Design of Fully-Diverse Multiple-Antenna Codes Based on $Sp(2)$

YINDI JING AND BABAK HASSIBI
Department of Electrical Engineering
California Institute of Technology, Pasadena, CA 91125
e-mail: yindi,hassibi@systems.caltech.edu

Abstract

Fully-diverse constellations, i.e., a set of unitary matrices whose pairwise differences are non-singular, are useful in multi-antenna communications, especially in multi-antenna differential modulation, since they have good pairwise error properties. Recently, group theoretic ideas, especially *fixed-point-free (fpf)* groups, have been used to design fully-diverse constellations of unitary matrices. Here we construct four-transmit-antenna constellations appropriate for differential modulation based on the symplectic group $Sp(2)$. They can be regarded as extensions of Alamouti's celebrated two-transmit-antenna orthogonal design which can be constructed from the group $Sp(1)$. We further show that the structure of the $Sp(2)$ code lends itself to efficient maximum likelihood (ML) decoding via the sphere decoding algorithm. Finally, the performance of the $Sp(2)$ code is compared with that of other existing codes including Alamouti's orthogonal design, a 4×4 complex orthogonal design, Cayley differential unitary space-time codes and group based codes.

Index Terms: Wireless communication systems, differential unitary space-time modulation, fully-diverse, Lie group, Lie algebra, sphere decode

1 Introduction

It is well known in theory that multiple antennas can greatly increase the data rate and the reliability of a wireless communication link in a fading environment. In practice, however, one needs to devise effective space-time transmission schemes. This is particularly challenging when the propagation environment is unknown to both the sender and the receiver, which is often the case for mobile applications when the channel changes rapidly.

A differential transmission scheme called *differential unitary space-time modulation* was proposed in [1, 2, 3], which is well-tailored for unknown continuously varying Rayleigh flat-fading channels. The signals transmitted are unitary matrices. In this scheme the probability of error of mistaking one signal S_i for another $S_{i'}$, at high SNR, is proved to be inversely proportional to $|\det(S_i - S_{i'})|$. Therefore the quality of the code is measured by its *diversity product*

$$\xi_{\mathcal{C}} = \frac{1}{2} \min_{S_i \neq S_{i'} \in \mathcal{C}} |\det(S_i - S_{i'})|^{\frac{1}{M}} \tag{1}$$

where M is number of transmitter antennas and $\mathcal{C}$ is the set of all possible signals. We therefore say that a code is *fully-diverse* or has *full diversity* if the determinants of the pairwise differences are all nonzero. The design problem is thus the following: "Given the number of transmitter antennas, M, and the transmission rate, R, find a set $\mathcal{C}$ of $L = 2^{MR}$ $M \times M$ unitary matrices, such that the diversity product as defined in (1) is as large as possible."

The design problem, as just stated, is not easy because first, since the *Stiefel* manifold of unitary matrices is not linear, it is not easy to design constellations of unitary matrices, and second, since both the signal set and the cost function are non-convex and the size of the problem can be huge, especially at high data rates, obtaining an exact solution of the design problem seems computationally intractable. Therefore, in [4, 5, 6], it was proposed to enforce a group structure on the constellation. This has several advantages as discussed in [4, 5, 6]. Moreover, it is shown that a constellation is fully-diverse if and only if the corresponding group is fixed-point-free (fpf), i.e. all non-identity matrices have no unit eigenvalues. In [4], all finite fully-diverse constellations that form a group are classified. And, in [5], it is proved that the only fpf infinite Lie groups are $U(1)$, the group of unit-modulus scalars, and $SU(2)$, the group of unit-determinant 2×2 unitary matrices.

However, no good constellations are obtained for very high rates from the finite fpf groups classified in [4], and constellations based on $U(1)$ and $SU(2)$ are constrained to one and two-transmit-antenna systems. In this paper, to get high-rate constellations which work for 4-transmit-antenna systems, we relax the fpf condition by considering Lie groups with non-identity elements having no more than k unit eigenvalues for some fixed $k > 0$. It can be shown that if a Lie group has rank n, then it has at least one element with $n - 1$ eigenvalues at 1. (The *rank* of a Lie group equals the maximum number of commuting basis elements of its Lie algebra, and it can be shown that fpf groups have rank 1. See [5, 7].) The lower the rank of a Lie group, the more possible it is to get a subset with no unit eigenvalue elements, that is, the more possible for us to find a fully-diverse subset of it. There are only three simple, simply-connected, compact Lie groups of rank 2: the Lie group of unit-determinant 3×3 unitary matrices $SU(3)$, the Lie group of unit-determinant 4×4 unitary, symplectic matrices $Sp(2)$, and $\mathcal{G}_2$ which is one of the 5 Cartan exceptional groups. In this paper, we focus on $Sp(2)$. The codes designed based on it are fully-diverse, can be used in four transmit antenna and any number of receive antenna systems, exist for almost any rate and lend themselves to polynomial-time ML decoding via the sphere decoder.

The paper is organized as follows. A brief review of Lie groups and Lie algebras, emphasizing the results we need, is given in Section 2. In Section 3, a method to parameterize the symplectic group $Sp(n)$ is proposed, based on which space-time $Sp(2)$ codes are constructed and the full diversity of the codes is proved. The decoding of the $Sp(2)$ codes is shown in Section 4. The sphere decoder can be used to achieve the ML solution with much less complexity. Simulation results, including comparisons with complex orthogonal designs and other existing schemes are presented in section 5. Section 6 is the conclusion. Some of the proofs of the theorems and lemmas in this paper are relegated to the Appendices.

1.1 Differential Unitary Space-Time Modulation

Before bringing in the design of the $Sp(2)$ constellations, we first present a brief introduction to the multiple antenna model and the differential unitary space-time signaling scheme. This follows [1, 2, 3].

Consider a wireless communication system with M transmit antennas and N receive antennas. We use a block-fading channel with coherence interval T (for more on this model, see [8, 9]). The system equations of the τ-th block can be written as:

$$X_\tau = \sqrt{\frac{\rho T}{M}} S_\tau H_\tau + V_\tau$$

Here, S denotes the $T \times M$ transmitted signal matrix with s_{tm} being the signal sent by the m-th transmit antenna at time t. The t-th row of S indicates the row vector of the transmitted values from all the transmit antennas at time t and the m-th column indicates the transmitted values of the m-th transmit antenna across the coherence interval T. H is the $M \times N$ complex-valued propagation matrix which remains constant during the coherence interval and h_{mn} is the propagation coefficient between the m-th transmit antenna and the n-th receive antenna. The h_{mn}s have a zero-mean unit-variance circularly-symmetric complex Gaussian distribution $\mathcal{CN}(0,1)$ and are independent of each other. We assume that the channel information is unknown to both the transmitter and the receiver. V is the $T \times N$ noise matrix with v_{tn} being the noise at the nth receive antenna at time t. The v_{tn}s are iid with $\mathcal{CN}(0,1)$ distribution. X is the $T \times N$ received signal matrix with x_{tn} being the received value by the n-th receive antenna at time t. The t-th row of X indicates the row vector of the received values at all the receivers at time t and the n-th column indicates the received values of the n-th transmit antenna across the coherence interval. We impose an extra power constraint on the transmitted signal

$$\frac{1}{M} \sum_{m=1}^{M} \mathrm{E}\, |s_{tm}|^2 = \frac{1}{T}, \quad t = 1, 2, ..., T$$

which means that the transmitted signal, s_{tm}, has an average expected power (over the M transmit antennas) equal to $\frac{1}{T}$ at each channel use. Therefore, ρ represents the expected SNR at each receive antenna.

One way to communicate with unknown channel information is to use the multiple-antenna differential unitary space-time modulation (USTM), which can be seen as a natural extension of the standard differential phase shift keying (DPSK) commonly used in signal-antenna unknown-channel systems. In differential USTM, the channel is used in blocks of M transmissions, which implies that the transmitted signals, S, are $M \times M$ unitary matrices. At the τ-th block, the transmitted $M \times M$ unitary matrix equals the product of the previously transmitted matrix and a unitary data matrix $V_{z_\tau}, z_\tau \in 0, ..., L-1$ taken from our signal set $\mathcal{C}$. In other words,

$$S_\tau = V_{z_\tau} S_{\tau-1} \tag{2}$$

with $S_0 = I_M$. Having V_{z_τ} unitary assures that the transmitted signal will not vanish or blow up to infinity. Since the channel is used M times, the corresponding transmission rate is $R = \frac{1}{M} \log_2 L$, where L is the cardinality of the code. If we further assume that the propagation environment keeps approximately constant for $2M$ consecutive channel uses, that is, $H_\tau \approx H_{\tau-1}$, then from the system equations for the τ-th and the $(\tau-1)$-th blocks, the following fundamental differential receiver equations are obtained [10],

$$X_\tau = V_{z_\tau} X_{\tau-1} + W'_\tau \tag{3}$$

where

$$W'_\tau = W_\tau - V_{z_\tau} W_{\tau-1}. \tag{4}$$

We can see that the channel matrix H does not appear in (3). This implies that, as long as the channel is approximately constant for $2M$ channel uses, differential transmission permits decoding without knowing the channel information.

Since V_{z_τ} is unitary, the additive noise term in (4) is statistically independent of V_{z_τ} and has independent complex Gaussian entries. Therefore, the maximum likelihood decoder of z_τ can be written as

$$\hat{z}_\tau = \arg \max_{l=0,...,L-1} \|X_\tau - V_l X_{\tau-1}\| \tag{5}$$

It is shown in [1, 3] that, at high SNR, the pairwise block probability of error (of transmitting V_l and erroneously decoding $V_{l'}$) has the upper bound

$$\mathrm{P} \leq \frac{1}{2}(\frac{8}{\rho})^{MN} \frac{1}{|\det(V_l - V_{l'})|^{2N}}$$

which is inversely proportional to the absolute value of the determinant of the pairwise difference to the power of $2N$. Therefore, most design schemes [1, 3, 4, 6] have focused on finding a constellation $\mathcal{V} = \{V_0, ..., V_L\}$ of $L = 2^{MR}$ unitary $M \times M$ matrices that maximizes $\xi_{\mathcal{C}}$ defined in (1). In general, the number of unitary

$M \times M$ matrices in $\mathcal{C}$ can be quite large. This huge number of signals calls into question the feasibility of computing $\xi_{\mathcal{C}}$ and also rules out the possibility of decoding via an exhaustive search. To design constellations that are huge, effective and yet still simple so that they can be decoded in real-time, some structure should be imposed upon the signal set. In this paper, we propose a well-structured code which is a subset of the symplectic Lie group $Sp(2)$ and lends itself to linear-algebraic decoding by using the sphere decoder.

2 Lie Groups and Lie Algebras

If our signal set has a group structure under matrix multiplication, its diversity product can be simplified as the follows,

$$\begin{aligned} \xi_{\mathcal{C}} &= \frac{1}{2}\min_{0\le l<l'<L} |\det(V_l - V_{l'})|^{\frac{1}{M}} \\ &= \frac{1}{2}\min_{0\le l<l'<L} |\det(V_l)\det(I - V_l^{-1}V_{l'})|^{\frac{1}{M}} \\ &= \frac{1}{2}\min_{I\neq V\in\mathcal{C}} |\det(I - V)|^{\frac{1}{M}} \end{aligned}$$

If we insist on a fully-diverse constellation, which means that $\xi_{\mathcal{C}} \neq 0$, then from the above equalities, the eigenvalues of all non-identity elements in the constellation must be different from from one. This leads to the following definition.

Definition 1 (Fix-Point-Free Group). *[5] A group $\mathcal{G}$ is called* **fixed-point-free (fpf)** *if and only if it has a faithful representation as unitary matrices with the property that the representation of each non-unit element of the group has no eigenvalue at unity.*

Note that the above definition does not require that in *every* representation of the group non-unit elements have no eigenvalue at unity, but rather that there exists one representation with this propriety. This is because any non-faithful representation cannot be fpf. The reason why we have defined fpf groups as those for which the representation of each non-unit element in the group, rather than each non-identity matrix in the representation, has no eigenvalue at unity is that had we not done so, all groups would have been fpf if represented as the identity matrix.

In [4], all finite fully-diverse constellations that form a group, i.e. all finite fixed-point-free groups, are classified by the following theorem.

Theorem 1 (Classification of finite fpf groups). *[4] A finite group is fixed-point-free if and only if it is isomorphic to either $G_{m,r}, D_{m,r,l}, E_{m,r}, F_{m,r,l}, J_{m,r}$ or $K_{m,r,l}$.*

(For definitions of the above $G_{m,r}, D_{m,r,l}, E_{m,r}, F_{m,r,l}, J_{m,r}$ or $K_{m,r,l}$ groups, see [4].)

As shown in [4], these finite fpf groups are few and far between. Although very good constellations are obtained for low to moderate rates, no good constellations are obtained for very high rates from these finite groups. This motivates the search for *infinite* fpf groups, in particular, their most interesting case, Lie groups.

Definition 2 (Lie Group). *[11] A* **Lie group** *is a differential manifold which is also a group such that the group multiplication (and also the inversion map) is a differential map.*

The above definition gives us the main reason for studying Lie groups in our constellation design. Since it has an underlying manifold, one may obtain a finite constellation, which is a subset of the infinite Lie group, by sampling appropriately the underlying continuous manifold.

Definition 3 (Lie algebra). *[7] A* **Lie algebra** $\mathbf{g}$ *is a vector space over a field F on which a product $[,]$, called the Lie bracket, is defined, with the properties*

1. $X, Y \in \mathbf{g}$ *implies* $[X, Y] \in \mathbf{g}$
2. $[X, \alpha Y + \beta Z] = \alpha[X, Y] + \beta[X, Z]$ *for* $\alpha, \beta \in F$ *and* $X, Y, Z \in \mathbf{g}$
3. $[X, Y] = -[Y, X]$
4. $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

It turns out that there is a close connection between Lie groups and Lie algebras.

Theorem 2 (Lie group and Lie algebra). *[7] Let G be a Lie group of matrices. Then* $\mathbf{g}$*, the set of tangent vectors to all curves in G at the identity, is a Lie algebra. Let* $\mathbf{g}$ *be a linear algebra generated by the basis $g_1, \cdots, g_n$, then $g(\theta) = e^{\theta_1 g_1 + \cdots + \theta_n g_n}$ is a local Lie group for small enough θ.*

Therefore, to obtain many, if not most, of its properties, one can study the Lie algebra, rather than the Lie group itself.

Here are some examples of Lie groups. $GL(n, \mathcal{C})$ is the group of non-singular $n \times n$ complex matrices, whose Lie algebra is the space of $n \times n$ complex matrices. $SL(n, \mathcal{C})$ is the group of unit-determinant non-singular $n \times n$ complex matrices whose Lie algebra is the space of $n \times n$ traceless matrices. $U(n)$ is the $n \times n$ complex unitary matrices group whose Lie algebra is the space of $n \times n$ skew-Hermitian matrices and $SU(n)$ is the group of unit-determinant $n \times n$ unitary matrices whose Lie algebra is the space of $n \times n$ traceless, skew-Hermitian matrices.

Theorem 3 (Lie groups with unitary representations). *[5] A Lie group has a representation as unitary matrices if and only of its Lie algebra is a compact semi-simple Lie algebra or the direct sum of $u(1)$ and a compact semi-simple Lie algebra.*

(For more on the definition of semi-simple and simple Lie algebras, see [5, 11, 12].)

Since any semi-simple Lie algebras is isomorphic to a product of simple Lie algebras [12] and a Lie algebra is simple if and only if its Lie group is simple, to design a set of unitary matrices, we will be interested in the compact, simple groups. And unsurprisingly, we want to design constellations that are fully-diverse. It is proved in [5] that the only fpf infinite Lie groups are $U(1)$, the group of unit-modulus scalars, and $SU(2)$, the group of unit-determinant 2×2 unitary matrices. Due to their dimensions, constellations based on these two Lie groups are constrained to one- and two-transmit-antenna systems. (Codes constructed based on higher dimensional representations of $SU(2)$ can be found in [13].) To obtain a 4-transmit-antenna constellation, we relax the fpf condition and consider those compact simple Lie groups whose non-unit elements have no more than $k > 0$ unit eigenvalues ($k = 0$ corresponds to fpf groups). We will obtain a constellation of finite size by sampling the Lie group's underlying manifold. When k is small, there is a good chance that, by sampling appropriately, the resulting code is fully-diverse. It follows from the exponential map relating Lie groups with Lie algebras that a matrix element of a Lie group has eigenvalues at 1 if and only if the corresponding matrix element (the logarithm of the element) in the corresponding Lie Algebra (tangent space at identity) has eigenvalues at zero and vice versa. Thus classifying Lie groups whose non-identity elements have no more than k eigenvalues at 1 is the same as classifying Lie algebras whose non-zero elements have no more than k eigenvalues at zero. Unfortunately, there does not appear to be a straightforward way of analyzing the number of zero eigenvalues of a matrix element of a Lie algebra. However, it turns out that the number of zero eigenvalues of a matrix element of a Lie algebra can be related to the *rank* of its Lie algebra, where the rank of a Lie algebra equals the maximum number of commuting basis elements of it.

