.d.}$$

As a remark, if $E_i \neq I$, then $\tilde{A}_i^H \tilde{B}_i + \tilde{B}_i^t \tilde{A}_i^*$ may not be 0, which is the reason why condition $i \neq j$ in (6) in Definition 1 for a Hurwitz family is required. On the other hand, by reviewing Proposition 1, condition $\tilde{A}_i^H \tilde{B}_i + \tilde{B}_i^t \tilde{A}_i^* = 0$ is crucial for a generalized complex orthogonal design as in (3).

Proof of Proposition 2.

Proposition 2 is a direct consequence of Lemma 3, 6, 7. **q.e.d.**

When we consider restricted generalized complex orthogonal designs, Lemma 1 can be sharpened as follows.

Lemma 8 *Let* $\mathcal{A}_1, \mathcal{A}_2, \cdots, \mathcal{A}_k$ *be* k *complex alphabet sets. Let* A, B, C *be three* $k \times k$ *complex matrices such that, for any* $\mathbf{z} = (z_1, z_2, \cdots, z_k)^t \in (\mathcal{A}_1 \times \mathcal{A}_2 \times \cdots \times \mathcal{A}_k)^t$, *the following holds:*

$$\mathbf{z}^H A\mathbf{z} + \mathbf{z}^H B\mathbf{z}^* + \mathbf{z}^t C\mathbf{z} = 0. \tag{51}$$

(i) If for any $1 \leq i \leq k$, $\mathcal{A}_i$ *satisfies the condition (i) in Theorem 3, i.e., none of the alphabet sets contains only collinear points, then matrices* A, $B + B^t$ *and* $C + C^t$ *are all diagonal;*

(ii) If for any $1 \leq i \leq k$, $\mathcal{A}_i$ *satisfies the conditions (i), (ii) and (iii) in Theorem 3, then*

$$A = B + B^t = C + C^t = 0; \tag{52}$$

(iii) If for any $1 \leq i \leq k$, $\mathcal{A}_i$ *satisfies the condition (i) in Theorem 3 with three non-zero points and* $0 \in \mathcal{A}_i$, *then, (52) holds.*

Proof. We first prove (i). Denote $A = (a_{ij})_{k\times k}, B = (b_{ij})_{k\times k}, C = (c_{ij})_{k\times k}$. For any $j \in \{1, 2, \cdots, k\}$, since $\mathcal{A}_j$ is admissible, there exist three points z_1^j, z_2^j, z_3^j, which are not on a straight line of the complex plane. Let $x_1^j = z_1^j - z_2^j, x_2^j = z_2^j - z_3^j, x_3^j = z_3^j - z_1^j$. Then, (28) and (29) hold for x_1^j, x_2^j, x_3^j. Furthermore, for any $s_j \in \{1,2,3\}$, if we let $\mathbf{z} = (z_{s_1}^1, z_{s_2}^2, \cdots, z_{s_k}^k)^t$, then by the condition of this lemma, we have

$$\mathbf{z}^H A\mathbf{z} + \mathbf{z}^H B\mathbf{z}^* + \mathbf{z}^t C\mathbf{z} = 0.$$

Expanding the above equation, we obtain

$$\sum_{i,j=1}^{k} a_{ij}(z_{s_i}^i)^* z_{s_j}^j + \sum_{i,j=1}^{k} b_{ij}(z_{s_i}^i)^*(z_{s_j}^j)^* + \sum_{i,j=1}^{k} c_{ij} z_{s_i}^i z_{s_j}^j = 0. \tag{53}$$