Lemma 1. *If a Lie algebra* **g** *of M by M matrices has rank r, then it has at least one non-zero element with $r-1$ eigenvalues at zero.*

Proof: Suppose $b_1, \cdots, b_r$ are the commuting basis elements of the Lie algebra. Since they commute, there exists a matrix T such that $b_1, \cdots, b_r$ can be diagonalized simultaneously. That is, there exists an $M \times M$ invertible matrix T, such that $b_i = TD_iT^{-1}$ where $D_i = \text{diag}\,(b_1^i, \cdots, b_M^i)$. Therefore, it is possible to design scalars $\alpha_1, \cdots, \alpha_r$ such that $\sum_{i=1}^r \alpha_i D_i$ is a diagonal matrix with the first $r-1$ diagonal elements being zero. The matrix $b = \sum_{i=1}^r \alpha_i b_i = T(\sum_{i=1}^r \alpha_i D_i)T^{-1}$, which is also an element in the Lie algebra, has therefore $r-1$ eigenvalues at zero. □

Therefore, instead of exploring Lie groups with non-identity elements having no more than k unit eigenvalues, we work on the compact simple Lie groups whose rank is no more than $k+1$, and obtain a finite subset of it as our constellation by sampling its underlying manifold. A beginning point is to look at those simple, compact Lie groups of rank 2. The following table is a complete list of the simple, simply connected, compact Lie groups [14].

Table 1: The Simple, Simply Connected Compact Lie Groups

Group	dimension	$\mathbf{Z}(G)$	Cartan Name	Rank
$SU(n); n \geq 2$	$n^2 - 1$	$\mathbf{Z}_n$	A_{n-1}	$n - 1$
$Sp(n); n \geq 2$	$n(2n+1)$	$\mathbf{Z}_2$	C_n	n
$Spin(2n+1); n \geq 3$	$n(2n+1)$	$\mathbf{Z}_2$	B_n	n
$Spin(2n); n \geq 4$	$n(2n-1)$	$\mathbf{Z}_4$ (n odd) $\mathbf{Z}_2 \times \mathbf{Z}_2$ (n even)	D_n	n
E_6	78	$\mathbf{Z}_3$	E_6	6
E_7	133	$\mathbf{Z}_2$	E_7	7
E_8	248	0	E_8	8
F_4	52	0	F_4	4
G_2	14	0	G_2	2

$Z(G)$ indicates the center of the group G. We can see that there are 3 groups with rank 2: the Lie group of unit-determinant 3×3 unitary matrices $SU(3)$, the Lie group of 4×4 unitary, symplectic matrices $Sp(2)$, and one of the five celebrated exceptional Lie groups of E. Cartan: $\mathcal{G}_2$. Since $Sp(1) = SU(2)$, and $SU(2)$ constitutes the orthogonal design of Alamouti [15], the symplectic group $Sp(2)$ can be regarded as a generalization of the orthogonal design. Therefore, the attention of this paper is focused on $Sp(2)$.

Definition 4 (Symplectic group). *[14] $Sp(n)$, the nth order symptectic group, is the set of complex $2n \times 2n$ matrices S obeying*

1. *Unitary condition: $S^*S = SS^* = I_{2n}$,*
2. *Symplectic condition: $S^t J_{2n} S = J_{2n}$,*

where $J_{2n} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}$.

In the following, subscripts which indicate dimensions are omitted when there is no confusion. S^t denotes the transpose of S and S^* denotes the conjugate transpose of S. I_n indicates the $n \times n$ identity matrix.

$Sp(n)$ has dimension $n(2n+1)$ and rank n. As mentioned before, we are most interested in the lowest rank case, which is also the simplest case of $n = 2$. Even though Lemma 1 claims that there exists an element of $Sp(2)$ with at least one unit eigenvalues, it can be shown that non-identity element of $Sp(2)$ can have up to 2 unit eigenvalues.

Lemma 2. *The multiplicity of the unit eigenvalue of any matrix in $Sp(2)$ is even.*

Proof: Assume S is a matrix in $Sp(2)$ and x is an eigenvector of S with eigenvalue 1. Then we have $Sx = x$. From the symplectic condition $S^tJS = J$, we have $S^*J\bar{S} = J$ and since S is unitary, $J\bar{S} = SJ$. Therefore, $SJ\bar{x} = J\bar{S}\bar{x} = J\overline{Sx} = J\bar{x}$. Therefore, $J\bar{x}$ is also an eigenvector of S with eigenvalue 1. We now argue that $x \neq J\bar{x}$. Assume $x = J\bar{x}$. Partition x as $[x_1, x_2]^t$, where x_1 and x_2 are 2-dimensional vectors. We have

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} \Rightarrow \begin{cases} x_1 = \bar{x}_2 \\ x_2 = -\bar{x}_1 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases},$$

from which we get $x = 0$, which contradicts the assumption that x is an eigenvector. Therefore, $x \neq J\bar{x}$, which means that the number of eigenvectors for any unit eigenvalue is even. Thus, the multiplicity of the unit eigenvalue is even. $\square$

From Lemma 2, if a matrix in $Sp(2)$ has a unit eigenvalue then its multiplicity must be 2 or 4. 4 unit eigenvalues means that the matrix is I_4. Therefore, non-identity element of $Sp(2)$ can have 0 or 2 unit eigenvalues.

3 Design of Fully-Diverse Codes Based on $Sp(2)$

We start by parametrizing $Sp(n)$.

3.1 Parametrization of $Sp(n)$

From condition 1 of Definition 4, we have $(S^t)^{-1} = \bar{S}$, where $\bar{S}$ is the conjugate of S, so condition 2 becomes

$$JS = \bar{S}J. \tag{6}$$

If we further partition the matrix S into a 2×2 block matrix of $n \times n$ sub-matrices $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$, from (6), we see that $C = \bar{B}$ and $D = \bar{A}$. Therefore, any matrices in $Sp(n)$ have the form

$$\begin{bmatrix} A & B \\ -\bar{B} & \bar{A} \end{bmatrix} \tag{7}$$

which is similar to Alamouti's 2-dimensional orthogonal design [15], but here instead of complex scalars, A and B are n by n complex matrices. The group, $Sp(n)$, can thus be identified as the subgroup of unitary matrices with generalized orthogonal design form. To get a more detailed structure of the Lie group, let us look at the conditions imposed on A and B for S to be unitary. From $SS^* = I_{2n}$ or $S^*S = I_{2n}$, we have

$$\begin{cases} AA^* + BB^* = I_n \\ BA^t = AB^t \end{cases} \text{or} \begin{cases} A^*A + B^t\bar{B} = I_n \\ A^*B = B^t\bar{A} \end{cases} \tag{8}$$

Lemma 3. *For any $n \times n$ complex matrices A and B that satisfy (8), there exist unitary matrices U and V such that $A = U\Sigma_A V$ and $B = U\Sigma_B\bar{V}$, where Σ_A and Σ_B are diagonal matrices whose diagonal elements are the singular values of A and B.*

Before the proof of this lemma, we first prove two intermediate lemmas.

Lemma 4. *For any $n \times n$ diagonal matrices D_1, D_2 with non-increasingly ordered diagonal entries, if $UD_1U^* = D_2$ for some unitary matrix U, then we have $D_1 = D_2 = D$ and U is a block diagonal matrix with the size of each block equal to the number of times that a diagonal element of D is repeated.*

Proof: Denote the i-th diagonal element of D_1 and D_2 as $d_{ii}^{(1)}$ and $d_{ii}^{(2)}$. Since UD_1U^* is a similarity transformation, which preserves the eigenvalues, the set of eigenvalues of D_1 is the same as the set of eigenvalues of D_2, or in other words, $d_{ii}^{(1)} = d_{jj}^{(2)}$ for some j. Noticing that the $d_{ii}^{(1)}s$ and $d_{ii}^{(2)}s$ are ordered non-increasingly, we have $d_{ii}^{(1)} = d_{ii}^{(2)}$ for $i = 1, 2, ..., n$, that is, $D_1 = D_2 = D$. Now write D as $\text{diag}\{d_1 I_{P_1}, \cdots, d_q I_{P_q}\}$, where P_i is the number of times the element d_i appears in D for $i = 1, \cdots, q$. It is obvious that U can be written as $\text{diag}\{U_1, \cdots, U_q\}$ where the size of U_i is P_q for $i = 1, \cdots, q$. □

Lemma 5. *If $UD^2 = D^2U$ for any $n \times n$ positive semi-definite diagonal matrices D and any $n \times n$ matrix U, then $UD = DU$.*

Proof: By looking at the (i, j)-th entries of UD^2 and D^2U, we have that $u_{ij}d_{jj}^2 = d_{ii}^2u_{ij}$. If $d_{ii}^2 \neq d_{jj}^2$, $u_{ij} = 0$ is obtained, and therefore $u_{ij}d_{jj} = d_{ii}u_{ij}$. If $d_{ii}^2 = d_{jj}^2$, since D is a positive semi-definite matrix, d_{ii} is non-negative, therefore, $d_{ii} = d_{jj}$ and so $u_{ij}d_{jj} = d_{ii}u_{ij}$ is obtained. Therefore, $UD = DU$. □

Now we are ready to prove Lemma 3.

Proof of Lemma 3: Suppose $A = U_A\Sigma_A V_A^*$ and $B = U_B\Sigma_B V_B^*$ are the singular value decompositions of A and B with the non-negative diagonal elements of Σ_A *non-decreasingly* ordered and the non-negative diagonal elements of Σ_B *non-increasingly* ordered. From the equation $AA^* + BB^* = I_n$ in (8), we have the following series of equations.

$$\begin{aligned} & U_A\Sigma_A V_A^* V_A \Sigma_A U_A^* + U_B\Sigma_B V_B^* V_B \Sigma_B U_B^* = I_n \\ \Rightarrow\ & U_A\Sigma_A^2 U_A^* + U_B\Sigma_B^2 U_B^* = I_n \\ \Rightarrow\ & \Sigma_A^2 + (U_A^*U_B)\Sigma_B^2(U_A^*U_B)^* = I_n \\ \Rightarrow\ & (U_A^*U_B)\Sigma_B^2(U_A^*U_B)^* = I_n - \Sigma_A^2 \end{aligned}$$

Since U_A and U_B are unitary, $U_A^* U_B$ is also unitary. Now since the diagonal entries of Σ_B^2 and $I_n - \Sigma_A^2$ are non-increasingly ordered, from Lemma 4, we have

$$\Sigma_B^2 = I_n - \Sigma_A^2. \tag{9}$$

Define $U_D = U_A^* U_B$. Therefore $U_B = U_A U_D$ and

$$U_D(I_n - \Sigma_A^2) = (I_n - \Sigma_A^2)U_D \Leftrightarrow U_D\Sigma_A^2 = \Sigma_A^2 U_D.$$

Since Σ_A is a positive semi-definite matrix, from Lemma 5, $U_D\Sigma_A = \Sigma_A U_D$ and therefore $U_D^*\Sigma_A = \Sigma_A U_D^*$. Further define $U_1 = U_B = U_A U_D$. Then,

$$A = U_A\Sigma_A V_A^* = U_A U_D U_D^* \Sigma_A V_A^* = U_A U_D \Sigma_A (V_A U_D)^* = U_1 \Sigma_A V_A'^*,$$

where we have defined $V_A' = V_A U_D$. Since $U_1 = U_B$, we also know that $B = U_1\Sigma_B V_B^*$. Thus A and B have the same left singular vectors. We now focus on the right singular vectors. From the equation $A^*A + B^t\bar{B} = I_n$ in (8), we can get the following series of equations,

$$\begin{aligned}
& V_A'\Sigma_A U_1^* U_1 \Sigma_A V_A'^* + \bar{V}_B \Sigma_B U_1^t \bar{U}_1 \Sigma_B V_B^t = I_n \\
\Rightarrow\quad & V_A'\Sigma_A^2 V_A'^* + \bar{V}_B\Sigma_B^2 V_B^t = I_n \\
\Rightarrow\quad & \Sigma_A^2 + (V_A'^*\bar{V}_B)\Sigma_B^2(V_A'^*\bar{V}_B)^* = I_n \\
\Rightarrow\quad & (V_A'^*\bar{V}_B)\Sigma_B^2(V_A'^*\bar{V}_B)^* = I_n - \Sigma_A^2
\end{aligned}$$

Therefore, we have $I_n - \Sigma_A^2 = \Sigma_B^2$ and $(V_A'^*\bar{V}_B)\Sigma_B^2 = \Sigma_B^2(V_A'^*\bar{V}_B)$ by using Lemma 4. Define $V_D = V_A'^*\bar{V}_B$, which is obviously a unitary matrix. Therefore $\bar{V}_B = V_A' V_D$, $V_D\Sigma_B^2 = \Sigma_B^2 V_D$, $V_D\Sigma_A^2 = \Sigma_A^2 V_D$, from which we get $V_D\Sigma_B = \Sigma_B V_D$ and $V_D\Sigma_A = \Sigma_A V_D$ by using Lemma 5. Now according to Lemma 3, V_D can be written as diag$\{V_1, \cdots, V_q\}$ with each V_i being a unitary matrix. Since V_i is unitary, there exists a Hermitian matrix G_i such that $V_i = e^{jG_i}$. Since G_i is Hermitian, so is $G_i/2$. Therefore, the matrix $e^{jG_i/2}$, which is the square root of V_i, is also a unitary matrix and $e^{jG/2}\Sigma_A = \Sigma_A e^{jG/2}, e^{jG/2}\Sigma_B = \Sigma_B e^{jG/2}$ can be obtained, where $G/2 = \text{diag}\{G_1/2, \cdots, G_q/2\}$. Therefore $e^{jG/2}$ is the square root of V_D. Thus, A and B can be written as

$$\begin{aligned}
A &= U_1\Sigma_A V_D V_1^* = U_1\Sigma_A e^{jG} V_1^* = U_1 e^{jG/2}\Sigma_A e^{jG/2} V_1^* = (U_1 e^{jG/2})\Sigma_A (V_1 e^{-jG/2})^* \\
B &= U_1\Sigma_B V_1^t = (U_1 e^{jG/2})\Sigma_B (V_1 e^{-jG/2})^t
\end{aligned}$$

where we have defined $V_1 = \bar{V}_B = V_A' V_D$. Therefore, we have $A = U\Sigma_A V^*$ and $Y = U\Sigma_B V^t$ for some unitary matrices $U = U_1 e^{jG/2}$ and $V = V_1 e^{-jG/2}$ or equivalently, $A = U\Sigma_A V$ and $Y = U\Sigma_B \bar{V}$ if we define $U = U_1 e^{jG/2}$and $V = (V_1 e^{-jG/2})^*$. $\square$

Lemma 3 indicates that A and B can be diagonalized by the same pair of unitary matrices. This leads to the following parametrization theorem of $Sp(n)$.

Theorem 4 (Parametrization of $Sp(n)$). *A matrix S belongs to $Sp(n)$ if and only if it can be written as*

$$S = \begin{bmatrix} U\Sigma_A V & U\Sigma_B \bar{V} \\ -\bar{U}\Sigma_B V & \bar{U}\Sigma_A \bar{V} \end{bmatrix} \tag{10}$$

where U and V are $n \times n$ unitary matrices, $\Sigma_A = diag(\cos\theta_1, \cdots \cos\theta_n), \Sigma_B = diag(\sin\theta_1, \cdots, \sin\theta_n)$ for some $\theta_1, \cdots, \theta_n \in [0, \pi/2]$.

Proof: Lemma 3 and formula (7) imply that any matrix in $Sp(2)$ can be written as the form in (10). Conversely, for any matrix S with the form of (10), it is easy to verify the unitary and symplectic conditions in definition 4. □

Now let us look at the dimension of S. It is known that a $n \times n$ unitary matrix has dimension $2n^2 - n - 2 \times \frac{n(n-1)}{2} = n^2$. Therefore, there are all together $2n^2$ degrees of freedom in the unitary matrices U and V. Together with the n real angles θ_i, the dimension of S is therefore $n(2n+1)$, which is exactly the same as that of $Sp(n)$. But from the discussion above, an extra condition is composed on the matrix S: the diagonal elements of Σ_A and Σ_B are non-negative and non-increasingly/non-decreasingly ordered. This might cause the dimension of S to be less than the matrices in $Sp(2)$ at first glance. However, the order and signs of the diagonal elements of Σ_A and Σ_B can be changed by multiplying U and V with elementary unitary matrices. Therefore, the constraint does not result in dimension reduction. Based on Theorem 4, the matrices in $Sp(n)$ can be parametrized by the elements of U, V and the real angles θ_is.

3.2 Design of $Sp(2)$ Codes

Let us now focus on the case of $n = 2$. We want to find a fully-diverse subset of $Sp(2)$. For simplicity, we first let $\Sigma_A = \Sigma_B = \frac{1}{\sqrt{2}} I_2$, by which 2 degrees of freedom are neglected. To get a finite subset of unitary matrices from the infinite Lie group, we further choose U and V as orthogonal designs with the entries of U chosen uniformly from the set of P-PSK signals: $\{1, e^{j\frac{2\pi}{P}}, \cdots, e^{j\frac{2\pi(P-1)}{P}}\}$ and the entries of V chosen uniformly from the set of Q-PSK signals shifted by an angle θ: $\{e^{j\theta}, e^{j(\frac{2\pi}{Q}+\theta)}, \cdots, e^{j(\frac{2\pi(Q-1)}{Q}+\theta)}\}$ [6, 15].

The following code is obtained.

$$\mathcal{C}_{P,Q,\theta} = \left\{ \frac{1}{\sqrt{2}} \left[\begin{array}{cc} UV & U\bar{V} \\ -\bar{U}V & \bar{U}\bar{V} \end{array} \right] \left| \begin{array}{c} U = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} e^{j\frac{2\pi k}{P}} & e^{j\frac{2\pi l}{P}} \\ -e^{-j\frac{2\pi l}{P}} & e^{-j\frac{2\pi k}{P}} \end{array} \right], \\ V = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} e^{j(\frac{2\pi m}{Q}+\theta)} & e^{j(\frac{2\pi n}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n}{Q}+\theta)} & e^{-j(\frac{2\pi m}{Q}+\theta)} \end{array} \right], \\ 0 \le k,l < P, 0 \le m,n < Q \end{array} \right. \right\} \tag{11}$$

for integers P and Q. $\theta \in [0, 2\pi)$ is an angle to be chosen later. There are P^2 possible U matrices and Q^2 possible V matrices. Since the channel is used in blocks of 4 transmissions, the rate of the code is therefore

$$\frac{1}{2}(\log_2 P + \log_2 Q). \tag{12}$$

Any transmission matrix in the code can be identified by the 4-tuple (k, l, m, n). The angle θ, an extra degree of freedom added to increase the diversity product, is used in the proof of the full diversity of the code, although simulation result indicates that the code is always fully-diverse no matter what value the θ takes.

We can also consider a similar code as follows.

$$\mathcal{C}'_{P,Q,\theta} = \left\{ \frac{1}{\sqrt{2}} \left[\begin{array}{cc} UV & U\bar{V} \\ -\bar{U}V & \bar{U}\bar{V} \end{array} \right] \left| \begin{array}{c} U = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} e^{j\frac{2\pi k}{P}} & e^{j\frac{2\pi l}{Q}} \\ -e^{-j\frac{2\pi l}{Q}} & e^{-j\frac{2\pi k}{P}} \end{array} \right], \\ V = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} e^{j(\frac{2\pi m}{P}+\theta)} & e^{j(\frac{2\pi n}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n}{Q}+\theta)} & e^{-j(\frac{2\pi m}{P}+\theta)} \end{array} \right], \\ 0 \le k,m < P, 0 \le l,n < Q \end{array} \right. \right\} \tag{13}$$

The rate of the code is the same as that of $\mathcal{C}_{P,Q,\theta}$ and its full diversity can be proved similarly. We, however, will focus on the code given in (11).

3.3 Full Diversity of the $Sp(2)$ Codes

In multiple antenna code design, the mostly used criterion is the full diversity of the code since the diversity product is directly related to the pairwise probability of error. We discuss this issue in this subsection.

To calculate the diversity product of the code $\mathcal{C}_{P,Q,\theta}$ given in (11), we first choose any 2 transmission matrices S_1 and S_2 in $\mathcal{C}_{P,Q,\theta}$ where

$$S_1 = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} U_1V_1 & U_1\bar{V}_1 \\ -\bar{U}_1V_1 & \bar{U}_1\bar{V}_1 \end{array} \right] \text{ and } S_2 = \frac{1}{\sqrt{2}} \left[\begin{array}{cc} U_2V_2 & U_2\bar{V}_2 \\ -\bar{U}_2V_2 & \bar{U}_2\bar{V}_2 \end{array} \right]. \tag{14}$$

and

$$U_i = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j\frac{2\pi k_i}{P}} & e^{j\frac{2\pi l_i}{P}} \\ -e^{-j\frac{2\pi l_i}{P}} & e^{-j\frac{2\pi k_i}{P}} \end{bmatrix} \text{ and } V_i = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j(\frac{2\pi m_i}{Q}+\theta)} & e^{j(\frac{2\pi n_i}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n_i}{Q}+\theta)} & e^{-j(\frac{2\pi m_i}{Q}+\theta)} \end{bmatrix} \tag{15}$$

where $k_i, l_i \in [0, P)$ and $m_i, n_i \in [0, Q)$ are integers for $i = 1, 2$. Before calculating the determinant of the difference of the two matrices, we first state some well-known facts about 2×2 orthogonal design [15].

Lemma 6. *For any non-zero* 2×2 *matrix*

$$M = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix},$$

we have

1. $\det M = |a|^2 + |b|^2$.
2. $M^*M = MM^* = (\det M)I_2$.
3. $\det M = 0$ *iff* $M = 0_2$.
4. $M^{-1} = \frac{M^*}{\det M}$ *and* $\bar{M}^{-1} = \frac{M^t}{\det M}$.

where $|a|$ *denotes the absolute value of the complex number* a. 0_2 *indicates the* 2×2 *zero matrix.*

Proof. Straightforward algebra. □

This lemma shows that the determinant of any non-zero matrix with an orthogonal design structure is real and positive.

Define

$$O_1 = U_1V_1 - U_2V_2 \quad \text{and} \quad O_2 = U_1\bar{V}_1 - U_2\bar{V}_2. \tag{16}$$

We have

$$\begin{aligned} \det(S_1 - S_2) &= \frac{1}{\sqrt{2}} \det \begin{bmatrix} U_1V_1 - U_2V_2 & U_1\bar{V}_1 - U_2\bar{V}_2 \\ -(\bar{U}_1V_1 - \bar{U}_2V_2) & \bar{U}_1\bar{V}_1 - \bar{U}_2\bar{V}_2 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \det \begin{bmatrix} O_1 & O_2 \\ -\bar{O}_2 & \bar{O}_1 \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} \det O_1 \det(\bar{O}_1 + \bar{O}_2 O_1^{-1} O_2) \end{aligned} \tag{17}$$

if O_1 is invertible. Since U_1, U_2, V_1 and V_2 are orthogonal designs and the addition, multiplication and conjugate operations preserve this property, O_1 and O_2 are also orthogonal designs. By taking advantage of this, when $\det O_1 \neq 0$ and $\det O_2 \neq 0$ the determinant of the difference can be calculated as follows,

$$
\begin{aligned}
\det(S_1 - S_2) &= \tfrac{1}{\sqrt{2}} \det O_1 \det(\bar{O}_1 + \tfrac{\bar{O}_2 O_1^* O_2}{\det O_1}) \\
&= \tfrac{1}{\sqrt{2}} \det O_1 \det(\tfrac{\bar{O}_1 O_2^* O_2}{\det O_2} + \tfrac{\bar{O}_2 O_1^* O_2}{\det O_1}) \\
&= \tfrac{1}{\sqrt{2}} \det O_1 \det(\tfrac{\bar{O}_1 O_2^*}{\det O_2} + \tfrac{\bar{O}_2 O_1^*}{\det O_1}) \det O_2 \\
&= \tfrac{1}{\sqrt{2}} \det(\tfrac{\sqrt{\det O_1 \det O_2}}{\det O_2} \bar{O}_1 O_2^* + \tfrac{\sqrt{\det O_1 \det O_2}}{\det O_1} \bar{O}_2 O_1^*) \\
&= \tfrac{1}{\sqrt{2}} \det(a \bar{O}_1 O_2^* + \tfrac{1}{a} (\bar{O}_1 O_2^*)^t) \\
&= \tfrac{1}{\sqrt{2}} \det(a \begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix} + \tfrac{1}{a} \begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}^t) \\
&= \tfrac{1}{\sqrt{2}} (|a\alpha + \tfrac{1}{a}\alpha|^2 + |a\beta - \tfrac{1}{a}\bar{\beta}|^2) \\
&= \tfrac{1}{\sqrt{2}} |\alpha|^2 (a + \tfrac{1}{a})^2 + \tfrac{1}{2} |a\beta - \tfrac{\bar{\beta}}{a}|^2
\end{aligned}
\tag{18}
$$

where $a = \sqrt{\frac{\det O_1}{\det O_2}}$ is a positive number and (α, β) is the first row of $\bar{O}_1 O_2^*$.

Lemma 7. *For any S_1 and S_2 given in (14) and (15) where $S_1 \neq S_2$, $\det(S_1 - S_2) = 0$ if and only if $O_1 = \pm J \bar{O}_2$, or equivalently*

$$
\left\{ \begin{array}{l} e^{2j\theta} w^+ = x^+ \\ e^{2j\theta} y^+ = z^+, \end{array} \right. \qquad \textit{or} \qquad \left\{ \begin{array}{l} e^{2j\theta} w^- = x^- \\ e^{2j\theta} y^- = z^-, \end{array} \right. \tag{19}
$$

where we defined

$$
\begin{aligned}
w^+ &= e^{2\pi j(\frac{k_1}{P} + \frac{m_1}{Q})} - e^{2\pi j(\frac{k_2}{P} + \frac{m_2}{Q})} + e^{2\pi j(\frac{l_1}{P} + \frac{m_1}{Q})} - e^{2\pi j(\frac{l_2}{P} + \frac{m_2}{Q})} \\
x^+ &= -e^{2\pi j(\frac{k_1}{P} - \frac{n_1}{Q})} + e^{2\pi j(\frac{k_2}{P} - \frac{n_2}{Q})} + e^{2\pi j(\frac{l_1}{P} - \frac{n_1}{Q})} - e^{2\pi j(\frac{l_2}{P} - \frac{n_2}{Q})} \\
y^+ &= e^{2\pi j(\frac{k_1}{P} + \frac{n_1}{Q})} - e^{2\pi j(\frac{k_2}{P} + \frac{n_2}{Q})} + e^{2\pi j(\frac{l_1}{P} + \frac{n_1}{Q})} - e^{2\pi j(\frac{l_2}{P} + \frac{n_2}{Q})} \\
z^+ &= e^{2\pi j(\frac{k_1}{P} - \frac{m_1}{Q})} - e^{2\pi j(\frac{k_2}{P} - \frac{m_2}{Q})} - e^{2\pi j(\frac{l_1}{P} - \frac{m_1}{Q})} + e^{2\pi j(\frac{l_2}{P} - \frac{m_2}{Q})}
\end{aligned}
\tag{20}
$$

and

$$
\begin{aligned}
w^- &= e^{2\pi j(\frac{k_1}{P} + \frac{m_1}{Q})} - e^{2\pi j(\frac{k_2}{P} + \frac{m_2}{Q})} - e^{2\pi j(\frac{l_1}{P} + \frac{m_1}{Q})} + e^{2\pi j(\frac{l_2}{P} + \frac{m_2}{Q})} \\
x^- &= e^{2\pi j(\frac{k_1}{P} - \frac{n_1}{Q})} - e^{2\pi j(\frac{k_2}{P} - \frac{n_2}{Q})} + e^{2\pi j(\frac{l_1}{P} - \frac{n_1}{Q})} - e^{2\pi j(\frac{l_2}{P} - \frac{n_2}{Q})} \\
y^- &= e^{2\pi j(\frac{k_1}{P} + \frac{n_1}{Q})} - e^{2\pi j(\frac{k_2}{P} + \frac{n_2}{Q})} - e^{2\pi j(\frac{l_1}{P} + \frac{n_1}{Q})} + e^{2\pi j(\frac{l_2}{P} + \frac{n_2}{Q})} \\
z^- &= -e^{2\pi j(\frac{k_1}{P} - \frac{m_1}{Q})} + e^{2\pi j(\frac{k_2}{P} - \frac{m_2}{Q})} - e^{2\pi j(\frac{l_1}{P} - \frac{m_1}{Q})} + e^{2\pi j(\frac{l_2}{P} - \frac{m_2}{Q})}
\end{aligned}
\tag{21}
$$

Proof: See Appendix B. ☐

We now present the main theorem.

Theorem 5 (Condition for full diversity). *There exists a θ such that the code in (11) is fully-diverse if and only if P and Q are relatively prime.*

This theorem provides the sufficient and necessary condition for the code set to be fully-diverse. Before proving this theorem, let us first state a few lemmas that are needed in the proof.

Lemma 8. *For any 4 points on the unit circle that add up (as complex numbers) to zero, it is always true that two of them add up to zero. (Clearly, the other two must also have a summation of zero.)*

Proof: See Appendix A. $\square$

Lemma 9. *If P and Q are relatively prime, then for any non-identical pair of (k_1, l_1, m_1, n_1) and (k_2, l_2, m_2, n_2), where $k_1, l_1, k_2, l_2 \in [0, P)$ and $m_1, n_1, m_2, n_2 \in [0, Q)$ are integers, the w^+, x^+, y^+, z^+ (as defined in (20)) cannot be all zeros. Also the w^-, x^-, y^-, z^- (as defined in (21)) cannot be all zeros.*

Proof: See Appendix C. $\square$

Now we are ready to prove Theorem 5.

Proof of Theorem 5: The proof has two steps. First, we prove the sufficiency of the condition, that is, assuming P and Q are relatively prime, we show that there exists a θ such that the code is fully-diverse. If P and Q are relatively prime, by Lemma 9, for any non-identical pair of signal matrices (k_1, l_1, m_1, n_1) and (k_2, l_2, m_2, n_2), the w^+, x^+, y^+, z^+ cannot all be zeros. (For definitions of w^+, x^+, y^+, z^+, see (20)). Define

$$\theta^+_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\arctan(\frac{x^+}{w^+}) \mod 2\pi & \text{if } \ w^+ \neq 0 \\ -\frac{1}{2}\arctan(\frac{w^+}{x^+}) \mod 2\pi & \text{if } \ w^+ = 0, x^+ \neq 0 \\ \frac{1}{2}\arctan(\frac{z^+}{y^+}) \mod 2\pi & \text{if } \ w^+ = x^+ = 0, y^+ \neq 0 \\ -\frac{1}{2}\arctan(\frac{y^+}{z^+}) \mod 2\pi & \text{if } \ w^+ = x^+ = y^+ = 0, z^+ \neq 0 \end{cases},$$

which is the same as

$$\theta^+_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\arctan(\frac{x^+}{w^+}) \mod 2\pi & \text{if } \ w^+ \neq 0 \\ 0 & \text{if } \ w^+ = 0, x^+ \neq 0 \\ \frac{1}{2}\arctan(\frac{z^+}{y^+}) \mod 2\pi & \text{if } \ w^+ = x^+ = 0, y^+ \neq 0 \\ 0 & \text{if } \ w^+ = x^+ = y^+ = 0, z^+ \neq 0 \end{cases}. \tag{22}$$

Also, by Lemma 9, the w^-, x^-, y^-, z^- cannot all be zeros. (For definitions of w^-, x^-, y^-, z^-, see (21)). Define

$$\theta^-_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\arctan(\frac{x^-}{w^-}) \mod 2\pi & \text{if } \ w^- \neq 0 \\ -\frac{1}{2}\arctan(\frac{w^-}{x^-}) \mod 2\pi & \text{if } \ w^- = 0, x^- \neq 0 \\ \frac{1}{2}\arctan(\frac{z^-}{y^-}) \mod 2\pi & \text{if } \ w^- = x^- = 0, y^- \neq 0 \\ -\frac{1}{2}\arctan(\frac{y^-}{z^-}) \mod 2\pi & \text{if } \ w^- = x^- = y^- = 0, z^- \neq 0 \end{cases},$$

which is the same as

$$\theta^{-}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\arctan(\frac{x^-}{w^-}) \mod 2\pi & \text{if} \quad w^- \neq 0 \\ 0 & \text{if} \quad w^- = 0, x^- \neq 0 \\ \frac{1}{2}\arctan(\frac{z^-}{y^-}) \mod 2\pi & \text{if} \quad w^- = x^- = 0, y^- \neq 0 \\ 0 & \text{if} \quad w^- = x^- = y^- = 0, z^- \neq 0 \end{cases} . \tag{23}$$

By choosing

$$\theta \notin \left\{\theta^{+}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2}, ||w^+| = |x^+|, |y^+| = |z^+|, 0 \leq k_1, l_1, k_2, l_2 < P, 0 \leq m_1, n_1, m_2, n_2 < Q\right\} \tag{24}$$

and

$$\theta \notin \left\{\theta^{-}_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2}, ||w^-| = |x^-|, |y^-| = |z^-|, 0 \leq k_1, l_1, k_2, l_2 < P, 0 \leq m_1, n_1, m_2, n_2 < Q\right\}, \tag{25}$$

(19) cannot be true at the same time. Therefore, by Lemma 7, $\det(S_1 - S_2) \neq 0$, which means that the code is fully-diverse. (we can always find an angle in $[0, 2\pi)$ that satisfies (24) since two sets at the right-hand side of (24) and (25) are finite.) This proves the sufficiency of the condition (P and Q are relatively prime) in Theorem 5.