Rewritting (53), we have

$$\begin{aligned}
& a_{11}(z_{s_1}^1)^* z_{s_1}^1 + \left(\sum_{j=2}^{k} a_{1j} z_{s_j}^j\right)(z_{s_1}^1)^* + \left(\sum_{j=2}^{k} a_{j1}(z_{s_j}^j)^*\right) z_{s_1}^1 + \sum_{i,j=2}^{k} a_{ij}(z_{s_i}^i)^* z_{s_j}^j \\
+ \;& b_{11}(z_{s_1}^1)^*(z_{s_1}^1)^* + \left(\sum_{j=2}^{k}(b_{1j}+b_{j1})(z_{s_j}^j)^*\right)(z_{s_1}^1)^* + \sum_{i,j=2}^{k} b_{ij}(z_{s_i}^i)^*(z_{s_j}^j)^* \\
+ \;& c_{11} z_{s_1}^1 z_{s_1}^1 + \left(\sum_{j=2}^{k}(c_{1j}+c_{j1}) z_{s_j}^j\right) z_{s_1}^1 + \sum_{i,j=2}^{k} c_{ij} z_{s_i}^i z_{s_j}^j = 0.
\end{aligned} \tag{54}$$

If we fix the indices $s_2, s_3, \cdots, s_k$, let $s_1 = 1, 2, 3$ and take the difference equations of three equations from (54) corresponding to $s_1 = 1, 2, 3$, we have

$$\begin{aligned}
& a_{11}(z_u^1)^* z_u^1 - a_{11}(z_v^1)^* z_v^1 + \left(\sum_{j=2}^{k} a_{1j} z_{s_j}^j\right)(z_u^1 - z_v^1)^* + \left(\sum_{j=2}^{k} a_{j1}(z_{s_j}^j)^*\right)(z_u^1 - z_v^1) \\
+ \;& b_{11}((z_u^1)^2 - (z_v^1)^2)^* + \left(\sum_{j=2}^{k}(b_{1j}+b_{j1})(z_{s_j}^j)^*\right)(z_u^1 - z_v^1)^* \\
+ \;& c_{11}((z_u^1)^2 - (z_v^1)^2) + \left(\sum_{j=2}^{k}(c_{1j}+c_{j1}) z_{s_j}^j\right)(z_u^1 - z_v^1) = 0,
\end{aligned} \tag{55}$$

where $u \neq v$ and $u, v \in \{1,2,3\}$. Therefore,

$$\begin{aligned}
& \sum_{j=2}^{k}\left(a_{j1}(z_{s_j}^j)^* + (c_{1j}+c_{j1}) z_{s_j}^j\right)(z_u^1 - z_v^1) + \sum_{j=2}^{k}\left(a_{1j} z_{s_j}^j + (b_{1j}+b_{j1})(z_{s_j}^j)^*\right)(z_u^1 - z_v^1)^* \\
= \;& -\left(a_{11}(z_u^1)^* z_u^1 - a_{11}(z_v^1)^* z_v^1 + b_{11}((z_u^1)^2 - (z_v^1)^2)^* + c_{11}((z_u^1)^2 - (z_v^1)^2)\right).
\end{aligned} \tag{56}$$

By letting $(u,v) = (1,2)$ and $(u,v) = (2,3)$ in (56) and the definitions of x_1^1, x_2^1 and x_3^1, we have, respectively,

$$\sum_{j=2}^{k}\left(a_{j1}(z_{s_j}^j)^* + (c_{1j}+c_{j1}) z_{s_j}^j\right) x_1^1 + \sum_{j=2}^{k}\left(a_{1j} z_{s_j}^j + (b_{1j}+b_{j1})(z_{s_j}^j)^*\right)(x_1^1)^*$$

$$= -\left(a_{11}(z_1^1)^* z_1^1 - a_{11}(z_2^1)^* z_2^1 + b_{11}((z_1^1)^2 - (z_2^1)^2)^* + c_{11}((z_1^1)^2 - (z_2^1)^2)\right), \tag{57}$$

and

$$\sum_{j=2}^{k}\left(a_{j1}(z_{s_j}^j)^* + (c_{1j}+c_{j1})z_{s_j}^j\right)x_2^1 + \sum_{j=2}^{k}\left(a_{1j}z_{s_j}^j + (b_{1j}+b_{j1})(z_{s_j}^j)^*\right)(x_2^1)^*$$
$$= -\left(a_{11}(z_2^1)^* z_2^1 - a_{11}(z_3^1)^* z_3^1 + b_{11}((z_2^1)^2 - (z_3^1)^2)^* + c_{11}((z_2^1)^2 - (z_3^1)^2)\right). \tag{58}$$

Since x_1^1, $(x_1^1)^*$, x_2^1 and $(x_2^1)^*$ are all functions of z_i^1 and $(z_i^1)^*$ for $i = 1, 2, 3$, and the coefficient matrix in linear system (57)-(58)

$$\begin{pmatrix} x_1^1 & (x_1^1)^* \\ x_2^1 & (x_2^1)^* \end{pmatrix}$$

has full rank from (29) from the admissibility of the alphabet set $\mathcal{A}_i$ with index $i = 1$, the solution of the linear system has the following form

$$\sum_{j=2}^{k}\left(a_{j1}(z_{s_j}^j)^* + (c_{1j}+c_{j1})z_{s_j}^j\right) = f(z_1^1, z_2^1, z_3^1, (z_1^1)^*, (z_2^1)^*, (z_3^1)^*), \tag{59}$$

$$\sum_{j=2}^{k}\left(a_{1j}z_{s_j}^j + (b_{1j}+b_{j1})(z_{s_j}^j)^*\right) = g(z_1^1, z_2^1, z_3^1, (z_1^1)^*, (z_2^1)^*, (z_3^1)^*), \tag{60}$$

where f and g are two functions. Equation (59) can be rewritten as

$$a_{21}(z_{s_2}^2)^* + (c_{12}+c_{21})z_{s_2}^2 + \sum_{j=3}^{k}\left(a_{j1}(z_{s_j}^j)^* + (c_{1j}+c_{j1})z_{s_j}^j\right)$$
$$= f(z_1^1, z_2^1, z_3^1, (z_1^1)^*, (z_2^1)^*, (z_3^1)^*). \tag{61}$$

Since the above equation holds for $s_2 = 1, 2, 3$, we fix s_j for $j \geq 3$ and let $s_2 = 1$, $s_2 = 2$ and $s_2 = 3$ and we then take the differences of these equations to obtain

$$a_{21}(x_1^2)^* + (c_{12}+c_{21})x_1^2 = 0, \tag{62}$$

$$a_{21}(x_2^2)^* + (c_{12}+c_{21})x_2^2 = 0, \tag{63}$$

where, as a remark, ^{2} is index not power. By (29) from the admissibility of the alphabet set $\mathcal{A}_2$ with index $i = 2$, we obtain $a_{21} = 0$ and $c_{12} + c_{21} = 0$. Similarly, we can obtain $a_{j1} = 0$ and $c_{1j} + c_{j1} = 0$ for $j \geq 3$.

By using equation (60), we can similarly derive $a_{1j} = 0$ and $b_{1j} + b_{j1} = 0$ for $j \geq 2$. This proves (i).

We next prove (ii). From (i), we only need to prove the diagonal elements of A, $B+B^t$ and $C+C^t$ are all zero.

For $\mathcal{A}_1$, from condition (ii) in Theorem 3, there are two points $z_1 = p + q\mathbf{j}$ with $pq \neq 0$ and $z_2 = z_1^* \neq z_1$ such that $z_1, z_2 \in \mathcal{A}_1$. Thus, from (56) and $a_{ij} = b_{ij} + b_{ji} = c_{ij} + c_{ji} = 0$ for $i \neq j$, by evaluating (56) for $z_{s_i}^1 = z_i$ for $i = 1, 2$, we have

$$a_{11} z_1^* z_1 - a_{11} (z_1^*)^* z_1^* + b_{11}((z_1)^2 - (z_1^*)^2)^* + c_{11}((z_1)^2 - (z_1^*)^2) = 0, \tag{64}$$

which implies $b_{11} = c_{11}$. Similarly, by evaluating (56) for $z_{s_i}^1 = z_i$ for $i = 1, 2, 3$ for the three points in the condition (iii) in Theorem 3, we have

$$a_{11}(z_u)^* z_u - a_{11}(z_v)^* z_v + b_{11}((z_u)^2 - (z_v)^2)^* + c_{11}((z_u)^2 - (z_v)^2) = 0, \tag{65}$$

for $u, v = 1, 2, 3$ and $u \neq v$. By using $b_{11} = c_{11}$ in the above equation we have

$$\begin{aligned} (|z_1|^2 - |z_2|^2) a_{11} + \mathrm{Re}((z_1)^2 - (z_2)^2) b_{11} &= 0, \\ (|z_1|^2 - |z_3|^2) a_{11} + \mathrm{Re}((z_1)^2 - (z_3)^2) b_{11} &= 0. \end{aligned}$$