In the second step, we prove the necessity of the condition, that is, assuming that P and Q are not relatively prime, we show that there exist two signals in the code such that the determinant of the difference of the two is zero for any θ. Assume that the greatest common divisor of P and Q is $G > 1$, then there exist positive integers P' and Q' such that $P = P'G$ and $Q = Q'G$. Consider the following two signal matrices S_1 and S_2 as given in (14) and (15), where we set the integers as $k_2 = k_1 - P', l_2 = l_1 - P', m_2 = m_1 + Q', n_2 = n_1 + Q'$ and $k_1 = l_1, k_2 = l_2$. (Here $k_1, l_1 \in [P', P), m_1, n_1 \in [0, Q - Q')$. Since $P > P'$and $Q > Q'$, we can always choose k_1, l_1, m_1, n_1 that satisfy the condition.) It is easy to verify that $w^+ = x^+ = y^+ = z^+ = 0$, which means that the first set of equations of (19) are true for any angle θ. Therefore, $O_1 = J\bar{O}_2$ and from Lemma 7, we have $\det(S_1 - S_2) = 0$, that is, the signal set in (11) is not fully-diverse. □

The following two plots show the diversity products of two $Sp(2)$ codes at different θ. Fig. 1 shows the diversity product of the $P = 7, Q = 3$ $Sp(2)$ code and Fig. 2 shows the diversity product of the $P = 11, Q = 7$ $Sp(2)$ code. Since the angles of the elements in the V matrix of (11) are chosen from Q-PSK signals shifted by an angle θ, we only need to set the changing region of θ to be $[0, 2\pi/Q)$. We can see from the two plots that the $P = 7, Q = 3$ $Sp(2)$ code gets its highest diversity product 0.1870 at $\theta = 0.0419$ and the $P = 7, Q = 3$ $Sp(2)$ code gets its highest diversity product 0.0937 at $\theta = 0$. Although the codes are fully-diverse at any θ.

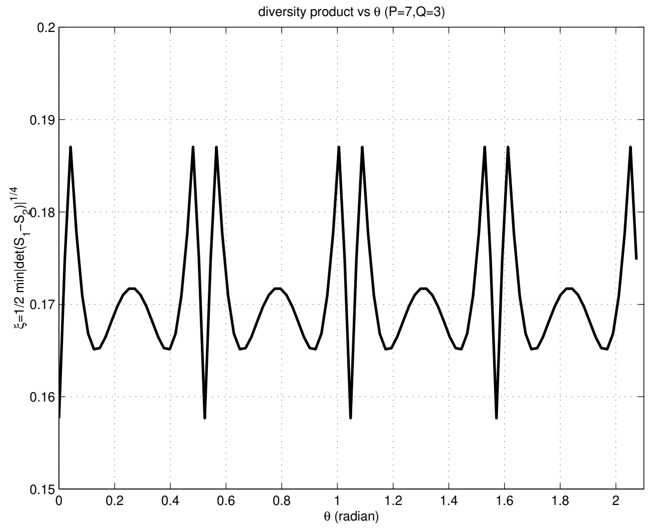

Figure 1: Diversity product of the $P = 7, Q = 3$ $Sp(2)$ code.

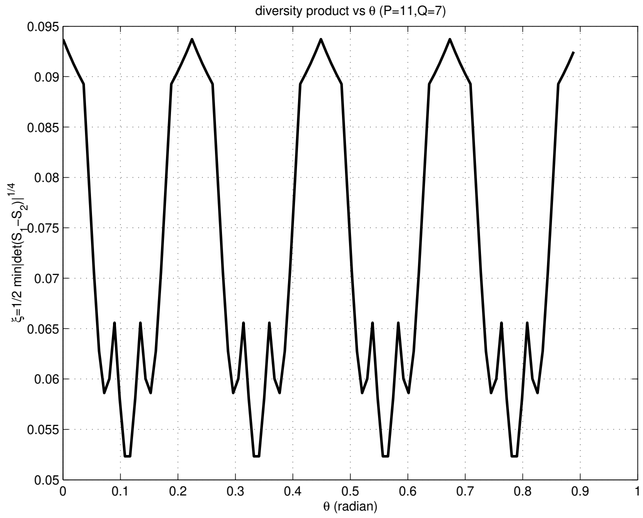

Figure 2: Diversity product of the $P = 11, Q = 7$ $Sp(2)$ code.

3.4 $Sp(2)$ Codes at Higher Rates

In section 3.2, the $Sp(2)$ code is designed with the 2 degrees of freedom in Σ_A and Σ_B unused. For higher rate code design, we can add one of the two degrees of freedom in by letting

$$\Sigma_A = \cos\gamma_i I_2, \Sigma_B = \sin\gamma_i I_2$$

where $\gamma_i \in \Gamma$ for some real set Γ. The code can be constructed as the follows.

$$\mathcal{C}_{P,Q,\theta,\Gamma} = \left\{ \begin{bmatrix} \cos\gamma_i UV & \sin\gamma_i U\bar{V} \\ -\sin\gamma_i \bar{U}V & \cos\gamma_i \bar{U}\bar{V} \end{bmatrix} \middle| \begin{array}{c} U = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j\frac{2\pi k}{P}} & e^{j\frac{2\pi l}{P}} \\ -e^{-j\frac{2\pi l}{P}} & e^{-j\frac{2\pi k}{P}} \end{bmatrix}, \\ V = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{j(\frac{2\pi m}{Q}+\theta)} & e^{j(\frac{2\pi n}{Q}+\theta)} \\ -e^{-j(\frac{2\pi n}{Q}+\theta)} & e^{-j(\frac{2\pi m}{Q}+\theta)} \end{bmatrix}, \\ 0 \le k, l < P, 0 \le m, n < Q, \gamma_i \in \Gamma \end{array} \right\} \tag{26}$$

where P and Q are integers and $\theta \in [0, 2\pi)$ is a constant to be chosen later. We can see that any signal matrix can be identified by the 5-tuple (k, l, m, n, γ_i). It is easy to see that the code in (11) is obtained by setting $\Gamma = \{\frac{\pi}{4}\}$. Since the code has all together $P^2Q^2|\Gamma|$ matrices and the channel is used in blocks of 4 transmissions, the rate of the code is

$$\frac{1}{2}(\log_2 P + \log_2 Q) + \frac{1}{4}\log_2|\Gamma|, \tag{27}$$

where $|\Gamma|$ indicates the cardinality of the set Γ. A conclusion similar to Theorem 5 about the full diversity of the code can be stated as follows.

Theorem 6 (Condition for full diversity). *If P and Q are relatively prime with P odd, and also the set Γ satisfies the following conditions,*

1. $\Gamma \subset (0, \frac{\pi}{2})$.

2. *For any $\gamma \in \Gamma$,*

$$\tan\gamma \neq \pm \frac{\sin\pi(\frac{l}{P}+\frac{m}{Q})}{\sin\pi(\frac{k}{P}+\frac{m}{Q})} \tag{28}$$

where $k, l \in (-P, P)$, $m \in (-Q, Q)$ are integers and $(k, m) \neq (0, 0)$.

3. *For any two different $\gamma_i, \gamma_j \in \Gamma$,*

$$\frac{\sin 2\gamma_i}{\sin 2\gamma_j} \neq \frac{\cos 2\pi\frac{s}{P}}{\cos 2\pi\frac{l}{P}}, \tag{29}$$

where $s, l \in [0, P)$ are integers.

then there exists a θ such that our signal set in (26) is fully-diverse.

Proof: First we need to show that the right hand side formula of the condition (28) is well-defined, that is, we need to show that $\sin\pi(\frac{k}{P}+\frac{m}{Q}) \neq 0$ for any integers $k \in (-P, P), m \in (-Q, Q)$ and $(k, m) \neq (0, 0)$. We prove this by contradiction. Assume that $\sin\pi(\frac{k}{P}+\frac{m}{Q}) = 0$. Therefore, $\pi(\frac{k}{P}+\frac{m}{Q}) = l\pi$, which is equivalent to $\frac{k}{P}+\frac{m}{Q} = l$, for some integer l. Since P and Q are relatively prime, we have $k|P$ and $m|Q$. Since $k \in (-P, P)$ and $m \in (-Q, Q)$, $k = 0$ and $m = 0$. This contradicts the condition that $(k, m) \neq (0, 0)$. We now prove that the right hand side formula of the condition (29) is well-defined, that is, we need to show that $\cos 2\pi\frac{l}{P} \neq 0$ for any $l \in [0, P)$. We prove this by contradiction. Assume that $\cos 2\pi\frac{l}{P} = 0$. Therefore, $2\pi\frac{l}{P} = r\pi + \frac{\pi}{2}$, that is, $\frac{l}{P} = \frac{2r+1}{4}$ for some integer r. Since $2r+1$ is odd, $4|P$ and this contradicts the condition that P is odd.

Now we prove this theorem. Assume that P and Q are relatively prime, P is odd and the set $\Gamma \subset (0, \frac{\pi}{2})$ satisfies the conditions (28) and (29). We want to show that there exists a θ such that the code is fully-diverse. It is equivalent to show that for any non-identical pair of signals S_i, S_j of the code, $\det(S_i - S_j) \neq 0$. Without loss of generality, assume

$$S_i = \begin{bmatrix} \cos\gamma_i U_i V_i & \sin\gamma_i U_i \bar{V}_i \\ -\sin\gamma_i \bar{U}_i V_i & \cos\gamma_i \bar{U}_i \bar{V}_i \end{bmatrix} \text{ and } S_j = \begin{bmatrix} \cos\gamma_i U_j V_j & \sin\gamma_i U_j \bar{V}_j \\ -\sin\gamma_i \bar{U}_j V_j & \cos\gamma_i \bar{U}_j \bar{V}_j \end{bmatrix}. \tag{30}$$

where U_i, V_i are 2×2 unitary matrices given by (15) and U_j, V_j are the two matrices in (15) by replacing i by j. The two signals are different indicates that the two 5-tuples $(k_i, l_i, m_i, n_i, \gamma_i)$ and $(k_j, l_j, m_j, n_j, \gamma_j)$ are not identical. From the proof of Lemma 7, $\det(S_i - S_j)$ is zero if and only if $O_1' = \pm J\bar{O}_2'$, where we define

$$O_1' = \cos\gamma_i U_i V_i - \cos\gamma_j U_j V_j \;\text{ and }\; O_2' = \sin\gamma_i U_i \bar{V}_i - \sin\gamma_j U_j \bar{V}_j. \tag{31}$$

By using (31) and (15), similar to the argument in Appendix B, $O_1' = \pm J\bar{O}_2'$ can be equivalently written as

$$\begin{cases} e^{2j\theta}(\cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{m_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{m_j}{Q})} \pm \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{m_i}{Q})} \mp \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{m_j}{Q})}) \\ \qquad = \mp\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{n_i}{Q})} \pm \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{n_j}{Q})} + \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{n_j}{Q})} \\ \text{or} \\ e^{2j\theta}(\cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{n_j}{Q})} \pm \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{n_i}{Q})} \mp \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{n_j}{Q})}) \\ \qquad = \pm\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{m_i}{Q})} \mp \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{m_j}{Q})} - \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{m_i}{Q})} + \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{m_j}{Q})} \end{cases} \tag{32}$$

Define

$$\begin{aligned} \tilde{w}^+ &= \cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{m_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{m_j}{Q})} + \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{m_i}{Q})} - \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{m_j}{Q})} \\ \tilde{x}^+ &= -\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{n_i}{Q})} + \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{n_j}{Q})} + \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{n_j}{Q})} \\ \tilde{y}^+ &= \cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{n_j}{Q})} + \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{n_i}{Q})} - \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{n_j}{Q})} \\ \tilde{z}^+ &= +\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{m_i}{Q})} - \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{m_j}{Q})} - \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{m_i}{Q})} + \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{m_j}{Q})} \end{aligned} \tag{33}$$

and

$$\begin{aligned} \tilde{w}^- &= \cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{m_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{m_j}{Q})} - \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{m_i}{Q})} + \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{m_j}{Q})} \\ \tilde{x}^- &= +\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{n_i}{Q})} - \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{n_j}{Q})} + \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{n_j}{Q})} \\ \tilde{y}^- &= \cos\gamma_i e^{2\pi j(\frac{k_i}{P}+\frac{n_i}{Q})} - \cos\gamma_j e^{2\pi j(\frac{k_j}{P}+\frac{n_j}{Q})} - \sin\gamma_i e^{2\pi j(\frac{l_i}{P}+\frac{n_i}{Q})} + \sin\gamma_j e^{2\pi j(\frac{l_j}{P}+\frac{n_j}{Q})} \\ \tilde{z}^- &= -\sin\gamma_i e^{2\pi j(\frac{k_i}{P}-\frac{m_i}{Q})} + \sin\gamma_j e^{2\pi j(\frac{k_j}{P}-\frac{m_j}{Q})} - \cos\gamma_i e^{2\pi j(\frac{l_i}{P}-\frac{m_i}{Q})} + \cos\gamma_j e^{2\pi j(\frac{l_j}{P}-\frac{m_j}{Q})} \end{aligned} \tag{34}$$

(32) is thus equivalent to,

$$\begin{cases} e^{2j\theta}\tilde{w}^+ = \tilde{x}^+ \\ e^{2j\theta}\tilde{y}^+ = \tilde{z}^+ \end{cases} \qquad \text{or} \qquad \begin{cases} e^{2j\theta}\tilde{w}^- = \tilde{x}^- \\ e^{2j\theta}\tilde{y}^- = \tilde{z}^-. \end{cases} \tag{35}$$

The following lemma can be proved.

Lemma 10. *For any non-identical $(k_i, l_i, m_i, n_i, \gamma_i)$ and $(k_j, l_j, m_j, n_j, \gamma_j)$, where $k_i, l_i, k_j, l_j \in [0, P)$, $m_i, n_i, m_j, n_j \in [0, Q)$ are integers and $\gamma_i, \gamma_j \in \Gamma$, if P and Q are relatively prime with P odd and if the set $\Gamma \subset (0, \frac{\pi}{2})$ satisfies conditions (28) and (29), then the $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ (as defined in (33)) cannot be zeros simultaneously. Also the $\tilde{w}^-, \tilde{x}^-, \tilde{y}^-, \tilde{z}^-$ (as defined in (34)) cannot be zeros simultaneously.*

Proof: See Appendix D. $\square$

Similar to the proof of Theorem 5, by Lemma 10, the $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ cannot be zeros simultaneously. Define

$$\tilde{\theta}^+_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\arctan(\frac{\tilde{x}^+}{\tilde{w}^+}) \mod 2\pi & \text{if } \tilde{w}^+ \neq 0 \\ -\frac{1}{2}\arctan(\frac{\tilde{w}^+}{\tilde{x}^+}) \mod 2\pi & \text{if } \tilde{w}^+ = 0, \tilde{x}^+ \neq 0 \\ \frac{1}{2}\arctan(\frac{\tilde{z}^+}{\tilde{y}^+}) \mod 2\pi & \text{if } \tilde{w}^+ = \tilde{x}^+ = 0, \tilde{y}^+ \neq 0 \\ -\frac{1}{2}\arctan(\frac{\tilde{y}^+}{\tilde{z}^+}) \mod 2\pi & \text{if } \tilde{w}^+ = \tilde{x}^+ = \tilde{y}^+ = 0, \tilde{z}^+ \neq 0 \end{cases},$$

which is the same as

$$\tilde{\theta}^+_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\arctan(\frac{\tilde{x}^+}{\tilde{w}^+}) \mod 2\pi & \text{if } \tilde{w}^+ \neq 0 \\ 0 & \text{if } \tilde{w}^+ = 0, \tilde{x}^+ \neq 0 \\ \frac{1}{2}\arctan(\frac{\tilde{z}^+}{\tilde{y}^+}) \mod 2\pi & \text{if } \tilde{w}^+ = \tilde{x}^+ = 0, \tilde{y}^+ \neq 0 \\ 0 & \text{if } \tilde{w}^+ = \tilde{x}^+ = \tilde{y}^+ = 0, \tilde{z}^+ \neq 0 \end{cases}.$$

Also, the $\tilde{w}^-, \tilde{x}^-, \tilde{y}^-, \tilde{z}^-$ cannot be zeros simultaneously. Define

$$\tilde{\theta}^-_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} \frac{1}{2}\arctan(\frac{\tilde{x}^-}{\tilde{w}^-}) \mod 2\pi & \text{if } \tilde{w}^- \neq 0 \\ -\frac{1}{2}\arctan(\frac{\tilde{w}^-}{\tilde{x}^-}) \mod 2\pi & \text{if } \tilde{w}^- = 0, \tilde{x}^- \neq 0 \\ \frac{1}{2}\arctan(\frac{\tilde{z}^-}{\tilde{y}^-}) \mod 2\pi & \text{if } \tilde{w}^- = \tilde{x}^- = 0, \tilde{y}^- \neq 0 \\ -\frac{1}{2}\arctan(\frac{\tilde{y}^-}{\tilde{z}^-}) \mod 2\pi & \text{if } \tilde{w}^- = \tilde{x}^- = \tilde{y}^- = 0, \tilde{z}^- \neq 0 \end{cases},$$

which is the same as

$$\tilde{\theta}^-_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2} = \begin{cases} -\frac{1}{2}\arctan(\frac{\tilde{x}^-}{\tilde{w}^-}) \mod 2\pi & \text{if } \tilde{w}^- \neq 0 \\ 0 & \text{if } \tilde{w}^- = 0, \tilde{x}^- \neq 0 \\ -\frac{1}{2}\arctan(\frac{\tilde{z}^-}{\tilde{y}^-}) \mod 2\pi & \text{if } \tilde{w}^- = \tilde{x}^- = 0, \tilde{y}^- \neq 0 \\ 0 & \text{if } \tilde{w}^- = \tilde{x}^- = \tilde{y}^- = 0, \tilde{z}^- \neq 0 \end{cases}.$$

By choosing

$$\theta \notin \left\{\tilde{\theta}^+_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2}, ||\tilde{w}^+| = |\tilde{x}^+|, |\tilde{y}^+| = |\tilde{z}^+|, 0 \leq k_1, l_1, k_2, l_2 < P, 0 \leq m_1, n_1, m_2, n_2 < Q\right\} \quad (36)$$

and

$$\theta \notin \left\{\tilde{\theta}^-_{k_1,l_1,m_1,n_1,k_2,l_2,m_2,n_2}, ||\tilde{w}^-| = |\tilde{x}^-|, |\tilde{y}^-| = |\tilde{z}^-|, 0 \leq k_1, l_1, k_2, l_2 < P, 0 \leq m_1, n_1, m_2, n_2 < Q\right\}, \quad (37)$$

(35) cannot be true. Therefore $\det(S_i - S_j) \neq 0$, which means that the code is fully-diverse. (we can always find an angle in $[0, 2\pi)$ that satisfies both (36) and (37) since the two sets at the right-hand side of (36) and (37) are finite.) $\square$

4 Decoding the $Sp(2)$ Code

One of the most prominent properties of our $Sp(2)$ code is that it is a generalization of the orthogonal design. In this section, we show how this property can be used to get linear-algebraic decoding, which means that the receiver can be made to form a system of linear equations in the unknowns.