The coefficient matrix is

$$\begin{aligned} \begin{pmatrix} |z_1|^2 - |z_2|^2 & \mathrm{Re}((z_1)^2 - (z_2)^2) \\ |z_1|^2 - |z_3|^2 & \mathrm{Re}((z_1)^2 - (z_3)^2) \end{pmatrix} &= \begin{pmatrix} p_1^2 - p_2^2 + (q_1^2 - q_2^2) & p_1^2 - p_2^2 - (q_1^2 - q_2^2) \\ p_1^2 - p_3^2 + (q_1^2 - q_3^2) & p_1^2 - p_3^2 - (q_1^2 - q_3^2) \end{pmatrix} \\ &= \begin{pmatrix} p_1^2 - p_2^2 & q_1^2 - q_2^2 \\ p_1^2 - p_3^2 & q_1^2 - q_3^2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \end{aligned}$$

Clearly, when the condition (iii) in Theorem 3 holds, the above coefficient matrix has full rank and therefore $a_{11} = b_{11} = 0$ and $c_{11} = b_{11} = 0$. The others $a_{jj} = b_{jj} = c_{jj} = 0$ for $j > 1$ can be similarly proved. This proves (ii).

We now prove (iii). Since $x_s^j = z_s^j - 0 \neq 0$, we let $k_s^j = (x_s^j)^* / x_s^j$. From (54) and the result in (i) we have

$$a_{jj} + b_{jj} k_s^j + c_{jj} (k_s^j)^{-1} = 0, \quad s = 1, 2, 3.$$

Therefore,

$$b_{jj} (k_s^j)^2 + a_{jj} k_s^j + c_{jj} = 0, \quad s = 1, 2, 3.$$

Thus, k_s^j, $s = 1, 2, 3$, can be regarded as the solutions of the equation $b_{jj}x^2 + a_{jj}x + c_{jj} = 0$. This equation is at most quadratic and therefore it has at most two distinct solutions unless all the coefficients of the equation are zero. If not all the coefficients are zero, then there exist $s_1, s_2 \in \{1, 2, 3\}$ such that $k_{s_1}^j = k_{s_2}^j$. Without loss of generality, we may assume $s_1 = 1, s_2 = 2$. Hence, $(x_1^j)^*/x_1^j = (x_2^j)^*/x_2^j$, which contradicts with (29). This proves that all the coefficients in equation $b_{jj}x^2 + a_{jj}x + c_{jj} \equiv 0$ are 0, i.e., $a_{jj} = b_{jj} = c_{jj} = 0$. This proves (iii). **q.e.d.**

References

[1] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.,* vol. 16, no. 8, pp. 1451-1458, 1998.

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory,* vol. 44, no. 2, pp. 744-765, 1998.

[3] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory,* vol. 45, no. 5, pp. 1456-1467, 1999.

[4] B. M. Hochwald, T. L. Marzetta, and C. B. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space-time spreading," *IEEE J. Selected Areas of Communications*, vol.19, pp.48-60, Jan. 2001.

[5] G. Ganesan and P. Stoica. "Space-time block codes: A maximum SNR approach," *IEEE Trans. Information Theory*, vol.47, pp.1650-1656, May 2001.

[6] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time block codes for complex signal constellations," *IEEE Trans. on Information Theory*, vol. 48, pp.384-395, Jan. 2002.

[7] A. R. Calderbank and A. F. Naguib, "Orthogonal designs and third generation wireless communication," *Surveys in Combinatorics*, Edited by J. W. P. Hirschfeld, pp. 75-107,

London Mathematical Society Lecture Note Series 288, Cambridge University Press, 2001.

[8] X.-B. Liang and X.-G. Xia, "Nonexistence of rate one space-time block codes from generalized complex linear processing orthogonal designs for more than two transmit antennas," *Tech. Report., Udel-ECE #01-3-1*, Department of Electrical and Computer Engineering, University of Delaware, Newark, Delaware, March 8, 2001. Also presented as recent results in the *International Symposium on Information Theory (ISIT'01)*, Washington DC, June 2001. *IEEE Trans. on Information Theory*, to appear.

[9] W. Su and X.-G. Xia, "Two generalized complex orthogonal space-time block codes of rates 7/11 and 3/5 for 5 and 6 transmit antennas," *IEEE Trans. on Information Theory*, vol.49, pp. 313-316, Jan. 2003.

[10] W. Su and X.-G. Xia, "On complex orthogonal space-time block codes from complex orthogonal designs," *Wireless Personal Communications*, vol. 25, pp. 1-26, April, 2003.

[11] X.-B. Liang, "A high-rate orthogonal space-time block code," *IEEE Communications Letters*, vol.7, pp. 222-223, May 2003.

[12] H-f. Lu, P. V. Kumar, and H. Chung, "On orthogonal designs and space-time codes," *Proceedings International Symposium on Information Theory (ISIT'02)*, PP. 418, Lausanne, Switzerland, 2002.

[13] H. Jafarkhani and F. Taherkhani, "Psuedo orthogonal designs as space-time block codes," *IEEE International Symposium on Advances in Wireless Communications (ISWC'02)*, Sept. 2002. Also available from `http://www.ece.uci.edu/~hamidj/publications.html`.

[14] L. A. Dalton and C. Georghiades, "A four transmit antenna orthogonal space-time block code with full diversity and rate," Proceedings of Allerton Conference, Oct. 2002.

[15] L. A. Dalton, "New orthogonal space-time block codes with full diversity," Master Thesis, Texas A&M University, Dec. 2002. Also available from `http://ee-wcl.tamu.edu/publ/projects.htm`.

[16] A. Hurwitz, "Über die Komposition der quadratischen Formen von beliebig vielen Variabeln," *Nachr. Ges. Wiss. Göttinggen*, (Math.-Phys. K1.), pp.309-316, 1898. Reprinted in *Math. Werke*, Bd. 2, Birkhäuser, pp.565-571, Basel 1963.

[17] A. Hurwitz, "Über die Komposition der quadratischen Formen," *Math. Ann.*, vol.88, pp.1-25, 1923.

[18] J. Radon, "Lineare scharen orthogonaler matrizen," in *Abhandlungen aus dem Mathematischen Seminar der Hamburgishen Universität*, vol. I, pp. 1-14, 1922.

[19] Y. C. Wong, "Isoclinic n-planes in Euclidean $2n$-space, Clifford parallels in Elliptic $(2n-1)$-space, and the Hurwitz matrix equations, (Part II)," *Men. Amer. Math. Soc.,* vol 41, 1961.

[20] T. Josefiak, "Realization of Hurwitz-Radon matrices," *Queen's Papers in Pure and Applied Mathematics*, no. 36, pp.346-351, 1976.

[21] W. Wolfe, "Amicable orthogonal designs - existence," *Canad. J. Math.*, vol.28, pp.1006-1020, 1976.

[22] A. V. Geramita and J. Seberry, *Orthogonal Designs, Quadratic Forms and Hadamard Matrices, Lecture Notes in Pure and Applied Mathematics,* vol.43, New York and Basel: Marcel Dekker, 1979.

[23] W. H. Greub, *Linear Algebra*, second edition, Springer-Verlag, 1963.

[24] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985, reprinted in 1999.

[25] D. B. Shapiro, *Compositions of Quadratic Forms*, New York, De Gruyter, 2000.

[26] D. B. Shapiro, Email communications, 2001.