4.1 Formulation

The ML decoder for differential USTM is given in (5), which, in our system, can be written as follows.

$$\begin{aligned} & \arg\max_{l=0,\ldots,L-1} \|X_\tau - V_l X_{\tau-1}\|_F^2 \\ = \; & \arg\max_{i,j} \left\| X_\tau - \frac{1}{\sqrt{2}} \begin{bmatrix} U_i & 0 \\ 0 & \bar{U}_i \end{bmatrix} \begin{bmatrix} I_2 & I_2 \\ -I_2 & I_2 \end{bmatrix} \begin{bmatrix} V_j & 0 \\ 0 & \bar{V}_j \end{bmatrix} X_{\tau-1} \right\|_F^2 \\ = \; & \arg\max_{i,j} \left\| \begin{bmatrix} U_i^* & 0 \\ 0 & U_i^t \end{bmatrix} X_\tau - \frac{1}{\sqrt{2}} \begin{bmatrix} V_j & \bar{V}_j \\ -V_j & \bar{V}_j \end{bmatrix} X_{\tau-1} \right\|_F^2 \end{aligned}$$

By writing the matrices in the norm column by column, we get

$$\arg\max_{i,j} \left\| \begin{bmatrix} \begin{bmatrix} U_i^* & 0 \\ 0 & U_i^t \end{bmatrix} X_1^{(\tau)} \\ \vdots \\ \begin{bmatrix} U_i^* & 0 \\ 0 & U_i^t \end{bmatrix} X_N^{(\tau)} \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} \begin{bmatrix} V_j & \bar{V}_j \\ -V_j & \bar{V}_j \end{bmatrix} X_1^{(\tau-1)} \\ \vdots \\ \begin{bmatrix} V_j & \bar{V}_j \\ -V_j & \bar{V}_j \end{bmatrix} X_N^{(\tau-1)} \end{bmatrix} \right\|_F^2$$

where $X_i^{(\tau)}$ denotes the ith column of X_τ and $X_i^{(\tau-1)}$ denotes the ith column of $X_{\tau-1}$. It is obvious that $X_i^{(\tau)}$ and $X_i^{(\tau-1)}$ are 4×1 column vectors. We further denote the (i,j)-th entry of X_τ as $X_{ij}^{(\tau)}$ and denote the (i,j)-th entry of X_τ as $X_{ij}^{(\tau-1)}$ for $i = 1,2,3,4$ and $j = 1,2,\cdots,N$. The ML decoder is equivalent to

$$\arg\max_{i,j} \left\| \begin{bmatrix} \begin{bmatrix} U_i^* \left[x_{11}^{(\tau)}, x_{21}^{(\tau)}\right]^t \\ U_j^t \left[x_{31}^{(\tau)}, x_{41}^{(\tau)}\right]^t \end{bmatrix} \\ \vdots \\ \begin{bmatrix} U_i^* \left[x_{N1}^{(\tau)}, x_{N2}^{(\tau)}\right]^t \\ U_j^t \left[x_{N3}^{(\tau)}, x_{N4}^{(\tau)}\right]^t \end{bmatrix} \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} \begin{bmatrix} V_j \left[x_{11}^{(\tau-1)}, x_{21}^{(\tau-1)}\right]^t + \bar{V}_j \left[x_{31}^{(\tau-1)}, x_{41}^{(\tau-1)}\right]^t \\ -V_j \left[x_{11}^{(\tau-1)}, x_{21}^{(\tau-1)}\right]^t + \bar{V}_j \left[x_{31}^{(\tau-1)}, x_{41}^{(\tau-1)}\right]^t \end{bmatrix} \\ \vdots \\ \begin{bmatrix} V_j \left[x_{1N}^{(\tau-1)}, x_{2N}^{(\tau-1)}\right]^t + \bar{V}_j \left[x_{3N}^{(\tau-1)}, x_{4N}^{(\tau-1)}\right]^t \\ -V_j \left[x_{1N}^{(\tau-1)}, x_{2N}^{(\tau-1)}\right]^t + \bar{V}_j \left[x_{3N}^{(\tau-1)}, x_{4N}^{(\tau-1)}\right]^t \end{bmatrix} \end{bmatrix} \right\|_F^2 .$$

From the design of the code, we know that the matrices U_i, V_j and their conjugates and transposes are orthogonal designs. For any orthogonal design $M = \begin{bmatrix} a & b \\ -\bar{b} & \bar{a} \end{bmatrix}$ and any 2-dimensional vector $X = [x_1, x_2]^t$,

MX can be written equivalently as

$$\begin{bmatrix} \Re x_1 & \Im x_1 & -\Re x_2 & \Im x_2 \\ \Re x_2 & -\Im x_2 & \Re x_1 & \Im x_1 \end{bmatrix} \begin{bmatrix} \Re a \\ \Im a \\ \Re b \\ \Im b \end{bmatrix} + i \begin{bmatrix} \Im x_1 & -\Re x_1 & -\Im x_2 & -\Re x_2 \\ \Im x_2 & \Re x_2 & \Im x_1 & -\Re x_1 \end{bmatrix} \begin{bmatrix} \Re a \\ \Im a \\ \Re b \\ \Im b \end{bmatrix},$$

i.e. the roles of M and X can be interchanged. Therefore, by careful calculation, the ML decoding can be shown to be equivalent to

$$\arg \max_{0\le k,l<P,0\le m,n<Q} \left\| \begin{bmatrix} \begin{bmatrix} \mathcal{A}_1(X_1^{(\tau)}) & -\mathcal{C}_1(X_1^{(\tau-1)}) \\ \mathcal{B}_1(X_1^{(\tau)}) & -\mathcal{D}_1(X_1^{(\tau-1)}) \end{bmatrix} \\ \cdots \\ \begin{bmatrix} \mathcal{A}_N(X_N^{(\tau)}) & -\mathcal{C}_N(X_N^{(\tau-1)}) \\ \mathcal{B}_N(X_N^{(\tau)}) & -\mathcal{D}_N(X_N^{(\tau-1)}) \end{bmatrix} \end{bmatrix} \begin{bmatrix} \cos\frac{2\pi k}{P} \\ \sin\frac{2\pi k}{P} \\ \cos\frac{2\pi l}{P} \\ \sin\frac{2\pi l}{P} \\ \cos(\frac{2\pi m}{Q}+\theta) \\ \sin(\frac{2\pi m}{Q}+\theta) \\ \cos(\frac{2\pi n}{Q}+\theta) \\ \sin(\frac{2\pi n}{Q}+\theta) \end{bmatrix} \right\|_F^2, \tag{38}$$

where

$$\mathcal{A}_i(X_i^{(\tau)}) = \frac{1}{\sqrt{2}} \begin{bmatrix} \Re x_{i1}^{(\tau)} & \Im x_{i1}^{(\tau)} & -\Re x_{i2}^{(\tau)} & \Im x_{i2}^{(\tau)} \\ \Re x_{i2}^{(\tau)} & -\Im x_{i2}^{(\tau)} & \Re x_{i1}^{(\tau)} & \Im x_{i1}^{(\tau)} \\ \Re x_{i3}^{(\tau)} & -\Im x_{i3}^{(\tau)} & -\Re x_{i4}^{(\tau)} & -\Im x_{i4}^{(\tau)} \\ \Re x_{i4}^{(\tau)} & \Im x_{i4}^{(\tau)} & \Re x_{i3}^{(\tau)} & -\Im x_{i3}^{(\tau)} \end{bmatrix} \tag{39}$$

$$\mathcal{B}_i(X_i^{(\tau)}) = \frac{1}{\sqrt{2}} \begin{bmatrix} \Im x_{i1}^{(\tau)} & -\Re x_{i1}^{(\tau)} & -\Im x_{i2}^{(\tau)} & -\Re x_{i2}^{(\tau)} \\ \Im x_{i2}^{(\tau)} & \Re x_{i2}^{(\tau)} & \Im x_{i1}^{(\tau)} & -\Re x_{i1}^{(\tau)} \\ \Im x_{i3}^{(\tau)} & \Re x_{i3}^{(\tau)} & -\Im x_{i4}^{(\tau)} & \Re x_{i4}^{(\tau)} \\ \Im x_{i4}^{(\tau)} & -\Re x_{i4}^{(\tau)} & \Im x_{i3}^{(\tau)} & \Re x_{i3}^{(\tau)} \end{bmatrix} \tag{40}$$

$$\mathcal{C}_i(X_i^{(\tau-1)}) = \frac{1}{2} \begin{bmatrix} \Re(x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Im(-x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Re(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(-x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Re(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) & \Re(-x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Im(-x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) \\ \Re(-x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Im(x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Re(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Re(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(-x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) & \Re(x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Im(x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) \end{bmatrix} \tag{41}$$

$$\mathcal{D}_i(X_i^{(\tau-1)}) = \frac{1}{2} \begin{bmatrix} \Im(x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Re(x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Im(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Re(x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Im(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Re(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(-x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Re(x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) \\ \Im(-x_{i1}^{(\tau-1)} + x_{i3}^{(\tau-1)}) & \Re(-x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Im(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Re(-x_{i2}^{(\tau-1)} - x_{i4}^{(\tau-1)}) \\ \Im(-x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Re(x_{i2}^{(\tau-1)} + x_{i4}^{(\tau-1)}) & \Im(x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) & \Re(-x_{i1}^{(\tau-1)} - x_{i3}^{(\tau-1)}) \end{bmatrix} \tag{42}$$

and $\left[\cos\frac{2\pi k}{P}, \sin\frac{2\pi k}{P}, \cos\frac{2\pi l}{P}, \sin\frac{2\pi l}{P}, \cos(\frac{2\pi m}{Q}+\theta), \sin(\frac{2\pi m}{Q}+\theta), \cos(\frac{2\pi n}{Q}+\theta), \sin(\frac{2\pi n}{Q}+\theta)\right]^t$ is the vector of unknowns. Notice that $\mathcal{B}_i(X_i^{(\tau)})$ can also be constructed as $\mathcal{A}_i(-jX_i^{(\tau)})$. We can see that formula (38) is quadratic in the $\sin$ and $\cos$ of the unknowns. Thus, it is possible to use fast decoding algorithms such as sphere decoding to achieve the ML solution in polynomial time.

Next we discuss the sphere decoding for the codes given in (26). For each of the angle γ_i, a sphere decoding is applied and one signal is gained. In doing the sphere decoding for each γ_i, matrices $\mathcal{A}_i, \mathcal{B}_i$ for $i = 1, 2, \cdots, N$ are the same as those in (40), but the $\mathcal{C}_i$ and $\mathcal{D}_i$ matrices should be modified to

$$\mathcal{C}_i = \begin{bmatrix} \Re(\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Re(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) \\ \Re(-\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Re(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) \end{bmatrix} \tag{43}$$

$$\mathcal{D}_i = \begin{bmatrix} \Im(\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Im(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) \\ \Im(-\cos\gamma_i x_{i1}^{(\tau-1)} + \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Im(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i2}^{(\tau-1)} - \sin\gamma_i x_{i4}^{(\tau-1)}) \\ \Im(-\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Re(\cos\gamma_i x_{i2}^{(\tau-1)} + \sin\gamma_i x_{i4}^{(\tau-1)}) & \Im(\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) & \Re(-\cos\gamma_i x_{i1}^{(\tau-1)} - \sin\gamma_i x_{i3}^{(\tau-1)}) \end{bmatrix}. \tag{44}$$

For each transmission, $|\Gamma|$ sphere decoding is used and therefore $|\Gamma|$ signal matrices are obtained in total. ML decoding (5) is used to get the optimal one. The complexity of this decoding algorithm is $|\Gamma|$ times the original one, but it is still cubic in the number of possible transmission signals.

4.2 Remarks on Sphere Decoding

Below are some remarks on the implementation of sphere decoding in our systems.

1. The main idea of sphere decoding is that instead of searching all possible lattice points, we restrict our search to a finite number of lattice points contained in a sphere of radius $\sqrt{C}$. In digital communication, the choice of the searching radius is very crucial to the speed of the algorithm. If the initial radius is chosen to be very large, then we are actually searching most of the points, by which not too much improvement on computational complexity can be gained over ML decoding. On the other hand, if the radius is chosen to be too small, then there may be no point in the sphere that we are searching. It is better to start with a small value then increase it gradually. In [16], the author proposed to chose the packing radius or the estimated packing radius to be the initial searching radius. In this paper, we use another initialization for the searching radius based on the noise level as in [17].

2. The noise of the system is given in (4). Since $W_\tau, W_{\tau-1}$ and the transmitted unitary matrix V_{z_τ} are independent, it is easy to prove that the noise has zero mean and variance $2NI_4$. Each component of the $4 \times N$-dimensional noise vector has zero means and variance 2. Therefore the random variable

$v = \|W'_\tau\|_F^2$ has χ^2 distribution with mean $4N$. We initialize the searching radius $\sqrt{C}$ in such a way that the probability that the correct signal is in the sphere is 0.9, that is, P $(\|v\|_F < \sqrt{C}) = 0.9$. If no point is found in the sphere, then we raise the searching radius to have the probability increased to 0.99 and so on. Using this algorithm, the probability that we can find a point during the first search is high.

3. From (38), it can be seen that the unknowns are in forms of $\sin$ and $\cos$. Notice that for any $\alpha \neq \beta \in [0, 2\pi)$, $\sin\alpha = \sin\beta$ if and only if $\beta = (2k+1)\pi - \alpha$ for some integer k. When P, Q are odd, we see that π cannot be in the set $\Theta_P = \{\frac{2\pi k}{P} | k = 0, 1, 2, ..., P-1\}$, which is the set of all possible angles of U_i's entries, and $\pi + \theta$ cannot be in the set $\Theta_Q = \{\frac{2\pi k}{Q} + \theta | k = 0, 1, 2, ..., P-1\}$, which is the set of all possible angles of V_i's entries. Therefore, the map $f_P : \Theta_P \to \{\sin x | x \in \Theta_P\}$ by $f_P(\theta) = \sin\theta$ and the map $f_Q : \Theta_Q \to \{\sin x | x \in \Theta_Q\}$ by $f_Q(\theta) = \sin\theta$ are one-to-one and onto maps. The independent unknowns k, l, m, n can thus be replaced equivalently by their sins: $\sin\frac{2\pi k}{P}, \sin\frac{2\pi l}{P}, \sin(\frac{2\pi m}{Q} + \theta), \sin(\frac{2\pi n}{Q} + \theta)$.

4. Notice that there are only 4 independent unknowns but 8 components in the unknown vector in (38). We combine the $2i$-th component (with the form of $\cos x$) and the $(2i+1)$-th component (with the form of $\sin x$) for $i = 1, 2, 3, 4$. From previous discussions we know that for any value in the set $\{\sin x | x \in \Theta_P\}$ or $\{\sin x | x \in \Theta_Q\}$, there is only one possible value in Θ_P or Θ_Q whose sin equals the value. Therefore, there is only one possible value of the $\cos$. In other words, for any possible value of the $2i$th component, there is one unique value for the $(2i+1)$th component. Therefore, it is natural to combine the $2i$-th and the $(2i+1)$-th components. To simplify the programming, while considering the searching range of each unknown variable, we skip the $2i$-th component and only consider the $(2i+1)$-th one. For example, instead of analyzing all the possible values of $\sin\frac{2\pi n}{Q}$ (the 8th component) satisfying $q_{88}\sin^2\frac{2\pi n}{Q} + q_{77}(\cos\frac{2\pi n}{Q} + S_8)^2 \leq C$, we consider all the possible values of $\sin\frac{2\pi n}{Q}$ satisfying $q_{88}\sin^2\frac{2\pi n}{Q} \leq C$ [16]. It may seem that we search more points than we need, but actually the extra points will be eliminated in the next step.

5 Simulation Results

In this section, we give examples of $Sp(2)$ codes and also the simulated performances of the codes for different rates. The fading coefficient from each transmit antenna to each receive antenna is modeled independently as a complex Gaussian variable with zero mean and unit variance and keeps constant for $2M$ channel uses. At each channel use, zero mean, unit variance complex Gaussian noise is added to each receive antenna. The block error rate (bler), which corresponds to errors in decoding the 4×4 transmitted matrices, is demonstrated as the error event of interest. We also compare the performances of the $Sp(2)$ codes with those of some of the group codes in [4], the differential Cayley codes [10], the 2×2 Alamouti's complex orthogonal designs [15] of the form

$$\mathcal{C}(z_1, z_2) = \begin{bmatrix} z_1 & z_2 \\ -z_2^* & z_1^* \end{bmatrix}, \tag{45}$$

and the 4×4 complex orthogonal design

$$\mathcal{C}(z_1, z_2, z_3) = \begin{bmatrix} z_1 & z_2 & z_3 & 0 \\ -z_2^* & z_1^* & 0 & -z_3 \\ -z_3^* & 0 & z_1^* & z_2 \\ 0 & z_3^* & -z_2^* & z_2 \end{bmatrix} \tag{46}$$

proposed in [18].

5.1 $Sp(2)$ Code vs. Cayley Code and Complex Orthogonal Designs

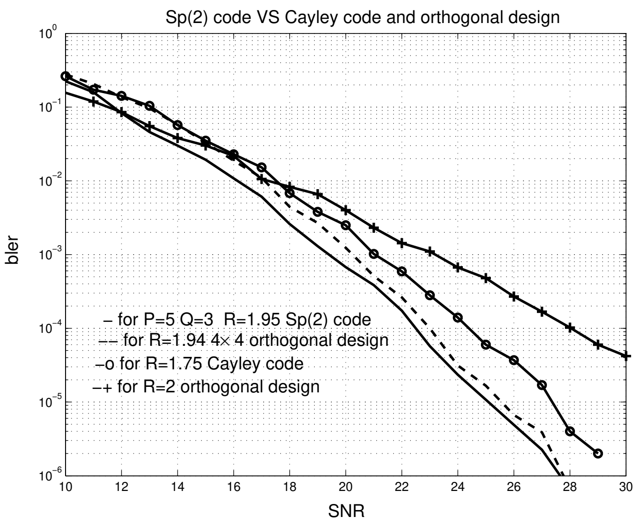

Figure 3: Comparison of the rate 1.95 $Sp(2)$ code with the rate 1.75 differential Cayley code, the rate 2, 2×2 and the rate 1.94, 4×4 complex orthogonal design with $N = 1$ receive antennas.

The first example is the $P = 5, Q = 3, \theta = 0$ $Sp(2)$ code, that is, the entries of the U matrix of the code are chosen from the 5-PSK signal set $\{1, e^{j\frac{2\pi}{5}}, e^{j\frac{4\pi}{5}}, e^{j\frac{6\pi}{5}}, e^{j\frac{8\pi}{5}}\}$, and the entries of the V matrix are chosen from the 3-PSK signal set $\{1, e^{j\frac{2\pi}{3}}, e^{j\frac{4\pi}{3}}\}$. Therefore by (12), the rate of the code is 1.95. We compare it with the rate 2, 2×2 complex orthogonal design given by (45), where z_1, z_2 are chosen from the 4-PSK signal set

$\{1, e^{j\frac{2\pi}{4}}, e^{j\frac{4\pi}{4}}, e^{j\frac{6\pi}{4}}\}$, a rate 1.75 differential Cayley code with parameters $Q = 7, r = 2$ [10] and also the rate 1.94, 4×4 complex orthogonal design given by (46), where z_1, z_2, z_3 are chosen from the 6-PSK signal set $\{1, e^{j\frac{2\pi}{6}}, e^{j\frac{4\pi}{6}}, e^{j\frac{6\pi}{6}}, e^{j\frac{8\pi}{6}}, e^{j\frac{10\pi}{6}}\}$. The number of receive antennas is 1. The performance curves are shown in Fig. 3. The solid line indicates the bler of the $Sp(2)$ code. The lines with circles indicates the bler of the differential Cayley code. The line with plus signs and the dashed line show the bler of the 2×2 and 4×4 complex orthogonal designs respectively. From the plot, we can see that the $Sp(2)$ code has the lowest bler at high SNR. For example, at a bler of 10^{-3}, the $Sp(2)$ code is 2dB better than the differential Cayley code, even though the Cayley code has a lower rate, 1dB better than the 4×4 complex orthogonal design and 4dB better than the 2×2 complex orthogonal design.

5.2 $Sp(2)$ Code vs. Group Constellations

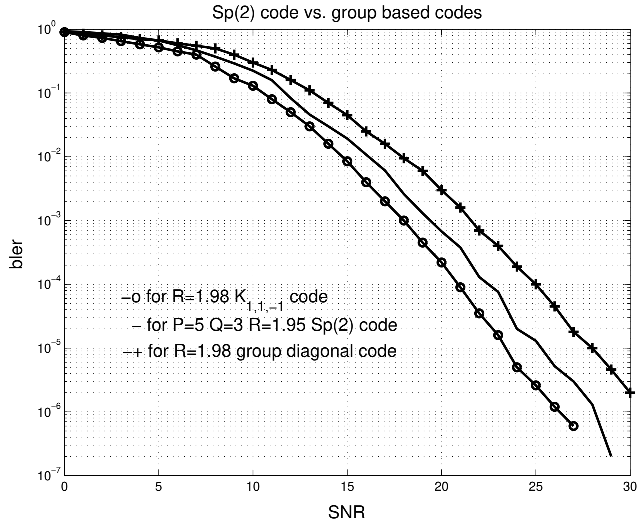

Figure 4: Comparison of the rate 1.95 $Sp(2)$ code with the rate 1.98 group-based $K_{1,1,-1}$ code and a rate 1.98 group-based diagonal code with $N = 1$ receive antennas.

In this subsection, we compare the same $Sp(2)$ code with a group-based diagonal code and the $K_{1,1,-1}$ code both at rate 1.98 [4]. The $K_{1,1,-1}$ code is in one of the 6 types of the finite fixed-point-free groups given in [4]. The number of receive antennas is 1. In Fig. 4, the solid line indicates the bler of the $Sp(2)$ code and the line with circles and plus signs show the bler of the $K_{1,1,-1}$ code and the diagonal code respectively. The plot indicates that the $Sp(2)$ code is better than the diagonal code but worse than the $K_{1,1,-1}$ code according to the bler. For example, at a bler of 10^{-3}, the $Sp(2)$ code is 2dB better than the diagonal code, but 1.5dB worse than the $K_{1,1,-1}$ group code. However, decoding the $K_{1,1,-1}$ code requires an exhaustive search over the entire constellation.

5.3 $Sp(2)$ Codes vs. Complex Orthogonal Designs

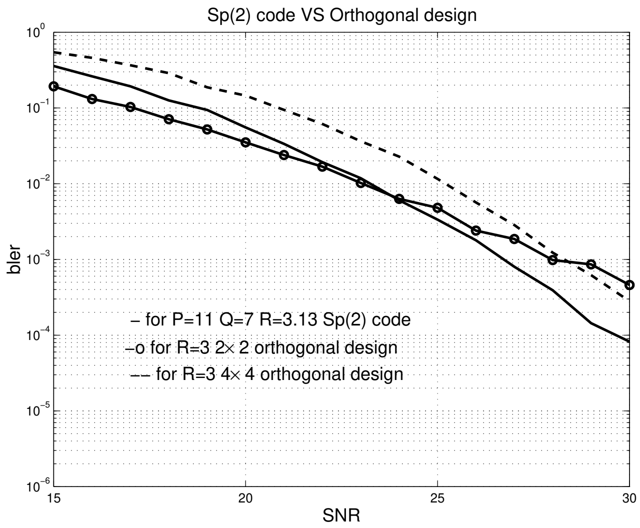

Figure 5: Comparison of the rate 3.13 $Sp(2)$ code with the rate 3, 2×2 and 4×4 complex orthogonal designs with $N = 1$ receive antenna.

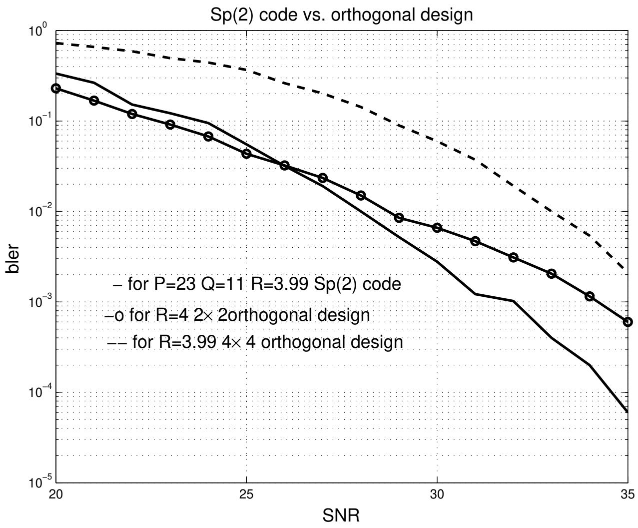

Figure 6: Comparison of the rate 3.99 $Sp(2)$ code with the rate 4, 2×2 and rate 3.99, 4×4 complex orthogonal designs with $N = 1$ receive antenna.

The comparisons of the $Sp(2)$ codes with complex orthogonal designs at rate approximately 3 and 4 are shown in the following two plots. In Fig. 5, the solid line indicates the bler of the $Sp(2)$ code of $P = 11, Q = 7, \theta = 0$. The line with circles shows the bler of the 2×2 complex orthogonal design (45) with z_1, z_2 chosen from 8-PSK. The dashed line indicates the bler of the rate 3, 4×4 complex orthogonal design (46) with z_1, z_2, z_3 chosen from 16-PSK. Therefore, the rate of the $Sp(2)$ code is 3.13 and the rate of the 2×2 and 4×4 orthogonal designs is 3. Similarly, in Fig. 6, the solid line indicates the bler of the $Sp(2)$ code of $P = 23, Q = 11, \theta = 0$. The line with circles shows the bler of the 2×2 complex orthogonal design (45) with z_1, z_2 chosen from 16-PSK and the dashed line indicates the bler of the 4×4 complex orthogonal design (46) with z_1, z_2, z_3 chosen from 40-PSK. Therefore, the rate of the $Sp(2)$ code is 3.99 and the rates of the 2×2 and 4×4 complex orthogonal designs are 4 and 3.99. The number of receive antennas is 1. We can see from the two figures that the $Sp(2)$ codes are better than the 4×4 complex orthogonal designs for all the SNRs and are better than the 2×2 complex orthogonal designs at high SNR.

5.4 Performance of $Sp(2)$ Codes at Higher Rates

In this subsection, simulated performances of the $Sp(2)$ codes at higher rates as given in (26) are shown for different Γ and are compared with the corresponding original codes given in (11), whose Γ is $\{\frac{\pi}{4}\}$.

The first example is the $P = 11, Q = 7, \theta = 0, \Gamma = \{\frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}\} + 0.012$ and $\Gamma = \{\frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\} + 0.02$ $Sp(2)$ codes. A small value is added to the set that is uniform on $(0, \frac{\pi}{2})$ to make the resulting Γ set satisfy conditions (28) and (29), that is, to guarantee the full diversity of the code. According to (27), the rates of the codes are 3.5296 and 3.7139. In Fig. 7, the dashed line and the line with plus signs indicate the bler of the $Sp(2)$ codes we just mentioned, which we call the new codes, and the solid line shows the bler of the $P = 11, Q = 7, \theta = 0$ $Sp(2)$ code with $\Gamma = \{\frac{\pi}{4}\}$ and rate 3.1334, which we call the original code. The figure shows that the new codes are about only 1dB and 2dB worse than the original one with rates 0.3962 and 0.5805 higher. The bler of the rate 4 non-group code given in [4] is also shown in the figure by the line with circles. We can see that performance of the new code is very close to that of the non-group code with rate 0.4704 lower. The result is actually encouraging since the design of the non-group is very difficult and its decoding needs exhaustive search over $2^{16} = 65,536$ possible signal matrices.

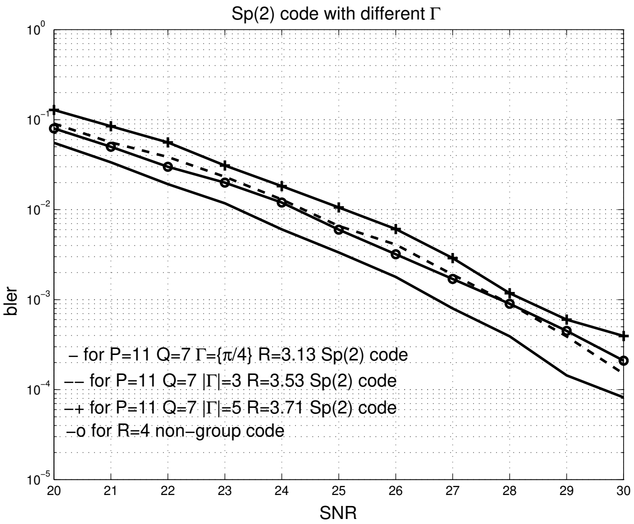

Figure 7: Comparison of $P = 11, Q = 7, \theta = 0$ $Sp(2)$ codes of $\Gamma = \{\frac{\pi}{4}\}, R = 3.1334$, $\Gamma = \{\frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}\} + 0.012, R = 3.5296$ and $\Gamma = \{\frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\} + 0.02, R = 3.7139$.

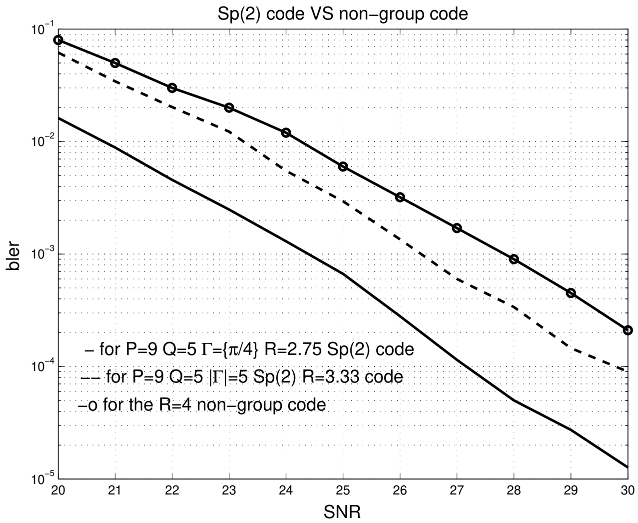

Figure 8: Comparison of $P = 9, Q = 5, \theta = 0.0377$ $Sp(2)$ code of $\Gamma = \{\frac{\pi}{4}\}, R = 2.7459$ and $\Gamma = \{\frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\} + 0.016, R = 3.3264$ with the non-group code.

The second example is the $P = 9, Q = 5, \theta = 0.0377, \Gamma = \{\frac{\pi}{12}, \frac{\pi}{6}, \frac{\pi}{4}, \frac{\pi}{3}, \frac{5\pi}{12}\} + 0.016$ $Sp(2)$ code. The rate of the code is therefore 3.3264 by formula (27). In Fig. 8, the dashed line indicates the bler of the rate

3.3264 $Sp(2)$ code we just mention, which we call the new code, and the solid line shows the bler of the $P = 9, Q = 5, \theta = 0.0377$ $Sp(2)$ with $\Gamma = \{\frac{\pi}{4}\}$ and rate 2.7459, which we call the original code. The figure shows that the new code is only about 2dB worse than the original one with rate 0.5805 higher. Also, the bler of the rate 4 non-group code given in [10] is shown in the figure by the line with circles. We can see that the new code is 1dB better than the non-group code with rate 0.6736 lower. As mentioned before, the result is actually encouraging since the design of the non-group is very difficult and its decoding needs exhaustive search over $2^{16} = 65,536$ possible signal matrices.

6 Conclusion

Our work can be regarded as an extension of [5]. In this paper, we work on the symplectic group $Sp(n)$ which has dimension $n(2n+1)$ and rank n and design codes based on $Sp(2)$. The group is not fpf, but we propose a method to design fully-diverse codes which are subsets of the group. The constellations designed are suitable for four transmit antennas and any number of receive antenna systems. The special symplectic structure of the codes lend themselves to decoding by linear-algebraic techniques, such as the sphere decoder. Simulation results show that they have better performances than the 2×2 and 4×4 complex orthogonal designs, a group-based diagonal code as well as differential Cayley codes at high SNR. Although they slightly underperform the $k_{1,1,-1}$ finite group code and the carefully designed non-group code, they do not need the exhaustive search (of exponentially growing size) required of such codes and therefore are far superior in term of decoding complexity. Our work shows the promise of studying constellations inspired by a group-theoretic considerations. Other Lie groups (such as G_2 and $SU(3)$) are currently under investigation.

A Proof of Lemma 7

Lemma 7. *For any S_1 and S_2 given in (14) and (15) where $S_1 \neq S_2$, $\det(S_1 - S_2) = 0$ if and only if $O_1 = \pm J\bar{O}_2$, or equivalently*

$$\begin{cases} e^{2j\theta}w^+ = x^+ \\ e^{2j\theta}y^+ = z^+, \end{cases} \qquad \text{or} \qquad \begin{cases} e^{2j\theta}w^- = x^- \\ e^{2j\theta}y^- = z^-, \end{cases}$$

where w^+, x^+, y^+, z^+ are defined in (20) and w^-, x^-, y^-, z^- are defined in (21).

Proof: First we assume that the determinant is zero and prove that $O_1 = \pm J\bar{O}_2$.

Assume $\det(S_1 - S_2) = 0$. If $\det O_1 = 0$, from Lemma 6, $O_1 = 0_2$. Therefore

$$\det(S_1 - S_2) = \frac{1}{\sqrt{2}} \det \begin{bmatrix} 0 & O_2 \\ -\bar{O}_2 & 0 \end{bmatrix} = \det{}^2 O_2 = 0$$

Thus, $O_2 = 0$ by Lemma 6. This indicates that $S_1 = S_2$, which contradicts $S_1 \neq S_2$. And the same for the case of $\det O_2 = 0$. Therefore, $\det O_1 \neq 0$ and $\det O_2 \neq 0$. From (18), $\det(S_1 - S_2)$ is always non-negative and equals 0 if and only if $|\alpha|^2(a + \frac{1}{a})^2 = 0$ and $|a\beta - \frac{\bar{\beta}}{a}|^2 = 0$. Since $(a + \frac{1}{a})^2 \geq 2$, the determinant equals zero if and only if $\alpha = 0$ and $a\beta = \frac{\bar{\beta}}{a}$, which can be written as $a^2\beta = \bar{\beta}$. By looking at the norm of each side of the equation, we get $a^2 = 1$. Since a is positive, $a = 1$ and thus $\beta = \bar{\beta}$, which means that β is real. Therefore, $\bar{O}_1 O_2^* = \begin{bmatrix} 0 & \beta \\ -\beta & 0 \end{bmatrix}$ with real β, which indicates that

$$\bar{O}_1 = \frac{\beta}{\det O_2} J O_2$$

or equivalently

$$O_1 = \frac{\beta}{\det O_2} J \bar{O}_2$$

since $O_2^{-1} = \frac{O_2^*}{\det O_2}$ by Lemma 6. Since $a = 1$ we have $\det O_1 = \det O_2$. Therefore, the following equations can be obtained.

$$\det O_2 = \det O_1 = \det(\frac{\beta}{\det O_2} J \bar{O}_2) = \left(\frac{\beta}{\det O_2}\right)^2 \overline{\det O_2} = \left(\frac{\beta}{\det O_2}\right)^2 \det O_2$$

Therefore,

$$\frac{\beta}{\det O_2} = \pm 1$$

Thus, $O_1 = \pm J\bar{O}_2$.

Now assume that $O_1 = \pm J\bar{O}_2$ and prove that $\det(S_1 - S_2) = 0$. First assume that O_1 is invertible. If $O_1 = \pm J\bar{O}_2$, we have $O_1^{-1} = (\pm J\bar{O}_2)^{-1} = \pm \bar{O}_2^{-1} \bar{J}^{-1} = \mp \bar{O}_2^{-1} J$. From (17),

$$\begin{aligned} \det(S_1 - S_2) &= \frac{1}{\sqrt{2}} \det O_1 \det(\pm \overline{JO_2} \mp \bar{O}_2 \bar{O}_2^{-1} J O_2) \\ &= \frac{1}{\sqrt{2}} \det O_1 \det(\pm \overline{JO_2} \mp J O_2) \\ &= 0 \end{aligned}$$

Secondly, assume that O_1 is not invertible, that is $\det O_1 = 0$. From Lemma 6, $O_1 = 0$. Therefore, from $O_1 = \pm J\bar{O}_2$, $O_2 = 0$. Thus, $S_1 - S_2 = 0_4$ and $\det(S_1 - S_2) = 0$. 0_4 indicates that 4×4 zero matrix.

Now what left for us to prove is that $O_1 = \pm J\bar{O}_2$ is equivalent to (19). By (16), it is equivalent to

$$U_1V_1 - U_2V_2 = \pm J(\bar{U}_1V_1 - \bar{U}_2V_2),$$

and thus,

$$(U_1 \mp J\bar{U}_1)V_1 = (U_2 \mp J\bar{U}_2)V_2.$$

Using (15), the following series of equations can be obtained.

$$\begin{aligned}
& \frac{1}{2}\left(\begin{bmatrix} e^{2\pi j\frac{k_1}{P}} & e^{2\pi j\frac{l_1}{P}} \\ -e^{-2\pi j\frac{l_1}{P}} & e^{-2\pi j\frac{k_1}{P}} \end{bmatrix} \mp \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} e^{-2\pi j\frac{k_1}{P}} & e^{-2\pi j\frac{l_1}{P}} \\ -e^{2\pi j\frac{l_1}{P}} & e^{2\pi j\frac{k_1}{P}} \end{bmatrix}\right) \begin{bmatrix} e^{2\pi j(\frac{m_1}{Q}+\theta)} & e^{2\pi j(\frac{n_1}{P}+\theta)} \\ -e^{-2\pi j(\frac{n_1}{Q}+\theta)} & e^{-2\pi j(\frac{m_1}{Q}+\theta)} \end{bmatrix} \\
= & \frac{1}{2}\left(\begin{bmatrix} e^{2\pi j\frac{k_2}{P}} & e^{2\pi j\frac{l_2}{P}} \\ -e^{-2\pi j\frac{l_2}{P}} & e^{-2\pi j\frac{k_2}{P}} \end{bmatrix} \mp \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} e^{-2\pi j\frac{k_2}{P}} & e^{-2\pi j\frac{l_2}{P}} \\ -e^{2\pi j\frac{l_2}{P}} & e^{2\pi j\frac{k_2}{P}} \end{bmatrix}\right) \begin{bmatrix} e^{2\pi j(\frac{m_2}{Q}+\theta)} & e^{2\pi j(\frac{n_2}{P}+\theta)} \\ -e^{-2\pi j(\frac{n_2}{Q}+\theta)} & e^{-2\pi j(\frac{m_2}{Q}+\theta)} \end{bmatrix} \\
\Leftrightarrow & \begin{bmatrix} e^{2\pi j(\frac{k_1}{P}+\frac{m_1}{Q}+\theta)} \pm e^{2\pi j(\frac{k_1}{P}-\frac{n_1}{Q}-\theta)} - e^{2\pi j(\frac{l_1}{P}+\frac{m_1}{Q}+\theta)} \pm e^{2\pi j(\frac{l_1}{P}-\frac{n_1}{Q}-\theta)} \\ \mp e^{2\pi j(\frac{k_1}{P}-\frac{m_1}{Q}-\theta)} + e^{2\pi j(\frac{k_1}{P}+\frac{n_1}{Q}+\theta)} + e^{2\pi j(\frac{l_1}{P}-\frac{m_1}{Q}-\theta)} \pm e^{2\pi j(\frac{l_1}{P}+\frac{n_1}{Q}+\theta)} \\ \pm e^{-2\pi j(\frac{k_1}{P}-\frac{m_1}{Q}-\theta)} - e^{-2\pi j(\frac{k_1}{P}+\frac{n_1}{Q}+\theta)} - e^{-2\pi j(\frac{l_1}{P}-\frac{m_1}{Q}-\theta)} \mp e^{-2\pi j(\frac{l_1}{P}+\frac{n_1}{Q}+\theta)} \\ e^{-2\pi j(\frac{k_1}{P}+\frac{m_1}{Q}+\theta)} \pm e^{-2\pi j(\frac{k_1}{P}-\frac{n_1}{Q}-\theta)} - e^{-2\pi j(\frac{l_1}{P}+\frac{m_1}{Q}+\theta)} \pm e^{-2\pi j(\frac{l_1}{P}-\frac{n_1}{Q}-\theta)} \end{bmatrix} \\
= & \begin{bmatrix} e^{2\pi j(\frac{k_2}{P}+\frac{m_2}{Q}+\theta)} \pm e^{2\pi j(\frac{k_2}{P}-\frac{n_2}{Q}-\theta)} - e^{2\pi j(\frac{l_2}{P}+\frac{m_2}{Q}+\theta)} \pm e^{2\pi j(\frac{l_2}{P}-\frac{n_2}{Q}-\theta)} \\ \mp e^{2\pi j(\frac{k_2}{P}-\frac{m_2}{Q}-\theta)} + e^{2\pi j(\frac{k_2}{P}+\frac{n_2}{Q}+\theta)} + e^{2\pi j(\frac{l_2}{P}-\frac{m_2}{Q}-\theta)} \pm e^{2\pi j(\frac{l_2}{P}+\frac{n_2}{Q}+\theta)} \\ \pm e^{-2\pi j(\frac{k_2}{P}-\frac{m_2}{Q}-\theta)} - e^{-2\pi j(\frac{k_2}{P}+\frac{n_2}{Q}+\theta)} - e^{-2\pi j(\frac{l_2}{P}-\frac{m_2}{Q}-\theta)} \mp e^{-2\pi j(\frac{l_2}{P}+\frac{n_2}{Q}+\theta)} \\ e^{-2\pi j(\frac{k_2}{P}+\frac{m_2}{Q}+\theta)} \pm e^{-2\pi j(\frac{k_2}{P}-\frac{n_2}{Q}-\theta)} \pm e^{-2\pi j(\frac{l_2}{P}+\frac{m_2}{Q}+\theta)} - e^{-2\pi j(\frac{l_2}{P}-\frac{n_2}{Q}-\theta)} \end{bmatrix} \\
\Leftrightarrow & \begin{cases} e^{2\theta j}(e^{2\pi j(\frac{k_1}{P}+\frac{m_1}{Q})} - e^{2\pi j(\frac{k_2}{P}+\frac{m_2}{Q})} \pm e^{2\pi j(\frac{l_1}{P}+\frac{m_1}{Q})} \mp e^{2\pi j(\frac{l_2}{P}+\frac{m_2}{Q})}) \\ \quad = \mp e^{2\pi j(\frac{k_1}{P}-\frac{n_1}{Q})} \pm e^{2\pi j(\frac{k_2}{P}-\frac{n_2}{Q})} + e^{2\pi j(\frac{l_1}{P}-\frac{n_1}{Q})} - e^{2\pi j(\frac{l_2}{P}-\frac{n_2}{Q})} \\ e^{2\theta j}(e^{2\pi j(\frac{k_1}{P}+\frac{n_1}{Q})} - e^{2\pi j(\frac{k_2}{P}+\frac{n_2}{Q})} \pm e^{2\pi j(\frac{l_1}{P}+\frac{n_1}{Q})} \mp e^{2\pi j(\frac{l_2}{P}+\frac{n_2}{Q})}) \\ \quad = \pm e^{2\pi j(\frac{k_1}{P}-\frac{m_1}{Q})} \mp e^{2\pi j(\frac{k_2}{P}-\frac{m_2}{Q})} - e^{2\pi j(\frac{l_1}{P}-\frac{m_1}{Q})} + e^{2\pi j(\frac{l_2}{P}-\frac{m_2}{Q})} \end{cases}
\end{aligned}$$

which is equivalent to (19). □

B Proof of Lemma 8

Lemma 8: *For any* 4 *points on the unit circle that add up (as complex numbers) to zero, it is always true that two of them add up to zero. (Clearly, the other two must also have a summation of zero.)*

Proof: Assume $e^{j\theta_1} + e^{j\theta_2} + e^{j\theta_3} + e^{j\theta_4} = 0$, then the following series of equations are true.

$$\begin{aligned}
& e^{j\frac{\theta_1+\theta_2}{2}}(e^{j\frac{\theta_1-\theta_2}{2}} + e^{-j\frac{\theta_1-\theta_2}{2}}) + e^{j\frac{\theta_3+\theta_4}{2}}(e^{j\frac{\theta_3-\theta_4}{2}} + e^{-j\frac{\theta_3-\theta_4}{2}}) = 0 \\
\Rightarrow \quad & 2\cos(\frac{\theta_1-\theta_2}{2})e^{j\frac{\theta_1+\theta_2}{2}} = -2\cos(\frac{\theta_3-\theta_4}{2})e^{j\frac{\theta_3+\theta_4}{2}} \\
\Rightarrow \quad & \begin{cases} \frac{\theta_1-\theta_2}{2} = \pm\frac{\theta_3-\theta_4}{2} + 2k\pi \\ \frac{\theta_1+\theta_2}{2} = \frac{\theta_3+\theta_4}{2} + (2l+1)\pi \end{cases} \quad \text{or} \quad \begin{cases} \frac{\theta_1-\theta_2}{2} = \pm\frac{\theta_3-\theta_4}{2} + (2k+1)\pi \\ \frac{\theta_1+\theta_2}{2} = \frac{\theta_3+\theta_4}{2} + 2l\pi \end{cases}
\end{aligned}$$

for some integers k and l. Without loss of generality, we only consider the first case. By adding the two equations, we have

$$\begin{cases} \theta_1 = \theta_3 + (2k+2l+1)\pi \\ \theta_2 = \theta_4 + (-2k+2l+1)\pi \end{cases} \Rightarrow \begin{cases} e^{j\theta_1} + e^{j\theta_3} = 0 \\ e^{j\theta_2} + e^{j\theta_4} = 0 \end{cases}$$

when plus sign is applied, or

$$\begin{cases} \theta_1 = \theta_4 + (2k+2l+1)\pi \\ \theta_2 = \theta_3 + (-2k+2l+1)\pi \end{cases} \Rightarrow \begin{cases} e^{j\theta_1} + e^{j\theta_4} = 0 \\ e^{j\theta_2} + e^{j\theta_3} = 0 \end{cases}$$

when minus sign is applied.

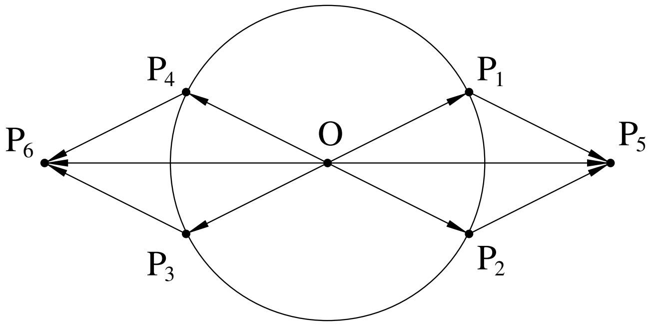

Figure 9: Figure for Lemma 8

This lemma can also be proved easily in a geometric way. As in Fig. 9, P_1, P_2, P_3, P_4 are the 4 points on the unit circle that add up to zero and O is the center of the unit circle, which is the origin. Without loss of generality, assume P_2 is the point that is closest to P_1. Since the 4 points add up to zero, $\overrightarrow{OP}_5$, which is the summation of $\overrightarrow{OP_1}$ and $\overrightarrow{OP}_2$, and $\overrightarrow{OP}_6$, which is the summation of $\overrightarrow{OP}_3$ and $\overrightarrow{OP}_4$, are on the same line with inverse directions and have the same length. Since $\overline{OP_1} = \overline{OP_3}$, $\overline{P_1P_5} = \overline{OP_2} = \overline{OP_4} = \overline{P_3P_6}$, $\triangle OP_1P_5 = \triangle OP_3P_6$. Therefore, $\angle P_1OP_5 = \angle P_3OP_6$. Thus, $\overrightarrow{OP}_1$ and $\overrightarrow{OP}_3$ are on the same line with inverse directions, which means that $\overrightarrow{OP}_1 + \overrightarrow{OP}_3 = 0$. $\square$

C Proof of Lemma 9

Lemma 9: *If P and Q are relatively prime, then for any integers $0 \le k_1, l_1, k_2, l_2 < P$ and $0 \le m_1, n_1, m_2, n_2 < Q$ such that $(k_1, l_1, m_1, n_1) \neq (k_2, l_2, m_2, n_2)$, the w^+, x^+, y^+, z^+ defined in (20) and the w^-, x^-, y^-, z^- defined in (21) cannot be all zeros.*

Proof: Here, we only prove that w^+, x^+, y^+, z^+ as defined in (20) cannot be all zeros. The proof of the other part (prove that w^-, x^-, y^-, z^- as defined in (21) cannot be all zeros) is very similar to it. We prove it by contradiction. Assume that P and Q are relatively prime and there exist integers k_1, l_1, k_2, l_2 in $[0, P-1]$ and m_1, n_1, m_2, n_2 in $[0, Q-1]$ such that $w^+ = x^+ = y^+ = z^+ = 0$. Since $x^+ = 0$, by Lemma 8, we have

$$\begin{cases} e^{2\pi j(\frac{k_1}{P}-\frac{n_1}{Q})} = e^{2\pi j(\frac{l_1}{P}-\frac{n_1}{Q})} \\ e^{2\pi j(\frac{k_2}{P}-\frac{n_2}{Q})} = e^{2\pi j(\frac{l_2}{P}-\frac{n_2}{Q})} \end{cases} \text{or} \begin{cases} e^{2\pi j(\frac{k_1}{P}-\frac{n_1}{Q})} = e^{2\pi j(\frac{k_2}{P}-\frac{n_2}{Q})} \\ e^{2\pi j(\frac{l_1}{P}-\frac{n_1}{Q})} = e^{2\pi j(\frac{l_2}{P}-\frac{n_2}{Q})} \end{cases} \text{or} \begin{cases} e^{2\pi j(\frac{k_1}{P}-\frac{n_1}{Q})} = -e^{2\pi j(\frac{l_2}{P}-\frac{n_2}{Q})} \\ e^{2\pi j(\frac{k_2}{P}-\frac{n_2}{Q})} = -e^{2\pi j(\frac{l_1}{P}-\frac{n_1}{Q})} \end{cases}$$

Without loss of generality and to simplify the proof, we only discuss the first case here. From the first set of equations, there exist integers i_1 and i_2 such that

$$\begin{cases} 2\pi(\frac{k_1}{P}-\frac{n_1}{Q}) = 2\pi(\frac{l_1}{P}-\frac{n_1}{Q}) + 2\pi i_1 \\ 2\pi(\frac{k_2}{P}-\frac{n_2}{Q}) = 2\pi(\frac{l_2}{P}-\frac{n_2}{Q}) + 2\pi i_2 \end{cases} \Rightarrow \begin{cases} \frac{k_1-l_1}{P} = i_1 \\ \frac{k_2-l_2}{P} = i_2 \end{cases}$$

Since $k_1, l_1, l_1, l_2 \in [0, P)$, $k_1 - l_1, k_2 - l_2 \in (-P, P)$. Therefore, $i_1 = i_2 = 0$, from which $k_1 = l_1$ and $k_2 = l_2$ are obtained. Using this result and $w^+ = 0$, we have $2e^{2\pi j(\frac{k_1}{P}+\frac{m_1}{Q})} = 2e^{2\pi j(\frac{k_2}{P}+\frac{m_2}{Q})}$, from which $\frac{k_1}{P} + \frac{m_1}{Q} = \frac{k_2}{P} + \frac{m_2}{Q} + i$ can be proved for some integer i. The equation is equivalent to $\frac{k_1-k_2}{P} = i - \frac{m_1-m_2}{Q}$. Since P and Q are relatively prime, we have $(k_1 - k_2)|P$ and $(m_1 - m_2)|Q$. Since $k_1, k_2 \in [0, P)$ and $m_1, m_2 \in [0, Q)$, $k_1 - k_2 \in (-P, P)$ and $m_1 - m_2 \in (-Q, Q)$. Then $k_1 - k_2 = 0$ and $m_1 - m_2 = 0$ which, combined with (C.1), means that $k_1 = k_2 = l_1 = l_2$ and $m_1 = m_2$. From $y^+ = 0$ and $e^{2\pi j\frac{k_1}{P}} \neq 0$, we have $2e^{2\pi j\frac{n_1}{Q}} = 2e^{2\pi j\frac{n_2}{Q}}$, which, similarly, leads to $\frac{n_1}{Q} = \frac{n_1}{Q} + i \Rightarrow n_1 - n_2 = iQ$, for some integer i. Since $n_1, n_2 \in [0, Q)$, $n_1 - n_2 \in (-Q, Q)$. Therefore, $i = 0$, that is $n_1 = n_2$. Therefore, $(k_1, l_1, m_1, n_1) = (k_2, l_2, m_2, n_2)$, which contradicts the condition of the lemma. □

D Proof of Lemma 10

Lemma 10: *For any non-identical pair $(k_i, l_i, m_i, n_i, \gamma_i)$ and $(k_j, l_j, m_j, n_j, \gamma_j)$ where $k_i, l_i, k_j, l_j \in [0, P)$, $m_i, n_i, m_j, n_j \in [0, Q)$ are integers and $\gamma_i, \gamma_j \in \Gamma$, if P and Q are relatively prime with P odd and also if the set Γ satisfies the following conditions,*

1. $\Gamma \subset (0, \frac{\pi}{2})$.
2. *For any $\gamma \in \Gamma$,*

$$\tan\gamma \neq \pm\frac{\sin\pi(\frac{l}{P}+\frac{m}{Q})}{\sin\pi(\frac{k}{P}+\frac{m}{Q})} \tag{D.1}$$

where $k, l \in (-P, P)$, $m \in (-Q, Q)$ are integers and $(k, m) \neq (0, 0)$.

3. For any two different $\gamma_i, \gamma_j \in \Gamma$,

$$\frac{\sin 2\gamma_i}{\sin 2\gamma_j} \neq \frac{\cos 2\pi \frac{s}{P}}{\cos 2\pi \frac{l}{P}}, \tag{D.2}$$

where $s, l \in [0, P)$ are integers.

then the $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ as defined in (33) cannot be zeros simultaneously. Also the $\tilde{w}^-, \tilde{x}^-, \tilde{y}^-, \tilde{z}^-$ as defined in (34) cannot be zeros simultaneously.

Proof: We prove this lemma by contradiction. We only prove that the $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ cannot be zeros simultaneously here since the proving that the $\tilde{w}^-, \tilde{x}^-, \tilde{y}^-, \tilde{z}^-$ cannot be zeros simultaneously is very similar.

Assume $\tilde{w}^+ = \tilde{x}^+ = \tilde{y}^+ = \tilde{z}^+ = 0$. From the definition of $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ in (33), we have

$$\begin{aligned}
e^{2\pi i \frac{m_i}{Q}}(\cos\gamma_i e^{2\pi i \frac{k_i}{P}} + \sin\gamma_i e^{2\pi i \frac{l_i}{P}}) &= e^{2\pi i \frac{m_j}{Q}}(\cos\gamma_j e^{2\pi i \frac{k_j}{P}} + \sin\gamma_j e^{2\pi i \frac{l_j}{P}}) && \text{(D.3)}\\
e^{-2\pi i \frac{n_i}{Q}}(\sin\gamma_i e^{2\pi i \frac{k_i}{P}} - \cos\gamma_i e^{2\pi i \frac{l_i}{P}}) &= e^{-2\pi i \frac{n_j}{Q}}(\sin\gamma_j e^{2\pi i \frac{k_j}{P}} - \cos\gamma_j e^{2\pi i \frac{l_j}{P}}) && \text{(D.4)}\\
e^{2\pi i \frac{n_i}{Q}}(\cos\gamma_i e^{2\pi i \frac{k_i}{P}} + \sin\gamma_i e^{2\pi i \frac{l_i}{P}}) &= e^{2\pi i \frac{n_j}{Q}}(\cos\gamma_j e^{2\pi i \frac{k_j}{P}} + \sin\gamma_j e^{2\pi i \frac{l_j}{P}}) && \text{(D.5)}\\
e^{-2\pi i \frac{m_i}{Q}}(\sin\gamma_i e^{2\pi i \frac{k_i}{P}} - \cos\gamma_i e^{2\pi i \frac{l_i}{P}}) &= e^{-2\pi i \frac{m_j}{Q}}(\sin\gamma_j e^{2\pi i \frac{k_j}{P}} - \cos\gamma_j e^{2\pi i \frac{l_j}{P}}) && \text{(D.6)}
\end{aligned}$$

The square of the norm of the left hand side of (D.3) equals to

$$\begin{aligned}
&\left|\cos\gamma_i e^{2\pi i \frac{k_i}{P}} + \sin\gamma_i e^{2\pi i \frac{l_i}{P}}\right|^2\\
= \;& (\cos\gamma_i \cos 2\pi \frac{k_i}{P} + \sin\gamma_i \cos 2\pi \frac{l_i}{P})^2 + (\cos\gamma_i \sin 2\pi \frac{k_i}{P} + \sin\gamma_i \sin 2\pi \frac{l_i}{P})^2\\
= \;& \cos^2\gamma_i \cos^2 2\pi \frac{k_i}{P} + \sin^2\gamma_i \cos^2 2\pi \frac{l_i}{P} + 2\cos\gamma_i \cos 2\pi \frac{k_i}{P} \sin\gamma_i \cos 2\pi \frac{l_i}{P} +\\
& \cos^2\gamma_i \sin^2 2\pi \frac{k_i}{P} + \sin^2\gamma_i \sin^2 2\pi \frac{l_i}{P} + 2\cos\gamma_i \sin 2\pi \frac{k_i}{P} \sin\gamma_i \sin 2\pi \frac{l_i}{P}\\
= \;& \cos^2 2\pi \frac{k_i}{P} + \sin^2 2\pi \frac{k_i}{P} + 2\cos\gamma_i \sin\gamma_i (\cos 2\pi \frac{k_i}{P} \cos 2\pi \frac{l_i}{P} + \sin 2\pi \frac{k_i}{P} \sin 2\pi \frac{l_i}{P})\\
= \;& 1 + \sin 2\gamma_i \cos 2\pi \frac{k_i - l_i}{P}
\end{aligned}$$

Similarly, the square of the norm of the right hand side of (D.3) equals to

$$1 + \sin 2\gamma_j \cos 2\pi \frac{k_j - l_j}{P}$$

Comparing the norms of both side of (D.3), we have that

$$1 + \sin 2\gamma_i \cos 2\pi \frac{k_i - l_i}{P} = 1 + \sin 2\gamma_j \cos 2\pi \frac{k_j - l_j}{P}$$

which is equivalent to

$$\sin 2\gamma_i \cos 2\pi \frac{k_i - l_i}{P} = \sin 2\gamma_j \cos 2\pi \frac{k_j - l_j}{P} \tag{D.7}$$

Define $r = |k_i - l_i|$ and $s = |k_j - l_j|$. Since $k_i, l_i, k_j, l_j \in [0, P)$, $k_i - l_i, k_j - l_j \in (-P, P)$. Therefore, $r, s \in [0, P)$ and $\cos 2\pi\frac{r}{P} = \cos 2\pi\frac{k_i - l_i}{P}$ and $\cos 2\pi\frac{s}{P} = \cos 2\pi\frac{k_j - l_j}{P}$. Thus,

$$\sin 2\gamma_i \cos 2\pi\frac{r}{P} = \sin 2\gamma_j \cos 2\pi\frac{s}{P}$$

From the first condition, $\sin\gamma \neq 0$ and $\cos\gamma \neq 0$ for any $\gamma \in \Gamma$. Therefore, when $\gamma_i \neq \gamma_j$, this contradictions (D.2). Therefore, when $\gamma_i \neq \gamma_j$ and (D.2) is satisfied, the $\tilde{w}^+, \tilde{x}^+, \tilde{y}^+, \tilde{z}^+$ cannot be zeros simultaneously.

Now look at the case of $\gamma_i = \gamma_j$. From (D.3), we have

$$\begin{aligned} \cos\gamma_i(e^{2\pi i(\frac{k_i}{P}+\frac{m_i}{Q})} - e^{2\pi i(\frac{k_j}{P}+\frac{m_j}{Q})}) &= -\sin\gamma_i(e^{2\pi i(\frac{l_i}{P}+\frac{m_i}{Q})} - e^{2\pi i(\frac{l_j}{P}+\frac{m_j}{Q})}) \\ 2i\cos\gamma_i \sin 2\pi(\frac{k_i - k_j}{2P} + \frac{m_i - m_j}{2Q})e^{2\pi i(\frac{k_i+k_j}{2P}+\frac{m_i+m_j}{Q})} &= -2\sin\gamma_i \sin 2\pi(\frac{l_i - l_j}{2P} + \frac{m_i - m_j}{2Q})e^{2\pi i(\frac{l_i+l_j}{2P}+\frac{m_i+m_j}{2Q})}. \end{aligned}$$

Therefore,

$$\cos\gamma_i \sin 2\pi(\frac{k_i - k_j}{2P} + \frac{m_i - m_j}{2Q}) = \pm\sin\gamma_i \sin 2\pi(\frac{l_i - l_j}{2P} + \frac{m_i - m_j}{2Q}) \tag{D.8}$$

From (D.1), we have

$$\cos\gamma_i \sin 2\pi(\frac{k_i - k_j}{2P} + \frac{m_i - m_j}{2Q}) \neq \pm\sin\gamma_i \sin 2\pi(\frac{l_i - l_j}{2P} + \frac{m_i - m_j}{2Q})$$

when $k_i - k_j, l_i - l_j \in (-P, P), m_i - m_j \in (-Q, Q)$ and $(k_i - k_j, m_i - m_j) \neq (0, 0)$. Therefore, for (D.8) to be true, $k_i - k_j = 0$ and $m_i - m_j = 0$. Thus, $\sin 2\pi(\frac{k_i - k_j}{2P} + \frac{m_i - m_j}{2Q}) = 0$. Sine $\sin\gamma_i \neq 0$, $\sin 2\pi(\frac{l_i - l_j}{2P} + \frac{m_i - m_j}{2Q}) = \sin\pi\frac{l_i - l_j}{2P} = 0$. Therefore, $l_i - l_j = 0$. So we get

$$\begin{cases} k_i = k_j \\ l_i = l_j \\ m_i = m_j \end{cases}$$

Similarly, from (D.4), we have

$$\begin{aligned} \sin\gamma_i(e^{2\pi i(\frac{k_i}{P}-\frac{n_i}{Q})} - e^{2\pi i(\frac{k_j}{P}-\frac{n_j}{Q})}) &= \cos\gamma_i(e^{2\pi i(\frac{l_i}{P}-\frac{n_i}{Q})} - e^{2\pi i(\frac{l_j}{P}-\frac{n_j}{Q})}) \\ 2i\sin\gamma_i \sin 2\pi(\frac{k_i - k_j}{2P} - \frac{n_i - n_j}{2Q})e^{2\pi i(\frac{k_i+k_j}{2P}-\frac{n_i+n_j}{Q})} &= 2i\cos\gamma_i \sin 2\pi(\frac{l_i - l_j}{2P} - \frac{n_i - n_j}{2Q})e^{2\pi i(\frac{l_i+l_j}{2P}-\frac{n_i+n_j}{2Q})}. \end{aligned}$$

Therefore,

$$\sin\gamma_i \sin 2\pi(\frac{k_i - k_j}{2P} - \frac{n_i - n_j}{2Q}) = \pm\cos\gamma_i \sin 2\pi(\frac{l_i - l_j}{2P} - \frac{n_i - n_j}{2Q}) \tag{D.9}$$

By similar argument,

$$\begin{cases} k_i = k_j \\ l_i = l_j \\ n_i = n_j \end{cases}$$

Therefore, $(k_i, l_i, m_i, n_i, \gamma_i) = (k_j, l_j, m_j, n_j, \gamma_j)$ and this contradicts the condition that they are different.

$\square$

Acknowledgment

This work was supported by an AFOSR for MURI "Mathematical Infrastructure for Robust Virtual Engineering," URI "Protecting Infrastructures from Themselves" and Caltech's Lee Center for Advanced Networking.

References

[1] B. Hochwald and W. Sweldens, "Differential unitary space time modulation," *IEEE Trans. Comm.*, vol. 48, pp. 2041–2052, Dec. 2000. Download available at `http://mars.bell-labs.com`.

[2] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *J. Sel. Area Comm.*, pp. 1169–1174, July 2000.

[3] B. Hughes, "Differential space-time modulation," *IEEE Trans. Info. Theory*, pp. 2567–2578, Nov. 2000.

[4] A. Shokrollahi, B. Hassibi, B. Hochwald, and W. Sweldens, "Representation theory for high-rate multiple-antenna code design," *IEEE Trans. Info. Theory*, vol. 47, pp. 2335–2367, Sep. 2001.

[5] B. Hassibi and M. Khorrami, "Fully-diverse multi-antenna constellations and fixed-point-free Lie groups," *submitted to IEEE Trans. Info. Theory*, 2000. Download available at `http://mars.bell-labs.com`.

[6] B. Hughes, "Optimal space-time constellations from groups," *submitted to IEEE Trans. Info. Theory*, 2000.

[7] Sattinger and Weaver, *Lie Groups and Algebras with applications to physics, geometry, and mechanics*. Springer-Verlag, 1986.

[8] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Trans. Info. Theory*, vol. 45, pp. 139–157, 1999.

[9] B. M. Hochwald and T. L. Marzetta, "Unitary space-time modulation for multiple-antenna communication in Rayleigh flat-fading," *IEEE Trans. Info. Theory*, vol. 46, pp. 543–564, Mar. 2000.

[10] B. Hassibi and B. Hochwald, "Cayley differential unitary spce-time codes," *IEEE Trans. Info. Theory*, pp. 1458–1503, 2002.

[11] Brocker and Dieck, *Representations of compact Lie Groups*. Springer-Verlag, 1995.

[12] J.-P. Serre, *Lie Algebras and Lie Groups*. Springer-Verlag, 1992.

[13] A. Shokrollahi, "Design of unitary space-time codes from representations of $su(2)$," *Preprint*, 2000.

[14] B. Simon, *Representations of finite and compact Lie Groups*. Providence, R.I. : American Mathematical Society, 1994.

[15] S. M. Alamouti, "A simple transmitter diversity scheme for wireless communications," *IEEE J. Sel. Area Comm.*, pp. 1451–1458, Oct. 1998.

[16] M. O. Damen, K. Abed-Meraim, and M. S. Lemdani, "Further results on the sphere decoder algorithm," *Submitted to IEEE Trans. Info. Theory*, 2000.

[17] B. Hassibi and H. Vikalo, "On the expected complexity of integer least-squares problems," in *Acoustics, Speech, and Signal Processing, 2002 IEEE International Conference*, April 2002.

[18] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time block codes for complex signal constellations," *IEEE Trans. Info. Theory*, pp. 384–395, 2002.